STEG WORKING PAPER

TECHNOLOGY ADOPTION AND LATE INDUSTRIALIZATION

Jaedo Choi and Younghun Shim

OCTOBER 2022
STEG WP033

Technology Adoption and Late Industrialization *

Jaedo Choi
University of Michigan, Ann Arbor

Younghun Shim
University of Chicago

October 2022

Abstract

We study how the adoption of foreign technology and local spillovers from such adoption contributed to late industrialization in a developing country during the postwar period. Using novel historical firm-level data for South Korea, we provide three empirical findings: (i) direct productivity gains to adopters, (ii) local productivity spillovers of the adoption, and (iii) complementarity in firms' adoption decisions. Based on these findings, we develop a dynamic spatial model with firms' technology adoption decisions and local spillovers. The spillovers induce dynamic complementarity in firms' technology adoption decisions. Because of this complementarity, the model potentially features multiple steady states. Temporary adoption subsidies can have permanent effects by moving an economy to a new transition path that converges to a higher-productivity steady state. We calibrate our model to the microdata and econometric estimates. We evaluate the effects of the South Korean government policy that temporarily provided adoption subsidies to heavy manufacturing firms in the 1970s. Had no adoption subsidies been provided, South Korea would have converged to a less industrialized steady state in which the heavy manufacturing sector's share of GDP would have been 15 percentage points lower and aggregate welfare would have been 10% lower compared to the steady state with successful industrialization. Thus, temporary subsidies for technology adoption had permanent effects.

Keywords: Technology adoption, industrialization, knowledge spillover, path dependence, big push
JEL Codes: O14, O33, O53, R12

*We are grateful to Andrei Levchenko, Dominick Bartelme, Jagadeesh Sivadasan, and Sebastian Sotelo for their guidance and invaluable suggestions. We also thank Costas Arkolakis, Maria Aristizabal-Ramírez, David Baqaee, Natalie Bau, Barthélémy Bonadio, Victor Couture, Javier Cravino, Alan Deardorff, Aaron Flaeen, Elisa Giannone, Kyle Handley, Ricardo Hausmann, Oleg Itskhoki, Yongjun Kim, Samuel Kortum, Mos Laoprapassorn, Kevin Lim, Kiminori Matsuyama, Peter Morrow, Nathan Nunn, Nishaad Rao, Ulrich Schetter, Michael Song, Linda Tesar, Daniel Trefler, Nick Tsivanidis, Jonathan Vogel, Xiaodong Zhu, and seminar participants at Berkeley Haas, CUHK, ETSG 2021, EWMES 2021, Federal Reserve Board, Harvard Growth Lab, HKU, HKUST, KDI, KIEP, KIF, KIPF, Michigan, MWIT 2021, NTU, PSU (New Faces in International Economics), Rochester, Toronto, UBC, UCLA, VEAMS, Yale, and Yonsei for helpful comments and discussions. We also thank the National Archives of Korea for sharing data. We gratefully acknowledge research support from the Kwanjeong Educational Foundation. E-mail: jaedohi@umich.edu.

1 Introduction

Large differences in cross-country total factor productivity (TFP) suggest that technology is fundamental to economic development.[1] Based on this observation, many economists and policy-makers have argued that the adoption of advanced technology that rich countries use can make poor countries richer (Parente and Prescott, 2002). Technology adoption can be an even more powerful driving force for economic development if and when technology is at least partially non-rival, and knowledge gained from adopting foreign technology can be spread to other local firms.[2]

In the postwar period, patterns of industrialization among developing countries diverged. The economic base of some developing countries such as South Korea, Taiwan, and Turkey transformed from agriculture to manufacturing, while the economies of many other developing countries remained stagnant. The countries whose base changed to manufacturing achieved industrialization by adopting foreign technology rather than developing their own technology.[3] Their adoption-driven industrialization is known as late industrialization, which differs from the earlier industrialization driven by invention or innovation in the Western countries (Amsden, 1989).[4] A look at what drove the rapid industrialization of these latecomers provides suggestive evidence about the potential importance of technology adoption for economic development. However, little is empirically and quantitatively known about the role of adoption due to the unavailability of detailed data about firms' adoption activities in countries that experienced late industrialization. The key challenge is that technology adoption is typically not observed directly but must be inferred from other equilibrium outcomes.

This paper answers the following question: How do the adoption of foreign technology and its local spillovers contribute to late industrialization? We study South Korea's transition toward heavy manufacturing sectors in the 1970s. South Korea is known for having the most successful and rapid industrialization among the latecomers.[5]

This paper makes three contributions. First, we overcome the empirical challenge in the literature by constructing a novel historical dataset that covers the universe of technology adoption contracts between South Korean and foreign firms. Most of the adopted technology during this period was related to knowledge about how to build and operate plants and capital equipment related to mass production. Using this dataset, we can measure firm-level technology adoption directly at the micro

[1] See Klenow and Rodríguez-Clare (1997) and Hall and Jones (1999).

[2] See Romer (1990). Recent studies provide empirical evidence about the existence of knowledge spillovers and find that knowledge spillovers tend to be highly localized (e.g., Jaffe et al., 1993; Kerr and Kominers, 2015; Kantor and Whalley, 2019; Moretti, 2021).

[3] "If industrialization first occurred in England on the basis of invention, and if it occurred in Germany and the United States on the basis of innovation, then it occurs now among "backward" countries on the basis of learning" (Amsden, 1989, p. 4). "Once South Korea reduced its barriers, thereby greatly increasing its TFP, it experienced a development miracle as it used more of the stock of available knowledge" (Parente and Prescott, 2002, p. 4).

[4] Building on Gerschenkron's (1952) insights on economic backwardness, Amsden (1989) defines late industrialization as the third wave of industrialization that occurred in a subset of developing countries in the twentieth century based on the adoption of foreign technology.

[5] See Lucas (1993).

level.

Second, using this novel dataset, we provide three empirical evidence on the firm-level effects of technology adoption. We provide the empirical evidence on the direct productivity gains using a winners vs. losers research design following Greenstone et al. (2010). An empirical challenge related to identifying the direct gains is the fact that firms make adoption decisions endogenously, which leads to the standard selection problem. We deal with this problem by comparing firms that successfully adopted technology and firms that received the approval from the government to pursue foreign technology and made a contract with a foreign firm but failed to adopt technology because the foreign firm canceled the contract due to circumstances unrelated to the South Korean firm. The first group of firms are the winners (the treated) in our winners vs. losers research design. The second group are the losers (the control). We construct pairs of winners and losers by matching each loser to a winner that is observationally similar and compare outcomes between these two groups. The identifying assumption is that the losers form a valid counterfactual for matched winners conditional on matched observables. We collect data about cancellations from historical contract documents. Our estimates imply that technology adoption increased adopters' sales and revenue total factor productivity by 40–50%.

Our second empirical finding is local productivity spillovers of the adoption. The key identification challenge when estimating the spillovers is that spatially correlated shocks affect both firms' performance and their neighbors' adoption decisions (Manski, 1993). We deal with this challenge by exploiting spatial variation at a fine level of geographic detail. The median land area of our geographic unit of analysis is the size of Manhattan, or almost 34 square miles. Within each region and sector, we construct a spillover measure for each firm as the weighted average of local firms' ever adoption status of the same sector where the weight is given by the inverse of distance to other firms. This measure varies at the firm-level within each region and sector depending on firms' geographical proximity to adopters. We then regress firms' sales while controlling for time-varying region-sector fixed effects. Because we control for these fixed effects, our results are driven by variation in distances to adopters of the same sector within regions instead of being driven by variation across regions and sectors, so the usual regional or sectoral unobservables are not a concern in our empirical analysis. We find that firms' sales grew faster when more neighboring firms had adopted foreign technology. Our estimates indicate that when the spillover measure increased by a one standard deviation, firms' sales increased by 10–15%.

We also provide empirical evidence consistent with complementarity in firms' technology adoption decisions, where firms' gains from adopting a new technology become larger when more neighboring firms have adopted technologies. We regress a dummy variable of adoption of a new technology on the spillover measure with the same set of granular fixed effects and controls as in the local spillover regression. We find that firms were more likely to adopt a new technology when more neighboring firms had adopted foreign technology. Our estimates imply that one standard deviation increase of

the spillover measure increased firms' probability of adopting a new technology by 1–1.5%.

Third, we construct a dynamic spatial general equilibrium model with heterogeneous firms' technology adoption decisions and local productivity spillovers. We use the model to evaluate the general equilibrium effect of the South Korean government policy that temporarily subsidized technology adoption by heavy manufacturing firms. Firms' adoption decisions and the spillover endogenously shape comparative advantage and export patterns at the regional and national levels. Firms can adopt a more productive modern technology after incurring a fixed adoption cost. The spillover operates with a one-period lag, where the current local productivity increases in the local share of adopters in the previous period. This time lag of the spillover is a source of dynamics in the model. Because of this time lag, the share of adopters becomes a time-varying state variable. The spillover generates dynamic complementarity in firms' adoption decisions. A higher share of adopters in the previous period leads to higher gains from adoption that in turn induces more firms to adopt technology in the current period. Because adopters do not internalize this spillover, the amount of adoption is suboptimal. This justifies appropriate policy interventions that promote adoption.

In a simplified model, we show analytically that dynamic complementarity can lead to multiple steady states. When multiple steady states exist, they can be Pareto-ranked based on the equilibrium share of adopters. We label the steady states with low and high shares of adopters pre-industrialized and industrialized, respectively. In this model, an initial condition determines which steady state is realized in the long run. If an economy begins with a sufficiently large share of adopters, it converges to the industrialized steady state, but if not, it converges to the pre-industrialized steady state. This is because when an economy begins with a sufficiently large share of adopters, dynamic complementarity induces more firms to adopt technology, which in turn magnifies the strength of the complementarity in subsequent periods and vice versa. A temporary adoption subsidy can have permanent effects by moving an economy that was converging to the pre-industrialized steady state to a new transition path that converges to the industrialized steady state.

We calibrate the model to both micro and regional data. The model delivers structural equations that can be mapped to our reduced-form regression specifications. Thus, we can use the reduced-form estimates to identify two parameters that govern the direct productivity gains and the spillover. Subsidies are modeled as input subsidies. We do not observe the subsidies directly, but the model delivers an identifying moment for the subsidies: increases in shares of adopters during the periods when subsidies were available relative to the initial period when the subsidies were not provided. We show that this moment uniquely identifies the input subsidy under simplifying assumptions. The intuition behind this moment is that given information on the direct and spillover gains from adoption identified by our reduced-form estimates, the relative increases in shares of adopters are attributable to a reduction in adoption costs induced by the subsidies. We estimate the subsidy rate by fitting this moment. Finally, we identify a fixed adoption cost by the shares of adopters in the initial period when the subsidies were not provided.

Using the calibrated model, we ask how the pattern of industrialization in South Korea would have evolved had the government not provided subsidies. Our results show that if subsidies had not been provided, South Korea would have converged to a less industrialized steady state. In the steady state of this counterfactual economy, the heavy manufacturing sector's share of GDP would have decreased by a 15 percentage point lower, exports would have been 22.5 percentage points lower, and employment would have been 3 percentage points lower than the steady state of the baseline economy where subsidies had been provided. Also, the aggregate welfare would have been 10% lower. The aggregate differences are driven by a few regions that become more productive because of subsidy-induced technology adoption.

Related Literature. Our paper contributes to four strands of the literature. The first is the empirical literature that studies firm-level effects of industrial technology adoption in developing countries (e.g., Atkin et al., 2017; Juhász, 2018; Giorcelli and Li, 2021; Juhász et al., 2020; de Souza, 2021; Hardy and McCasland, 2021). Credible empirical evidence on firm-level effects of industrial technology in developing countries is scarce. We contribute to this literature by providing new empirical evidence on the direct productivity gains to adopters.

Second, this paper contributes to the empirical literature on local knowledge spillovers (see, among many others, Jaffe et al., 1993; Keller, 2002; Arzaghi and Henderson, 2008; Greenstone et al., 2010; Bloom et al., 2013; Kerr and Kominers, 2015; Kantor and Whalley, 2019; Moretti, 2021). While previous papers have focused on the local spillovers of R&D or innovation activities in developed countries, we provide new empirical evidence on local productivity spillovers of technology adoption in a developing country context and show that it was an important driving factor behind industrialization in South Korea.

Third, we contribute to the quantitative literature on multiple equilibria and the big push. According to the big push literature that dates to Rosenstein-Rodan (1943) and Hirschman (1958), underdevelopment results from complementarity and coordination failures (e.g., Murphy et al., 1989; Redding, 1996; Rodríguez-Clare, 1996; Ciccone, 2002; Kline and Moretti, 2014). We contribute to this literature by quantifying the aggregate consequences of coordination failure in firms' technology adoption decisions, multiple equilibria induced by this failure, and effects of the temporary subsidies provided by the South Korean government. While Crouzet et al. (2020) studied complementarity in technology adoption decisions of firms caused by network externalities and Buera et al. (2021) studied complementarity caused by higher intermediate intensities of the adoption goods, we study the local productivity spillovers of the adoption. The modeling framework of our paper is most closely related to that of Allen and Donaldson (2020) who study the role of history in determining spatial distribution of economic activity. Technology adoption choices are also determined by history in our model. Unlike the macroeconomic literature on barriers to technology adoption (e.g., Parente and Prescott, 1994; Comin and Hobijn, 2010; Cole et al., 2016), we study the coordination failure.

Finally, this paper contributes to the trade literature on the evolution of comparative advantage.

Aggregate data show that comparative advantage evolves (Cai et al., 2022; Hausmann and Klinger, 2007; Hanson et al., 2015; Levchenko and Zhang, 2016; Schetter, 2019; Atkin et al., 2021), but the understanding of what drives this evolution has been limited so far. Using detailed microdata, Pellegrina and Sotelo (2021) document how knowledge diffusion through migration shaped the comparative advantage of Brazil, and Arkolakis et al. (2019) study the role immigrants played in diffusing knowledge in the United States in the nineteenth century. We contribute to this literature by quantifying how technology adoption shaped South Korea's comparative advantage in heavy manufacturing sectors using the novel data.

The rest of this paper is organized as follows. Section 2 describes the data we used for our empirical and quantitative analysis. Section 3 describes the historical background of South Korea's late industrialization and the South Korean government policy that promoted technology adoption. Section 4 presents reduced-form evidence on direct productivity gains to adopters, local productivity spillovers, and complementarity in firms' technology adoption decisions. In Section 5, we build the quantitative model. Section 6 describes how the model can be mapped to the data and reduced-form estimates. Section 7 presents quantitative analysis of the South Korean government policy. Section 8 concludes the paper.

2 Data

We construct our main dataset by merging firm balance sheet data with data on firms' technology adoption activities. We link these two datasets based on firms' names. The resulting dataset includes only firms in the manufacturing sectors. We classified firms into 10 manufacturing sectors, 4 of which are heavy manufacturing. The sample period of the constructed dataset is 1970 to 1982. The final dataset has 7,223 unique firms of which 49% are heavy manufacturing.

The final dataset includes 1,698 contracts made by 628 unique firms. Of these, 1,361 contracts and 457 firms were in heavy manufacturing sectors. Most of the adopted technologies were related to know-how about how to install or operate capital equipment or turnkey plants.[6] Firm balance sheet information is representative at the national level. On average, the dataset covers 75% of sectoral gross output from the input-output (IO) tables and 66% of the gross national output. We describe our data and its construction procedure in Section A in more detail.

Firm-Level Technology Adoption Contracts. We hand-collected and digitized firm-level data on technology adoption from official documents related to domestic firms' technology contracts with foreign firms from the National Archives of Korea and from the Korea Industrial Technology Association (1988). These documents had information about names of domestic and foreign contractors and contract years from 1966 to 1988. The law required domestic firms to submit related documents

[6]Specifically, about 74% of technology adoption contracts provided the know-how, 21.2% granted licenses, and 4% permitted the use of trademarks. For example, Figure A1 is one page of the contract document between Kolon (South Korean) and Mitsui Toatsu (Japanese), both of which are chemical manufacturers. The contract shows that Mitsui Toatsu had to provide technical assistance and blueprints to Kolon.

when they signed technology adoption contracts with foreign firms.[7]

Balance Sheet Data. We obtain firm balance sheet data by digitizing the Annual Reports of Korean Companies published by the Korea Productivity Center. Their publications cover firms with more than 50 employees. The data has information on sales, assets, fixed assets, and addresses of locations of establishments for the sample period between 1970 and 1982. Employment is not available until 1972. Using the addresses of plants and factories, we map firms' adoption activities to their location of production. We convert addresses to the 2010 administrative divisions of South Korea.

3 Historical Background of Late Industrialization in South Korea

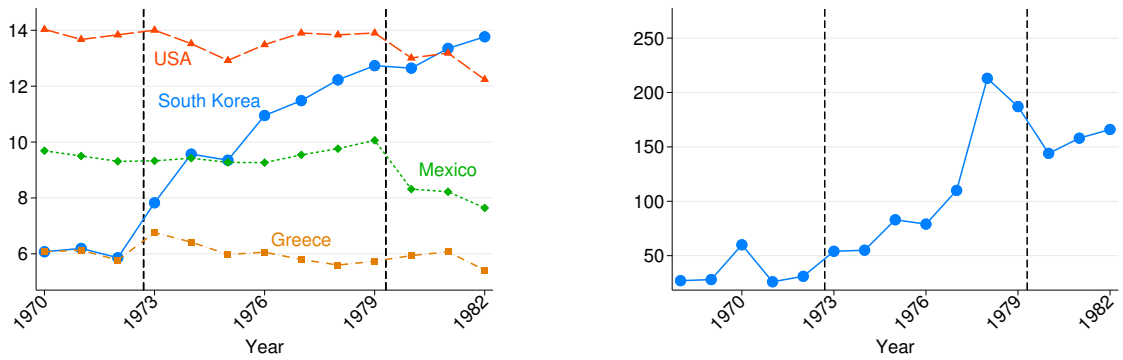
In late 1972, the South Korean government launched the Heavy and Chemical Industry (HCI) Drive to modernize and promote heavy manufacturing sectors, including chemicals, electronics, machinery, steel, non-ferrous metal, and transport equipment. One of the main policy instruments was subsidies for adopting foreign industrial technology.[8] In the 1970s, the adoption of foreign technologies and imported capital equipment related to those technologies were the main means of technology transfer from foreign developed economies to South Korea.[9]

The timing of the policy that subsidized technology adoption and the selection of the targeted sectors were driven by a political shock rather than economic conditions (Lane, 2019; Choi and Levchenko, 2021; Kim et al., 2021). After the Vietnam War, President Nixon changed the diplomatic policy of the United States toward its East Asian allies. In the Nixon Doctrine (1969), he declared that the East Asian allies of the United States, including South Korea, should take primary responsibility for their self-defense instead of relying on the United States military. He also planned the complete withdrawal of the United States military from South Korea. However, at this time, military tension between South and North Korea was rising. Because South Korea was heavily reliant on the United States military, the Nixon Doctrine posed a threat to the national defense of South Korea. In late 1972,

[7] Any domestic firms' transactions with foreign firms, including technology adoption contracts, were strictly regulated under the Foreign Capital Inducement Act, which was first enacted in 1966. According to the law, once a domestic firm got approval from the government for the adoption, it had to report the related information to the Economic Planning Board that played a central role in the economic policy-making process in South Korea during the sample period.

[8] For example, Hyundai Motors, the largest automotive company in South Korea, did not have its own models until 1972. It merely reassembled the existing car model developed by Ford and imported most of the automobile parts. Hyundai Motors did not start to produce its own models until 1972, when it became possible because of technology adoption. In 1974, Hyundai Motors hired George Turnbull, the former director at British Leyland as a new vice-president in order to improve its management technology. In 1976, Hyundai Motors adopted engine technology from Perkins Engine, design from Ital Design, and transmission technology from Mitsubishi, which are British, Italian, and Japanese firms, respectively. The government subsidized Hyundai Motors to enable it to import new capital equipment and construct new turnkey plants related to the technologies it had adopted. See Choi and Levchenko (2021) for how the South Korean government subsidized firms during the 1970s.

[9] Another commonly used means of technology transfer in developing countries is the foreign direct investment (FDI) (Keller, 2004). In South Korea, however, FDI did not play a big role. The South Korean government strictly regulated FDI, and the total value of the technologies and capital equipment domestic firms imported was 22 times greater than that of FDI. Moreover, when compared to other developing countries, South Korea had a lower stock of FDI. For example, the value of South Korea's stock of FDI was only 7 percent of the value of Brazil's stock in 1983 (Kim, 1997, p.42-43).



A. Heavy mfg. share of GDP (%)

B. # of new foreign technology adoption contracts made by South Korean heavy mfg. firms, 1968-1982

Figure 1. Late Industrialization and Technology Adoption in South Korea

Notes. The two dotted vertical lines represent the start and end of the South Korean government policy that subsidized technology adoption from 1973 to 1979. We obtain data on heavy manufacturing's share of GDP across countries from the OECD's STAN Structural Analysis Database and the OECD National Accounts Statistics database.

in order to modernize South Korea's military forces and achieve self-reliant defense against North Korea, President Park of South Korea announced the drive to promote the heavy and chemical manufacturing sectors that are related to the arms industry. The government considered South Korea's underdeveloped technology in heavy manufacturing sectors as one of the national threats, and given South Korea's large technology gap with the world frontier, the government deemed technology adoption to be the most effective way to catch up with the frontier.[10] The HCI Drive was temporary because it ended in 1979 after President Park was assassinated.

The left panel of Figure 1 plots the GDP share of the heavy manufacturing sector in South Korea and other selected economies. While at the beginning of the period of our analysis, South Korea's heavy manufacturing share was only 6%, it achieved a remarkable takeoff during the sample period, surpassing Mexico by the mid-1970s and the United States by 1982. Consistent with the GDP shares, employment and export shares of the heavy manufacturing sectors also increased from 4 to 8% and 13.7 to 35% between 1972 and 1982. The right panel plots the yearly number of new adoption contracts between South Korean and foreign firms. Our novel data reveals that the yearly number of contracts between South Korean and foreign firms for new technology quadrupled in the period between 1970 and 1982. This sudden and rapid increase in the rate of adoption coincided

[10] "Without rapidly improving our underdeveloped technology, our nation will be unable to secure an independent national defense system ... Inevitably, we will face a decline in our competitiveness of exports goods in international markets and national power, which bodes ill for our chance of a peaceful reunification with North Korea . . . Considering our nation's current technological state, adopting foreign advanced technologies and continuously adapting them to our needs seem to be the most effective catching-up strategy." (Ministry of Science and Technology, 1972, p. 3–4)

with temporary government subsidies for technology adoption in South Korea from 1973 to 1979. Even after the policy ended in 1979, the South Korean economy continued to specialize in the heavy manufacturing sectors.

4 Empirical Evidence on Technology Adoption

In this section, we examine how technology adoption benefited South Korean firms. We provide econometric evidence on (i) direct productivity gains for adopters, (ii) local productivity spillovers, and (iii) complementarity in firms' adoption decisions. According to the historical narrative, large-sized South Korean firms tend to rely on foreign sources to acquire advanced technologies, whereas small-sized firms relied on reverse engineering of technologies adopted by neighboring firms or on hiring experienced engineers from local adopters to obtain new technologies.[11] Our econometric evidence on the direct gains and the local spillovers capture the former and the latter, respectively.

Also, the local spillovers can further incentivize firms to adopt more technologies if the spillovers increase the profitability of modern technologies, which generates the pattern consistent with our third empirical evidence. Many previous papers have studied various channels through which the spillovers can increase the profitability of modern technologies, such as complementarity between gains from the adoption and firm productivity and higher intermediate input intensities of modern technologies.[12]

4.1 Direct Productivity Gains to Adopters

Empirical Strategy: Winners vs. Losers Research Design. When estimating the direct productivity gains to adopters, one of the key econometric challenges is that the adoption decisions firms make are endogenous. Unobservable systematic differences between adopters and non-adopters may result in a spurious correlation between adoption status and adopters' performance, leading to the standard selection bias problem. An ideal empirical scenario would be a random assignment of adoption status across firms. To approximate an ideal random assignment, we implement a winners vs. losers research design, drawing on Greenstone et al. (2010) that generates quasi-experimental variation in adoption status.

We define winners (the treated) as firms that successfully adopted technology from foreign firms. We define losers (the comparison) as non-adopters that made contracts with foreign firms that got approved by the government but were not able to adopt foreign technology because the foreign firm

[11] See Kim and Kim (1985) and Kim (1997). For instance, during the 1970s, there were 15 firms producing black-and-white TV producers. The first four large firms started producing TV after adopting foreign technologies, but the other 11 acquired technologies by hiring experienced engineers from the first four adopters (Kim, 1997, p. 156). See B.3 for historical case studies.

[12] For example, see Yeaple (2005), Verhoogen (2008), Lileeva and Trefler (2010), and Bustos (2011) for the complementarity between gains from new technologies and firm productivity through market size effects. Matsuyama (1995), Ciccone (2002), and Buera et al. (2021) show that when setting-up or production of modern technologies has higher intermediate input intensities, more adoption implies lower costs of production or set up costs, which increases firms' adoption incentives.

canceled the contract for reasons that had nothing to do with the South Korean firm. Examples include cancellations due to bankruptcy or to changes in the management team of the foreign firm. We exclude cancellations by domestic firms. The reasons for these cancellations include a domestic firm's sudden decreases in cash flow. See Figure A3 for an example of a cancellation by a loser. When contracts were canceled after approval from the government, domestic firms had to report the related documents on the reason for the cancellation. We collect data on contract cancellations by reading thousands of historical documents from the archives.

After identifying losers, we match each loser with an adopter using the exact Mahalanobis matching algorithm. The matching proceeds in two steps. First, we exactly match on region and sector in order to absorb shocks within regions and sectors, such as market size or local wages. Second, within regions and sectors, we choose a winner that was most similar to a loser in terms of firm size measured by log assets, where the similarity is measured by the Mahalanobis distance. We match losers and winners with replacement, so we can match one winner to multiple losers in a given year if they were in the same sector and region. The matching procedure gives us 34 pairings among 57 unique firms. All the matched pairs consist of heavy manufacturing firms. See Section C.6 for more detail on the matching procedure.

Using the matched pairs of winners and losers, we estimate the following event study specification, which is a generalized difference-in-differences (diff-in-diffs) design where a matched winner adopted in different periods and a loser was the control group. For firm i of pair p in period t,

$$y_{ipt} = \sum_{\tau=\underline{T}}^{\bar{T}} \beta_\tau \times D_{pt}^\tau + \sum_{\tau=\underline{T}}^{\bar{T}} \beta_\tau^{diff} \times D_{pt}^\tau \times \mathbb{1}[Adopt_{it}] + \delta_i + \delta_p + \delta_t + \epsilon_{ipt}, \tag{4.1}$$

where i denotes firm, p pair, and t time. D_{pt}^τ are event-study variables defined as $D_{pt}^\tau := \mathbb{1}[t-\tau = t(p)]$, where $t(p)$ is event year of pair p.[13] $\mathbb{1}[Adopt_{it}]$ is a dummy variable for adoption status. δ_i, δ_p, and δ_t are firm, pair, and year fixed effects. ϵ_{ipt} is an error term. Dependent variables y_{ipt} are log sales, log revenue TFP estimated based on Wooldridge (2009), and labor productivity defined as value added per worker. Matching with replacement introduces mechanical correlation across residuals, because of the possible appearance of the same firm. Thus, we two-way cluster standard errors at the level of both firms and pairs. Section C.7 describes our revenue TFP estimation procedure in more detail.

Identifying Assumption. Our identifying assumption is that losers form valid counterfactuals for winners. For this assumption to hold, (i) losers and winners should be ex-ante similar in terms of both observables and unobservables prior to an event conditional on matched controls, and (ii) cancellations by foreign firms should be uncorrelated with domestic firms' unobservables.

[13] This specification is robust to possible issues of a staggered diff-in-diffs design with heterogeneous treatment effects. First, our event study specification allows for dynamic treatment effects. Second, 31 out of 34 losers did not adopt technology after the cancellation up to 5 years, so they can be considered as clean controls. Third, because we are controlling event dummies, we do not use past treated winners as controls.

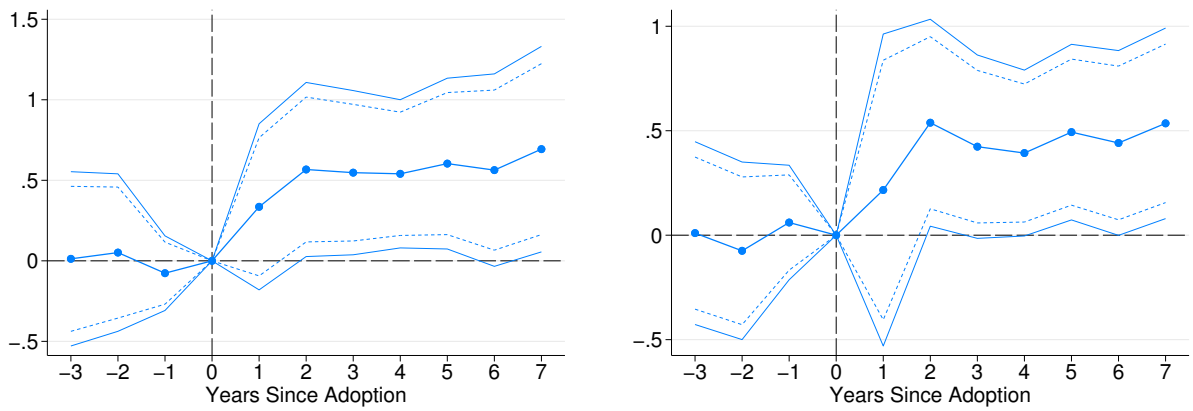
Our matching procedure makes it likely that the first condition would hold. It ensures that losers and winners are well-balanced in terms of observable covariates. Also, because we are comparing winners and losers that both wanted to adopt technology, we are indirectly controlling for underlying unobservables that made these firms self-select into the adoption. Finally, although unobservable political favors or subsidies provided during the periods when subsidies were available could have affected firms' adoption decisions, we expect that winners and losers had a similar level of political favor from the government when they made contracts because our definition of losers required government-approved contracts.

Because we do not find differential pre-trends between winners and losers (which will be shown below), the second condition of our identifying assumption would be violated only by unobservable shocks that affected losers' performance after the event and were correlated with foreign firms' cancellations, but did not affect losers' performance before the event. One example would be a negative shock of losers at the time of the event that caused losers to be matched with a bad foreign contractor that experienced a change in its management teams or went into bankruptcy. We can directly test this using firm-to-firm structure of our technology adoption contract data. If our results are driven by matching based on negative shocks, we would expect the characteristics of foreign firms that made contracts winners and losers to be different.

Balance. To assess covariate balance between two groups, we report descriptive statistics of the matched pairs and covariate balance test results. The descriptive statistics (Table C1) show that none of the t-statistics of tests that the mean of sales, employment, fixed assets, assets, and labor productivity of two groups are equal are statistically significant.[14] In Table C2, we report the results of covariate balance tests where we estimate a linear probability model of the effects of pre-event firm observables on adoption status. Across all specifications, none of the estimated coefficients of firm observables are statistically different from zero both individually and jointly once we control for pair fixed effects. These results indicate that firm observables cannot predict the cancellations of losers, which supports our identifying assumption that cancellations by foreign firms were exogenous shocks to domestic firms.

We compare two groups of foreign firms that made contracts with winners and losers based on their patenting activities in the United States. We obtain data on patenting activities in the United States from the United States Patent and Trademark Office (USPTO). We use firms' patenting activities in the United States as a proxy for how these firms are close to the world technology frontier. When these foreign firms made contracts, Table C3 shows that none of the t-statistics of tests that various measures of patent activities of two groups are equal are statistically significant. This rules out an alternative story that negative shocks made losers be matched with bad foreign

[14] Both winners and losers were larger than the average of all heavy manufacturing firms. For example, the average log sales of all heavy manufacturing firms were 15.54, but the averages of winners of losers were 17.80 and 18.46, respectively (column (2) of Table A2). Therefore, non-adopters may not represent a valid counterfactual for adopters, and naive comparison between them may lead to biased estimates.



A. Log sales

B. Log revenue TFP

Figure 2. Direct Productivity Gains to Adopters: Winners vs. Losers Design

Notes. This figure illustrates the estimated β_{τ}^{diff} in Equation (4.1) based on winners vs. losers research design. In Panels A, B, and C, the dependent variables are log sales and revenue TFP. We estimate revenue TFP based on Wooldridge (2009). All specifications control for event time dummies, and firm, pair, and calendar year fixed effects. The plotted coefficients correspond to columns (1)-(2) of Table 1. The figure reports 90 and 95 percent confidence intervals based on standard errors two-way clustered at the levels of pairs and firms.

firms.

Baseline Results. Table 1 and Figure 2 report the estimated coefficients in Equation (4.1). There are no pre-trend. Winners' sales, revenue TFP, and labor productivity did not begin to increase until adoptions occurred. 4 years after the adoption, winners' sales, revenue TFP, and labor productivity increased by 47%, revenue TFP by 42%, and labor productivity by 62%, and these effects were persistent.

Robustness. Increases in firms' sales or revenue TFP measures may reflect increases in demand shocks or mark-ups of the domestic market rather than productivity (Syverson, 2011). To deal with this issue, we merge our data set with KIS-VALUE that covers firms' export data after 1980. We find that the winners were 29 percentage points more likely to be an exporter and increased amounts of exports 7 or 8 years after the event when compared to the losers. These increases in exports in foreign markets are unlikely to be driven by demand shocks or mark-ups of the domestic market. See Section C.3 for more detail and other alternative hypotheses.

We compare our estimates from the winners vs. losers research design to estimates from a standard two-way fixed effects event study design that does not correct the selection problem. We find that the estimates from the standard event-study design is downward biased. The magnitude of the estimated coefficients from the standard event-study design is roughly 50% smaller than our estimates. This

Table 1: Event Study Estimates of Direct Productivity Gains to Adopters: Winners vs. Losers Research Design

Research Design	Winners vs. losers					
Dep. Var.	Log sales	Log labor productivity	Log revenue TFP			
			W. (2009)	ACF (2015)	LP (2003)	OLS
	(1)	(2)	(3)	(4)	(5)	(6)
3 years before event	0.00	–0.09	0.01	0.06	0.04	0.00
	(0.27)	(0.41)	(0.24)	(0.30)	(0.24)	(0.29)
2 years before event	0.07	–0.36	–0.11	–0.18	–0.08	–0.19
	(0.24)	(0.46)	(0.24)	(0.34)	(0.24)	(0.34)
1 year before event	–0.10	–0.02	0.04	0.10	0.06	0.08
	(0.12)	(0.23)	(0.15)	(0.19)	(0.15)	(0.19)
Year of event						
1 year after event	0.31	0.28	0.22	0.37	0.23	0.33
	(0.25)	(0.41)	(0.37)	(0.38)	(0.37)	(0.39)
2 years after event	0.53*	0.64**	0.56**	0.71**	0.56**	0.67**
	(0.27)	(0.30)	(0.26)	(0.30)	(0.26)	(0.29)
3 years after event	0.47*	0.62**	0.41*	0.66**	0.43*	0.63**
	(0.26)	(0.29)	(0.23)	(0.28)	(0.23)	(0.27)
4 years after event	0.48**	0.62**	0.42*	0.67**	0.45**	0.63**
	(0.23)	(0.27)	(0.21)	(0.25)	(0.21)	(0.24)
5 years after event	0.58**	0.43	0.52**	0.64**	0.52**	0.57*
	(0.26)	(0.36)	(0.21)	(0.29)	(0.23)	(0.29)
6 years after event	0.54*	0.55*	0.46**	0.59**	0.46*	0.56**
	(0.29)	(0.28)	(0.23)	(0.29)	(0.24)	(0.27)
7 years after event	0.66**	0.56*	0.57**	0.69**	0.58**	0.67**
	(0.31)	(0.32)	(0.23)	(0.29)	(0.23)	(0.28)
Firm FE	✓	✓	✓	✓	✓	✓
Pair FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Adj. R^2	0.88	0.61	0.86	0.94	0.90	0.60
# cluster (pair)	34	34	34	34	34	34
# cluster (firm)	57	57	57	57	57	57
N	951	835	827	827	827	827

Notes. This table reports the estimated event study coefficients β_τ^{diff} in Equation (4.1) based on the winners vs. losers research design. β_0^{diff} is normalized to zero. The dependent variables are log sales, log revenue TFP, and log labor productivity defined as value added divided by employment. Value added is obtained as sales multiplied by the value added shares obtained from input-output tables corresponding to each year. In columns (3), (4), (5), and (6), we estimate log revenue TFP based on Wooldridge (2009), Ackerberg et al. (2015), Levinsohn and Petrin (2003), and OLS, respectively. All specifications control for event time dummies, firm fixed effects, pair fixed effects, and calendar year fixed effects. Robust standard errors in parenthesis are two-way clustered at the pair and firm levels. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

shows that correcting the selection problem is important for understanding the impact of technology adoption. See Section C.4 for more detail.

We run the same regressions using different revenue TFP measures based on Ackerberg et al. (2015), Levinsohn and Petrin (2003), and OLS. The results are reported in Figure C1 and columns (4)-(6) of Table 1. Even though we use different measures, the estimated event study shows no pre-trend, and the estimated coefficients are within a standard error of the estimates of column (3) of Table 1.

4.2 Local Productivity Spillover of Technology Adoption

In this subsection, we provide empirical evidence on local productivity spillovers of technology adoption. Our measure for the spillover is a weighted mean of the local adoption status of firms within the same sector, where the weight is given by the inverse of distance between firms. We define the spillover experienced by firm i in region n and sector j at time t as follows:

$$\text{Spill}_{inj(t-h)} = \sum_{k \in nj/\{i\}} \left\{ \frac{(1/dist_{ik})\mathbb{1}[\text{Adopt}_{knj(t-h)}]}{\sum_{k' \in nj/\{i\}} (1/dist_{ik'})} \right\}, \tag{4.2}$$

where $nj/\{i\}$ is a set of sector j firms in region n excluding firm i, $dist_{ik}$ is the distance between firms i and k, and $\mathbb{1}[\text{Adopt}_{knj(t-h)}]$ is a dummy variable for firm k's adoption status lagged by h years. Lagging the variable allows for the possibility that it took some time for new knowledge from adopted technologies to diffuse locally. When we construct the spillover measure for firm i, we exclude firm i to rule out mechanical correlation. For our baseline specification, we set the value of h as 4 and conduct robustness checks for different values of h. Each firm within the same region and sector has different values for spillover depending on its distance from adopters. Distance from adopters is the main variation we use for our empirical analysis.

The spillover measure can be interpreted as the probability that firm i's manager would meet other managers who worked in firms that had adopted foreign technologies. Each manager is endowed with a unit of time and can randomly meet at most one manager from other firms. The probability that a manager would meet a manager from firm k is given by the inverse of the distance between firms i and k. The inverse of the distance is a proxy for spatial frictions that would have impeded local interaction between managers of two firms.[15] The spillover measure captures the fact that

[15] By taking the weighted average, we implicitly assume that the spillover measure is invariant to the total number of firms. As far as we know, there is no consensus about the functional form of knowledge spillovers (Combes and Gobillon, 2015). However, we think the weighted average is more suitable in our setting. First, this is consistent with our theoretical interpretation, which is also widely adopted in growth and knowledge diffusion literature (Lucas and Moll, 2014; Buera and Oberfield, 2020; Perla et al., 2021). Given that managers' time is a limited resource in the real world, this theoretical interpretation seems to be more natural than an alternative scenario where a manager can interact with all firms in the same local area. In this alternative scenario, the spillover varies depending on the total number of adopters rather than the shares. Second, the literature on externalities has commonly used averages to capture agglomeration forces, such as local shares of skilled labor and population density.

knowledge spillovers are highly localized and quickly decay with distance. This is supported by the recent empirical evidence on knowledge spillovers (e.g., Jaffe et al., 1993; Kerr and Kominers, 2015; Kantor and Whalley, 2019; Moretti, 2021).

Using this spillover measure, we consider the following long-difference regression model:

$$\triangle y_{injt} = \beta^S \triangle \text{Spill}_{inj(t-4)} + \mathbf{X}'_{injt}\boldsymbol{\beta} + \triangle \delta_{njt} + \triangle \epsilon_{injt}, \tag{4.3}$$

where $\triangle$ is a time difference operator and i denotes firm, j sector, n region, and t time. y_{injt} are dependent variables, $\mathbf{X}_{injt}$ are controls, and δ_{njt} represent time-varying region-sector fixed effects that absorb time-varying shocks within each region and sector. Firm time-invariant factors are differenced out. Standard errors are two-way clustered at the levels of regions and conglomerates. In South Korea, large conglomerate groups known as chaebols own multiple firms across sectors and regions. Clustering at the conglomerate level allows for arbitrary correlation of error terms between firms within the same conglomerate group.

To use the data more efficiently, we use overlapping 8-year long-differences: 1971-1979 and 1972-1980. Each set covers the period between 1973 and 1979 when the temporary subsidies were provided. Because we cluster firms at both region and conglomerate level, this is innocuous. We add dummies for each set of differences and for interaction terms between these dummies and δ_{njt}.

Identifying Assumption. Our identifying assumption for a causal interpretation is that distance to adopters within regions and sectors ($\text{Spill}_{inj(t-4)}$) is uncorrelated with the error term ϵ_{injt} conditional on δ_{njt}, δ_i, and other controls. There are two main identification concerns highlighted by Manski (1993). First, neighborhood shocks within regions and sectors that are correlated across firms can affect both firm i's outcomes and the adoption decisions of neighboring firms, leading to spurious correlation. Second, adopters tend to be larger than non-adopters and omitting other effects of being close to large firms can lead to omitted variable bias.

We deal with the first concern by controlling for time-varying region-sector fixed effects at a fine level of geographic detail. The median size of our geographical unit of analysis for the sample is about Manhattan-sized, or almost 34 square miles. This is much finer than the unit of analysis in many previous studies. Our identifying variation comes purely from distance to adopters within the same sector and region, but not from variation across regions or sectors. Variation in $\text{Spill}_{inj(t-4)}$ mainly comes from two sources: (i) adoption decisions by non-adopters operating at the start of the sample period, and (ii) entry and adoption decisions of new firms entering between the start and the end of the sample period.[16] Because we control for δ_{njt} and difference out δ_i, neighboring firms' adoption decisions based on time-varying region-sector factors do not bias our estimates. Only adoption or entry decisions based on time-varying firm-specific factors that are spatially correlated

[16] The firms that enter began production between the start and the end of the sample period affected the spillover measure of firms that were in production at the beginning of the sample period, but we did not include them in the sample because we restrict the sample to firms that were operating at the start of the period.

at the neighborhood level would bias our estimates. For example, infrastructure improvement at the neighborhood level that affected both firms' outcomes and adoption decisions would bias our estimates. Exploiting spatial variation at a fine level mitigates potential spatial correlation at the neighborhood level within each region and sector.

We deal with the second concern by isolating variation in proximity to adopters from proximity to large firms by controlling for other potential means of local spatial interactions between firms. We control for the average sales of local firms by inversely weighting distances:

$$\ln\left(\text{Spill-Sales}_{injt}\right) = \ln\left(\sum_{k\in nj/\{i\}}\left\{\frac{(1/dist_{ik})\text{Sales}_{kt}}{\sum\limits_{k'\in nj/\{i\}}(1/dist_{ik'})}\right\}\right). \tag{4.4}$$

This weighted average sale proxies other agglomeration or competition forces of being close to large firms within the same region and sector. We also control for a measure of access to local markets attributable to local input sourcing by taking the weighted sum of neighbors' sales period t input-output coefficients, where the weight is given by the inverse of the distances (Donaldson and Hornbeck, 2016):

$$\ln\left(\text{Input-MA}_{injt}\right) = \ln\left(\sum_{j'}\sum_{k\in nj'/\{i\}}\gamma_j^{j'}(1/dist_{ik})\text{Sales}_{kt}\right), \tag{4.5}$$

where $\gamma_j^{j'}$ represent shares of sector j' intermediate inputs used by sector j. This measure of market access is a proxy for differential market size attributable to localized input sourcing. Because we do not have information on commodity or service sector firms, we sum j' only across manufacturing sectors.

Baseline Results. Table 2 reports the OLS estimates for β^S when the dependent variables are sales and revenue TFP. We report the estimation results for the sample of firms that were operating before 1973 and after 1979 and did not adopt foreign technologies between these periods. All specifications include the initial dependent variable. Column (1) of Panel A is our baseline estimate.[17] The estimated coefficient is statistically significantly positive. One standard deviation increase in the spillover (0.033) contributes to 14.5% increases in sales. β^s can also be interpreted as a semi-elasticity of non-adopters' sales to local shares of adopters in a hypothetical region when all firms are equally

[17] The magnitude of the estimated coefficients is consistent with the estimates in the literature on local knowledge spillover. The estimates in column (1) of Panel A indicate that the elasticity of firms' sales to the spillover at the mean and the 90th and 95th percentiles is 0.05, 0.13, and 0.26, respectively. We calculate the elasticity of the adoption spillover as follows. The mean level of the local share of adopters is 0.011. An increase of 1% of the mean level (0.00011) increases firms' sales by 0.05% ($= 100 \times 0.00011 \times 4.39$). The elasticities at the 90th and 95th percentiles are calculated similarly. For example, estimates from Bloom et al. (2013) imply that the elasticity of firms' sales to their spillover measure based on patents is 0.19–0.26. We calculate elasticity above the 90th percentile because shares of adopters are highly skewed; where the 75th, 90th, 95th, and 99th percentiles were 0, 0.03, 0.06, and 0.18, respectively.

Table 2: Local Spillovers from Technology Adoption

Dep. Var.	Log sales					Log revenue TFP				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Spill	4.39***	3.79**	4.94***	4.23***	4.07**	5.55***	5.41***	5.81***	5.34***	5.11**
	(1.54)	(1.64)	(1.70)	(1.50)	(1.76)	(1.84)	(1.62)	(2.08)	(1.78)	(1.92)
ln(Spill-Sales)			–0.02		–0.02			–0.02		–0.01
			(0.01)		(0.01)			(0.02)		(0.02)
ln(Input-MA)				–0.03	–0.02				–0.04**	–0.03
				(0.02)	(0.02)				(0.02)	(0.02)
Region-Sector FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Conglomerate FE		✓			✓		✓			✓
Adj. R^2	0.18	0.22	0.19	0.19	0.22	0.44	0.42	0.44	0.44	0.42
# clusters (region)	53	53	53	53	53	41	36	41	41	36
# clusters (conglomerate)	636	630	636	636	630	324	275	324	324	275
N	1079	1073	1079	1079	1073	344	292	344	344	292

Notes. This table reports the OLS estimates of Equation (4.3). When we construct the spillover measure defined in Equation (4.2), we lag the adoption status of firms by four years. In Panel A, we use the subsample that include only firms that did not adopt any technology during the sample period. In Panel B, we use the full sample of adopters and non-adopters and control for adoption status. The dependent variables are log sales in columns (1)-(5) and revenue TFP in columns (6)-(10). We estimate revenue TFP based on Wooldridge (2009). The additional controls ln(Spill-Sales) and ln(Input-MA) are additional controls defined in Equations (4.4) and (4.5). In all specifications, we control for region-sector fixed effects and for the initial dependent variable at the start of the sample period. Standard errors are two-way clustered at both region and conglomerate level and are reported in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

distanced. In this interpretation, a one percentage point increase in the local share of adopters leads to a 4.39% increase in non-adopters' sales in the hypothetical region. In columns (2), (3), and (4), we also control for conglomerate fixed effects, ln(Spill-Sales), and ln(Input-MA), respectively. In column (5), we control for all other variables. The estimates with additional controls all stay within a standard error of the baseline estimate. The estimated coefficients of ln(Spill-Sales) and ln(Input-MA) are not statistically significant and do not take positive values. If knowledge spillovers decay more quickly with distances than other spatial interactions captured by the other two controls and the other two spatial interactions operate at a broader spatial scale than knowledge spillovers, it is possible to have the null results of ln(Spill-Sales) and ln(Input-MA). In columns (6)-(10), we use log revenue TFP as a dependent variable. The number of samples for revenue TFP was smaller because employment data was not available until 1972. The estimates for log revenue TFP are about 20% larger than the estimates for log sales.

Robustness. We provide a battery of robustness checks. In Table C4, we run the same regression for the full sample, including both adopters and non-adopters. For the full sample, we control for a dummy variable for own adoption status. Because they are likely to be correlated with the error term, we do not meaningfully interpret this variable. The estimates based on the full sample are within a standard error of the baseline estimates in column (1) of Panel A.

Instead of using the spillover measure with a four year lag, we use the spillover measure with three or five year lags. The results are reported in Tables C5 and C6. The estimated coefficients from these robustness checks remain within a standard error of the baseline estimates.

It is possible that the local spillovers were operating at a broader level than our geographical unit of analysis. To check this, we aggregate our geographical unit to 42 regions based on industrial structure and electoral districts. We define the spillover similarly to Equation (4.2) at the broader regional level. Then, we run the same regression while controlling the same set of region and sector fixed effects with the baseline specification. Thus, we absorb the same time-varying shocks with the baseline specification while allowing the spillovers to operate at the broader level. We two-way cluster at the broader regional level and the conglomerate level. The results are reported in Tables C9. The estimated coefficients remain within a standard error of the baseline estimates.

We consider cross-sector spillovers in Section C.5. Following Ellison et al. (2010), we construct the local cross-sector spillover measure based on the expression in Equation (4.2) and the input-output table coefficients. We do not find statistically significant results for the local cross-sector spillovers.

Instead of using log sales or revenue TFP, we use log fixed assets, assets, and employment, labor productivity. The results are reported in Tables C11 and C12. The estimated coefficients are statistically significant and are positive for different dependent variables except for employment.

4.3 Complementarity in Firms' Technology Adoption Decisions

In this subsection, we provide empirical evidence on complementarity in firms' technology adoption decisions. We run the same regression as Equation (4.3) while using a dummy variable of adoption of a new technology as a new dependent variable and controlling for the same set of fixed effects and additional controls.[18] By using this dependent variable, we can examine how neighboring firms' adoption status affect firms' new adoption decisions. The identifying assumptions for the causal interpretation are the same as those of the spillover regression model.

Baseline Results. Table 3 reports the results. Across the specifications in columns (1)-(5), the estimated coefficients of the spillover measure are positive and statistically significant. These positive estimates imply that firms are more likely to adopt a new technology if more neighboring firms

[18] A dummy variable of adoption of a new technology as a new dependent variable $\mathbb{1}[\text{New Contract}_{injt}]$ differs from the dummy variable of firms' ever adoption status $\mathbb{1}[Adopt_{injt}]$ that is used to construct the spillover measure. For example, if a firm has not adopted any foreign technologies and makes a new contract in time t, both $\mathbb{1}[\text{New Contract}_{injt}]$ and $\mathbb{1}[Adopt_{injt}]$ become 1 in t. If a firm had made a contract before t but did not make a new contract in t, then only $\mathbb{1}[Adopt_{injt}]$ takes a value of 1.

Table 3: Complementarity in Firms' Technology Adoption Decisions

Dep. Var.	$\mathbb{1}$[New Contract]					*asinh*(# New Contract)				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Spill	0.49***	0.46***	0.49***	0.49**	0.47***	0.46**	0.42***	0.46***	0.45**	0.41**
	(0.18)	(0.15)	(0.18)	(0.19)	(0.15)	(0.17)	(0.15)	(0.17)	(0.18)	(0.16)
ln(Spill-Sales)			–0.00		–0.00			–0.00		–0.00
			(0.00)		(0.00)			(0.00)		(0.00)
ln(Input-MA)				–0.00	0.00				–0.00	–0.00
				(0.00)	(0.00)				(0.00)	(0.00)
Region-Sector FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Conglomerate FE		✓			✓		✓			✓
Adj. R^2	0.02	0.07	0.02	0.02	0.07	0.09	0.19	0.09	0.09	0.19
# cluster (region)	61	61	61	61	61	61	61	61	61	61
# cluster (conglomerate)	1414	1413	1414	1414	1413	1414	1413	1414	1414	1413
N	2689	2688	2689	2689	2688	2689	2688	2689	2689	2688

Notes. This table reports the OLS estimates of Equation (4.3). When we construct the spillover measure defined in Equation (4.2), we lag firms' adoption status by four years. In columns (1)-(5), the dependent variables are a dummy variable of whether a firm makes a new technology adoption contracts made in a given year. In columns (6)-(11), the dependent variables are the inverse hyperbolic sine transformation of the number of new technology adoption contracts made in a given year. ln(Spill-Sales) and ln(Input-MA) are additional controls defined in Equations (4.4) and (4.5). In all specifications, we control for region-sector fixed effects and the initial dependent variable at the start of the sample period. Standard errors are two-way clustered at both region and conglomerate level and are reported in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

have adopted technologies. This is consistent with complementarity in firms' adoption decisions where gains from adopting a new technology are larger if more neighboring firms have adopted foreign technologies. In terms of economic magnitude, one standard deviation increase of the adoption spillover measure increases a firm's probability of making new technology adoption contracts by 1.5 percentage points on average. This magnitude of 1.5% is about half of the average share of firms that make a new contract in a given year (3%) and about 10% of shares of firms that ever adopted technologies from foreign firms in 1982 (12%). In columns (5)-(10), the dependent variables are inverse hyperbolic sine transformation of the number of newly adopted technologies (Burbidge et al., 1988). The estimated coefficients imply that one standard deviation increase of the spillover measure increases a 0.28 standard deviation of the dependent variable.

Robustness. We consider similar robustness checks in Section 4.2. See Tables C7 and C8 for different lags, Table C10 for the spillover measure defined at the broader level, and Table C16 for the cross-sector spillover. Across robustness checks, the estimates remain within a standard error of

the baseline results.

5 Theoretical Framework

In this section, we present a dynamic spatial model with firms' endogenous adoption decisions and local productivity spillovers.

5.1 Setup

We consider a small open economy Home with N regions and J sectors. We divide the world into Home and Foreign. We assume that Home is small and it cannot affect Foreign aggregates. However, its domestic prices are determined by domestic supply and demand conditions, and Home firms face downward sloping demands from Foreign. Subscripts $n, m \in \mathcal{N}$ index Home regions, and $j, k \in \mathcal{J}$ sectors, where $\mathcal{N}$ and $\mathcal{J}$ are the sets of Home regions and sectors. Time is discrete and indexed by $t \in \{1, 2, \ldots\}$.

There are two types of goods: intermediate and final goods. Intermediate goods are produced by intermediate goods producers. There is a fixed mass of firms (M_{nj}) in each region and sector. Sectors are either tradable ($j \in \mathcal{J}^x$) or non-tradable ($j \notin \mathcal{J}^x$). For $j \in \mathcal{J}^x$, intermediate goods are tradable across regions and can be exported to Foreign. Both internal and international trade of sector j are subject to iceberg trade costs $\tau_{nmj} \geq 1$ and $\tau^x_{nj} \geq 1$, respectively. When exporting to Foreign, firms additionally incur fixed export costs (Melitz, 2003). In a subset of sectors ($\mathcal{J}^T \subset \mathcal{J}$), firms in these sectors can adopt advanced technology from foreign sources after incurring fixed adoption costs.

In each region, there is a competitive labor market. We normalize the total population of the Home regions to 1: $L_t = \sum_{n \in \mathcal{N}} L_{nt} = 1$, where L_{nt} is population in region n.

5.2 Firms

Production. Each intermediate variety is produced by intermediate goods producers, which we call firms. Each firm is indexed by subscript i. Firms are heterogeneous in productivity. Firm i's output y_{it} is

$$y_{it} = z_{it} L_{it}^{\gamma_j^L} \prod_{k \in \mathcal{J}} M_{it}^{\gamma_j^k}, \qquad \gamma_j^L + \sum_{k \in \mathcal{J}} \gamma_j^k = 1, \tag{5.1}$$

where z_{it} is firm i's productivity, L_{it} are labor inputs, M_i^k are sector k intermediate inputs, and γ_j^k are Cobb-Douglas shares. A unit cost of an input bundle is $c_{njt} = (w_{nt}/\gamma_j^L)^{\gamma_j^L} \prod_{k \in \mathcal{J}} (P_{nkt}/\gamma_j^k)^{\gamma_j^k}$, where w_{nt} is wage and P_{njt} is a price of intermediate inputs.

In each region and sector, a final goods producer produces nontradable local sectoral aggregate goods used for final consumption and for intermediate inputs. They are perfectly competitive. A final goods producer aggregates all available varieties from all regions and countries using a constant

elasticity of substitution (CES) aggregator:

$$Q_{njt} = \left[\sum_{m \in \mathcal{N}} \int_{\omega \in \Omega_{mj}} q_{it}(\omega)^{\frac{\sigma-1}{\sigma}} d\omega + \int_{\omega \in \Omega_j^f} q_{it}^f(\omega)^{\frac{\sigma-1}{\sigma}} d\omega \right]^{\frac{\sigma}{\sigma-1}}. \tag{5.2}$$

Q_{njt} are the quantities of local aggregate sectoral goods produced. Ω_{mj} is the set of available sector j varieties in region m. q_{it} and q_{it}^f are the quantities demanded of an intermediate variety ω produced by a domestic and a foreign firm, respectively. We assume that the available set of foreign varieties Ω_j^f is exogenously given to the Home regions and is the same across regions. Because there are no fixed export costs for internal trade, each region faces the same set of available varieties.

The exact CES price index is

$$P_{njt}^{1-\sigma} = \sum_{m \in \mathcal{N}} \left[\int_{\omega \in \Omega_{mj}} p_{it}(\omega)^{1-\sigma} \right] + (\tau_{nj}^x)^{1-\sigma} \underbrace{\int_{\omega \in \Omega_j^f} p_{it}^f(\omega)^{1-\sigma}}_{=(c_{jt}^f)^{1-\sigma}}, \tag{5.3}$$

where p_{it} is the price of Home variety and p_{it}^f is a FOB price of an imported variety from Foreign. Because we have assumed a small open economy, Home takes the FOB prices of foreign firms as given and therefore c_{jt}^f is exogenous to Home.

Technology Adoption and Exports. In each period, firms make two static decisions: (i) whether to adopt advanced technology and (ii) whether to export. Both adopting technology and exporting incur adoption and export fixed costs in units of input bundles (F_j^T and F_j^x). The fact that both adoption and export costs are fixed costs make firms' decisions static. This static nature of firms' problems makes the model computable while preserving rich cross-sectional regional heterogeneity, and connecting the model to the data and econometric estimates. Once firms decide to adopt technology and pay fixed adoption costs, they can increase their productivity.[19]

Firm productivity z_{it} is composed of three terms:

$$z_{it} = \underbrace{\eta^{T_{it}}}_{\substack{\text{Direct} \\ \text{productivity gains}}} \times \underbrace{f(\lambda_{njt-1}^T)}_{\substack{\text{Local} \\ \text{spillover}}} \times \underbrace{\phi_{it}}_{\substack{\text{Exogenous} \\ \text{productivity}}},$$

where $\eta > 1$ is direct productivity gains from adoption, T_{it} is a binary adoption decision, $f(\lambda_{njt-1}^T)$ is a local adoption spillover that increases in the share of adopters in the previous period λ_{njt-1}^T, and ϕ_{it} is

[19] F_j^T is a reduced-form parameter that includes direct payment to foreign sources, the costs of installing a new structure or capital equipment related to a newly adopted technology, and any barriers to adoption. Many previous papers have studied sources of adoption barriers in developing countries (see, among many others, Parente and Prescott, 1994; Banerjee and Duflo, 2005; Acemoglu et al., 2007; Atkin et al., 2017). Also, South Korea's political context in the 1970s might have affected F_j^T. Due to the Cold War, the United States government wanted the South Korean economy to be self-sustaining and promoted South Korea's economic growth. Therefore, it did not block transfers of technology to South Korean firms (Vogel, 1991, p.8).

exogenous productivity. We allow the spillover to operate with a one-period lag (Allen and Donaldson, 2020), which is more realistic given that our focus is the transformation of the South Korean economy within 10 years instead of the long-run outcomes that have been studied more frequently in the trade literature.[20] When making adoption decisions, adopters internalize the direct productivity gain η but not the spillover $f(\lambda^T_{njt-1})$. These externalities mean that social returns to adoption are larger than private returns. This leads to adoption rates that are lower than the socially optimum level. Because of firms' endogenous technology adoption decisions, z_{it} is endogenously determined in the equilibrium.[21] For sectors where technology adoption is not available, firms' productivity consists of only exogenous productivity: $z_{it} = \phi_{it}$.

$f(\lambda^T_{njt-1})$ captures local knowledge spillovers from newly adopted technologies. We parametrize $f(\lambda^T_{njt-1})$ as follows:

$$f(\lambda^T_{njt-1}) = exp(\delta\lambda^T_{njt-1}),$$

where $\delta > 0$ is the semi-elasticity of firm productivity with respect to a local share of adopters. Under this parametrization, we show that δ can be mapped to the reduced-form spillover estimate in Section 4.2. The spillover can be micro-founded based on (1) local diffusion of new engineering knowledge; and (2) learning externalities and labor mobility across firms. These two sets of microfoundations are based on historical case studies of South Korea in the 1970s.[22] Complete derivations of the microfoundations and related historical cases are described in Appendix B.3.

ϕ_{it} is drawn from a distribution $G_{njt}(\phi)$, which varies across regions, sectors, and periods. Each draw is independent across firms, regions, sectors, and time. We assume that exogenous productivity ϕ_{it} follows a bounded Pareto distribution (Chaney, 2008; Helpman et al., 2008):

$$\phi_{it} \sim \frac{1 - (\phi_{it}/\phi^{min}_{njt})^{-\theta}}{1 - (\phi^{max}_{njt}/\phi^{min}_{njt})^{-\theta}},$$

[20] Allowing the spillover to operate with a lag gives an economy a deterministic static equilibrium in each period (Adserà and Ray, 1998). This is a desirable theoretical property for two reasons. First, we can rule out unrealistic situations where an economy swings from one equilibrium to another in a different period depending on agents' self-fulfilling beliefs. Second, because there is a unique static equilibrium for each period, the model can be easily mapped to cross-sectional data. In general, multiple static equilibria models suffer from identification issues due to multiplicity. Similar to our setting, Kline and Moretti (2014) and Allen and Donaldson (2020) also allowed agglomeration to operate with some lags.

[21] Recent studies (see, among many others, Desmet and Rossi-Hansberg, 2014; Desmet et al., 2018; Walsh, 2019; Nagy, 2020; Peters, 2021) also present dynamic spatial model with endogenous local productivity. In our setup, local productivity is endogenously determined because of firms' technology adoption decisions.

[22] The historical evidence shows that new ideas and knowledge about adopted technologies were frequently transmitted to local capital goods producers through reverse engineering of capital equipment related to adopted technologies. Also, technical personnel of adopters moved frequently to other firms and their movement played an important role in diffusing knowledge about adopted technologies. It is further supported by higher aggregate labor mobility rates in South Korea in the 1970s than those of Japan and the United States (Kim and Topel, 1995). Also, learning externalities and knowledge spillovers through labor mobility have been widely studied in the literature. For example, see Lucas (1988) for learning externalities of human capital; see Stoyanov and Zubanov (2012) and Serafinelli (2019) for empirical evidence on effects of labor mobility across firms on knowledge diffusion.

which is parametrized by ϕ_{njt}^{max}, ϕ_{njt}^{min}, and θ. We also assume that the gap between the lower and upper bounds of the distribution is the same across regions, sectors, and periods: $\phi_{njt}^{max} = \kappa\phi_{njt}^{min}$, parametrized by κ. The lower bound of the distribution may vary across regions, sectors, and periods, but the upper bound is always proportional to the lower bound by κ. This distributional assumption gives us analytical expressions for aggregate variables and rationalizes zeros observed in the data.[23]

Adoption Subsidy. We model the adoption subsidies in Section 3 as input subsidies because the South Korean government provided subsidies to large adopters so they could purchase intermediate inputs and new capital equipment related to the technologies they adopted. Adopters are potentially eligible for input subsidies $0 < s_{njt} < 1$ that can vary across regions, sectors, and periods. Therefore, firm i's unit cost of production, $\tilde{c}_{it}$, is $\frac{c_{njt}}{\phi_{it} f(\lambda_{njt-1}^T)}$ if firm i adopts technology or $\frac{1-s_{njt}}{\eta} \times \frac{c_{njt}}{\phi_{it} f(\lambda_{njt-1}^T)}$ if it did not. Adopters have a lower unit cost of production than non-adopters because of higher productivity (η) and input subsidies (s_{njt}).

The government imposes a labor tax (τ_t^w) to finance these subsidies.[24] We assume that the labor tax rate is constant across regions, so the after-tax wages in region n are $(1-\tau_t^w)w_{nt}$. The government budget is balanced every period.

A Firm's Maximization Problem. Each firm faces a CES demand and is monopolistic for its own variety. Firm i's quantities demanded from region m are $q_{inmjt} = (\tilde{p}_{it})^{-\sigma} P_{mjt}^{\sigma-1} E_{mjt}$ and when firm i charges price $\tilde{p}_{it}$. The demanded from foreign markets at that price is $q_{injt}^x = (\tilde{p}_{it})^{-\sigma} D_{jt}^f$. A firm optimally charges a constant mark-up $\mu = \sigma/(\sigma-1)$ over its marginal cost. Thus, the prices charged by firm i in region n of sector j charged to buyers in region m are $p_{inmjt} = \mu\tau_{nmj}\tilde{c}_{it}$ and export prices are $p_{injt}^x = \mu\tau_{nj}^x\tilde{c}_{it}$.

[23] If $\kappa \to \infty$, the bounded Pareto distribution becomes unbounded Pareto. However, the unbounded Pareto distributional assumption cannot rationalize zeros because as productivity is unbounded, there is always a small share of firms that adopt technology regardless of the values of F_j^T. Helpman et al. (2008) also uses a bounded Pareto distributional assumption to rationalize zero trade flows across countries.

[24] The assumption that the government finances its adoption subsidies through a labor tax is based on the labor market policies and the pro-business attitude of the authoritarian South Korean government in the 1970s. The government restricted firms' nominal wage growth to below 80% of the sum of inflation and aggregate productivity growth and enacted temporary provisions in 1971 to prohibit labor union activities (Kim and Topel, 1995). Also, see footnote 3 of Itskhoki and Moll (2019).

A firm's profit is obtained after maximizing over T_{it} and x_{it}:

$$\begin{aligned}
\pi_{it} = \pi(\phi_{it}) &= \max_{x_{it}, T_{it} \in \{0,1\}} \{\pi(T_{it}, x_{it}; \phi_{it})\} \\
&= \max_{x_{it}, T_{it} \in \{0,1\}} \Bigg\{ \underbrace{\sum_{m \in \mathcal{N}} \left[\frac{1}{\sigma} \left(\mu \frac{\tau_{nmj}(1 - s_{njt})^{T_{it}} c_{njt}}{\phi_{it} \eta^{T_{it}} f(\lambda^T_{njt-1})} \right)^{1-\sigma} P^{\sigma-1}_{mjt} E_{mjt} \right]}_{:= \pi^d(T_{it}; \phi_{it}) = \sum_{m \in \mathcal{N}} \pi^m(T_{it}; \phi_{it})} \\
&\qquad + x_{it} \Bigg[\underbrace{\frac{1}{\sigma} \left(\mu \frac{\tau^x_{nj}(1 - s_{njt})^{T_{it}} c_{njt}}{\phi_{it} \eta^{T_{it}} f(\lambda^T_{njt-1})} \right)^{1-\sigma} D^f_{jt}}_{:= \pi^x(T_{it}; \phi_{it})} - c_{njt} F^x_j \Bigg] - T_{it} c_{njt} F^T_j \Bigg\},
\end{aligned} \tag{5.4}$$

where x_{it} and T_{it} are binary export and adoption decisions, E_{mjt} are region m's total expenditures on sector j goods, and D^f_{jt} are exogenous foreign demands. $\pi^m(T_{it}; \phi_{it})$ are operating profits conditional on adoption status obtained from region m, and $\pi^d(T_{it}; \phi_{it}) = \sum_{m \in \mathcal{N}} \pi^m(T_{it}; \phi_{it})$ are the sum of all these profits from domestic regions. $\pi^x(T_{it}; \phi_{it})$ are operating profits in foreign markets conditional on adoption status.

Adoption and Export Cutoff Productivities. Firms adopt technology and export their goods when the gains from these activities are larger than their fixed costs. Because these gains from adoption and exporting are higher when firms are more productive, firms' adoption and export decisions are characterized by cutoff productivities. Only firms with productivity above these cutoffs participate in adoption and exporting. We assume that fixed adoption costs are higher enough than fixed export costs that adopters always export to foreign markets.

The export cutoff $\bar{\phi}^x_{njt}$ is determined at where operating profits in foreign markets are equal to fixed export costs:

$$\bar{\phi}^x_{njt} = \frac{\mu c_{njt} (\sigma c_{njt} F^x_j)^{\frac{1}{\sigma-1}}}{f(\lambda^T_{njt-1}) \left((\tau^x_{nj})^{1-\sigma} D^f_{jt} \right)^{\frac{1}{\sigma-1}}}. \tag{5.5}$$

The adoption cutoff $\bar{\phi}^T_{njt}$ is determined at where profits when adopting technology and profits when not adopting are equalized:

$$\bar{\phi}^T_{njt} = \frac{\mu c_{njt} (\sigma c_{njt} F^T_j)^{\frac{1}{\sigma-1}}}{\left(\left(\frac{\eta}{1 - s_{njt}} \right)^{\sigma-1} - 1 \right)^{\frac{1}{\sigma-1}} f(\lambda^T_{njt-1}) \left(\sum_{m \in \mathcal{N}} \tau^{1-\sigma}_{nmj} P^{\sigma-1}_{mjt} E_{mjt} + (\tau^x_{nj})^{1-\sigma} D^f_{jt} \right)^{\frac{1}{\sigma-1}}}. \tag{5.6}$$

Under the distributional assumption, a share of adopters is expressed as:

$$\lambda^T_{njt} = 1 - G_{njt}(\bar{\phi}^T_{njt}) = \begin{cases} 1 & \text{if} \quad \bar{\phi}^T_{njt} \leq \phi^{min}_{njt} \\ \frac{(\bar{\phi}^T_{njt}/\phi^{min}_{njt})^{-\theta} - \kappa^{-\theta}}{1-\kappa^{-\theta}} & \text{if} \quad \phi^{min}_{njt} < \bar{\phi}^T_{njt} \leq \kappa\phi^{min}_{njt} \\ 0 & \text{if} \quad \bar{\phi}^T_{njt} > \kappa\phi^{min}_{njt}, \end{cases} \tag{5.7}$$

and a mass of adopters is obtained as $M^T_{njt} = M_{nj} \times \lambda^T_{njt}$. Similarly, a share of exporters is $\lambda^x_{njt} = 1 - G_{njt}(\bar{\phi}^x_{njt})$ and a mass of exporters is expressed as $M^x_{njt} = M_{nj} \times \lambda^x_{njt}$.

Dynamic Complementarity. Spillovers generate dynamic complementarity in firms' adoption decisions: a higher share of adopters in the previous period increases gains from adoption in the current period, which is consistent with our third empirical evidence. The dynamic complementarity operates in two ways. The first way drives from complementarity between market size and productivity increases from adoption (Yeaple, 2005; Verhoogen, 2008; Bustos, 2011; Lileeva and Trefler, 2010). Because stronger spillovers increase the productivity of one region relative to other regions, firms in the more productive region will have a larger market and larger gains from adoption due to scale effects. This complementarity further incentivizes firms in the more productive region to adopt technology in the current period, which in turn magnifies spillover in subsequent periods.

The second form of dynamic complementarity derives from reduced fixed adoption costs. Because adoption costs are in units of input bundles, local sectoral aggregate goods are used for fixed adoption costs. Overall increases in productivity due to the spillover lower the costs of local sectoral aggregate goods, which in turn lowers fixed adoption costs (Matsuyama, 1995; Buera et al., 2021). Lower fixed adoption costs induce more firms to adopt technology in the current period, which in turn strengthens the spillover in subsequent periods. The spillover in one region also lowers fixed adoption costs in other regions through trade linkages.

5.3 Households.

Households make decisions of migration and consumption. For tractability, we assume that households are myopic and maximize per-period utility. Households in region n supply labor inelastically and earn wage w_{nt}. Because of the fixed entry assumption, the net profits of firms are redistributed back to households. Each household owns w_{nt} shares of a fund that collects profits from all firms across regions and sectors and redistributes back to households each period (Chaney, 2008).

Households have Cobb-Douglas preferences over final consumption baskets:

$$u(\{C_{jt}\}_{j\in\mathcal{J}}) = \prod_{j=1}^{J} C^{\alpha_j}_{njt}, \quad \sum_{j\in\mathcal{J}} \alpha_j = 1, \tag{5.8}$$

where C_{njt} is the consumption of local sector j aggregate goods and α_j is the final good consumption shares. Households are subject to their budget constraints in each period: $\sum_{j\in\mathcal{J}} P_{njt}C_{njt} = (1-\tau^w_t +$

$\bar{\pi}_t)w_{nt}$, where P_{njt} is the price index of local sector j goods and $(1-\tau_t^w+\bar{\pi}_t)w_{nt}$ is the total income of households, which is the sum of after-tax wages $(1-\tau_t^w)w_{nt}$ and income from dividends $\bar{\pi}_t^h w_{nt}$. We denote the ideal price index for households in region n using $P_{nt}=\prod_{j=1}^{J}P_{njt}^{\alpha_j}$.

At the end of the period, households choose which region to work and live in the next period. After making migration decisions, households supply labor and earn wages. The utility of a household h that lived in region m and moved to region n in period t is

$$\mathcal{U}_{mnt}^h(\epsilon_{nt}^h)=V_{nt}u(\{C_{jt}\}_{j\in\mathcal{J}})d_{mn}\epsilon_{nt}^h=V_{nt}\frac{(1-\tau_t^x+\bar{\pi}_t^h)w_{nt}}{P_{nt}}d_{mn}\epsilon_{nt}^h, \tag{5.9}$$

where V_{nt} is an exogenous amenity in region n, d_{mn} are the utility costs of moving from m to n, and ϵ_{nt}^h is an idiosyncratic preference shock that is independent across households, regions, and periods.

We assume that ϵ_{nt}^h follows a Fréchet distribution with the shape parameter ν: $\epsilon_t^h\sim F(\epsilon)=exp(\epsilon^{-\nu})$, where $\epsilon_t^h=\{\epsilon_{nt}^h\}_{n\in\mathcal{N}}$ (Eaton and Kortum, 2002). Then a share of households moving to region n from region m in period t is given by

$$\mu_{mnt}=\frac{\left(V_{nt}\frac{(1-\tau_t^x+\bar{\pi}_t^h)w_{nt}}{P_{nt}}d_{mn}\right)^{\nu}}{\sum\limits_{n'=1}^{N}\left(V_{n't}\frac{(1-\tau_t^x+\bar{\pi}_t^h)w_{nt}}{P_{n't}}d_{mn'}\right)^{\nu}}. \tag{5.10}$$

The shape parameter ν is migration elasticity that governs the responsiveness of migration flows to real income changes of destination.[25] The population of region n in period t is the sum of all migrants to region n from all other regions in time $t-1$. Therefore, the spatial distribution of population evolves according to the following law of motion:

$$L_{nt}=\sum_{m\in\mathcal{N}}\mu_{mnt}L_{mt-1}. \tag{5.11}$$

Welfare. In each period, the expected utility of each household of region n, prior to realizing idiosyncratic taste shocks ϵ_{nt}^h, is equal to

$$U_{nt}=\mathbb{E}\left[\max_m\left\{\mathcal{U}_{mn,t}^h(\epsilon_{nt}^h)\right\}\right]=\left[\sum_{m\in\mathcal{N}}\left(V_{nt}\frac{(1-\tau_t^x+\bar{\pi}_t^h)w_{nt}}{P_{nt}}d_{mn}\right)^{\nu}\right]^{\frac{1}{\nu}}. \tag{5.12}$$

We define aggregate welfare as the average of U_{nt} weighted by population:

$$U_t^{agg}=\sum_{n\in\mathcal{N}}\frac{L_{nt-1}}{\sum\limits_{m\in\mathcal{N}}L_{mt-1}}U_{nt}. \tag{5.13}$$

[25] Higher ν implies less heterogeneity of preference shocks across households, which makes the utility of households more sensitive to amenity-adjusted real income. Therefore, with higher ν, migration flows will be more sensitive to real income.

Aggregate Variables. For notational convenience, we define the average productivity including subsidies for all firms as follows:

$$\bar{\phi}^{avg}_{njt} = f(\lambda^T_{njt-1})\left[\int_{\phi^{min}_{njt}}^{\bar{\phi}^T_{njt}} \phi_{it}^{\sigma-1} dG_{njt}(\phi_{it}) + \int_{\bar{\phi}^T_{njt}}^{\kappa\phi^{min}_{njt}} \left(\frac{\eta}{1-s_{njt}}\phi_{it}\right)^{\sigma-1} dG_{njt}(\phi_{it})\right]^{\frac{1}{\sigma-1}}.$$

$\bar{\phi}^{avg}_{njt}$ captures the average cost advantage of sector j firms in region n. $\bar{\phi}^{avg}_{njt}$ decreases in $\bar{\phi}^T_{njt}$ (higher λ^T_{njt}), higher s_{njt}, or higher λ^T_{njt-1}. The expression for the average productivity including subsidies for exporters ($\bar{\phi}^{avg,x}_{njt}$) can be defined similarly, but the lower bound is replaced with $\bar{\phi}^x_{njt}$ instead of ϕ^{min}_{njt} because of selection effects induced by fixed exporting costs.

Aggregate variables in this economy can be expressed as a function of $\bar{\phi}^{avg}_{njt}$ and $\bar{\phi}^{avg,x}_{njt}$. Because of the distributional assumptions, these aggregate variables allow for closed-form expressions. See Section B.1 for detailed closed-form expressions for aggregate variables and their derivations. The price index is expressed as

$$P^{1-\sigma}_{njt} = \sum_{m\in\mathcal{N}} \left[M_{mj}\left(\frac{\mu\tau_{mnj}c_{mjt}}{\bar{\phi}^{avg}_{mjt}}\right)^{1-\sigma}\right] + (\tau^x_{nj}c^f_{jt})^{1-\sigma}. \tag{5.14}$$

Region n's share of the total sector j expenditure on goods from domestic region m and from Foreign are expressed as:

$$\pi_{mnjt} = \left(\frac{\tau_{mnj}c_{mjt}/\bar{\phi}^{avg}_{mjt}}{P_{njt}}\right)^{1-\sigma} \qquad \text{and} \qquad \pi^f_{njt} = \left(\frac{\tau^x_{nj}c^f_{jt}}{P_{njt}}\right)^{1-\sigma}. \tag{5.15}$$

Regional gross output for domestic expenditures R^d_{njt} and the total value of exports R^x_{njt} are expressed as:

$$R^d_{njt} = M_{nj}\left(\frac{\mu c_{njt}}{\bar{\phi}^{avg}_{njt}}\right)^{1-\sigma} \sum_{m\in\mathcal{N}} \tau^{1-\sigma}_{nmj} P^{\sigma-1}_{mjt} E_{mjt} \qquad \text{and} \qquad R^x_{njt} = M^x_{njt}\left(\frac{\mu\tau^x_{nj}c_{njt}}{\bar{\phi}^{avg,x}_{njt}}\right)^{1-\sigma} D^f_{jt}. \tag{5.16}$$

The total regional gross output R_{njt} is the sum of R^d_{njt} and R^x_{njt}.

5.4 Equilibrium

Timing. We denote the geographic fundamentals and subsidies across regions and sectors as

$$\mathbf{\Psi}_t = \{\phi^{min}_{njt}, V_{nt}, D^f_{jt}, c^f_{jt}\} \qquad \text{and} \qquad \mathbf{s}_t = \{s_{njt}\}.$$

The timing of this model is as follows. Given $\{\lambda^T_{njt-1}, L_{nt-1}\}$, $\mathbf{\Psi}_t$, and $\mathbf{s}_t$, households make static consumption and migration decisions and firms make static adoption and export decisions in t. These decisions, production, consumption, and wages are determined by the static equilibrium in t, in which households maximize their utility, firms maximize their profits, and market clearing conditions are

satisfied. $\{\lambda^T_{njt}, L_{nt}\}$, which are the outcomes of the static equilibrium in t, become the state variables in $t+1$ and so on.

Static Equilibrium. Given $\{\lambda^T_{njt-1}\}$, $\{L_{nt-1}\}$, $\mathbf{\Psi}_t$, and $\mathbf{s}_t$, firms maximize profits (Equation (5.4)), households maximize utility (Equation (5.8)), and the following market clearing conditions are satisfied each period.

Labor market clearing implies that labor supply is equal to labor demand in each region:

$$w_{nt}L_{nt} = \left[\sum_{j\in\mathcal{J}} \gamma^L_j \left(\frac{\sigma-1}{\sigma}R_{njt} + M^T_{njt}c_{njt}F^T_j + M^x_{njt}c_{njt}F^x_j\right)\right], \tag{5.17}$$

where the right hand side is the sum of labor used for production, fixed export costs, and fixed adoption costs.

The government budget is balanced each period:

$$\tau^w_t \sum_{n\in\mathcal{N}} w_{nt}L_{nt} = \sum_{n\in\mathcal{N}}\sum_{j\in\mathcal{J}^T}\left[\frac{\sigma-1}{\sigma}\frac{s_{njt}-1}{s_{njt}}M_{nj}\int_{\bar{\phi}^T_{njt}}^{\kappa\phi^{min}_{njt}} r(\phi_{it})dG_{njt}(\phi)\right], \tag{5.18}$$

where $r(\phi_{it})$ are firm i's revenues. The left hand side of this equation is total government tax revenue and the right hand side is total government spending.

Region n's total expenditure on sector j goods is the sum of the total expenditure on intermediate inputs and final consumption goods in sector j:

$$E_{njt} = \sum_{k\in\mathcal{J}} \gamma^j_k \left(\frac{\sigma-1}{\sigma}R_{nkt} + M^T_{nkt}c_{nkt}F^T_k + M^x_{nkt}c_{nkt}F^x_k\right) + \alpha_j(1-\tau^w_t+\bar{\pi}^h_t)w_{nt}L_{nt}. \tag{5.19}$$

Goods market clearing implies that region n's total sector j gross output is the sum of the value of total exports and the value of the total demand for region n's sector j goods across the Home regions:

$$R_{njt} = R^x_{njt} + \sum_{m\in\mathcal{N}} \pi_{nmjt}E_{mjt}. \tag{5.20}$$

Labor and goods market clearing conditions imply that trade is balanced.

Dynamic Equilibrium. In this economy, $\{\lambda^T_{njt}, L_{nt}\}$ are dynamic state variables that follow the laws of motions in Equations (5.7) and (5.11), respectively. The law of motion of λ^T_{njt} is the key equation of this model. This equation establishes a relationship between λ^T_{njt-1} to λ^T_{njt} and introduces dynamics in this economy, although all decisions made by agents are static.

We define the dynamic equilibrium of this economy as follows:

Definition 1. *Given initial shares of adopters $\{\lambda^T_{njt_0}\}$ and the path of the geographic fundamentals $\mathbf{\Psi}_t$ and subsidies $\{s_{njt}\}$, a dynamic equilibrium is a path of wages $\{w_{nt}\}$, price indices $\{P_{njt}\}$, a set*

of functions $\{p_{inmjt}(\omega), q_{inmjt}(\omega), p^x_{injt}(\omega), q^x_{injt}(\omega), T_{it}(\omega), x_{it}(\omega)\}$*, labor tax* $\{\tau^w_t\}$*, population* $\{L_{nt}\}$*, and shares of adopters* $\{\lambda^T_{njt}\}$ *such that*

- ***(Static Equilibrium)*** *for each period* t*, (i) firms maximize profits (Equation (5.4)); (ii) households maximize utility by making consumption decisions (Equation (5.8)); (iii) labor markets clear (Equation (5.17)); (iv) goods markets clear (Equation (5.20)); (v) trade is balanced, and (vi) the government budget is balanced (Equation (5.18));*
- ***(Law of Motion of Dynamic State Variables)*** *(vii)* $\{L_{nt}\}$ *follows the law of motion in Equation (5.11); and (viii)* $\{\lambda^T_{njt}\}_{j\in\mathcal{J}^T}$ *follows the law of motion in Equation (5.7).*

Equilibrium conditions (i)-(vi) determine the static equilibrium allocation in each period. Conditions (vii) and (viii) determine the laws of motion for the dynamic state variables.

5.5 Analytical Results: Multiple Steady States

In this subsection, we show that multiple steady states can arise in a simplified model. We consider a closed economy with one sector and one region where labor is the only factor of production. We drop subscripts n and j for notational convenience. We make the following simplifying assumptions:

Assumption 1. *(i)* $|\mathcal{N}| = |\mathcal{J}| = 1$ *and* $\tau^x_{nj} \to \infty$ *(closed economy with one region and one sector); (ii)* $M = 1$ *(normalization); (iii)* $\kappa \to \infty$ *and* $\phi^{min}_t = 1$ *(unbounded Pareto); (iv)* F^T *is in units of final goods (dynamic complementarity); and (v)* $\sigma > 2$ *(uniqueness).*

Assumptions 1(i)-(iii) are imposed for analytical tractability. Under these assumptions, exogenous productivity follows an unbounded Pareto distribution with a normalized location parameter and firm mass is normalized to be one. Assumption 1(iv) is a source of dynamic complementarity in firms' adoption decisions in this environment. With only one region and the CES demand structure, the complementarity between market size and gains from adoption does not operate in this environmen and the dynamic complementarity comes only from fixed adoption costs in units of final goods.[26] Assumption (v) is a sufficient condition to guarantee a unique static equilibrium each period.[27]

[26] With only one region, each firm's increase in productivity due to spillovers is exactly canceled out by competition due to other firms' increases in productivity under the CES demand structure. Therefore, the overall increase in productivity through spillovers does not change the relative market size of each firm, which nullifies the complementarity between market size and gains from adoption. However, because fixed adoption costs are in units of final goods, the spillover from the previous period lowers fixed adoption costs in the current period, which further incentivizes more firms to adopt technology in the current period and generates dynamic complementarity. At the other extreme, when fixed adoption costs are in units of labor, there is no dynamic complementarity, and the equilibrium share of adopters λ^{T*}_t is not affected by λ^T_{t-1}, because overall productivity increase induced by the spillover increases labor demand, which in turn increases wages and the total fixed adoption cost $w_t F^T$. This is formally stated in Section B.2.3. Although we assume that fixed adoption costs are only in units of final goods for simplicity, as long as parts of fixed adoption costs are in units of final goods, the model generates dynamic complementarity.

[27] When a fixed adoption cost is in units of final goods, and $\sigma \leq 2$ holds, multiple static equilibria can arise each period regardless of the existence of the spillover. This is because firms do not take the aggregate price index into account when making adoption decisions. When the share of adopters becomes larger, the aggregate price index becomes lower and this, in turn, decreases fixed adoption costs and vice versa. This degree of responsiveness of the price index to the share of adopters decreases in the values of the elasticity of substitution σ. When σ is sufficiently low, two static

By combining Equations (5.6) and (5.7), we can derive the analytical expression of the short-run equilibrium share of adopters $\lambda_t^{T*} = \lambda_t^{T*}(\lambda_{t-1}^T; \eta, \delta)$ conditional on a share of adopters in the previous period λ_{t-1}^T. The equilibrium share of adopters is determined at $\lambda_t^{T*}(\lambda_{t-1}^T; \eta, \delta) = \min\{\hat{\lambda}_t^T(\lambda_{t-1}^T; \eta, \delta), 1\}$, where $\hat{\lambda}_t^T(\lambda_{t-1}^T; \eta, \delta)$ is implicitly defined by the following equation:

$$\hat{\lambda}_t^T(\lambda_{t-1}^T; \eta, \delta) = \Bigg[\underbrace{A(\hat{\lambda}_t^T(\lambda_{t-1}^T; \eta, \delta))^{2-\sigma} \frac{(\eta^{\sigma-1} - 1)}{\sigma F T} f(\lambda_{t-1}^T)}_{\text{Marginal adopters' net gains from adoption}}\Bigg]^{\frac{\theta}{\sigma-1}},$$

$$\text{where} \quad A(\lambda^T) = \left[\frac{\theta}{\tilde{\theta}}\Big((\eta^{\sigma-1} - 1)(\lambda^T)^{1-\frac{\sigma-1}{\theta}} + 1\Big)\right]^{\frac{1}{\sigma-1}} \quad \text{and} \quad f(\lambda^T) = exp(\delta\lambda^T). \tag{5.21}$$

The equilibrium share is characterized by the cutoff productivity level which is determined by the point at which the net gains from adoption for marginal adopters are equal to zero. Similarly, the time-invariant steady state share of adopters satisfies $\lambda^{T*} = \lambda_t^{T*} = \lambda_{t-1}^{T*}$ and is determined by $\lambda^{T*} = \lambda^{T*}(\lambda^{T*}; \eta, \delta)$.

Given any initial shares of adopters, this economy has a unique deterministic equilibrium path to the steady state due to Assumption 1(v). Because static equilibrium is unique each period, there exists a unique sequence of static equilibrium that forms a unique deterministic dynamic path. λ_t^{T*} increases in λ_{t-1}^T due to dynamic complementarity. λ_t^{T*} also increases in two parameters: η and δ. η and δ increase λ_t^{T*} by increasing the net gains for marginal adopters and magnifying dynamic complementarity, respectively.[28] Most importantly, we show that multiple steady can arise due to the dynamic complementarity in this economy. When multiple steady states exist, these steady states can be Pareto-ranked based on the steady state share of adopters, and the initial share of adopters determines which steady state is realized in the long-run. Proposition 1 summarizes these results.

Proposition 1. *Under Assumption 1,*

(i) (Uniqueness) Given any initial shares of adopters $\lambda_{t_0}^T$, there exists a unique dynamic equilibrium;

equilibria can arise where one has a higher share of adopters and the other has a lower share. By imposing $\sigma > 2$, we are ruling out the possibility of these multiple static equilibria. Matsuyama (1995) and Buera et al. (2021) have studied these types of multiple equilibria. Our model differs from their models because our multiple long-run steady states are generated because of the local spillover. Also, it is natural to assume $\sigma > 2$ because commonly calibrated parameter values for σ are larger than 2 (Broda and Weinstein, 2006).

[28] Note that two terms are related to the direct productivity gains governed by η in Equation (5.21): $(\eta^{\sigma-1} - 1)$ and $A(\lambda_t^T)^{2-\sigma}$. The term $(\eta^{\sigma-1} - 1)$ captures marginal adopters' net gains from adoption, which is internalized by their adoption decisions. The term $A(\lambda_t^T)^{2-\sigma} = A(\lambda_t^T)^{1-\sigma} \times A(\lambda_t^T)$ captures two composite general equilibrium effects of direct productivity gains. These two general equilibrium effects work in the opposite directions in the net gains for marginal adopters. First, $A(\lambda_t^T)^{1-\sigma}$ captures competition effects, which decrease in λ_t^T. As more firms adopt technology (increases in λ_t^T), the productivity of competitors increases, which in turn intensifies competition across firms and decreases marginal adopters' net gains due to increased competition (or decreased market size). The second general equilibrium effect is that a higher share of adopters decreases the price of final goods and therefore fixed adoption costs (Assumption 1(iv)). Assumption 1(v) ensures that the first general equilibrium effect dominates this second general equilibrium effect.

(ii) (Dynamic Complementarity) $\frac{\partial \hat{\lambda}_t^T(\lambda_{t-1}^T;\eta,\delta)}{\partial \lambda_{t-1}^T} > 0$*;*

(iii) (Comparative Statistics) $\frac{\partial \hat{\lambda}_t^T(\lambda_{t-1}^T;\eta,\delta)}{\partial \eta} > 0$ *and* $\frac{\partial \hat{\lambda}_t^T(\lambda_{t-1}^T;\eta,\delta)}{\partial \delta} > 0$*;*

(iv) (Multiple Steady States) There exist intervals $[\underline{\delta}, \bar{\delta}]$ *and* $[\underline{\eta}, \bar{\eta}]$ *such that holding other parameters constant, multiple steady states arise only for* $\delta \in [\underline{\delta}, \bar{\delta}]$ *and* $\eta \in [\underline{\eta}, \bar{\eta}]$*;*

and (v) (Pareto-Ranked) If multiple steady states exist, these steady states can be Pareto-ranked based on the equilibrium share of adopters.

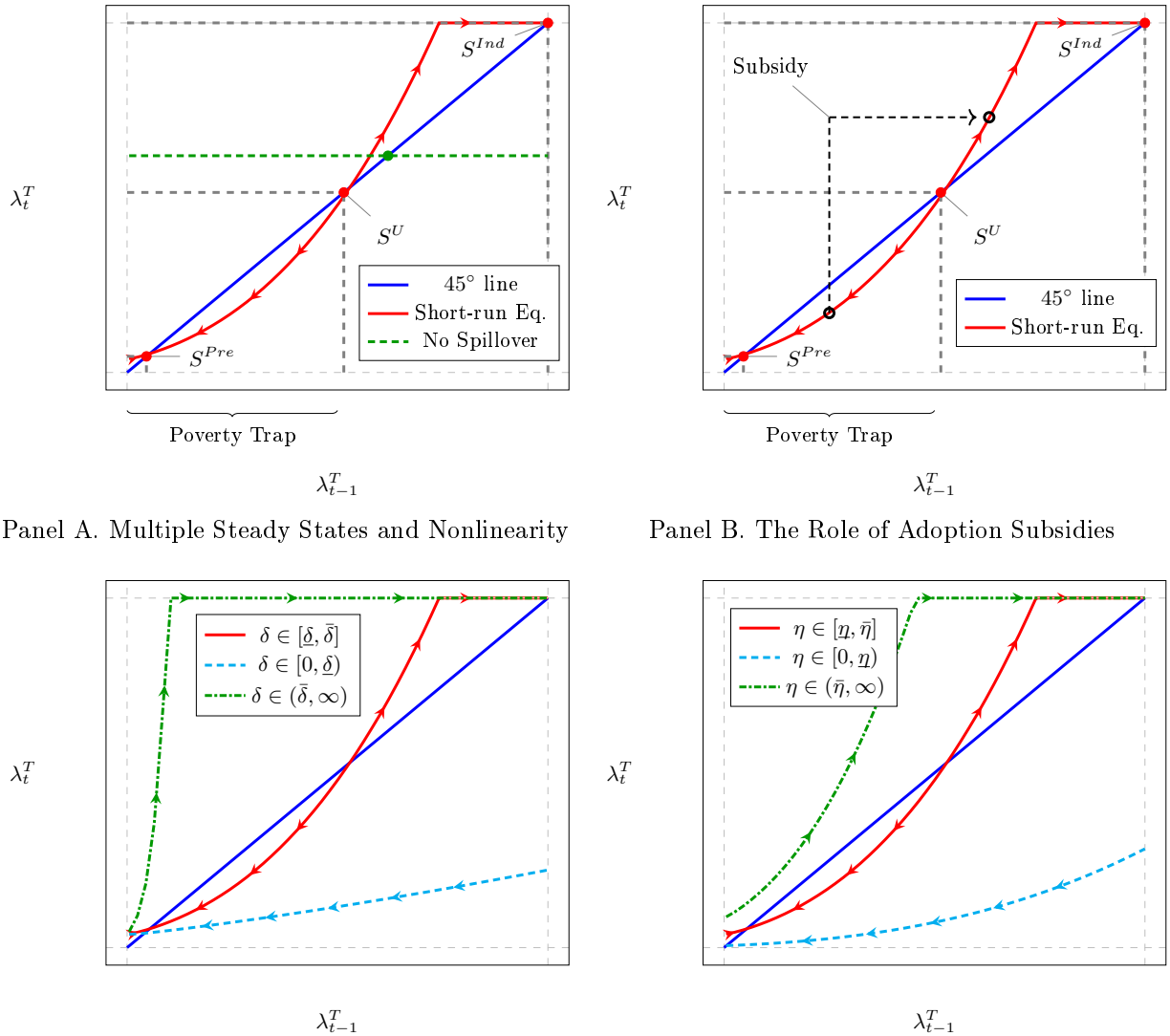
The case of multiple steady states is illustrated in Panel A of Figure 3, where there are three different steady states with two basins of attraction.[29] The red locus is defined by Equation (5.21), where each point on the locus is a short-run equilibrium given λ_{t-1}^T. Given λ_{t-1}^T, λ_t^{T*} is determined in period t; and then given λ_t^{T*}, λ_{t+1}^{T*} is determined in the next period $t+1$; and so on. Therefore, the equilibrium moves along the red locus as time passes. The steady state is determined at the point where $\lambda_{t-1}^{T*} = \lambda_t^{T*}, \forall t$ holds; that is, where the red locus intersects with the 45-degree blue line. There are three intersection points: S^{Pre}, S^U, and S^{Ind}, which we label as the pre-industrialized, unstable, and industrialized steady states, respectively.

Because technology adoption increases firms' productivity, these steady states can be Pareto-ranked depending on the steady state share of adopters. At S^{Ind}, all firms adopt technology, and at S^{Pre} there is a smaller share of adopters than the other two, so S^{Ind} Pareto-dominates the other two steady states and S^{Pre} is Pareto-dominated by the other two. S^U is unstable in that the economy converges to this steady state only when the initial condition is given by the value of S^U. The nonlinearity of the red locus means that it intersects with the 45-degree line multiple times and generates the multiple steady states, where the spillover ($\delta > 0$) generates such nonlinearity. For example, if there is no spillover ($\delta = 0$), there is always a unique steady state illustrated by the intersection of the green dashed horizontal line and the 45-degree line. When there is no spillover, the equilibrium share of adopters is determined regardless of the share of adopters in the previous period, which gives the horizontal line in the graph.

For the initial conditions given by $\lambda_{njt_0}^T \geq S^U$, the economy converges to S^{Pre}, and for $\lambda_{njt_0}^T \geq S^U$, the economy converges to S^{Ind}. Because firms do not internalize the spillover, if an economy is locked into the region $\lambda_{njt_0}^T \geq S^U$, it converges to S^{Pre}, although an economy has the potential to reach S^{Ind}. This region is known as a poverty trap in the literature (Azariadis and Stachurski, 2005).

Multiple Steady States and the Permanent Effects of Temporary Subsidies. When multiple steady states exist, temporary subsidies for technology adoption in the initial period can have permanent effects that move an economy that is initially in the poverty trap to a new transition

[29] In this economy, there are at most three multiple steady states because of the functional form assumption imposed on the spillover: $f(\lambda_{t-1}^T) = exp(\delta \lambda_{t-1}^T)$. The imposed spillover functional form makes λ_t^T strictly convex in λ_{t-1}^T so that the red locus in Figure 3 intersects with the 45-degree line two times at the most. With a functional form assumption that generates a larger degree of nonlinearity, it is possible to have greater or fewer multiple steady states than three.



Panel A. Multiple Steady States and Nonlinearity

Panel B. The Role of Adoption Subsidies

Panel C. Comparative Statistics of δ

Panel D. Comparative Statistics of η

Figure 3. Multiple Steady States and Comparative Statistics

Notes. In Panel A, the multiple steady states arise only when the short-run equilibrium curve is sufficiently nonlinear. In Panel B, temporary subsidies can have permanent effects by moving an economy to a new transition path that converges to the higher-productivity steady state S^{Ind}. In Panels C and D, the multiple steady states arise only for the medium range of values of η and δ.

path that converges to an industrialized steady state. This is illustrated in Panel A of Figure 3. Suppose the initial condition is given as $\lambda^T_{njt_0} < S^U$, so that an economy converges to S^{Pre}. However, if the government implements a one-time policy that subsidizes technology adoption, this can shift the share of adopters above the S^U level, which causes an economy to converge to the industrialized steady state S^{Ind}. This can rationalize South Korea's pattern of industrialization toward heavy manufacturing sectors and the temporary policy between 1973 and 1979.

In this model, only multiple steady states can rationalize the permanent effects of temporary subsidies.[30] When there is only one steady state, subsidies temporarily shift the short-run equilibrium curve while they are provided, but the curve moves back to the original position after subsidies end and the economy converges to its original steady state.

Comparative Statistics. Proposition 1(iv) implies that multiple steady states arise only for the medium ranges of $\delta \in [\underline{\delta}, \bar{\delta}]$ and $\eta \in [\underline{\eta}, \bar{\eta}]$; that is when spillovers or direct productivity gains are neither too strong nor too weak. If these values are too high or too low, the dynamic complementarity becomes too strong or too weak and cannot generate enough nonlinearity of the short-run equilibrium locus, which means that it intersects with the 45-degree line only once. This is graphically illustrated in Panels C and D of Figure 3.

The comparative statistics offer one potential explanation for why the South Korean economy experienced remarkable transformation toward heavy manufacturing sectors when other developing countries did not. Both η and δ can be country-specific and depend on specific features of each country. η is generally related to the absorptive capacity of new technology and δ is related to degree of barriers to knowledge diffusion. For instance, countries with lower amounts of skilled labor endowments and higher language barriers may have lower values of η and δ. Compared to other developing countries, South Korea had higher amounts of skilled labor and used the same language (Rodrik, 1995), which can make South Korea have higher values of η and δ. South Korea could have been a special case because its values of η and δ were in a range that generated multiple steady states.

6 Taking the Model to the Data

In our quantitative exercises, we aggregate sectors into four categories: commodity, light manufacturing, heavy manufacturing, and service sectors. Given that most of the adoption occurred in the heavy manufacturing sectors, we assume that technology adoption is available only for the heavy manufacturing sector. The service sector is nontradable across regions and countries. We also aggregate the data to 42 regions so that each region has at least two firms in each sector based on the

[30] Even if a unique steady state exists, there is room for policy interventions because of externalities. However, when a unique steady state exists, these policy interventions should be implemented in every single period in order to produce permanent effects. This point is graphically illustrated in Figure B1. Kline and Moretti (2014), who studied the Tennessee Valley Authority program in the United States, did not detect nonlinearities in the agglomeration elasticity, so they concluded that the program did not have permanent effects because the agglomeration elasticity was not nonlinear enough to generate multiple steady states.

administrative divisions in the 1970s and on electoral districts. One period in the model corresponds to 4 years in the data, so the timing of the spillover in the model is consistent with the spillover estimates in Section 4.2.

The model is fully parametrized by subsidies $\mathbf{s}_t$, geographic fundamentals $\mathbf{\Psi}_t$, and the following structural parameters

$$\mathbf{\Theta} = \{ \underbrace{M_{nj}}_{\substack{\text{Fixed}\\ \text{firm mass}}} , \underbrace{\theta, \kappa}_{\substack{\text{Pareto}\\ \text{distribution}}} , \underbrace{\eta, \delta, F_j^T}_{\substack{\text{Technology}\\ \text{adoption}}} , \underbrace{\sigma, \gamma_j^k, \gamma_j^L}_{\text{Production}} , \underbrace{\tau_{nmj}, F_j^x, \tau_{nj}^x}_{\text{Trade costs}}, \underbrace{\nu, d_{nm}}_{\substack{\text{Spatial}\\ \text{mobility}}}, \underbrace{\alpha_j}_{\text{Preferences}} \}.$$

We divide the set of structural parameters Θ into two subgroups depending on whether they are externally or internally calibrated:

$$\mathbf{\Theta}^E = \{ \underbrace{\underbrace{\eta, \delta,}_{\substack{\text{Reduced-form}\\ \text{estimates}}} \underbrace{M_{nj}, \theta, \sigma, \gamma_j^L, \gamma_j^k, \nu, d_{nm}, \tau_{nmj}, \tau_{nj}^x, \alpha_j}_{\text{Standard in the literature}}}_{\text{Externally calibrated parameters}} \} \quad \text{and} \quad \mathbf{\Theta}^M = \underbrace{\{\kappa, F_j^x, F_j^T\}}_{\substack{\text{Internally}\\ \text{calibrated}\\ \text{parameters}}}.$$

Our calibration procedure proceeds in two steps. First, we externally calibrate $\mathbf{\Theta}^E$, of which η and δ can be mapped to the reduced-form estimates in Section 4 and the remaining parameters are standard in the literature. Second, we internally calibrate $\mathbf{\Theta}^M$, subsidy $\mathbf{s}_t$, and geographic fundamentals $\mathbf{\Psi}_t$ using the method of moments.

6.1 Externally Calibrated Parameters

Technology Adoption $\{\eta, \delta\}$. η and δ are parameters that govern the magnitude of direct productivity gains and spillovers. The reduced-form estimates of direct productivity gains and spillovers in Section 4 can be mapped to η and δ of the model. Taking the log of adopters' sales, we can derive the following regression model:

$$\log Sales_{it} = (\sigma - 1)\log(\eta) T_{it} + \underbrace{\delta \lambda_{njt}^T + \log\Big(\sum_{m \in \mathcal{N}} \tau_{nmj} P_{mjt}^{\sigma-1} E_{mjt} + \tau_{nj}^x D_{jt}^f \Big)}_{\text{Absorbed out by exactly matching on region and sector}} + (\sigma - 1)\log \phi_{it},$$

which can be mapped to our winners vs. losers specification (Equation (4.1)). By exactly matching on region and sector, we absorb out the spillover, the unit cost of production, and the market size that are common across firms within regions and sectors. Exogenous cancellations by foreign firms can be interpreted as a shock to the fixed adoption cost F_j^T in our model framework that is orthogonal to firms' productivity $\log \phi_{it}$. We calibrate η using the point estimate of β_4^{diff} in Equation (4.1). β_4^{diff} is consistent with one period of the model. Based on column (1) of Table 1, we set $\eta = exp(0.50)/(\sigma - 1)$.

Similarly, taking the log on non-adopters' sales, we can derive the following regression model:

$$\log Sales_{it} = (\sigma - 1)\delta\lambda^T_{njt} + \log\Big(\sum_{m\in\mathcal{N}} \tau_{nmj} P^{\sigma-1}_{mjt} E_{mjt} + \tau^x_{nj} D^f_j\Big) + (\sigma - 1)\log\phi_{it}.$$

Although this is similar to our reduced-form specification for spillover (Equation (4.3)), they differ in terms of variation in spillovers within regions and sectors. λ^T_{njt} is common within regions and sectors in the model, whereas the spillover ($\text{Spill}_{inj(t-h)}$) in Equation (4.3) differs across firms within region-sector depending on their distances from adopters. To connect the model to the data, we rely on the fact that the reduced-form estimates of spillovers can be interpreted as the semi-elasticity of the local share of adopters when distances between firms are equal. We rely on this interpretation and assume that firms in the model are equally distant from each other in a finite set of regions. We set δ to be $4.5/(\sigma - 1)$, which is the average value of estimates of spillovers in columns (1)-(5) of Table 2.

Spatial Mobility $\{\nu, d_{mn}\}$. We parametrize migration costs as a function of geographic distance: $d_{nm} = (dist_{nm})^\zeta \times \epsilon^d_{nm}$, where $dist_{nm}$ is the distance between regions n and m and ϵ^d_{nm} is a residual that is not explained by distance. We set ν to be 2, which is the estimate from Peters (2021). Using Equation (5.10), we derive a gravity equation for migration flows. ζ is externally calibrated by estimating this gravity equation. Using migration flows from 1990 to 1995, we run the following regression model:

$$\log \mu^{1995}_{nm1990} = -\nu\zeta \log dist_{nm} + \delta_m + \delta_n + \epsilon^d_{mn}, \tag{6.1}$$

where μ^{1995}_{nm1990} represent shares of migrants from region n to region m and δ_n and δ_m are region fixed effects.[31] To address attenuation bias arising from statistical zeros in the gravity models, we estimate the equation using the PPML. Under the assumed value for ν, we obtain the value for ζ from the estimated coefficients. The estimated coefficient is statistically significant at 1% level and implies that $\hat{\zeta} = 1.39/\nu$.

Variable Trade Costs $\{\tau_{nmj}, \tau^x_{nj}\}$. We parametrize variable internal trade costs as a function of the geographic distance $\tau_{nmj} = (dist_{nm})^\xi$ and assume that ξ is the same across different sectors. We do not observe internal trade flows, so we borrow the estimates from the literature. We use the distance elasticity estimate from Monte et al. (2018) and set $\xi = 1.29/(\sigma - 1)$.

For international trade costs, we assume that firms have to ship their products to the nearest port and then pay both variable and fixed international trade costs at the port when they export or import. Under this assumption, international variable trade costs can be parametrized as $\tau^x = \tilde{\tau}^x \times (dist^{port}_n)^\xi$.

[31] The estimation procedure is described in detail in Section D.5. The data on migration shares comes from the 1995 Population and Housing Census, which was the closest to our sample periods among the accessible population census data. Because of data availability, regions are aggregated up to 35 regions. μ^{1995}_{nm1990} is obtained as the total number of migrants who moved from region n to region m in the period 1990 to 1995 divided by the total population of region n in 1990. When we compute the total number of population and migrants, we restrict our sample age to 20 to 55.

$\tilde{\tau}^x$ represent variable costs incurred at the port. Any variable trade costs that are common across regions are not separately identifiable from D^f_{jt}, so we normalize $\tilde{\tau}^x$ to be 1. $(dist^{port}_n)^\xi$ represent variable costs associated with shipping from region n to the nearest port among the seven largest ports in South Korea, where $dist^{port}_n$ is the distance between region n and the nearest port and ξ is the same parameter of the parametrization of internal trade costs.

The Remaining Parameters $\{\sigma, \theta, M_{nj}, \alpha_j, \gamma^L_j, \gamma^k_j\}$. The remaining parameters are the elasticity of substitution, Pareto shape parameter, exogenous firm mass, and Cobb-Douglas shares of preference and production function. Following Broda and Weinstein (2006), we set the elasticity of substitution σ to be 4. We set the Pareto shape parameter θ to be $1.06 \times (\sigma - 1)$ (Axtell, 2001).[32] We set M_{nj} to be proportional to the GDP share of each region and sector and set $\sum_{n \in \mathcal{N}} M_{nj} = 1$ following Chaney (2008). The Cobb-Douglas shares of preference (α_j) and production function (γ^k_j and γ^L_j) are taken from the input-output table for 1972.

6.2 Internally Calibrated Parameters

$\mathbf{\Theta}^M = \{F^x_j, F^T_j, \kappa\}$, $\mathbf{s}_t = \{s_{njt}\}$, and $\mathbf{\Psi}_t = \{\phi^{min}_{njt}, V_{nt}, D^f_{jt}, c^f_{jt}\}$ are calibrated using the method of moments. Our calibration procedure requires moments from firm-level data and a set of cross-sectional aggregate variables in 1972, 1976, and 1980 which cover the periods when the subsidies were provided between 1973 and 1979. The required set of aggregate variables include region-sector level gross output $\{R_{njt}\}$, regional population distribution $\{L_{nt}\}$, aggregate export and import shares, initial shares of adopters $\{\lambda^T_{nj-1}\}$ and initial population distribution $\{L_{n,-1}\}$. $\{\lambda^T_{nj-1}\}$ and $\{L_{n,-1}\}$ are taken as given when solving the model for $t = 1$. Section D.2 explains the algorithm of the calibration procedure and how we construct the data inputs in detail.

Constrained Minimization Problem. We calibrate $\mathbf{\Theta}^M$, $\mathbf{\Psi}_t$, and $\mathbf{s}_t$ by solving the following constrained minimization problem:

$$\{\hat{\mathbf{\Theta}}^M, \hat{\mathbf{s}}_t\} \equiv \underset{\mathbf{\Theta}^M, \mathbf{s}_t}{\arg\min}\{L(\mathbf{\Theta}^M, \mathbf{s}_t)\} = \underbrace{(\bar{m} - m(\mathbf{\Theta}^M, \mathbf{s}_t, \mathbf{\Psi}_t))'\mathbf{W}(\bar{m} - m(\mathbf{\Theta}^M, \mathbf{s}_t, \mathbf{\Psi}_t))}_{\text{Micro moments}}$$
$$\text{subject to} \quad \underbrace{\mathbf{C}(\mathbf{\Theta}^M, \mathbf{s}_t, \mathbf{\Psi}_t) = \mathbf{C}_t}_{\text{Aggregate data}}, \quad t \in \{1, 2, 3\}. \tag{6.2}$$

$\mathbf{W}$ is a weighting matrix. $\bar{m}$ and $m(\mathbf{\Theta}^M, \mathbf{s}_t, \mathbf{\Psi}_t)$ are the model moments and data counterparts of the objective function. $\mathbf{C}(\mathbf{\Theta}^M, \mathbf{s}_t, \mathbf{\Psi}_t) = \mathbf{C}_t$ are the imposed constraints. $\mathbf{C}(\mathbf{\Theta}^M, \mathbf{s}_t, \mathbf{\Psi}_t)$ and $\mathbf{C}_t$ are the model moments and data counterparts of the constraints. For the weighting matrix, we use the identity matrix.

[32] Under the Pareto distributional assumption with shape parameter θ, the distribution of firm sales follows the Pareto distribution with shape parameter $\theta/(\sigma - 1)$. Many previous studies that have estimated θ using firm sales distribution have found that $\theta/(\sigma - 1)$ is close to 1 (Axtell, 2001; di Giovanni et al., 2011; di Giovanni and Levchenko, 2012, 2013).

Identification of Subsidies. We do not observe subsidies directly in the data. Following the historical narrative, subsidies are provided only in $t = 2, 3$ in the model, which corresponds to 1976 and 1980 in the data. Given the lack of information on the distribution of subsidies across regions, we assume that the government provided the same subsidy level $\bar{s}$ across regions in $t = 2, 3$:

$$s_{njt} = \begin{cases} \bar{s} & \text{if} \quad t \in \{2, 3\}, \quad \forall n \in \mathcal{N}, \quad \forall j \in \mathcal{J}^T \cap \mathcal{J}^{pol} \\ 0 & \text{otherwise,} \end{cases} \tag{6.3}$$

where $\mathcal{J}^{pol}$ is a subset of sectors that the government targeted.

Despite the lack of data on subsidies, with the above parametrization of subsidies, we can identify $\bar{s}$ using the model structure and the reduced-form estimates that measure direct and spillover benefits from adoption. Our intuition is that given information about the benefits from adoption (direct productivity gains and spillovers), the increases in shares of adopters in 1976 or 1980 relative to 1972 are attributable to the policy. Based on this intuition, we calibrate the subsidy level by indirect inference. In particular, using both actual and model-generated data, we estimate the following regression for only the heavy manufacturing sector in 1972 and 1980 using the Poisson pseudo-maximum likelihood (PPML) to incorporate zeros (Silva and Tenreyro, 2006):

$$\ln \lambda^T_{n,heavy,t} = \alpha + \beta^{pol} \times D^{pol}_t + \beta_1 \lambda^T_{n,heavy,t-1} + \epsilon_{n,heavy,t}. \tag{6.4}$$

Then, we use the estimated β^{pol} from the actual data as the identifying moment for $\bar{s}$ and fit the estimated β^{pol} from the model-generated data to this moment (Nakamura and Steinsson, 2018). We formally describe the intuition behind the identifying moment and calibration procedure in Section D.4.

Objective Function: Micro Moments, $\{\boldsymbol{\Theta}^M, \bar{s}\}$. We identify $\boldsymbol{\Theta}^M = \{F^T_j, F^x_j, \kappa\}$ and subsidy rate $\bar{s}$ using micro moments. $\bar{s}$ is identified by the identifying moment discussed above. We identify fixed adoption costs F^T_j using the median of shares of adopters across regions in 1972 and 1980. We identify κ using the share of regions with zero adoption in 1972 and 1980. κ rationalizes zero adoption in some regions observed in the data. If κ is sufficiently low; that is, if the gap between the Pareto lower and upper bounds becomes narrower, the cutoff adoption productivity becomes larger than the Pareto upper bound $\kappa\phi^{min}_{njt}$ for some regions, resulting in zero adoption in these regions. We calibrate fixed export costs of the light and heavy manufacturing sectors F^x_j to match the median shares of exporters across regions in 1972.[33] Because we do not have detailed data on firms in the commodity sector, we set the fixed export costs of the commodity sector to be the same as those of

[33] Our firm balance sheet data has information on exports. However, many observations were missing. Given that export data are very noisy, we do not use export information for our reduced-form empirical analysis, but only for computing the moment on shares of exporters for our quantitative analysis.

Table 4: Calibration Strategy

	Description	Value	Identification / Moments
			External calibration
	Structural parameters		
η	Direct productivity gains	1.3	Winners vs. Losers, Table 1 col. 1, $(\sigma-1)\log(\eta)=0.5$
δ	Spillover semi-elasticity	2.25	Spillover estimate, Table 2, $4.5=(\sigma-1)\delta$
σ	Elasticity of substitution	3	Broda and Weinstein (2006)
θ	Pareto shape parameter	2.12	Axtell (2001), $\theta/(\sigma-1)=1.06$
ν	Migration elasticity	2	Peters (2021)
ζ	Migration cost, $d_{mn}=(dist_{nm})^{\zeta}$	0.78	Gravity estimates
ξ	Internal and international trade costs	0.43	Monte et al. (2018)
α_j	Preferences		IO table, 1972
γ_j^k	Production		IO table, 1972
M_{nj}	Exogenous firm mass		Value added, 1972 (Chaney, 2008)
			Internal calibration: Method of moment
	Structural parameters		
F_j^T	Fixed adoption cost	0.28	Share of adopters, heavy mfg.
F_j^x	Fixed export cost, commodity & light mfg.	0.06	Share of exporters, light mfg.
F_j^x	Fixed export cost, heavy mfg.	0.05	Share of exporters, heavy mfg.
κ	Pareto upper bound	4.42	# of regions with zero adoption
	Geographical fundamentals		
ϕ_{nj}^{min}	Natural advantage (Pareto lower bound)		Dist. region & sector sales, 1972, 1976, 1980
D_j^f	Foreign market size		Sectoral export intensity, 1972, 1976, 1980
c_j^f	Foreign price of imported inputs		Sectoral import intensity, 1972, 1976, 1980
V_{nt}	Amenity		Pop. dist., 1972, 1976, 1980
	Subsidy		
$\bar{s}$	Subsidy rate	0.11	Identifying moment $\hat{\beta}^{pol}$, Equation (6.4)

Notes. This table reports calibrated objects of the model, their values, and their identifying moments.

the light manufacturing sector.

Constraints: Aggregate Data, $\boldsymbol{\Psi}_t$. The constraints in Equation (6.2) identify geographic fundamentals $\boldsymbol{\Psi}_t$. We impose the constraints such that shares of gross output at the region and sector levels, aggregate export and import shares, and regional population distribution of the model (Equations (5.10), (5.15), (5.16)) are exactly fitted to the counterpart of the data in 1972, 1976, and 1980. The number of constraints is the same with the dimension of the geographic fundamentals. Therefore, for any given parameters $\boldsymbol{\Theta}^M$ and subsidy rate $\bar{s}$, the geographic fundamentals are exactly identified by these constraints and there exists a set of geographic fundamentals that rationalizes the data.

Table 5: Model Fit

Moment	Model	Data
Identifying moment $\hat{\beta}^{pol}$, Equation (6.4)	0.65	0.83
Med. shares of exporters in 1972, light mfg.	0.22	0.21
Med. shares of exporters in 1972, heavy mfg.	0.14	0.18
Med. shares of adopters in 1972	0.06	0.07
Med. shares of adopters in 1982	0.12	0.19
Share of zero adoption regions in 1972	0.59	0.53
Share of zero adoption regions in 1982	0.83	0.94

Notes. This table presents the values of the internally calibrated parameters and their identifying moments in the data.

Because geographic fundamentals are exactly identified, we can identify the average productivity including subsidies $\bar{\phi}^{avg}_{njt}$ following the model-inversion logic from Allen and Arkolakis (2014).[34] However, we cannot identify what portion of $\bar{\phi}^{avg}_{njt}$ is attributable to natural advantages ϕ^{min}_{njt}, shares of adopters λ^{T}_{njt}, or subsidies $\bar{s}$ from aggregate data alone. To isolate ϕ^{min}_{njt}, λ^{T}_{njt}, and $\bar{s}$ from $\bar{\phi}^{avg}_{njt}$, we need information on fixed adoption costs F^{T}_{j} and subsidies $\bar{s}$ from the micro moments.

6.3 Calibration Results and Model Fit

Table 4 presents the summary of our calibration strategy and the values of the externally and internally calibrated parameters. The estimated adoption cost is 5.6 times larger than the estimated fixed export cost. The estimated subsidy rate is 0.11, which indicates that adopters are subsidized with 11% of input expenditures. Table 5 reports the model fit. The data moments are well-approximated in the model.

Non-targeted Moments. Our calibration strategy only fits the cross-sectional data for 1972, 1976, and 1980 and does not fit the evolution of variables after 1980. Also, we do not target employment directly. However, our model fits the evolution of heavy manufacturing's share of GDP quite well even after 1980 and the evolution of its share of employment between 1972 and 1980 (Panels A and B of Figure 4), which are non-targeted moments.

In Figure D1, we compare regional shares of the heavy manufacturing sector's gross output computed from the 2004 Mining and Manufacturing Survey and those calculated from the model of

[34] By fitting the input-output tables, we can only identify relative productivity differences across regions and sectors, but we cannot identify aggregate shifters of productivity. Thus, when we fit gross output shares at regional and sectoral levels, we normalize ϕ^{min}_{njt} of one region and sector pair to 1 for each period. This is not a big concern because our interest is the comparison between the baseline economy and the counterfactual economy, which differences out the common aggregate components.

the corresponding model period. Although we do not directly target the spatial distribution of the gross output of the heavy manufacturing sector, the spatial distribution computed from the model is qualitatively and quantitatively very similar to that observed in the data.

7 The Aggregate and Regional Effects of the Temporary Adoption Subsidy

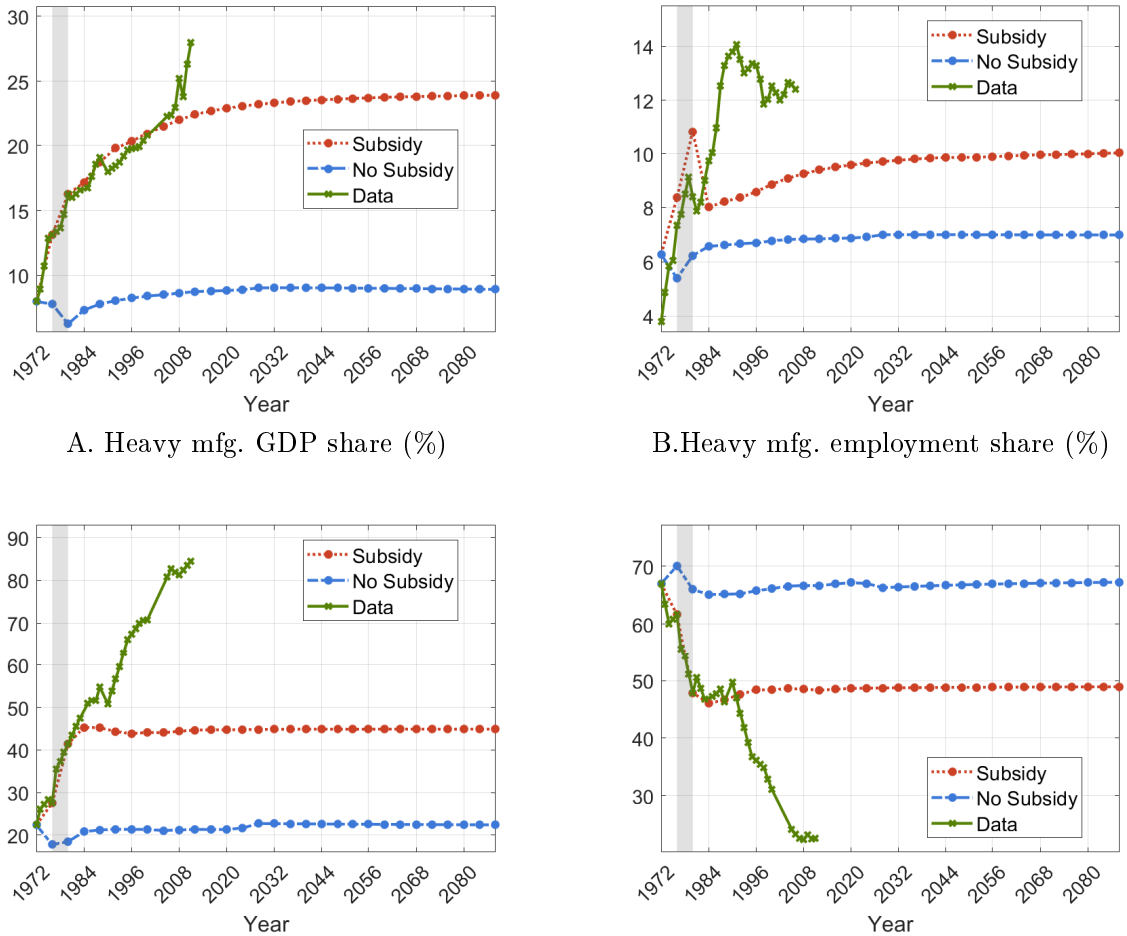
In this section, we ask how the aggregate and regional patterns of industrialization in South Korea would have evolved differently if the temporary subsidies had not been provided. In the baseline economy, the subsidies are provided, whereas the subsidies are not provided in the counterfactual economy. We compare these baseline and the counterfactual economies. Unlike the simplified model in Section 5.5 where there is a maximum of three steady states, the full quantitative model potentially admits a larger number of steady states. Which steady state will be reached in the long-run is of computational question, given calibrated values of $\{\mathbf{\Psi}_t, \bar{s}, \mathbf{\Theta}\}$ that are chosen to match cross-sectional data in 1972, 1976, and 1980 rather than chosen arbitrarily.

Figure 4 reports our main counterfactual results. In Panels A, B, C, and D, we compare the heavy manufacturing's shares of GDP, employment, and exports, and the light manufacturing's shares of exports. Had temporary adoption subsidies not been provided, South Korea's pattern of industrialization and its comparative advantage would have evolved differently. When compared to the steady state of the baseline economy, heavy manufacturing's share of GDP would have decreased by 15 percentage points, its share of employment would have decreased by 3 percentage points, and its share of exports would have decreased by 22.5 percentage points, and these changes would have been permanent in the steady state of the counterfactual economy. The reason why our model does not explain the evolution of the shares of employment and export after 1980 well is that we do not directly target evolution of the model after 1980.

Panel A of Figure 5 reports the average productivity of each region under the baseline and counterfactual economies. We define the average productivity as $M_{nj}[\int z_{it}(\phi)^{\sigma-1} dG_{njt}(\phi)]^{1/(\sigma-1)}$.[35] The x and y axes are the average productivity of the heavy manufacturing sector in each region under the baseline and counterfactual economies, respectively. Dots located below the 45 degree line, denoted as red stars, represent regions that have higher levels of productivity in the baseline when compared to the counterfactual. The figure shows that only five regions have higher productivity levels in the steady state of the baseline economy when compared to that of the counterfactual economy. Most of the regions have the same level of productivity in both steady states. This implies that the aggregate industrialization toward the heavy manufacturing sector in the baseline economy is driven by large productivity increases of these five regions.

Panel B of Figure 5 plots the regional welfare gains in the steady states. The x and y axes are the

[35] Because M_{nj} and ϕ_{njt}^{min} are not separately identifiable under the fixed entry, $M_{nj}[\int z_{it}(\phi)^{\sigma-1} dG_{njt}(\phi)]^{1/(\sigma-1)}$ can be considered to be the average productivity when $M_{nj} = 1, \forall n, j$.



A. Heavy mfg. GDP share (%)

B.Heavy mfg. employment share (%)

C. Heavy mfg. export share (%)

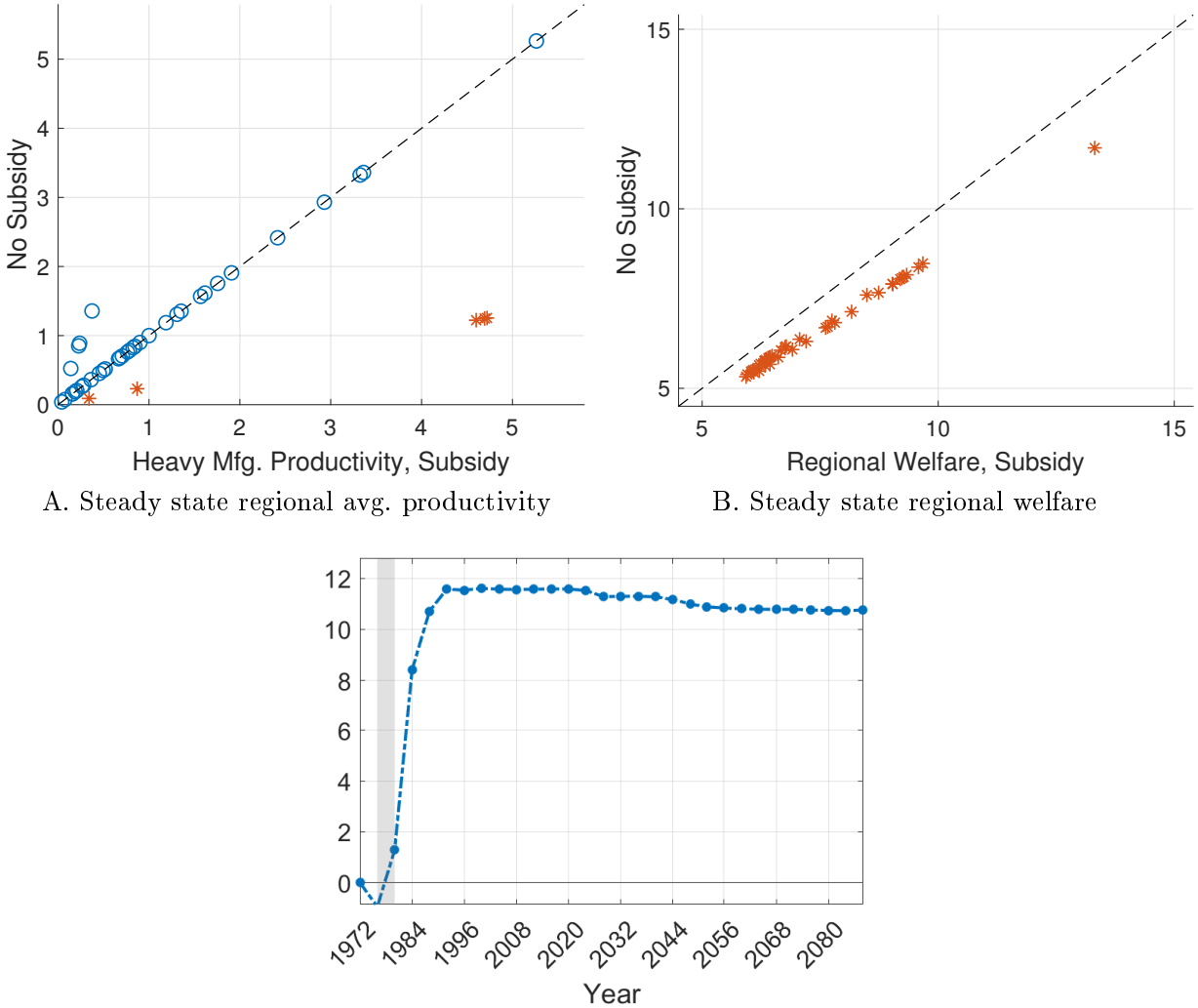
D. Light mfg. export share (%)

Figure 4. Results of the Counterfactual Analysis

Notes. This figure plots the counterfactual results. The green solid line plots the actual data computed from the input-output tables. The red dotted line plots the outcomes of the baseline economy and the blue dotted line plots the outcomes of the counterfactual economy.

regional welfare in the steady state under the baseline and counterfactual economies, respectively. In the steady states, all regions have higher welfare levels in the baseline than the counterfactual. Large productivity increases of a few regions and their specialization into the heavy manufacturing sector led to increases in welfare across all regions through trade linkages.

Panel C of Figure 5 plots the aggregate welfare gains in the baseline economy over the counterfactual economy. The aggregate welfare of the baseline is 10.7% permanently higher than the counterfactual once the economies reach steady states. The discounted utility ($\sum_{t=1}^{\infty} U_t^{agg}$) is also 10% higher in the baseline than the counterfactual. At the beginning of the implementation of the



A. Steady state regional avg. productivity

B. Steady state regional welfare

C. Aggregate welfare gain (%)

Figure 5. Counterfactual Results. Productivity and Welfare

Notes. This figure plots the counterfactual results. Panel A and B report the regional average productivity and regional welfare. The x and y axes plot each region's average productivity and welfare under the baseline and counterfactual economies. In Panels A and B, each dot is colored red if a corresponding region experienced increases in the average productivity and regional welfare. Panel C reports the ratio of the aggregate welfare of the baseline economy to that of the counterfactual economy.

subsidies, the aggregate welfare of the baseline first decreases in the short run compared to the counterfactual because calibrated subsidies are not optimally designed.[36]

[36] Analyzing the optimal subsidy in this economy is outside the purview of this paper. For the optimal policy, see Bartelme et al. (2020), Fajgelbaum and Gaubert (2020) and Lashkaripour and Lugovskyy (2020) in the static setting.

Roundabout Production. We find that a roundabout production structure plays an important role in generating permanent effects of subsidies. A roundabout production structure amplifies the impact of subsidies through cost and demand linkages (Krugman and Venables, 1995). Heavy manufacturing sectors had disproportionally larger own-cost shares of production γ_j^j than other sectors. Own-cost shares of production were 0.09 for commodities, 0.26 for light manufacturing, 0.46 for heavy manufacturing, and 0.13 for service. Because of these linkages, complementarity between firm-scale and gains from technology adoption causes more firms to adopt technology. We do the same counterfactual exercises with a new production structure where labor is the only factor of production, and there are no intermediate inputs. The results are reported in Figure D3. Holding other parameters, subsidies, and geographic fundamentals constant, we find that both the baseline and the counterfactual economies converge to the same steady state.

Geography: Foreign Market Size and Migration Costs. We examine how geography interacts with the effects of temporary adoption subsidies. When we compare the baseline and counterfactual economies with and without the subsidies, we change geographical features of the South Korean economy to examine how its long-term effects differ from the main results in Figure 4. We specifically examine the role of foreign market size and migration costs. We focus on foreign market size because of the large increase in the volume of South Korea's exports in the 1960s and 1970s and narratives that suggest that export expansion played an important role in South Korea's economic development.[37] We study migration costs because there were dramatic increases in migration flows from rural to urban areas in South Korea in the 1970s. This migration pattern is a common feature during industrialization.[38]

We examine how the effects of subsidies would have been if foreign market size had been lower. We decrease the foreign market size of the heavy manufacturing sector D_{jt}^f so that export shares in the heavy manufacturing sector in 1972 was 6.6%. This is the 1966 level; it replaced the 1972 level of 22%. The results are reported in Figure D5. The gap between the heavy manufacturing GDP shares in the two steady states is about 5 percentage points, which is 10 percentage points smaller than the main results in Figure 4. These results provide suggestive evidence that exports and subsidies together might have played an important role in shaping South Korea's economic development.

[37] Dramatic rapid increases in South Korea's exports were the outcomes of the government's export-promotion policy and increases in foreign demand shocks. South Korea joined the General Agreements on Tariff and Trade (GATT) in 1967 during the Kennedy Round and eliminated tariffs on imported inputs for exports (Connolly and Yi, 2015). It also devalued its over-valued currency in 1964, which boosted its exports (Irwin, 2021). Also, the United States' demand for foreign imports increased significantly in the period 1960 to 1980. During that period, shares of the United States' imports in the total gross national product rose from 6% to 22%.

[38] According to the World Development Indicators (World Bank), the rural population of South Korea decreased from 60 to 40% between 1970 and 1982. Migrants as a percent of the total population increased from 12.6% in 1970 to 21.9% in 1982. Many developing countries underwent rapid transitions from rural to urban during industrialization in the twentieth century. See Table 1 of Lucas (2004). Young (1995) also finds that labor reallocation into manufacturing played a significant role in manufacturing growth in the East Asian countries. Higher levels of migration costs may hinder labor reallocation into manufacturing.

We next examine how the effects of subsidies would have been if migration costs had been higher. We set migration costs to be 10% higher than the baseline calibrated value. Because of higher migration costs, fewer workers move toward regions with higher productivity brought about by technology adoption, which in turn increases wages and the cost of production. Because of the complementarity between firm scale and gains from adoption, fewer firms would adopt technology. These results are reported in Figure D4. The difference between the heavy manufacturing sector's share of GDP for the two steady states is around 9 percentage points, which is 6 percentage points smaller than the main results in Figure 4.

Comparative Statistics. We conduct the comparative statistics of δ and η to examine how the parameters we choose drive these multiple steady states. In Figure D2, we show that the differences between the outcomes of the baseline and counterfactual economies in the steady states become negligible when either δ or η is too low, consistent with the comparative static results of Proposition 1(iv) in the simplified model.

8 Conclusion

We find that the impact of technology adoption on late industrialization in South Korea was significant both empirically and quantitatively. Our finding confirms a widely held belief by economists and policymakers that technology adoption can foster economic development of developing countries. We find that technology adoption not only directly benefited adopters but also had large local spillover effects. Based on these findings, we build a dynamic spatial model in order to conduct a counterfactual analysis of the South Korean government policy that provided temporary subsidies for technology adoption in the heavy manufacturing sectors. Using a quantitative model calibrated to firm-level data and to econometric estimates, we show that temporary adoption subsidies can have a permanently large impact on an economy by moving it to a new transition path that converges to a more industrialized steady state.

We believe that our empirical findings and quantitative results are important for two reasons. First, they highlight that externalities may explain why technologies diffuse slowly to developing countries and why appropriate policy interventions might be necessary to boost productivity. Second, we show that knowledge flows from developed countries to developing countries can be an important source of economic development.

Although we have mainly focused on the spatial spillover of technology adoption, there might be many other sources of externalities and frictions that hinder firms in developing countries from adopting more advanced technology. We abstracted from both uncertainties about future technology and forward-looking technology adoption decisions by firms. Incorporating more realistic assumptions on agents' beliefs in the model and how these beliefs interact with multiple equilibria would be an interesting extension. We leave these questions for future research.

References

Acemoglu, Daron, "A Microfoundation for Social Increasing Returns in Human Capital Accumulation," *Quarterly Journal of Economics*, 1996, *111* (3), 779–804.

— , Pol Antràs, and Elhanan Helpman, "Contracts and Technology Adoption," *American Economic Review*, 2007, *97* (3), 916–943.

Ackerberg, Daniel A., Kevin Caves, and Garth Frazer, "Identification Properties of Recent Production Function Estimators," *Econometrica*, 2015, *83* (6), 2411–2451.

Adserà, Alícia and Debraj Ray, "History and Coordination Failure," *Journal of Economic Growth*, 1998, *3* (3), 267–276.

Allen, Treb and Costas Arkolakis, "Trade and the Topography of the Spatial Economy," *Quarterly Journal of Economics*, 2014, *129* (2014), 1085–1139.

— and Dave Donaldson, "The Geography of Path Dependence in the Spatial Economy," *NBER Working Paper 28059*, 2020.

Amsden, Alice H., *Asia's Next Giant: South Korea and Late Industrialization*, Oxford University Press, 1989.

Anderson, James E. and Eric Van Wincoop, "Trade Costs," *Journal of Economic Literature*, 2004, *42* (3), 691–751.

Arkolakis, Costas, Michael Peters, and Sun Kyoung Lee, "European Immigrants and the United States' Rise to the Technological Frontier," *Working Paper*, 2019.

Arzaghi, Mohammad and J. Vernon Henderson, "Networking off Madison Avenue," *Review of Economic Studies*, 2008, *75* (4), 1011–1038.

Atkin, David, Arnaud Costinot, and Masao Fukui, "Globalization and the Ladder of Development: Pushed to the Top or Held at the Bottom?," *Working Paper*, 2021.

— , Azam Chaudhry, Shamyla Chaudry, Amit K. Khandelwal, and Eric Verhoogen, "Organizational Barriers to Technology Adoption: Evidence from Soccer-Ball Producers in Pakistan," *Quarterly Journal of Economics*, 2017, *132* (3), 1101–1164.

Axtell, Robert L., "Zipf Distribution of U.S. Firm Sizes," *Science*, 2001, *293* (5536), 1818–1820.

Azariadis, Costas and John Stachurski, "Poverty Traps," in Phillippe Aghion and Steven Durlauf, eds., *Handbook of Economic Growth*, Vol. 1, Amsterdam: Elsevier, 2005, chapter 5, pp. 295–384.

Banerjee, Abhijit V. and Esther Duflo, "Growth Theory through the Lens of Development Economics," in Philippe Aghion and Steven Durlauf, eds., *Handbook of Economic Growth*, 7 ed., Vol. 1, Amsterdam: Elsevier, 2005, chapter 7, pp. 473–552.

Bartelme, Dominick, Arnaud Costinot, Dave Donaldson, and Andres Rodriguez-Clare, "The Textbook Case for Industrial Policy: Theory Meets Data," *NBER Working Paper 26193*, 2020.

Bena, Jan, Miguel A. Ferreira, Pedro Matos, and Pedro Pires, "Are Foreign Investors Locusts? The Long-Term Effects of Foreign Institutional Ownership," *Journal of Financial Economics*, 2017, *126* (1), 122–146.

Bloom, Nicholas, Mark Schankerman, and John Van Reenen, "Identifying Technology Spillovers and Product Market Rivalry," *Econometrica*, 2013, *81* (4), 1347–1393.

Broda, Christian and David E. Weinstein, "Globalization and the Gains from Variety," *Quarterly Journal of Economics*, 2006, *121* (2), 541–585.

Buera, Francisco J. and Ezra Oberfield, "The Global Diffusion of Ideas," *Econometrica*, 2020, *88* (1), 83–114.

— , Hugo Hopenhayn, Yongseok Shin, and Nicholas Trachter, "Big Push in Distorted Economies," *NBER Working Paper 28561*, 2021.

Burbidge, John B., Lonnie Magee, and A. Leslie Robb, "Alternative Transformations to Handle Extreme Values of the Dependent Variable," *Journal of the American Statistical Association*, 1988, *83* (401), 123–127.

Bustos, Paula, "Trade Liberalization, Exports, and Technology Upgrading: Evidence on the Impact of MERCOSUR on Argentinian Firms," *American Economic Review*, 2011, *101* (1), 304–340.

Cai, Jie, Nan Li, and Ana Maria Santacreu, "Knowledge Diffusion, Trade, and Innovation across Countries and Sectors," *American Economic Journal: Macroeconomics*, 2022, *14* (1), 104–145.

Cameron, Colin A., Jonah B. Gelbach, and Douglas L. Miller, "Bootstrap-Based Improvements for Inference with Clustered Errors," *Review of Economics and Statistics*, 2008, *90* (3), 414–427.

Chaney, Thomas, "Distorted Gravity: The Intensive and Extensive Margins of International Trade," *American Economic Review*, 2008, *98* (4), 1707–1721.

Choi, Jaedo and Andrei A. Levchenko, "The Long-Term Effects of Industrial Policy," *NBER Working Paper 29263*, 2021.

Ciccone, Antonio, "Input Chains and Industrialization," *Review of Economic Studies*, 2002, *69* (3), 565–587.

__ **and Robert E. Hall**, "Productivity and the Density of Economic Activity," *American Economic Review*, 1996, *86* (1), 54–70.

Cole, Harold L., Jeremy Greenwood, and Juan M. Sanchez, "Why Doesn't Technology Flow From Rich to Poor Countries?," *Econometrica*, 2016, *84* (4), 1477–1521.

Combes, Pierre Philippe and Laurent Gobillon, "The Empirics of Agglomeration Economies," in Gilles Duranton, J. Vernon Henderson, and WIlliam C. Strange, eds., *Handbook of Regional and Urban Economics*, elsevier ed., Vol. 5, North Holland, Amsterdam: Elsevier, 2015, chapter 5, pp. 247–338.

Comin, Diego and Bart Hobijn, "An Exploration of Technology Diffusion," *American Economic Review*, 2010, *100* (5), 2031–2059.

Connolly, Michelle and Kei-Mu Yi, "How Much of South Korea's Growth Miracle Can Be Explained by Trade Policy?," *American Economic Journal: Macroeconomics*, 2015, *7* (4).

Crouzet, Nicolas, Apoorv Gupta, and Filippo Mezzanotti, "Shocks and Technology Adoption: Evidence from Electronic Payment Systems," *Working Paper*, 2020.

de Souza, Gustavo, "The Labor Market Consequences of Appropriate Technology," *Working Paper*, 2021.

Desmet, Klaus and Esteban Rossi-Hansberg, "Spatial Development," *American Economic Review*, 2014, *104* (4), 1211–1243.

__ **, Dávid Krisztián Nagy, and Esteban Rossi-Hansberg**, "The Geography of Development," *Journal of Political Economy*, 2018, *126* (3), 903–983.

di Giovanni, Julian and Andrei A. Levchenko, "Country Size, International Trade, and Aggregate Fluctuations in Granular Economies," *Journal of Political Economy*, 2012, *120* (6), 1083–1182.

__ **and** __ , "Firm Entry, Trade, and Welfare in Zipf's World," *Journal of International Economics*, 2013, *89* (2), 283–296.

__ **,** __ **, and Isabelle Mejean**, "Foreign Shocks as Granular Fluctuations," *NBER Working Paper 28123*, 2020.

__ **,** __ **, and Romain Rancière**, "Power Laws in Firm Size and Openness to Trade: Measurement and Implications," *Journal of International Economics*, 2011, *85* (1), 42–52.

Donaldson, Dave and Richard Hornbeck, "Railroads and American Economic Growth: A ''Market Access" Approach," *Quarterly Journal of Economics*, 2016, *131* (2), 799–858.

Eaton, Jonathan and Samuel Kortum, "Technology, Geography, and Trade," *Econometrica*, 2002, *70* (5), 1741–1749.

Ellison, Glenn, Edward L. Glaeser, and William R. Kerr, "What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns," *American Economic Review*, 2010, *100* (3), 1195–1213.

Enos, John L. and Woo-Hee Park, *The Adoption and Diffusion of Imported Technology: The Case of Korea*, Routledge, 1988.

Fajgelbaum, Pablo D. and Cecile Gaubert, "Optimal Spatial Policies, Geography, and Sorting," *Quarterly Journal of Economics*, 2020, *135* (2), 959–1036.

Gerschenkron, Alexander, *Economic Backwardness in Historical Perspective*, Cambridge, MA: Belknap Press, 1952.

Gibbons, Steve, Henry G. Overman, and Eleonora Patacchini, "Spatial Methods," in Gilles Duranton, J. Vernon Henderson, and William C. Strange, eds., *Handbook of Regional and Urban Economics*, gilles dur ed., Vol. 5, North Holland, Amsterdam: Elsevier, 2015, chapter 3, pp. 115–168.

Giorcelli, Michela and Bo Li, "Technology Transfer and Early Industrial Development: Evidence from the Sino-Soviet Alliance," *NBER Working Paper 29455*, 2021.

Greenstone, Michael, Richard Hornbeck, and Enrico Moretti, "Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings," *Journal of Political Economy*, 2010, *118* (3), 536–598.

Hall, Robert E. and Charles I. Jones, "Why Do Some Countries Produce so Much More Output per Worker than Others?," *Quarterly Journal of Economics*, 1999, *114* (1), 83–116.

Hanlon, W. Walker and Antonio Miscio, "Agglomeration: A Long-Run Panel Data Approach," *Journal of Urban Economics*, 2017, *99*, 1–14.

Hanson, Gordon H., Nelson Lind, and Marc-Andreas Muendler, "The Dynamics of Comparative Advantage," *NBER Working Paper 21753*, 2015.

Hardy, Morgan and Jamie McCasland, "It Takes Two: Experimental Evidence on the Determinants of Technology Diffusion," *Journal of Development Economics*, 2021, *149*.

Hausmann, Ricardo and Bailey Klinger, "The Structure of the Product Space and the Evolution of Comparative Advantage," *CID Working Paper Series*, 2007.

Helpman, Elhanan, Marc J. Melitz, and Yona Rubinstein, "Estimating Trade Flows: Trading Partners and Trading Volumes," *Quarterly Journal of Economics*, 2008, *123* (2), 441–487.

Hirschman, Albert O., *The Strategy of Economic Development*, Yale University Press, 1958.

Irwin, Douglas A., "From Hermit Kingdom to Miracle on the Han: Policy Decisions that Transformed South Korea into an Export Powerhouse," *National Bureau of Economic Research 29299*, 2021.

Itskhoki, Oleg and Benjamin Moll, "Optimal Development Policies With Financial Frictions," *Econometrica*, 2019, *87* (1), 139–173.

Jaffe, Adam B., Manuel Trajtenberg, and Rebecca Henderson, "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations," *Quarterly Journal of Economics*, 1993, *108* (3), 577–598.

Jovanovic, Boyan, "Observable Implications of Models with Multiple Equilibria," *Econometrica*, 1989, *57* (6), 1431–1437.

— **and Rafael Rob**, "The Growth and Diffusion of Knowledge," *Review of Economic Studies*, 1989, *56* (4), 569–582.

Juhász, Réka, "Temporary Protection and Technology Adoption: Evidence from the Napoleonic Blockade," *American Economic Review*, 2018, *108* (11), 3339–3376.

__ , **Mara P. Squicciarini, and Nico Voigtländer**, "Technology Adoption and Productivity Growth: Evidence from Industrialization in France," *NBER Working Paper 27503*, 2020.

Kantor, Shawn and Alexander Whalley, "Research Proximity and Productivity: Long-Term Evidence from Agriculture," *Journal of Political Economy*, 2019, *127* (2), 819–854.

Keller, Wolfgang, "Geographic Localization of International Technology Diffusion," *American Economic Review*, 2002, *92* (1), 120–142.

__ , "International Technology Diffusion," *Journal of Economic Literature*, 2004, *42* (3), 752–782.

Kerr, William R. and Scott Duke Kominers, "Agglomerative Forces and Cluster Shapes," *Review of Economics and Statistics*, 2015, *97* (4), 877–899.

Kim, Dae-Il and Robert H. Topel, "Labor Markets and Economic Growth: Lessons from Korea's Industrialization, 1970-1990," in "Differences and Changes in Wage Structures," Cambridge: National Bureau of Economic Research, 1995, pp. 227–264.

Kim, Linsu, *Imitation to Innovation: The Dynamics of Korea's Technological Learning*, Boston: Harvard Business School Press, 1997.

__ **and Youngbae Kim**, "Innovation in a Newly Industrializing Country: A Multiple Discriminant Analysis," *Management Science*, 1985, *31* (3), 312–322.

Kim, Minho, Munseob Lee, and Yongseok Shin, "The Plant-Level View of an Industrial Policy: The Korean Heavy Industry Drive of 1973," *NBER Working Paper 29252*, 2021.

Klenow, Peter J. and Andrés Rodríguez-Clare, "The Neoclassical Revival in Growth Economics: Has It Gone Too Far?," *NBER Macroeconomics Annual*, 1997, *12*, 73–103.

Kline, Patrick and Enrico Moretti, "Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority," *Quarterly Journal of Economics*, 2014, *129* (1), 275–331.

Korea Industrial Technology Association, *Records of Technology Acquisition Contracts: 1962-1988*, Seoul: Korea Industrial Technology Association, [in Korean], 1988.

Krugman, Paul, "The Narrow Moving Band, the Dutch Disease, and the Competitive Consequences of Mrs. Thatcher. Notes on Trade in the Presence of Dynamic Scale Economies," *Journal of Development Economics*, 1987, *27* (1-2), 41–55.

__ , "History versus Expectations," *Quarterly Journal of Economics*, 1991, *106* (2), 652–667.

__ **and Anthony J. Venables**, "Globalization and the Inequality of Nations," *Quarterly Journal of Economics*, 1995, *110* (4), 857–880.

Lane, Nathan, "Manufacturing Revolutions Industrial Policy and Networks in South Korea," *Working Paper*, 2019.

Lashkaripour, Ahmad and Volodymyr Lugovskyy, "Scale Economies and the Structure of Trade and Industrial Policy," *Working Paper*, 2020.

Levchenko, Andrei A. and Jing Zhang, "The Evolution of Comparative Advantage: Measurement and Welfare Implications," *Journal of Monetary Economics*, 2016, *78*, 96–111.

Levinsohn, James and Amil Petrin, "Estimating Production Functions using Inputs to Control for Unobservables," *Review of Economic Studies*, 2003, *70* (2), 317–341.

Lileeva, Alla and Daniel Trefler, "Improved Access to Foreign Markets Raises Plant-Level Productivity. . .for Some Plants," *Quarterly Journal of Economics*, 2010, *125* (3), 1051–1099.

Lucas, Robert E., "On the Mechanics of Economic Development," *Journal of Monetary Economics*, 1988, *22* (1), 3–42.

— , "Making a Miracle," *Econometrica*, 1993, *61* (2), 251–272.

— , "Life Earnings and Rural-Urban Migration," *Journal of Political Economy*, 2004, *112* (1), 29–59.

— **and Benjamin Moll**, "Knowledge Growth and the Allocation of Time," *Journal of Political Economy*, 2014, *122* (1), 1–51.

Luedde-Neurath, Richard, *Import Controls and Export-Oriented Development: A Reassessment of the South Korean Case*, Routledge, 1986.

Magee, Chris, "U.S. Tariffs at the SIC Level, 1974-1988," 1986.

Malmendier, Ulrike, Enrico Moretti, and Florian S. Peters, "Winning by Losing: Evidence on the Long-Run Effects of Mergers," *Review of Financial Studies*, 2018, *31* (8), 3212–3264.

Manski, Charles F., "Identification of Endogenous Social Effects: the Reflection Problem," *Review of Economic Studies*, 1993, *60* (3), 531–542.

Matsuyama, Kiminori, "Increasing Returns, Industrialization, and Indeterminacy of Equilibrium," *Quarterly Journal of Economics*, 1991, *106* (2), 617–650.

— , "Agricultural Productivity, Comparative Advantage, and Economic Growth," *Journal of Economic Theory*, 1992, *58* (2), 317–334.

— , "Complementarities and Cumulative Processes in Models of Monopolistic Competition," *Journal of Economic Literature*, 1995, *33* (2), 701–729.

Melitz, Marc J., "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 2003, *71* (6), 1695–1725.

Ministry of Science and Technology, *Science and Technology Annual* 1972.

Monte, Ferdinando, Stephen J. Redding, and Esteban Rossi-Hansberg, "Commuting, Migration, and Local Employment Elasticities," *American Economic Review*, 2018, *108* (12), 3855–3890.

Moretti, Enrico, "Workers' Education, Spillovers, and Productivity: Evidence from Plant-Level Production Functions," 2004.

— , "The Effect of High-Tech Clusters on the Productivity of Top Inventors," *American Economic Review*, 2021, *111* (10), 3328–3375.

Murphy, Kevin M., Andrei Shleifer, and Robert W. Vishny, "Industrialization and the Big Push," *Journal of Political Economy*, 1989, *97* (5), 1003–1026.

Nagy, Dávid Krisztián, "Hinterlands, City Formation and Growth: Evidence from the U.S. Westward Expansion," *Working Paper*, 2020.

Nakamura, Emi and Jón Steinsson, "Identification in Macroeconomics," *Journal of Economic Perspectives*, 2018, *32* (3), 59–86.

Parente, Stephen L. and Edward C. Prescott, "Barriers to Technology Adoption and Development," *Journal of Political Economy*, 1994, *102* (2), 298–321.

— **and** — , *Barriers to Riches*, MIT Press, 2002.

Pellegrina, Heitor and Sebastian Sotelo, "Migration, Specialization, and Trade: Evidence from the Brazilian March to the West," *NBER Working Paper 28421*, 2021.

Perla, Jesse, Christopher Tonetti, and Michael E. Waugh, "Equilibrium Technology Diffusion, Trade, and Growth," *American Economic Review*, 2021, *111* (1), 73–128.

Peters, Michael, "Market Size and Spatial Growth - Evidence from Germany's Post-War Population Expulsions," *NBER Working Paper 29329*, 2021.

Redding, Stephen, "The Low-Skill, Low-Quality Trap: Strategic Complementarities Between Human Capital and R&D," *Economic Journal*, 1996, *106* (435), 458–470.

Rodríguez-Clare, Andrés, "The Division of Labor and Economic Development," *Journal of Development Economics*, 1996, *49* (1), 3–32.

Rodrik, Dani, "Getting Interventions Right: How South Korea and Taiwan Grew Rich," *Economic Policy*, 1995, *20*, 53–107.

Romer, Paul M., "Endogenous Technological Change," *Journal of Political Economy*, 1990, *98* (5), S71–S102.

Rosenstein-Rodan, Paul N., "Problems of Industrialisation of Eastern and South-Eastern Europe," *Economic Journal*, 1943, *53*, 202–211.

Schetter, Ulrich, "A Structural Ranking of Economic Complexity," *Working Paper*, 2019.

Serafinelli, Michel, ""Good" Firms, Worker Flows, and Local Productivity," *Journal of Labor Economics*, 2019, *37* (3), 747–792.

Silva, J. M. C. Santos and Silvana Tenreyro, "The Log of Gravity," *Review of Economics and Statistics*, 2006, *88* (4), 641–658.

Stoyanov, Andrey and Nikolay Zubanov, "Productivity Spillovers across Firms through Worker Mobility," *American Economic Journal: Applied Economics*, 2012, *4* (2), 168–198.

Syverson, Chad, "What Determines Productivity?," 2011.

Verhoogen, Eric A., "Trade, Quality Upgrading, and Wage Inequality in the Mexican Manufacturing Sector," *Quarterly Journal of Economics*, 2008, *123* (2), 489–530.

Vogel, Erza F., *The Four Little Dragons: The Spread of Industrialization in East Asia*, Harvard University Press, 1991.

Walsh, Conor, "Firm Creation and Local Growth," *Working Paper*, 2019.

Wooldridge, Jeffrey M., "On Estimating Firm-Level Production Functions Using Proxy Variables to Control for Unobservables," *Economics Letters*, 2009, *104* (3), 112–114.

Yeaple, Stephen R., "A Simple Model of Firm Heterogeneity, International Trade, and Wages," *Journal of International Economics*, 2005, *65* (1), 1–20.

Young, Alwyn, "The Tyranny of Numbers: Confronting the Statistical Realities of the East Asian Growth Experience," *Quarterly Journal of Economics*, 1995, *110* (3), 641–680.

APPENDIX

Appendix A Data

A.1 Data on Technology Adoption

ARTICLE III. SUPPLY OF TECHNICAL ASSISTANCE

1. MITSUI TOATSU shall transmit in documentary form to KOLON, TECHNICAL INFORMATION.

2. MITSUI TOATSU shall provide, upon the request of KOLON, the services of its technical personnel to assist KOLON in the engineering, construction and operation of the PLANT and in the quality and production control of LICENSED PRODUCT.
KOLON shall, for such services of technical personnel, pay the reasonable salaries, travelling and living expenses of such technical personnel while away from their own factories and offices.
The number of such technical personnel, the period of the services and the payment shall be discussed and decided separately between the parties.

3. MITSUI TOATSU shall receive KOLON's technical trainees at a plant designated by MITSUI TOATSU in order to train them

Figure A1. Example. A Contract between Kolon and Mitsui Toatsu

Institutional Background of Technology Adoption Contract Documents. After Chung-Hee Park came to power through a military coup, he created the Economic Planning Board (EPB) in 1961 to promote economic development and design better economic policies. President Park was in power for 19 years. He was the chairman of the military junta for 1961 and 1962. In 1963 and 1967, he was elected a president of the civilian government. In 1971, he was re-elected for what was supposed to be his last presidency. In 1972, President Park declared martial law and amended the country's constitution into an authoritarian document, called the Yushin constitution, which extended his term of office as president indefinitely. After 1961 and until President Park was assassinated in 1979, the EPB was at the center of South Korea's economic policy making process.

During his presidency, the Foreign Capital Act strictly regulated domestic firms' transactions with foreign firms, including technology adoption contracts. The law required South Korean firms to obtain approval from the EPB before they made contracts to adopt new technology from foreign firms. They also had to submit documentation of their plans for using the technologies they adopted and copies of the contracts. Beginning in 1961 and continuing until the mid-1980s, the EPB met every month. In each meeting, they examined new contracts between domestic and foreign firms. The National Archives of Korea collected and preserved the documents the EPB examined in its monthly meetings. Most of our technology adoption data mainly comes from historical contract documents from the National Archives of Korea.

Figure A1 is one page from a contract document between Kolon (South Korean) and Mitsui Toatsu (Mitsui) (Japanese). Most of the adopted technologies involved the transfer of knowledge about how to build and operate plants and capital equipment related to mass production. For example, Figure A1 specifies that Mitsui had to provide blueprints, send skilled engineers to train South Korean workers, and provide training service by inviting South Korean engineers to its plants in Japan.

One may wonder why foreign firms were selling technology to Korean firms in the 1970s although South Korean firms that adopted technologies from these foreign firms could have been a future competitor in international markets. An example of technology contracts between Pohang Iron and Steel Company (POSCO), a South Korean company, and Nippon Steel Company (NSC), a Japanese company, might explain this. POSCO made a technology contract about construction and operation of integrated steel mills. NSC sent its skilled engineers to teach Korean engineers of POSCO how to run integrated steel mills.

First, NSC could earn a lot of profits from this contract. The fixed fee that POSCO had to pay for the contract accounted for 20% of the total annual export of plant engineering of NSC. Second, NSC did not transfer state of the art technology but more standardized technology that were widely used in developed countries. For example, NSC refused to transfer technology related to the computerization of production system, which was considered to be state of the art at that time. In the early 1980s when POSCO grew fast and became a big competitor in international steel markets, NSC refused to make further official technology contracts with POSCO. Third, foreign firms did not expect that South Korean firms would absorb technology within such a short period of time. The CEO of NSC, Eishiro Saito, said that he did not expect remarkably high rates of POSCO's technology absorption and said in his interview that technology adoption contracts between the two firms hit NSC like a boomerang (Chosun-ilbo, 1976. 11. 23).

Available Information. From these contracts, we obtained three main pieces of information: names of domestic firms, names of foreign firms, and contract years. We use the information on names of domestic firms and contract years to construct a dummy variable of firms' adoption status. We use information on the names of foreign firms to match them to the USPTO.

A.2 Firm Balance Sheet Data.

We match firm balance sheet data obtained from the Annual Reports of Korean Companies between 1970 and 1982. These reports are published by the Korea Productivity Center. We obtain firms' balance sheet variables and locations of production from these reports.

Balance Sheet Variables. The information from balance sheet includes sales, assets, fixed assets, and employment. Employment data does not begin until 1972. We convert all monetary values into 2015 US dollars. The dataset covers firms with more than 50 employees. The dataset also includes information on firms' start years. We use this start year information to trace changes in firm names.

Location of Production. The dataset includes detailed information on the address of the location of production. We convert addresses to the 2010 administrative divisions of South Korea up to the town level. (We classify firms' location of production into villages (li) and neighborhood (dong) levels. Then, using distance between towns, we calculate distance between firms within the same district.

Sector Groupings. We classify firms into 10 manufacturing sectors. We classify four as heavy manufacturing sectors, largely following the sector classification in Lane (2019). Table A1 reports the classification. It is similar to the classification in Choi and Levchenko (2021), who used the same firm balance sheet data. The numbers inside the parenthesis are ISIC Rev. 3.1 (ISIC) codes. We use these ISIC codes to map our firm data to other trade or tariff data.

Coverage. Figure A2 reports the average coverage of the firm-level data across different sectors. We report the ratio between the sum of all firms in each year divided by gross output from the input-output table for corresponding years. When we compute this coverage, we impute using the information on assets for some observations that lack information about sales. For each sector j, we run the following regression model:

$$\ln(\text{Sales}_{it}) = \beta_j \ln(\text{Assets}_{it}) + \delta_t + \epsilon_{it}.$$

Using the estimated coefficient of β_j, we impute missing sales using $\hat{\beta}_j \ln(\text{Assets}_{it})$.

Across sectors, our dataset covers about 70% of gross output from the input-output table. However, there is some heterogeneity across sectors. Machinery and Transportation Equipment and Petrochemical and Chemical have higher coverage rates, whereas Food, Beverage, and Tobacco and Apparel, Leather, and Textile have relatively less coverage than other sectors.

A.3 Other Data Sets

United States Patent and Trademark Office (USPTO). We use the USPTO data to measure foreign firms' patenting activities.[39] We match the USPTO with foreign contractors in our dataset using their names. Our matching procedure proceeds as follows.

- Step 1: Clean firms' names.
  - For example, we erase words like "Inc" or "Comp."
- Step 2: Fuzzy match firms' names from our dataset and the USPTO. We use the fuzzmatcher package in Python.
- Step 3: Hand-match firms that are not matched in the first step based on names.
- Step 4: For foreign firms that have different assignee IDs in the UPSTO but with the same ID (gvkey) in the Global Compustat, we give them a unique assignee ID and sum the numbers of patents and citations up to the Compustat ID level.
  - When we merge assignee IDs and gvkey, we use the matching constructed by Bena et al. (2017).

Input-Output Tables. We obtain input-output tables from the Bank of Korea. [40] Input-output tables are available for 1970, 1973, 1975, 1978, 1980, 1983, and 1985 during the sample period. We convert codes of the input-output tables into ISIC Rev. 3.1 (ISIC) codes.

[39] We download the dataset from https://patentsview.org.

[40] We download the data from Economic Statistics of the Bank of Korea, https://ecos.bok.or.kr/EIndex_en.jsp.

Table A1: Classification of Sectors

	Aggregated Industry	Industry
Heavy Mfg.	(i) Chemicals, Petrochemicals, Rubber, & Plastic Products	Coke oven products (231) Refined petroleum products (232) Basic chemicals (241) Other chemical products (242) Man-made fibres (243) except for pharmaceuticals and medicine chemicals (2423) Rubber products (251) Plastic products (252)
	(ii) Electrical Equipment	Office, accounting, & computing machinery (30) Electrical machinery and apparatus n.e.c. (31) Ratio, television and communication equipment and apparatus (32) Medical, precision, and optical instruments, watches and clocks (33)
	(iii) Basic & Fabricated Metals	Basic metals (27) Fabricated metals (28)
	(iv) Machinery & Transport Equipment	Machinery and equipment n.e.c. (29) Motor vehicles, trailers and semi trailers (34) Building and repairing of ships and boats (351) Railway and tramway locomotives and rolling stock (352) Aircraft and spacecraft (353) Transport equipment n.e.c. (359)
Light Mfg.	(v) Food, Beverages, & Tobacco	Food products and beverages (15) Tobacco products (16)
	(vi) Textiles, Apparel, & Leather	Textiles (17) Apparel (18) Leather, luggage, handbags, saddlery, harness, and footwear (19)
	(vii) Manufacturing n.e.c.	Manufacturing n.e.c. (369)
	(viii) Wood, Paper, Printing, & Furniture	Wood and of products, cork (20) Paper and paper products (21) Publishing and printing (22) Furniture (361)
	(ix) Pharmaceuticals & Medicine Chemicals	pharmaceuticals and medicine chemicals (2423)
	(x) Other Nonmetallic Mineral Products	Glass and glass products (261) on-metallic mineral products n.e.c. (269)

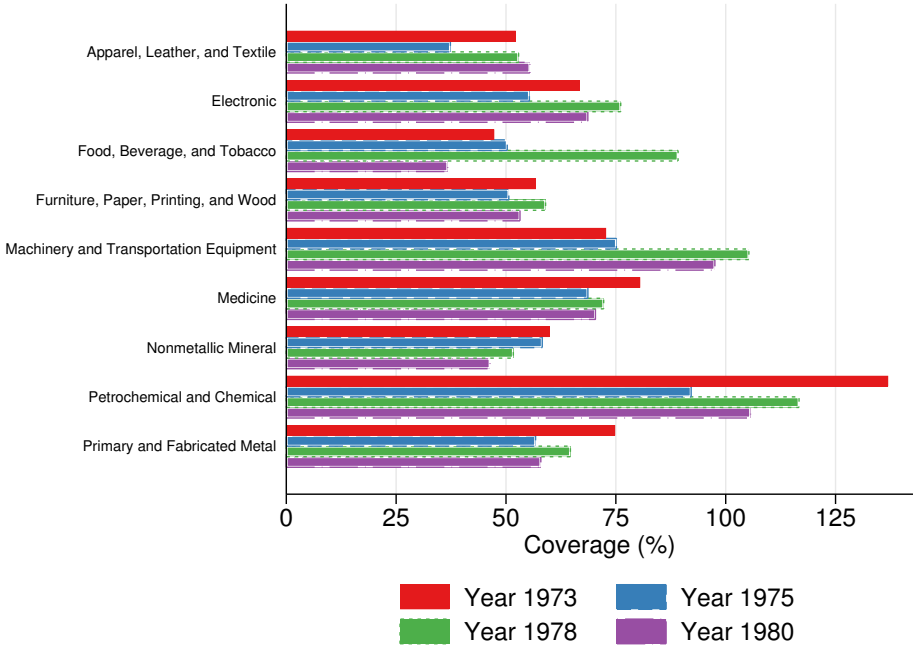


Figure A2. Coverage of Manufacturing Sectors in Our Dataset

Notes. This figure plots the ratio of sectoral gross output from the input-output tables to the sum of firms' sales in corresponding sectors.

OECD Stan Database. We obtain the cross-country data on heavy manufacturing's contributions to GDP in Figure 1 from OECD Stan database, which has sectoral GDP information at the two-digit ISIC3 level.[41]

For Mexico, the total GDP was not available for the early 1970s, but the total value of light and heavy manufacturing shares was available. Therefore, from the OECD Stan database, we can calculate heavy manufacturing sector's contribution to the manufacturing sector's GDP, but we could not calculate the sector's contribution to the national GDP. Thus, we supplement the Mexico sample with data on manufacturing's contribution to total GDP obtained from the World Bank Indicators.[42] We then obtain heavy manufacturing's share of GDP as that sector's share of the total value manufacturing added to the Mexican economy multiplied by the manufacturing sector's contribution to total GDP.

A.4 Merging Technology Adoption and Firm Balance Sheet Data Sets

We match technology adoption and firm balance sheet datasets using firms' names and information about start year and sector. We match the two datasets based on the following criteria:

[41] We download the data from OECD, "STAN Database for Structural Analysis," https://stats.oecd.org/Index.aspx?DataSetCode=STAN.

[42] We download the data from World Bank, "Manufacturing, Value Added (%)," https://data.worldbank.org/indicator/NV.IND.MANF.ZS.

1. Firms should have the same name in a given year.
2. Firms should have begun operation before the years they adopted new technology.
   - Even if we observe the same names in both datasets, if adoption activities happened before start year information in the balance sheet data, we do not match those firms.
3. Firms should be in the same sector.
   - Each contract document has a brief description about the technology firms adopted
   - Even if we observe the same names in both datasets, if these descriptions do not align with the recorded sector in the balance sheet data, we do not match those firms.

One of the key challenges when merging two datasets based on firms' names is that many firms changed their names during the sample period. We tracked each firm's name in the Annual Reports of Korean Companies and in the history sections of the firms' websites. We also searched for firm names at https://www.jobkorea.co.kr and https://www.saramin.co.kr, which are the two largest job posting sites. We identified firm names as the same firm only if the information in the Annual Reports of Korean Companies matched information obtained on the Internet. We also searched in newspapers from the 1970s, which sometimes had articles that announced a firm's change of name. When a firm merged with another firm, we counted that as an exit.

A.5 An Example of a Loser

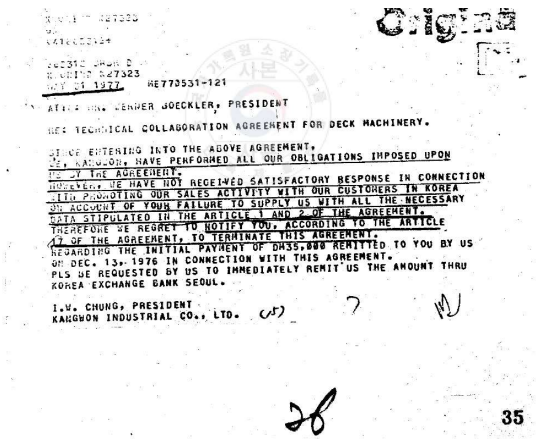
We identify losers from contract documents. The Foreign Capital Act required firms to submit related documents when their contracts failed if the EPB had approved the contract. They had to submit official cancellation contract documents and documents that described why the contract had failed.

Figure A3 reports an example of a loser. Kangwon Industrial Co. (Kangwon) and the German firm Broehl Maschinen Fabric GmbH (Broehl) made a contract regarding deck machinery. Although Kangwon paid a fixed fee in advance, Broehl did not send a blueprint. Panel A is the English document related to the termination of the contract between two firms. Panel B is the Korean document in which Kangwon reported why the contract had failed. The document says that the contract failed because although Kangwon asked Broehl several times to fulfill the contract after Kangwan paid the fee, Broehl did not respond.

A.6 Descriptive Statistics.

Table A2 reports the descriptive statistics of the constructed dataset. The table reports firm balance sheet variables, including log sales, assets, fixed assets, and employment, and variables related to firms' adoption activities. In columns (1), (2), and (3), we include samples of all manufacturing, heavy manufacturing, and light manufacturing firms. $\mathbb{1}$[Adopt] is a dummy variable that equals 1 if a firm is in a contractual relationship with any foreign firms. From the contract data, we observe when firms made adoption contracts and what years the contracts were made. The dummy variable equals 1 if a firm was under contract with a foreign firm. $\mathbb{1}$[Adopt] is a dummy variable that equals 1 if a firm ever adopted foreign technology during the sample period. Consistent with the historical narrative, adoption activities were concentrated among heavy manufacturing firms. In the period 1970 to 1982, an average of 13% of heavy manufacturing firms adopted technology at least once. Only 4.2% of light manufacturing firms adopted technology during that period.

In Panel A of Figure A4, we have plotted the evolution of the size of the heavy and light manu-



Original

MAY 31 1977 KE770531-121

ATTN: MR. WERNER GOECKLER, PRESIDENT

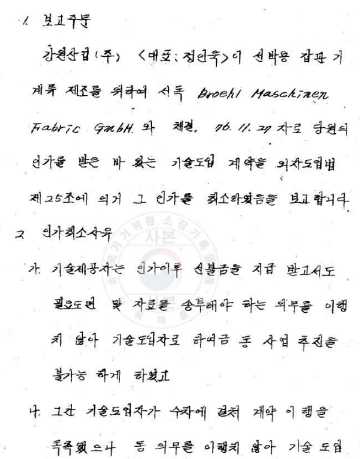
RE: TECHNICAL COLLABORATION AGREEMENT FOR DECK MACHINERY.

SINCE ENTERING INTO THE ABOVE AGREEMENT,
WE, KANGWON, HAVE PERFORMED ALL OUR OBLIGATIONS IMPOSED UPON
US BY THE AGREEMENT.
HOWEVER, WE HAVE NOT RECEIVED SATISFACTORY BESPONSE IN CONNECTION
WITH PROMOTING OUR SALES ACTIVITY WITH OUR CUSTOMERS IN KOREA
ON ACCOUNT OF YOUR FAILURE TO SUPPLY US WITH ALL THE NECESSARY
DATA STIPULATED IN THE ARTICLE 1 AND 2 OF THE AGREEMENT.
THEREFORE WE REGRET TO NOTIFY YOU, ACCORDING TO THE ARTICLE
17 OF THE AGREEMENT, TO TERMINATE THIS AGREEMENT.
REGARDING THE INITIAL PAYMENT OF DM35,000 REMITTED TO YOU BY US
ON DEC. 13, 1976 IN CONNECTION WITH THIS AGREEMENT.
PLS BE REQUESTED BY US TO IMMEDIATELY REMIT US THE AMOUNT THRU
KOREA EXCHANGE BANK SEOUL.

I.W. CHUNG, PRESIDENT
KANGWON INDUSTRIAL CO., LTD.

35

A. Documentation of the termination of the contract between Kangwan and Broehl



1. 보고구분

강원산업(주) <대표: 정인욱>이 선박용 갑판 기계류 제조를 위하여 서독 Broehl Maschinen Fabric GmbH. 와 체결, 76. 11. 29 자로 당원의 인가를 받은 바 있는 기술도입 계약을 외자도입법 제25조에 의거 그 인가를 취소하였음을 보고합니다.

2. 인가취소사유

가. 기술제공자는 인가이후 선불금을 지급 받고서도 필요도면 및 자료를 송부해야 하는 의무를 이행치 않아 기술도입자로 하여금 동 사업 추진을 불가능 하게 하였고

나. 그간 기술도입자가 수차에 걸쳐 계약 이행을 독촉했으나 동 의무를 이행치 않아 기술 도입

B. Document that explains the reason for the termination

Figure A3. Example of a Loser Firm

Table A2: Descriptive Statistics.

	All mfg. (1)	Heavy mfg. (2)	Light mfg. (3)
	Firm Balance Sheet		
ln(Sales)	15.65	15.54	15.75
	(1.925)	(1.938)	(1.910)
ln(Assets)	15.14	15.10	15.18
	(1.766)	(1.764)	(1.767)
ln(Fixed Assets)	13.96	13.94	13.98
	(1.966)	(1.933)	(1.992)
ln(Emp)	5.166	5.028	5.285
	(1.321)	(1.319)	(1.311)
	Technology Adoption		
$\mathbb{1}$[Adopt]	0.0587	0.0951	0.0267
	(0.235)	(0.293)	(0.161)
$\mathbb{1}$[Ever Adopt]	0.0841	0.132	0.0418
	(0.278)	(0.339)	(0.200)
N	43720	20497	23223

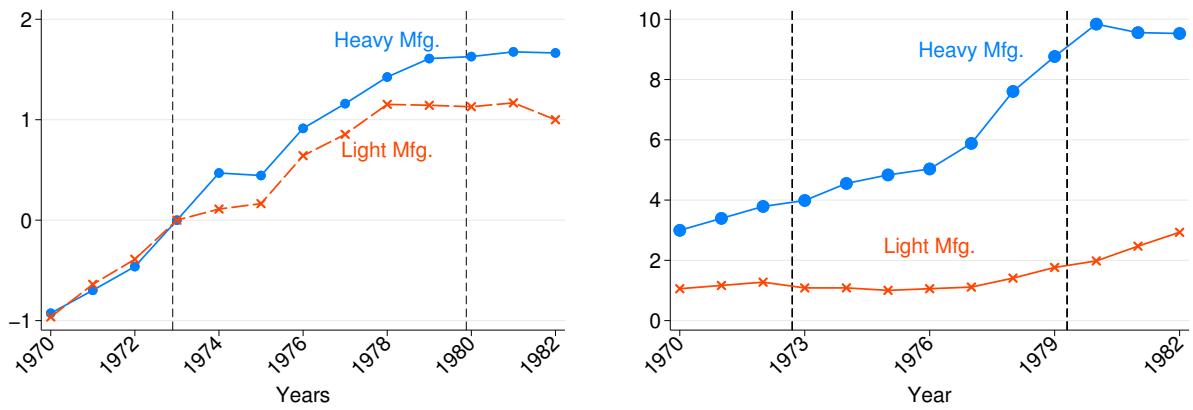
Notes. This table reports the descriptive statistics. All monetary values are in 2015 US dollars. $\mathbb{1}$[Adopt] is a dummy variable which equals one if a firm was in a technology adoption contract relationship with foreign firms in a given year. $\mathbb{1}$[Ever Adopt] is a dummy variable which equals one if a firm ever had technology adoption contracts with foreign firms.

facturing sectors. We measure the size of sector j as follows:

$$\ln Size_{jt} = \ln\left(\sum_{i \in j} Sale_{it}\right), \qquad j \in \{\text{Light}, \text{Heavy}\}.$$

We normalize the size of each sector by their 1973 level so we can track how the heavy and light manufacturing sectors evolved differently after the adoption subsidy policy was implemented in 1973. In Panel B of Figure A4, we have plotted shares of adopters in heavy and light manufacturing sectors. The shares are defined as firms that were in contractual relationships with foreign firms as a percentage of the total number of firms in a given year.

The patterns from the firm-level data reveal a similar pattern in Figure 1. The total size of heavy manufacturing sectors began increasing faster than that of the light manufacturing sectors after 1973, and this rapid increase coincided with increases in the amount of new technology that heavy



A. log sum of sales

B. shares of adopters

Figure A4. Evolution of Size of Manufacturing Sectors and Shares of Adopters from the Firm-Level Data

Notes. Panels A and B of this figure plot evolution of the size of manufacturing sectors and shares of adopters, respectively. The size of each sector is measured as a log of the total sum of firms' sales in each sector. We normalize the size of each sector by their levels in 1973. Shares of adopters are computed as shares of firms that were in a technology adoption contract with foreign firms in a given year. The two dotted vertical lines represent the start and the end of the South Korean government policy that subsidized technology adoption between 1973 and 1979.

manufacturing firms adopted.

Appendix B Model

B.1 Closed-Form Expressions for Regional Variables

In this section, we derive closed-form expressions for price index, regional gross output for domestic expenditures, and regional exports. Given optimal adoption and export decisions and the bounded Pareto distributional assumption, regional-level variables summed across firms within regions and sectors can be expressed as a function of shares of adopters, shares of exporters, subsidies, and natural advantage.

Price Index. A price index of sector j in region n is

$$P_{njt}^{1-\sigma} = \sum_{m\in\mathcal{N}} M_{mj} \Bigg\{ \underbrace{\int_{\phi_{mjt}^{min}}^{\bar{\phi}_{mjt}^{T}} \left(\frac{\sigma}{\sigma-1} \frac{\tau_{mnj} c_{mjt}}{f(\lambda_{mjt-1}^{T})\phi_{it}} \right)^{1-\sigma} dG_{mjt}(\phi_{it})}_{\text{Non-adopters' varieties}} + \underbrace{\int_{\bar{\phi}_{mjt}^{T}}^{\phi_{mjt}^{max}} \left(\frac{\sigma}{\sigma-1} \frac{\tau_{mnj}(1-s_{mjt}) c_{mjt}}{f(\lambda_{mjt-1}^{T})\phi_{it}} \right)^{1-\sigma} dG_{mjt}(\phi_{it})}_{\text{Adopters' varieties}} \Bigg\} + \underbrace{(\tau_{nj}^{x} c_{jt}^{f})^{1-\sigma}}_{\text{Foreign varieties}} .$$

That equation can be rewritten as:

$$\begin{aligned}
P_{njt}^{1-\sigma} &= \sum_{m\in\mathcal{N}} \Bigg\{ M_{mj} (\mu \tau_{mnj} c_{mjt})^{1-\sigma} f(\lambda_{mjt-1}^{T})^{\sigma-1} \frac{\theta}{\tilde{\theta}} \frac{1}{1-\kappa^{-\theta}} (\phi_{mjt}^{min})^{\theta} \\
&\quad \times \left[\left((\phi_{mjt}^{min})^{-\tilde{\theta}} - (\bar{\phi}_{mjt}^{T})^{-\tilde{\theta}} \right) + \left(\frac{\eta}{1-s_{mjt}} \right)^{\sigma-1} \left((\bar{\phi}_{mjt}^{T})^{-\tilde{\theta}} - (\phi_{mjt}^{max})^{-\tilde{\theta}} \right) \right] \Bigg\} + (\tau_{nj}^{x} c_{jt}^{f})^{1-\sigma} \\
&= \sum_{m\in\mathcal{N}} \Bigg\{ M_{mj} (\mu \tau_{mnj} c_{mjt})^{1-\sigma} f(\lambda_{mjt-1}^{T})^{\sigma-1} \frac{\theta}{\tilde{\theta}} \frac{1}{1-\kappa^{-\theta}} (\phi_{mjt}^{min})^{\sigma-1} \\
&\quad \times \left[\left(\left(\frac{\eta}{1-s_{mjt}} \right)^{\sigma-1} - 1 \right) \left(\frac{\bar{\phi}_{mjt}^{T}}{\phi_{mjt}^{min}} \right)^{-\tilde{\theta}} + \left(1 - \left(\frac{\eta}{1-s_{mjt}} \right)^{\sigma-1} \kappa^{-\tilde{\theta}} \right) \right] \Bigg\} + (\tau_{nj}^{x} c_{jt}^{f})^{1-\sigma} \\
&= \sum_{m\in\mathcal{N}} \Bigg\{ M_{mj} (\mu \tau_{mnj} c_{mjt})^{1-\sigma} \\
&\quad \underbrace{\times f(\lambda_{mjt-1}^{T})^{\sigma-1} \frac{\theta}{\tilde{\theta}} \frac{1}{1-\kappa^{-\theta}} (\phi_{mjt}^{min})^{\sigma-1} \left[\left(\left(\frac{\eta}{1-s_{mjt}} \right)^{\sigma-1} - 1 \right) (\tilde{\lambda}_{mjt}^{T})^{\frac{\tilde{\theta}}{\theta}} + \left(1 - \left(\frac{\eta}{1-s_{mjt}} \right)^{\sigma-1} \kappa^{-\tilde{\theta}} \right) \right]}_{=\bar{\phi}_{mjt}^{avg}} \Bigg\} \\
&\quad + (\tau_{nj}^{x} c_{jt}^{f})^{1-\sigma},
\end{aligned}$$

where $\tilde{\theta} = \theta - (\sigma - 1)$ and $\tilde{\lambda}_{njt}^{T}$. The last equality comes from Equation (5.7).

From the algebra above, a price index can be re-expressed as:

$$P_{njt}^{1-\sigma} = \sum_{m\in\mathcal{N}} \left[M_{mj} \underbrace{(\mu\tau_{mnj}c_{mjt})^{1-\sigma}}_{\text{Unit cost}} \times \underbrace{(\bar{\phi}_{mjt}^{avg})^{\sigma-1}}_{\substack{\text{Average productivity}\\ \text{including subsidies}}} \right] + \underbrace{(\tau_{nj}^{x}c_{jt}^{f})^{1-\sigma}}_{\substack{\text{Consumer foreign}\\ \text{market access}}}, \tag{B.1}$$

where

$$\begin{aligned} \bar{\phi}_{njt}^{avg} &= \bar{\phi}^{avg}(\lambda_{njt-1}^{T}, \lambda_{njt}^{T}, s_{njt}, \phi_{njt}^{min}) \\ &= \frac{\theta f(\lambda_{njt-1}^{T})(\phi_{njt}^{min})^{\sigma-1}}{\tilde{\theta}(1-\kappa^{-\theta})} \left\{ \left(\left(\frac{\eta}{1-s_{njt}} \right)^{\sigma-1} - 1 \right) (\tilde{\lambda}_{njt}^{T})^{\frac{\tilde{\theta}}{\theta}} + \left(1 - \left(\frac{\eta}{1-s_{njt}} \right)^{\sigma-1} \kappa^{-\tilde{\theta}} \right) \right\}, \end{aligned} \tag{B.2}$$

$\tilde{\lambda}_{njt}^{T} = (1-\kappa^{-\theta})\lambda_{njt}^{T} + \kappa^{-\theta}$ and $\tilde{\theta} = \theta - (\sigma - 1)$.[43] Price index depend on the three terms: unit cost, average productivity including subsidies $\bar{\phi}_{njt}^{avg}$, and consumer foreign market access $(\tau_{nj}^{x}c_{jt}^{f})^{1-\sigma}$. $\bar{\phi}_{njt}^{avg}$ captures how region n can produce sector j intermediate varieties at cheaper cost than other regions. Region n can produce at cheaper costs if it has technological advantages (λ_{njt}^{T}, λ_{njt-1}^{T}, ϕ_{njt}^{min}) or higher subsidies (s_{njt}). Holding other variables constant, the price index is lower when (i) neighboring regions have lower unit costs (either lower τ_{nmj} or c_{mjt}), (ii) neighboring regions have higher productivity or obtain more subsidies (higher $\bar{\phi}_{njt}^{avg}$), or (iii) the price of imported inputs is lower (lower τ_{nj}^{x} or c_{jt}^{f}).

The average productivity including subsidies (Equation (B.2)) increases in the share of adopters in the previous period λ_{njt-1}^{T}, the share of adopters in the current period λ_{njt}^{T}, subsidies s_{njt}, and the natural advantage captured by the Pareto lower bound ϕ_{njt}^{min}. The share of adopters in $t-1$ increases average productivity directly through spillover and indirectly by inducing more firms to adopt technology in period t (Equation (5.7)). The current share of adopters increases the average productivity through direct productivity gains. Subsidies increase the average productivity directly by lowering the cost of production for adopters and indirectly by inducing more firms to become adopters in t. Finally, a natural advantage is an exogenous productivity shifter.

Gross Output and Export. Region n's sector j gross output R_{njt} is the sum of gross output for domestic expenditures R_{njt}^{d} and the total value of export R_{njt}^{x}: $R_{njt} = R_{njt}^{d} + R_{njt}^{x}$.

Regional exports can be written as

$$\begin{aligned} R_{njt}^{x} = M_{nj} \Bigg[&\int_{\bar{\phi}_{njt}^{T}}^{\bar{\phi}_{njt}^{max}} \left(\frac{\sigma}{\sigma-1} \frac{\tau_{nj}^{x}(1-s_{njt})c_{njt}}{\eta f(\lambda_{njt-1}^{T})\phi_{it}} \right)^{1-\sigma} dG_{njt}(\phi_{it}) \\ &+ \int_{\bar{\phi}_{njt}^{x}}^{\bar{\phi}_{njt}^{T}} \left(\frac{\sigma}{\sigma-1} \frac{\tau_{nj}^{x}c_{njt}}{f(\lambda_{njt-1}^{T})\phi_{it}} \right)^{1-\sigma} dG_{njt}(\phi_{it}) \Bigg] D_{jt}^{f}, \end{aligned} \tag{B.3}$$

where the first and the second terms inside the brackets are the total sum of exports by adopters and non-adopters in sector j of region n.

[43] When $\lambda_{njt}^{T} \to 0$, the average productivity becomes $\bar{\phi}_{njt}^{avg} = \frac{\theta}{\tilde{\theta}(1-\kappa^{-\theta})} f(\lambda_{njt-1}^{T})(\phi_{njt}^{min})^{\sigma-1}(1-\kappa^{-\tilde{\theta}})$. When $\lambda_{njt}^{T} \to 1$, the average productivity becomes $\bar{\phi}_{njt}^{avg} = \frac{\theta}{\tilde{\theta}(1-\kappa^{-\theta})} f(\lambda_{njt-1}^{T})(\phi_{njt}^{min})^{\sigma-1}\left(\frac{\eta}{1-s_{njt}}\right)^{\sigma-1}(1-\kappa^{-\tilde{\theta}})$.

The first term inside the bracket can be expressed as:

$$\begin{aligned}
&\int_{\bar{\phi}^T_{njt}}^{\bar{\phi}^{max}_{njt}} \left(\frac{\sigma}{\sigma-1}\frac{\tau^x_{nj}s_{njt}c_{njt}}{\eta f(\lambda^T_{njt-1})\phi_{it}}\right)^{1-\sigma} dG_{njt}(\phi_{it}) \\
&= \frac{\theta f(\lambda^T_{njt-1})^{\sigma-1}}{\tilde{\theta}(1-\kappa^{-\theta})}(\mu c_{njt})^{1-\sigma}\left(\frac{\eta}{1-s_{njt}}\right)^{\sigma-1}(\phi^{min}_{njt})^{-\theta}\left((\bar{\phi}^T_{njt})^{-\tilde{\theta}} - (\kappa\phi^{min}_{njt})^{-\tilde{\theta}}\right) \\
&= \frac{\theta f(\lambda^T_{njt-1})^{\sigma-1}}{\tilde{\theta}(1-\kappa^{-\theta})}(\mu c_{njt})^{1-\sigma}\left(\frac{\eta}{1-s_{njt}}\right)^{\sigma-1}(\phi^{min}_{njt})^{\sigma-1}\left((\tilde{\lambda}^T_{njt})^{\frac{\tilde{\theta}}{\theta}} - \kappa^{-\tilde{\theta}}\right),
\end{aligned} \tag{B.4}$$

where $\tilde{\lambda}^T_{njt} = (1-\kappa^{-\theta})\lambda^T_{njt} + \kappa^{-\theta}$. The last equality comes from Equation (5.7).

The second term can be re-expressed as:

$$\begin{aligned}
&\int_{\bar{\phi}^x_{njt}}^{\bar{\phi}^T_{njt}} \left(\frac{\sigma}{\sigma-1}\frac{\tau^x_{nj}c_{njt}}{f(\lambda^T_{njt-1})\phi_{it}}\right)^{1-\sigma} dG_{njt}(\phi_{it}) \\
&= \frac{\theta f(\lambda^T_{njt-1})^{\sigma-1}}{\tilde{\theta}(1-\kappa^{-\theta})}(\mu c_{njt})^{1-\sigma}(\phi^{min}_{njt})^{-\theta}\left((\bar{\phi}^x_{njt})^{-\tilde{\theta}} - (\bar{\phi}^T_{njt})^{-\tilde{\theta}}\right) \\
&= \frac{\theta f(\lambda^T_{njt-1})^{\sigma-1}}{\tilde{\theta}(1-\kappa^{-\theta})}(\mu c_{njt})^{1-\sigma}(\phi^{min}_{njt})^{\sigma-1}\left((\tilde{\lambda}^x_{njt})^{\frac{\tilde{\theta}}{\theta}} - (\tilde{\lambda}^T_{njt})^{\frac{\tilde{\theta}}{\theta}}\right),
\end{aligned} \tag{B.5}$$

where $\tilde{\lambda}^x_{njt} = (1-\kappa^{-\theta})\lambda^x_{njt} + \kappa^{-\theta}$. The last equality comes from the fact that $\lambda^x_{njt} = 1 - G_{njt}(\bar{\phi}^x_{njt})$.

Using Equations (B.3), (B.4), and (B.5), and $M^x_{njt} = M_{nj} \times \lambda^x_{njt}$, regional exports can be expressed as:

$$R^x_{njt} = M^x_{njt}(\mu c_{njt})^{1-\sigma} \times \underbrace{(\bar{\phi}^{avg,x}_{njt})^{\sigma-1}}_{\substack{\text{Exporters'}\\ \text{average productivity}\\ \text{including subsidies}}} \times \underbrace{(\tau^x_{nj})^{1-\sigma}D^f_{jt}}_{\substack{\text{Firm foreign}\\ \text{market access}}},$$

where

$$\begin{aligned}
\bar{\phi}^{avg,x}_{njt} &= \bar{\phi}^{avg,x}(\lambda^T_{njt-1}, \lambda^T_{njt}, \lambda^x_{njt}, s_{njt}, \phi^{min}_{njt}) \\
&= \frac{\theta f(\lambda^T_{njt-1})(\phi^{min}_{njt})^{\sigma-1}}{\tilde{\theta}(1-\kappa^{-\theta})}\frac{(\tilde{\lambda}^x_{njt})^{\frac{\tilde{\theta}}{\theta}}}{\lambda^x_{njt}} \\
&\quad \times \left\{\left(\left(\frac{\eta}{1-s_{njt}}\right)^{\sigma-1} - 1\right)\left(\frac{\tilde{\lambda}^T_{njt}}{\tilde{\lambda}^x_{njt}}\right)^{\frac{\tilde{\theta}}{\theta}} + \left(1 - \left(\frac{\eta}{1-s_{njt}}\right)^{\sigma-1}\kappa^{-\tilde{\theta}}(\tilde{\lambda}^x_{njt})^{-\frac{\tilde{\theta}}{\theta}}\right)\right\},
\end{aligned} \tag{B.6}$$

$\tilde{\lambda}^x_{njt} = (1-\kappa^{-\theta})\lambda^x_{njt} + \kappa^{-\theta}$ and $\bar{\phi}^{avg,x}_{njt}$ represent the exporters' average productivity including subsidies.

Gross output for domestic expenditures and regional exports are written as:

$$R^d_{njt} = M_{nj}(\mu c_{njt})^{1-\sigma} \times \underbrace{(\bar{\phi}^{avg}_{njt})^{\sigma-1}}_{\substack{\text{Average productivity}\\ \text{including subsidies}}} \times \underbrace{\sum_{m\in\mathcal{N}} \tau^{1-\sigma}_{nmj}P^{\sigma-1}_{mjt}E_{mjt}}_{\substack{\text{Firm domestic}\\ \text{market access}}}.$$

and

$$R^x_{njt} = M^x_{njt}(\mu c_{njt})^{1-\sigma} \times \underbrace{(\bar{\phi}^{avg,x}_{njt})^{\sigma-1}}_{\substack{\text{Exporters'}\\ \text{average productivity}\\ \text{including subsidies}}} \times \underbrace{(\tau^x_{nj})^{1-\sigma} D^f_{jt}}_{\substack{\text{Firm foreign}\\ \text{market access}}}.$$

Average productivity increases for both total domestic sales and export, as do access to markets and subsidies. The cost of production also decreases for domestic sales and exports.

One difference between $\bar{\phi}^{avg,x}_{njt}$ (Equation (B.6)) and $\bar{\phi}^{avg}_{njt}$ (Equation (B.2)) is that $\bar{\phi}^{avg,x}_{njt}$ also depends on shares of exporters λ^x_{njt}. λ^x_{njt} captures selection induced by fixed export costs. Because of fixed export costs, only more productive firms self-select into exporting, which makes the average productivity of exporters higher than the average productivity of all firms: $\bar{\phi}^{avg,x}_{njt} > \bar{\phi}^{avg}_{njt}$. The average productivity of exporters decreases in shares of exporters λ^x_{njt} because larger shares of exporters implies that less productive firms participate in exporting, which in turn leads to weaker selection effects and lowers the average productivity of exporters. At one extreme where all firms are exporting ($\lambda^x_{njt} = 1$), there is no selection effect and $\bar{\phi}^{avg,x}_{njt}$ becomes equal to $\bar{\phi}^{avg}_{njt}$.

B.2 Analytical Results: Multiple Steady States

B.2.1 Derivation of the Equilibrium Share of Adopters in the Simplified Model.

In the simplified model, the cutoff for adoption is expressed as

$$(\bar{\phi}^T_t)^{\sigma-1} = \frac{\sigma P_t F^T}{(\eta^{\sigma-1}-1)(\mu w_t)^{1-\sigma} f(\lambda^T_{t-1})^{\sigma-1} P^\sigma_t Q_t} \tag{B.7}$$

and the probability of adoption is $\lambda^T_t = (\bar{\phi}^T_t)^{-\theta}$, which can be re-written as

$$(\lambda^T_t)^{-\frac{1}{\theta}} = \bar{\phi}^T_t \tag{B.8}$$

First, we show that

$$Q_t = \left[\frac{\theta}{\tilde{\theta}}\Big((\eta^{\sigma-1}-1)(\lambda^T_t)^{1-\frac{\sigma-1}{\theta}}+1\Big)\right]^{\frac{1}{\sigma-1}} f(\lambda^T_{t-1})$$

and

$$\frac{w_t}{P_t} = \frac{\sigma-1}{\sigma}\left[\frac{\theta}{\tilde{\theta}}\Big((\eta^{\sigma-1}-1)(\lambda^T_t)^{1-\frac{\sigma-1}{\theta}}+1\Big)\right]^{\frac{1}{\sigma-1}} f(\lambda^T_{t-1}),$$

where $\tilde{\theta} = \theta - (\sigma-1)$. Note that

$$\frac{L_t}{Q_t} = \frac{\int l(\omega)d\omega}{Q_t} = \int \frac{y(\omega)}{Q}\frac{1}{z(\omega)}d\omega = \int \frac{1}{z(\omega)}\left(\frac{p(\omega)}{P_t}\right)^{-\sigma} d\omega,$$

where $z(\omega) = \eta(\omega) f(\lambda^T_{t-1})\phi(\omega)$ for adopters and $z(\omega) = f(\lambda^T_{t-1})\phi(\omega)$ for non-adopters. After substituting $L_t = 1$ and $(p(\omega)/P)^{-\sigma} = \frac{\sigma}{\sigma-1}\frac{w_t}{z(\omega)}$ which holds under assumption of monopolistic competition in the above equation, we obtain $Q_t = [\int z(\omega)^{\sigma-1}d\omega]^{\frac{1}{\sigma-1}}$. Using the assumption of Pareto distribution

and the cutoff property, we can further derive that

$$Q_t = \left[\frac{\theta}{\tilde{\theta}}\Big((\eta^{\sigma-1}-1)(\bar{\phi}_t^T)^{\theta-(\sigma-1)}+1\Big)\right]^{\frac{1}{\sigma-1}} f(\lambda_{t-1}^T)$$
$$= \underbrace{\left[\frac{\theta}{\tilde{\theta}}\Big((\eta^{\sigma-1}-1)(\lambda_t^T)^{1-\frac{\sigma-1}{\theta}}+1\Big)\right]^{\frac{1}{\sigma-1}}}_{=A(\lambda_t^T)} \times f(\lambda_{t-1}^T), \quad \text{(B.9)}$$

where the second equality is derived from Equation (B.8). Similarly, using

$$P_t = [\mu w_t \int z(\omega)^{\sigma-1} d\omega]^{\frac{1}{1-\sigma}},$$

we can derive that

$$\frac{w_t}{P_t} = \frac{w_t}{[\int(\mu w_t/z_{it}(\omega))^{1-\sigma}]^{\frac{1}{1-\sigma}}} = \frac{\sigma-1}{\sigma}\left[\frac{\theta}{\tilde{\theta}}\Big((\eta^{\sigma-1}-1)(\lambda_t^T)^{1-\frac{\sigma-1}{\theta}}+1\Big)\right]^{\frac{1}{\sigma-1}} f(\lambda_{t-1}^T), \quad \text{(B.10)}$$

where the second equality is also derived from Equation (B.8).

Substituting Equations (B.8), (B.9), and (B.10) into Equation (B.7), we can obtain that

$$\lambda_t^T = \left(\frac{(\eta^{\sigma-1}-1)}{\sigma F^T} \times A(\lambda_t^T)^{2-\sigma} \times f(\lambda_{t-1}^T)\right)^{\frac{\theta}{\sigma-1}}. \quad \text{(B.11)}$$

Let $\hat{\lambda}_t^T$ be the solution of Equation (B.11). Because the equilibrium share is bounded by 1, the equilibrium share is defined as follows:

$$\lambda_t^T = \begin{cases} \hat{\lambda}_t^T & \text{if} \quad A(\hat{\lambda}_t^T)^{2-\sigma} f(\lambda_{t-1}^T)\frac{\eta^{\sigma-1}-1}{\sigma F^T} < 1 \\ 1 & \text{if} \quad A(\hat{\lambda}_t^T)^{2-\sigma} f(\lambda_{t-1}^T)\frac{\eta^{\sigma-1}-1}{\sigma F^T} \geq 1. \end{cases}$$

B.2.2 Proofs of Proposition 1: Multiple Steady States

Proof of Proposition 1(i). We defined equilibrium using the following equation:

$$\lambda_t^T = \left[A(\lambda_t^T)^{2-\sigma}\frac{(\eta^{\sigma-1}-1)}{\sigma F^T} \times f(\lambda_{t-1}^T)\right]^{\frac{\theta}{\sigma-1}}.$$

Because the left hand side strictly increases in λ_t^T but the right hand side strictly decreases in λ_t^T due to Assumption 1(v), there exists a unique value of λ_t^T that satisfies this equation. If the obtained λ_t^T from this equation is greater than 1, $\lambda_t^T = 1$.

□

Proof of Proposition 1(ii) and (iii). We prove Proposition 1(ii) and (iii) using the implicit function theorem. Let

$$G(\lambda_t^T; \eta, \delta, \lambda_{t-1}^T) = A(\lambda_t^T)^{2-\sigma} \times f(\lambda_{t-1}^T)\frac{(\eta^{\sigma-1}-1)}{\sigma F^T} - (\lambda_t^T)^{\frac{\sigma-1}{\theta}} \quad \text{(B.12)}$$

where

$$A(\lambda_t^T) = \left[\frac{\theta}{\theta-(\sigma-1)}\Big((\eta^{\sigma-1}-1)(\lambda_t^T)^{\frac{\theta-(\sigma-1)}{\theta}}+1\Big)\right]^{\frac{1}{\sigma-1}} \quad \text{and} \quad f(\lambda_{t-1}^T) = exp(\delta\lambda_{t-1}^T).$$

Note that in period t, firms take $f(\lambda_{t-1}^T)$ as given, so $f(\lambda_{t-1}^T)$ is just a constant in the above equation.

Taking the derivative of Equation (B.12) with respect to λ_t^T, we obtain

$$\frac{\partial G}{\partial \lambda_t^T} = \left(\frac{2-\sigma}{\sigma-1}\right)A(\lambda_t^T)^{3-2\sigma}(\eta^{\sigma-1}-1)\frac{\theta-(\sigma-1)}{\theta}(\lambda_t^T)^{-\frac{\sigma-1}{\theta}}f(\lambda_{t-1}^T)\frac{(\eta^{\sigma-1}-1)}{\sigma F^T} - \frac{\sigma-1}{\theta}(\lambda_t^T)^{\frac{-\theta+(\sigma-1)}{\theta}} < 0, \quad \text{(B.13)}$$

where the last inequality comes from $\sigma > 2$ (Assumption 1).

Taking the derivative of Equation (B.12) with respect to λ_{t-1}^T, we obtain

$$\frac{\partial G}{\partial \lambda_{t-1}^T} = A(\lambda_t^T)^{2-\sigma}\frac{\eta^{\sigma-1}-1}{\sigma F^T}exp(\delta\lambda_{t-1}^T)\delta > 0. \quad \text{(B.14)}$$

Applying the implicit function theorem and using the signs of Equations (B.13) and (B.14), we obtain

$$\frac{\partial \lambda_t^T}{\partial \lambda_{t-1}^T} = -\frac{\partial G/\partial \lambda_t^T}{\partial G/\partial \lambda_{t-1}^T} > 0,$$

which proves that λ_t^T strictly increases in λ_{t-1}^T. This proves Proposition 1(ii).

Taking the derivative of Equation (B.12) with respect to η, we obtain

$$\begin{aligned}\frac{\partial G}{\partial \eta} &= \left(\frac{2-\sigma}{\sigma-1}\right)A(\lambda_t^T)^{3-2\sigma}f(\lambda_{t-1}^T)\frac{\theta}{\theta-(\sigma-1)}(\lambda_t^T)^{\frac{\theta-(\sigma-1)}{\theta}}(\sigma-1)\eta^{\sigma-2}\frac{(\eta^{\sigma-1}-1)}{\sigma F^T} \\ &\quad + A(\lambda_t^T)^{2-\sigma}f(\lambda_{t-1}^T)\frac{(\sigma-1)\eta^{\sigma-2}}{\sigma F^T} \qquad \text{(B.15)} \\ &= A(\lambda_t^T)^{3-2\sigma}f(\lambda_{t-1}^T)\frac{(\sigma-1)\eta^{\sigma-2}}{\sigma F^T}\frac{\theta}{\theta-(\sigma-1)}\left[\frac{1}{\sigma-1}(\eta^{\sigma-1}-1)(\lambda_t^T)^{\frac{\theta}{\theta-(\sigma-1)}}+1\right] > 0.\end{aligned}$$

Taking the derivative of Equation (B.12) with respect to δ, we obtain

$$\frac{\partial G}{\partial \delta} = A(\lambda_t^T)^{2-\sigma}\frac{\eta^{\sigma-1}-1}{\sigma F^T}exp(\delta\lambda_{t-1}^T)\lambda_{t-1}^T > 0. \quad \text{(B.16)}$$

Applying the implicit function theorem and using the signs of Equations (B.13), (B.16), and (B.15),

$$\frac{\partial \lambda_t^T}{\partial \eta} = -\frac{\partial G/\partial \lambda_t^T}{\partial G/\partial \eta} > 0$$

and

$$\frac{\partial \lambda_t^T}{\partial \delta} = -\frac{\partial G/\partial \lambda_t^T}{\partial G/\partial \delta} > 0.$$

This proves Proposition 1(iii). □

Proof of Proposition 1(iv). First, we show that λ_t^T is strictly convex in λ_{t-1}^T. To show the strict convexity, we have to show that $\frac{\partial^2 \lambda_t^T}{\partial(\lambda_{t-1}^T)^2} > 0$. We show this by applying the implicit function theorem and doing some tedious algebra. Applying the implicit function theorem,

$$\frac{\partial^2 \lambda_t^T}{\partial(\lambda_{t-1}^T)^2} = -\frac{1}{(\partial G/\partial \lambda_t^T)^3} \times \left[\frac{\partial G}{\partial \lambda_{t-1}^T} \times \left(\frac{\partial G}{\partial \lambda_t^T}\right)^2 - \left(\frac{\partial^2 G}{\partial \lambda_t^T \partial \lambda_{t-1}^T} + \frac{\partial^2 G}{\partial \lambda_{t-1}^T \partial \lambda_t^T}\right) \times \frac{\partial G}{\partial \lambda_{t-1}^T} \times \frac{\partial G}{\partial \lambda_t^T} + \frac{\partial^2 G}{\partial(\lambda_t^T)^2} \times \left(\frac{\partial G}{\partial \lambda_{t-1}^T}\right)^2\right]. \tag{B.17}$$

We examine the sign of each term in the above equation.

$$\frac{\partial^2 G}{\partial(\lambda_{t-1}^T)^2} = A(\lambda_t^T)^{2-\sigma} \frac{(\eta^{\sigma-1}-1)}{\sigma F^T} exp(\delta \lambda_{t-1}^T)\delta^2 > 0. \tag{B.18}$$

$$\frac{\partial^2 G}{\partial \lambda_t^T \partial \lambda_{t-1}^T} = \frac{\partial^2 G}{\partial \lambda_{t-1}^T \partial \lambda_t^T} = \frac{2-\sigma}{\sigma-1} A(\lambda_t^T)^{3-2\sigma} \times \left[\frac{\theta-(\sigma-1)}{\theta}(\eta^{\sigma-1}-1)(\lambda_t^T)^{-\frac{\sigma-1}{\theta}}\right] exp(\delta \lambda_{t-1}^T)\lambda_{t-1}^T < 0. \tag{B.19}$$

$$\begin{aligned}\frac{\partial^2 G}{\partial(\lambda_t^T)^2} = {} & \frac{(2-\sigma)(3-\sigma)}{(\sigma-1)^2} A(\lambda_t^T)^{2-2\sigma)} \left[\frac{\theta-(\sigma-1)}{\theta}(\lambda_t^T)^{-\frac{\sigma-1}{\theta}}(\eta^{\sigma-1}-1)\right]^2 exp(\delta \lambda_{t-1}^T)\frac{(\eta^{\sigma-1}-1)}{\sigma F^T} \\ & + \frac{\sigma-2}{\theta} A(\lambda_t^T)^{3-2\sigma}(\eta^{\sigma-1}-1)\frac{\theta-(\sigma-1)}{\theta}(\lambda_t^T)^{-\frac{\sigma-1}{\theta}-1} exp(\delta \lambda_{t-1}^T)\frac{(\eta^{\sigma-1}-1)}{\sigma F^T} \\ & + \frac{\sigma-1}{\theta}\frac{\theta-(\sigma-1)}{\theta}(\lambda_t^T)^{-\frac{\theta-(\sigma-1)}{\theta}-1} > 0.\end{aligned} \tag{B.20}$$

Substituting Equations (B.13), (B.14), (B.18), (B.19), and (B.20) in Equation (B.17), we obtain $\frac{\partial^2 \lambda_t^T}{\partial(\lambda_{t-1}^T)^2} > 0$, which proves strict convexity.

Because the intercept of λ_t^T-axis is always positive and λ_t^T is strictly increasing and strictly convex in λ_{t-1}^T, the locus defined by $(\lambda_{t-1}^T, \lambda_t^T)$ that satisfies Equation (5.21) can intersect with the 45-degree line two times at most.[44] Because $\lambda_t^T(\delta, \eta)$ strictly increases in δ and η, there exists $\underline{\delta}$ and $\underline{\eta}$ such that the 45 degree line and Equation (B.12) meet at $\lambda_{t-1}^T = 1$. Also, by the same logic, there exists $\bar{\delta}$ and $\bar{\eta}$ such that the 45 degree line is tangent to Equation (B.12). The two lines meet at least twice for $\delta \in [\underline{\delta}, \bar{\delta}]$ and $\eta \in [\underline{\eta}, \bar{\eta}]$.

□

[44] The intercept is always positive because of the assumption of unbounded Pareto distribution which guarantees a positive share of adopters at $\lambda_{t-1}^T = 0$.

Proof of Proposition 1(v). The welfare of household is $\frac{w_t+\Pi_t}{P_t}$ where Π_t are the aggregate profits summed across all firms in the economy.[45] This can be expressed as $\frac{w_t}{P_t}+\frac{\Pi_t}{P_t}$. Using Equations (B.9) and (B.10) and the following expression

$$\frac{\Pi_t}{P_t}=\frac{1}{\sigma}\mu^{1-\sigma}(w_t/P_t)^{1-\sigma}\left[\int_{\omega\in\Omega}z(\omega)^{\sigma-1}d\omega\right]Q_t,$$

we can derive that the welfare can be expressed as $f(\lambda_{t-1}^T)A(\lambda_t^T)$. The welfare in the steady state is $f(\lambda^{T*})A(\lambda^{T*})$, which strictly increases in λ^{T*}. Therefore, the equilibrium with a larger mass of adopters Pareto-dominates the equilibrium with a smaller mass of adopters.

□

B.2.3 Source of Dynamic Externality

In this subsection, we use the simplified model to show that dynamic externalities are generated because fixed adoption costs are in units of final goods. We show that when fixed adoption costs are in units of labor, there are no dynamic externalities.

Suppose fixed adoption costs are in units of labor. The cutoff for adoption is defined as

$$(\bar{\phi}_t^T)^{\sigma-1}=\frac{\sigma w_t F^T}{(\eta^{\sigma-1}-1)(\mu w_t)^{1-\sigma}f(\lambda_{t-1}^T)^{\sigma-1}P_t^\sigma Q_t},$$

which is similar to Equation (B.7), but P_tF^T is replaced with w_tF^T. $\frac{w_t}{P_t}$ and Q_t are defined analogously to Equations (B.9) and (B.10) regardless of the fact that fixed adoption costs are in units of labor. Substituting Equations (B.9) and (B.10) into the above cutoff, we can derive that

$$\lambda_t^T=\left(\frac{(\eta^{\sigma-1}-1)}{\sigma F^T}\times\mu\times A(\lambda_t^T)^{1-\sigma}\right)^{\frac{\theta}{\sigma-1}}. \tag{B.21}$$

This expression differs from the expression of Equation (B.11) in that μ replaces $f(\lambda_{t-1}^T)$.

The equilibrium share of adopters in Equation (B.21) shows that the static short-run equilibrium is uniquely determined regardless of values of λ_{t-1}^T. This is because a fixed adoption cost is in units of labor. If there were a higher share of adopters in the previous period, that would increase the overall productivity in t. The increase in productivity would lead to increases in the overall demand for labor. As labor demands increase the equilibrium wage, fixed adoption costs (w_tF^T) would become higher. In the equilibrium, increases in fixed adoption costs would exactly cancel out increases in overall productivity, which in turn would mean that the equilibrium share of adopters would not be affected by λ_{t-1}^T (Equation (B.21)).

B.2.4 Temporary Subsidies Can Have Permanent Effects Only When Multiple Steady States Exist

We show that temporary subsidies cannot have permanent effects when multiple steady states do not exist in the simplified model in Section 5.5. Suppose temporary subsidies are provided temporarily for periods $t\in\{t_0,\ldots,t_1\}$, where $0<t_0<t_1$. Between t_0 and $t_1<\infty$, adopters are subject to an input subsidy rate $\bar{s}<1$. Also suppose that the short-run equilibrium curve is not sufficiently nonlinear enough to generate multiple steady states and there is only a unique steady-state. For

[45] Note that $L_t=1$.

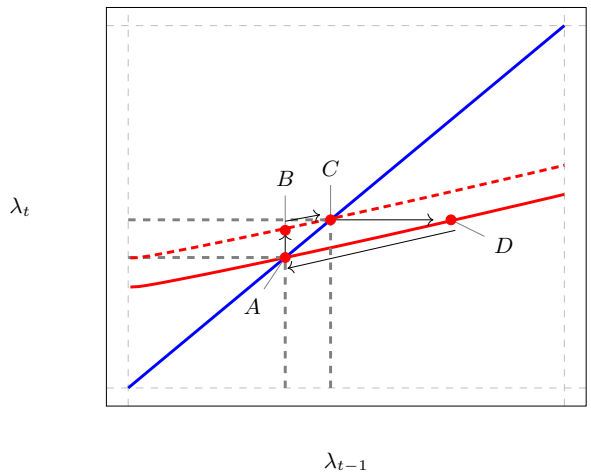


Figure B1. Temporary Subsidies Can Have Permanent Effects Only When Multiple Steady States Exist.

Notes. This figure illustrates that when multiple steady states do not exist, temporary adoption subsidies cannot have permanent effects.

simplicity, we assume that the economy starts at the original steady state in the initial time period.

Figure B1 graphically illustrates that temporary subsidies have temporary effects when there is a unique steady state. The solid red locus and the dashed red loci are the short-run equilibrium curves when adoption subsidies are not provided and provided permanently, respectively. In this economy, the strength of the spillover is not large enough to generate multiple steady states. At t_0, an economy jumps up from the original steady state A to a new point B, which is on the new short-run equilibrium curve when subsidy $\bar{s}$ is permanently provided. Point C is the steady state of this new short-run equilibrium curve. Therefore, between t_0 and t_1, it converges to the new steady state C. However, after the end of the temporary subsidies at t_1, the short-run equilibrium curve moves back to the original short-run equilibrium curve and the economy jumps to D and starts converging to the original steady state A.

Even if there is a unique steady state, there is still room for policy interventions due to externalities. However, these policy interventions have to be provided permanently to have permanent effects. For example, the new steady state in Figure B1 can have a higher level of welfare than the original steady state, and this new steady state can be sustained when $\bar{s}$ is permanently provided each period. This would be similar to the static setting with externalities. However, these permanent policies are inconsistent with the industrialization pattern in South Korea, where adoption subsidies were only provided from 1973 to 1979.

B.3 Possible Microfoundations for Adoption Spillovers

B.3.1 Local Diffusion of Knowledge

Setup. Consider a closed economy with one sector and N regions. For notational convenience, we omit a subscript j that denotes sectors. Each firm faces a CES demand and is monopolistic for its

own variety. Goods are freely tradable across regions.

Firms' Maximization Problem. A firm receives exogenous productivity $\tilde{\phi}_{it}$, which is independent and identically distributed across firms. Given this exogenous productivity, firms make two static decisions each period: (1) whether to adopt advanced foreign technology T_{it}; and (2) a level of innovation a_{it} as in Desmet and Rossi-Hansberg (2014).

Given $\tilde{\phi}_{it}$, a firm optimally chooses (1) whether to adopt technology T_{it} and (2) a level of innovation a_{it}:

$$\pi_{it} = \max_{T_{it}\in\{0,1\}, a_{it}\in[0,\infty)} \left\{ \frac{1}{\sigma}\left(\frac{\sigma}{\sigma-1}\frac{w_{nt}}{\tilde{\eta}^{T_{it}} a_{it}^{\gamma_1}\tilde{\phi}_{it}}\right)^{1-\sigma} P_t^{\sigma-1}E_t - T_{it}P_tF^T - w_{nt}a_{it}^{\alpha_1}g(\lambda_{nt-1}^T)B_t \right\}, \quad \text{(B.22)}$$

where $T_{it} \in \{0,1\}$ is a dummy variable for adoption status, $\tilde{\eta}$ is direct productivity gains from adoption, w_{nt} are local wages, $P_t^{\sigma-1}E_t$ is market size, F^T is the total fixed adoption cost in units of labor, and $a_{it}^{\alpha_1}g(\lambda_{nt-1}^T)B_t$ is the cost of innovation in units of labor. $\alpha_1 > 0$ holds so that the cost of adoption increases in a_{it}. To simplify the algebra, we assume that B_t is proportional to market size $P_t^{\sigma-1}E_t$; that is, $B_t = b_1P_t^{\sigma-1}E_t$ with a constant term b_1 .

The positive externalities of adoption come from $g(\lambda_{nt-1}^T)$ of the cost of innovation. We assume that $\frac{\partial g(\lambda_{nt-1}^T)}{\partial \lambda_{nt-1}^T} < 0$ holds, so a larger share of adopters in the previous period decreases the cost of innovation in the current period. This cost specification captures local diffusion knowledge from newly adopted technologies. With more firms adopting advanced technologies, other local firms are more likely to learn new ideas from these adopters and can use this knowledge for their own innovation. $g(\lambda_{nt-1}^T)$ captures the local diffusion of ideas in a reduced-form. We assume that $\gamma_1(\sigma-1)-\alpha_1+1 < 0$ holds.[46]

A firm's optimal choice of a_{it} is characterized by the following first-order condition:

$$\gamma_1(\sigma-1)a_{it}^{\gamma_1(\sigma-1)}\frac{1}{\sigma}\left(\frac{\sigma}{\sigma-1}\frac{w_{nt}}{\tilde{\eta}^{T_{it}}\tilde{\phi}_{it}}\right)^{1-\sigma} - b_1w_{nt}\alpha_1a_{it}^{\alpha_1-1}g(\lambda_{nt-1}^T) = 0,$$

which gives the optimal level of own innovation a_{it}^*

$$a_{it}^* = \bar{C}_{nt}^1 g(\lambda_{nt-1}^T)^{\frac{-1}{\alpha_1-1-\gamma_1(\sigma-1)}}(\tilde{\eta}^{T_{it}}\tilde{\phi}_{it})^{\frac{1-\sigma}{\alpha_1-1-\gamma_1(\sigma-1)}},$$

where $\bar{C}_{nt}^1$ is a collection of constants and variables that are common within region n.[47] Note that both $\frac{\delta_{-1}}{\alpha_1-1-\gamma_1(\sigma-1)} > 0$ and $\frac{1-\sigma}{\alpha_1-1-\gamma_1(\sigma-1)} > 0$ hold. This implies that the optimal amount of innovation increases in a share of adopters in the previous period λ_{nt-1}^T, increases if $T_{it} = 1$, and increases in exogeonus productivity $\tilde{\phi}_{it}$. Substituting the optimal a_{it}^* into Equation (B.22), a firm's maximization

[46] This parameter restriction guarantees the second-order condition of a firm's maximization problem.

[47] Specifically, $\bar{C}_{nt}^1 = \left[\frac{\sigma b_1\alpha_1}{\gamma_1(\sigma-1)}\left(\frac{\sigma}{\sigma-1}\right)^{\sigma-1}\right]^{\frac{1}{\gamma_1(\sigma-1)-\alpha_1+1}} w_{nt}^{\frac{\sigma}{\gamma_1(\sigma-1)-\alpha_1+1}}$.

problem can be rewritten as:

$$\pi_{it} = \max_{T_{it}\in\{0,1\}} \left\{ \frac{1}{\sigma}\left(\frac{\sigma}{\sigma-1} \frac{w_{nt}}{(\bar{C}^1_{nt})^{\gamma_1} g(\lambda^T_{nt-1})^{\frac{-1}{\alpha_1-1-\gamma_1(\sigma-1)}} (\tilde{\eta}^{\frac{\alpha_1-\sigma-\gamma_1(\sigma-1)}{\alpha_1-1-\gamma_1(\sigma-1)}})^{T_{it}} \tilde{\phi}_{it}^{\frac{\alpha_1-\sigma-\gamma_1(\sigma-1)}{\alpha_1-1-\gamma_1(\sigma-1)}}} \right)^{1-\sigma} \times P_t^{\sigma-1} E_t - T_{it} P_t F^T \right\}.$$

Note that $g(\lambda^T_{nt-1})^{\frac{-1}{\alpha_1-1-\gamma_1(\sigma-1)}}$ can be mapped to $f(\lambda^T_{nt-1})$, $\tilde{\phi}_{it}^{\frac{\alpha_1-\sigma-\gamma_1(\sigma-1)}{\alpha_1-1-\gamma_1(\sigma-1)}}$ can be mapped to ϕ_{it}, and $\tilde{\eta}^{\frac{\alpha_1-\sigma-\gamma_1(\sigma-1)}{\alpha_1-1-\gamma_1(\sigma-1)}}$ can be mapped to η in Equation (5.4) in the main text.

Historical Case Study. The case study comes from (Kim, 1997, p. 182-184). Wonil Machinery Work (henceforth Wonil) started its business as a small hot and cold rolling mill producer. One local firm imported a more sophisticated 4-high nonreverse cold rolling mill, which was a technology widely used in developed countries. Wonil's engineers had an opportunity to see how the local firm was operating the state of the art mills, and could obtain technical information indirectly from this local firm. From this opportunity, Wonil could develop its own 4-high cold rolling mill blueprints and start producing them without adopting from foreign countries. This development of own blueprints was considered to be a milestone in the firms' history.

B.3.2 Learning Externalities and Labor Mobility in an Imperfect Labor Market

Setup. Consider a closed economy with one sector and N regions. For notational convenience, we omit a subscript j that denotes sectors. Each firm faces a CES demand and is monopolistic for its own variety. Goods are freely tradable across regions.

In each region, there is a unit measure of engineers and firms. Engineers live two periods, child and adult. They only consume and work in their adulthood. They cannot move to new locations. Once engineers become adults in the second period, they give birth to a child. Engineers who work in firms that adopted technologies pass their knowledge to their children. This learning from parents increases the engineering skills of children when they grow up, which increases engineering skills by $\gamma_1 > 1$. If parents do not work in firms with foreign technology, their children's engineering skills are 1. We assume that the engineering skills of newborn children are 1 if the parents work for non-adopter firms and $\gamma_1 > 1$ if the parents work for adopter firms.

Following Acemoglu (1996), we assume that engineers and firms are randomly matched one to one. The surplus this match generates–that is, the profits generated–is divided among engineers and firms based on Nash bargaining. Managers take a proportion of $\tilde{\beta}$. Once engineers and firms are randomly matched within a region, they jointly maximize profits.

Because the firm makes decisions about adopting technology before the matching happens, it must make these decisions based on anticipated profits. A firm's overall productivity depends on (1) exogenous productivity $\tilde{\phi}_{it}$ that is iid drawn in each period, (2) the engineering skills of matched engineers, and (3) adoption decisions.

Firms' Maximization Problem. Because of the random matching process, firms are matched with engineers with higher engineering skills γ_1 with a probability of λ^T_{nt-1} and they are matched with engineers with lower skills 1 with a probability of $1 - \lambda^T_{nt-1}$.

A firm's maximization problem can be written as

$$\pi_{it} = \max_{T_{it}\in\{0,1\}} (1-\tilde{\beta})\Bigg\{\lambda^T_{nt-1}\frac{1}{\sigma}\left(\frac{\sigma}{\sigma-1}\frac{w_{nt}}{\tilde{\eta}^{T_{it}}\gamma_1\tilde{\phi}_{it}}\right)^{1-\sigma} P_t^{\sigma-1}E_t$$
$$+ (1-\lambda^T_{nt-1})\frac{1}{\sigma}\left(\frac{\sigma}{\sigma-1}\frac{w_{nt}}{\tilde{\eta}^{T_{it}}\tilde{\phi}_{it}}\right)^{1-\sigma} P_t^{\sigma-1}E_t - P_tF^TT_{it}\Bigg\},$$

where λ_{nt-1} is a local share of adopters in the previous period, $\tilde{\phi}_{it}$ is exogenous productivity, w_{nt} is a local wage, T_{it} is a binary adoption decision, F^T is a fixed adoption cost in units of final goods, γ_1 is engineering skills of engineers whose parents worked in adopter firms, and $\tilde{\eta}$ is the direct productivity gain from adoption. Doing some algebra, the maximization problem above can be rewritten as

$$\pi_{it} = \max_{T_{it}\in\{0,1\}} (1-\tilde{\beta})\left\{\frac{1}{\sigma}\left(\frac{\sigma}{\sigma-1}\frac{w_{nt}}{\tilde{f}(\lambda^T_{nt-1})\tilde{\eta}^{T_{it}}\tilde{\phi}_{it}}\right)^{1-\sigma} P_t^{\sigma-1}E_t - P_tF^TT_{it}\right\},$$

where

$$\tilde{f}(\lambda^T_{nt-1}) = [\lambda^T_{nt-1}(\gamma_1^{\sigma-1}-1)+1]^{\frac{1}{\sigma-1}}.$$

$\tilde{f}(\lambda^T_{nt-1})$ increases in the local share of adopters in the previous period, and corresponds to $f(\lambda^T_{njt-1})$ in Equation (5.4) in the main text.

Historical Case Study. In the 1970s, labor mobility across firms was high in South Korea (Kim and Topel, 1995). The average duration of a job in the manufacturing sector in South Korea was around 4 years, which was less than half of the average of a job in the United States (9 years).

Consistent with the aggregate statistics from Kim and Topel (1995), Enos and Park (1988, Chapter 7) provides a historical case study on the diffusion of knowledge through labor mobility in steel industry. The Pohang Iron and Steel Company Ltd. (POSCO), the nation's first integrated steel mill, began operation in 1973. Given South Korea's lack of technology, imported technology played a significant role for POSCO when it began operating. The government heavily subsidized POSCO for the adoption of technology and installation of imported capital equipment associated with the imported technologies. Some of the technicians who left POSCO got jobs in firms located near POSCO that produced capital goods. The technicians helped those firms produce capital equipment that POSCO used, such as equipment for treating water and collecting dust and a large magnetic crane. In the early 1970s, this capital equipment was all imported, but it started to be produced by local suppliers because of knowledge spillover from technicians who had worked at POSCO.

Enos and Park (1988, p. 166) provides another example about the role of labor mobility flows between big firms. Daewoo Heavy Industries Ltd (henceforth Daewoo) built the first diesel engine plant in South Korea after adopting technology from MAN in West Germany. However, one year after Daewoo began operating the plant, Hyundai Heavy Industries (henceforth Hyundai) adopted technology from Perkins in the United States and began producing diesel engines. When it began operations, Hyundai lured skilled engineers who had acquired technological knowledge away from Daewoo by offering them higher salaries. Daewoo lost 33% of its skilled workers as a result.

Both aggregate statistics on labor mobility and two historical case studies support one potential mode of knowledge diffusion through labor mobility.

Appendix C Empirics

C.1 Additional Tables

Table C1: Descriptive Statistics: Winners vs. Losers Design Samples from the Year of the Cancellation to 5 Years before the Cancellation

	Winner				Loser				t-Statistics
	Mean	Med.	SD	Obs.	Mean	Med.	SD	Obs.	(Col. 1 - Col. 5)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
log sales	17.80	18.21	2.22	133	18.46	18.45	1.78	131	2.36 [0.13]
log employment	7.34	7.60	1.23	109	7.07	7.19	1.54	130	0.23 [0.64]
log fixed assets	17.15	17.10	2.26	162	17.19	17.64	2.26	158	0.01 [0.93]
log assets	18.00	17.99	2.10	162	18.12	18.40	2.08	158	0.07 [0.80]
log value added/emp	9.57	9.70	1.26	102	9.95	9.62	1.35	122	1.55 [0.22]

Notes. This table reports the descriptive statistics of the winners vs. losers design samples from the year of the cancellation to 5 Years before the cancellation. Column (9) reports the t-statistics of the mean difference between winners and losers with its p value in brackets. Standard errors are two-way clustered by pair and firm and reported in parenthesis. The number of pairs and firms are 34 and 57. All monetary values are measured in 2015 US dollars.

Table C2: Covariate Balance Test: Winners vs. Losers Design Samples from the Year of the Cancellation to 5 Years before the Cancellation

Dep. Var. $\mathbb{1}[Adopt_{it}]$	Bivariate		Multivariate	
	(1)	(2)	(3)	(4)
log sales	-0.04 (0.03)	-0.1 (0.07)	-0.49 (0.14)***	0.14 (0.47)
N	264	262		
Log employment	0.04 (0.03)	0.05 (0.07)	0.29 (0.15)*	-0.36 (0.5)
N	239	238		
Log fixed assets	0.00 (0.02)	0.02 (0.07)	-0.02 (0.16)	0.16 (0.22)
N	319	319		
Log assets	0.00 (0.02)	0.00 (0.08)	0.22 (0.21)	0.03 (0.33)
N	213	212		
Log labor productivity	-0.06 (0.03)	-0.06 (0.06)	0.27 (0.14)*	-0.36 (0.49)
N	224	221	224	221
F-test [p val]			4.55 [0.00]	0.72 [0.61]
Year FE	✓	✓	✓	✓
Pair FE		✓		✓

Notes. This table reports the covariate balance tests of the winners vs. losers design samples from the year of the cancellation to 5 years before the cancellation. The dependent variable is a dummy variable that equals 1 if a firm adopted technology in the event time. Each cell in columns (1) and (2) reports estimates from a separate bivariate regression. F statistics of joint significance are reported for multivariate regressions, and their p-values are reported in brackets. Standard errors are two-way clustered by pair and firm and reported in parenthesis. This dataset has 33 pairs and 55 firms.

Table C3: Descriptive Statistics of Patenting Activities by Foreign Contractors: Winners vs. Losers Design Samples

	Winner				Loser				t-Statistics	
	Mean	Med.	SD	Obs.	Mean	Med.	SD	Obs.	(Col. 1 - Col. 5)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	
	Panel A. Yearly Measures									
ln(Patent + 1)	1.54	0.00	2.11	34	1.73	0.00	2.55	34	0.14	[0.71]
ln(Citation + 1)	1.71	0.00	2.36	34	2.06	0.00	2.88	34	0.34	[0.57]
$\mathbb{1}$[Patent > 0]	0.44	0.00	0.50	34	0.39	0.00	0.49	34	0.24	[0.63]
$\mathbb{1}$[Citation > 0]	0.42	0.00	0.50	34	0.42	0.00	0.50	34	0.00	[1.00]
	Panel B. Cumulative Measures									
ln(Cum. Patent + 1)	2.20	0.00	2.72	34	2.57	1.15	3.13	34	0.35	[0.56]
ln(Cum. Citation + 1)	2.39	0.00	2.94	34	2.85	1.50	3.41	34	0.46	[0.50]
$\mathbb{1}$[Cum. Patent > 0]	0.47	0.00	0.51	34	0.56	1.00	0.50	34	0.58	[0.45]
$\mathbb{1}$[Cum. Citation > 0]	0.47	0.00	0.51	34	0.56	1.00	0.50	34	0.52	[0.48]

Notes. This table reports the descriptive statistics of patenting activites of two groups of foreign firms that made contracts with winners and losers. Column (9) reports t-statistics of the mean difference between two groups with its p-value in brackets. Patent and Citation are the number of patents made in an event year and the number of citations by other patents in an event year. Cum. Patent and Cum. Citation are the cumulative number of patents made up to an event year and the number of citations by other patents up to an event year. Standard errors are clustered by pair and reported in parenthesis.

Table C4: Local Productivity Spillovers from Technology Adoption - Robustness: Full Sample

Dep. Var.	Log sales					Log revenue TFP				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Spill	4.23***	3.93***	4.45***	3.86***	3.72**	4.75***	3.99**	4.72***	4.45***	3.44*
	(1.18)	(1.43)	(1.31)	(1.19)	(1.52)	(1.63)	(1.90)	(1.73)	(1.58)	(1.82)
$\mathbb{1}[Adopt]$	0.32**	0.26	0.32**	0.31**	0.25	0.15*	0.14	0.15*	0.14	0.12
	(0.15)	(0.20)	(0.15)	(0.15)	(0.19)	(0.09)	(0.10)	(0.09)	(0.09)	(0.10)
ln(Spill-Sales)			−0.01		−0.01			0.00		0.00
			(0.01)		(0.01)			(0.02)		(0.02)
ln(Input-MA)				−0.05***	−0.04*				−0.05***	−0.05**
				(0.02)	(0.02)				(0.02)	(0.02)
Region-Sector FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Conglomerate FE		✓			✓		✓			✓
Adj. R^2	0.19	0.24	0.19	0.19	0.24	0.37	0.43	0.37	0.38	0.43
# clusters (region)	54	54	54	54	54	45	41	45	45	41
# clusters (conglomerate)	702	697	702	702	697	381	338	381	381	338
N	1264	1259	1264	1264	1259	431	387	431	431	387

Notes. This table reports the OLS estimates of Equation (4.3). When we construct the spillover measure defined in Equation (4.2), we lag the adoption status of firms by four years. In Panel A, we use the subsample that include only firms that did not adopt any technology during the sample period. In Panel B, we use the full sample of adopters and non-adopters and control for adoption status. The dependent variables are log sales in columns (1)-(5) and revenue TFP in columns (6)-(10). We estimate revenue TFP based on Wooldridge (2009). The additional controls ln(Spill-Sales) and ln(Input-MA) are additional controls defined in Equations (4.4) and (4.5). In all specifications, we control for region-sector fixed effects and for the initial dependent variable at the start of the sample period. Standard errors are two-way clustered at both region and conglomerate level and are reported in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C5: Local Productivity Spillovers from Technology Adoption: Robustness - 3 Year Lag

Dep. Var.	Log sales					Log revenue TFP				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Panel A: Never-Adopter Sample									
Spill	3.67***	2.96**	4.17***	3.59***	3.23**	2.59*	2.24	2.77*	2.60*	2.11
	(1.25)	(1.40)	(1.43)	(1.20)	(1.55)	(1.41)	(1.43)	(1.45)	(1.36)	(1.43)
ln(Spill-Sales)			–0.02		–0.02			–0.01		–0.01
			(0.02)		(0.01)			(0.02)		(0.02)
ln(Input-MA)				–0.03	–0.02				–0.04**	–0.03
				(0.02)	(0.02)				(0.02)	(0.03)
Adj. R^2	0.18	0.22	0.19	0.19	0.22	0.43	0.41	0.43	0.43	0.41
# clusters (region)	53	53	53	53	53	41	36	41	41	36
# clusters (conglomerate)	636	630	636	636	630	324	275	324	324	275
N	1079	1073	1079	1079	1073	344	292	344	344	292
	Panel B: Full Sample									
Spill	3.48***	3.27***	3.67***	3.22***	3.12**	2.67*	2.05	2.63*	2.51*	1.68
	(1.15)	(1.22)	(1.27)	(1.10)	(1.27)	(1.36)	(1.24)	(1.36)	(1.29)	(1.10)
$\mathbb{1}[Adopt]$	0.31**	0.26	0.31**	0.30*	0.24	0.11	0.11	0.11	0.11	0.09
	(0.15)	(0.20)	(0.15)	(0.15)	(0.19)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)
ln(Spill-Sales)			–0.01		–0.01			0.00		0.01
			(0.01)		(0.01)			(0.02)		(0.02)
ln(Input-MA)				–0.05***	–0.04*				–0.06***	–0.05**
				(0.02)	(0.02)				(0.02)	(0.02)
Adj. R^2	0.19	0.23	0.19	0.19	0.24	0.36	0.42	0.36	0.37	0.43
# clusters (region)	54	54	54	54	54	45	41	45	45	41
# clusters (conglomerate)	702	697	702	702	697	381	338	381	381	338
N	1264	1259	1264	1264	1259	431	387	431	431	387
Region-Sector FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Conglomerate FE		✓			✓		✓			✓

Notes. This table reports the OLS estimates of Equation (4.3). When we construct the spillover measure defined in Equation (4.2), we lag the adoption status of firms by three years. In Panel A, we use the subsample that include only firms that did not adopt any technology during the sample period. In Panel B, we use the full sample of adopters and non-adopters and control for adopters' adoption status. The dependent variables are log sales in columns (1)-(5) and revenue TFP in columns (6)-(10). We estimate revenue TFP based on Wooldridge (2009). ln(Spill-Sales) and ln(Input-MA) are additional controls defined in Equations (4.4) and (4.5). In all specifications, we control for region-sector fixed effects and the initial dependent variable at the start of the sample period. Standard errors are two-way clustered at both region and conglomerate level and are reported in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C6: Local Productivity Spillovers from Technology Adoption: Robustness - 5 Year Lag

Dep. Var.	Log sales					Log revenue TFP				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Panel A: Never-Adopter Sample									
Spill	3.84**	3.48*	4.19**	3.69**	3.63**	4.88***	5.12***	5.03***	4.69***	4.78***
	(1.78)	(1.84)	(1.76)	(1.73)	(1.80)	(1.72)	(1.16)	(1.84)	(1.64)	(1.35)
ln(Spill-Sales)			–0.02		–0.02			–0.01		–0.01
			(0.01)		(0.01)			(0.02)		(0.02)
ln(Input-MA)				–0.03	–0.02				–0.04**	–0.03
				(0.02)	(0.02)				(0.02)	(0.02)
Adj. R^2	0.18	0.22	0.18	0.18	0.22	0.44	0.42	0.44	0.44	0.42
# clusters (region)	53	53	53	53	53	41	36	41	41	36
# clusters (conglomerate)	636	630	636	636	630	324	275	324	324	275
N	1079	1073	1079	1079	1073	344	292	344	344	292
	Panel B: Full Sample									
Spill	4.12***	3.50**	4.28***	3.75***	3.26**	3.86**	3.47*	3.82**	3.53**	2.88
	(1.35)	(1.56)	(1.35)	(1.32)	(1.55)	(1.64)	(2.01)	(1.71)	(1.59)	(1.91)
$\mathbb{1}[Adopt]$	0.32**	0.26	0.32**	0.31*	0.25	0.13	0.13	0.13	0.12	0.11
	(0.16)	(0.20)	(0.16)	(0.16)	(0.20)	(0.09)	(0.10)	(0.09)	(0.09)	(0.10)
ln(Spill-Sales)			–0.01		–0.01			0.00		0.00
			(0.01)		(0.01)			(0.02)		(0.02)
ln(Input-MA)				–0.05***	–0.04*				–0.05***	–0.05**
				(0.02)	(0.02)				(0.02)	(0.02)
Adj. R^2	0.19	0.23	0.19	0.19	0.24	0.36	0.42	0.36	0.38	0.43
# clusters (region)	54	54	54	54	54	45	41	45	45	41
# clusters (conglomerate)	702	697	702	702	697	381	338	381	381	338
N	1264	1259	1264	1264	1259	431	387	431	431	387
Region-Sector FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Conglomerate FE		✓			✓		✓			✓

Notes. This table reports the OLS estimates of Equation (4.3). When we construct the spillover measure defined in Equation (4.2), we lag firms' adoption status by five years. In Panel A, we use the subsample that include only firms that did not adopt any technology during the sample period. In Panel B, we use the full sample of adopters and non-adopters and control for adopters' adoption status. The dependent variables are log sales in columns (1)-(5) and revenue TFP in columns (6)-(10). We estimate revenue TFP based on Wooldridge (2009). ln(Spill-Sales) and ln(Input-MA) are additional controls defined in Equations (4.4) and (4.5). In all specifications, we control for region-sector fixed effects and the initial dependent variable at the start of the sample period. Standard errors are two-way clustered at both region and conglomerate level and are reported in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C7: Local Productivity Spillovers of Technology Adoption - Dependent Variable: Dummy Variable of Adoption of a New Technology - Robustness: 3 Year Lag

Dep. Var.	$\mathbb{1}$[New Contract]					*asinh*(# New Contract)				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Spill	0.42***	0.38***	0.42***	0.42***	0.39***	0.58**	0.52**	0.58**	0.58**	0.52**
	(0.15)	(0.12)	(0.16)	(0.15)	(0.13)	(0.24)	(0.24)	(0.24)	(0.24)	(0.24)
ln(Spill-Sales)			0.00		–0.00			0.00		0.00
			(0.00)		(0.00)			(0.00)		(0.00)
ln(Input-MA)				–0.00	–0.00				–0.00	–0.00
				(0.00)	(0.00)				(0.00)	(0.00)
Region-Sector FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Conglomerate FE		✓			✓		✓			✓
Adj. R^2	0.15	0.21	0.15	0.15	0.21	0.17	0.27	0.17	0.17	0.27
# cluster (region)	60	60	60	60	60	60	60	60	60	60
# cluster (conglomerate)	1423	1422	1423	1423	1422	1423	1422	1423	1423	1422
N	2706	2705	2706	2706	2705	2706	2705	2706	2706	2705

Notes. This table reports the OLS estimates of Equation (4.3). When we construct the spillover measure defined in Equation (4.2), we lag firms' adoption status by three years. In columns (1)-(5), the dependent variables are a dummy variable of whether a firm makes a new technology adoption contracts made in a given year. In columns (6)-(11), the dependent variables are the inverse hyperbolic sine transformation of the number of new technology adoption contracts made in a given year. ln(Spill-Sales) and ln(Input-MA) are additional controls defined in Equations (4.4) and (4.5). In all specifications, we control for region-sector fixed effects and the initial dependent variable at the start of the sample period. Standard errors are two-way clustered at both region and conglomerate level and are reported in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C8: Local Productivity Spillovers of Technology Adoption - Dependent Variable: Dummy Variable of Adoption of a New Technology - Robustness: 5 Year Lag

Dep. Var.	$\mathbb{1}$[New Contract]					*asinh*(# New Contract)				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Spill	0.15	0.18	0.15	0.15	0.18	0.18	0.22	0.18	0.18	0.21
	(0.19)	(0.14)	(0.18)	(0.19)	(0.14)	(0.17)	(0.15)	(0.17)	(0.17)	(0.15)
ln(Spill-Sales)			0.00		0.00			0.00		0.00
			(0.00)		(0.00)			(0.00)		(0.00)
ln(Input-MA)				–0.00	–0.00				–0.00	–0.00
				(0.00)	(0.00)				(0.00)	(0.00)
Region-Sector FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Conglomerate FE		✓			✓		✓			✓
Adj. R^2	0.14	0.20	0.14	0.14	0.20	0.16	0.27	0.16	0.16	0.27
# cluster (region)	60	60	60	60	60	60	60	60	60	60
# cluster (conglomerate)	1423	1422	1423	1423	1422	1423	1422	1423	1423	1422
N	2706	2705	2706	2706	2705	2706	2705	2706	2706	2705

Notes. This table reports the OLS estimates of Equation (4.3). When we construct the spillover measure defined in Equation (4.2), we lag firms' adoption status by five years. In columns (1)-(5), the dependent variables are a dummy variable of whether a firm makes a new technology adoption contracts made in a given year. In columns (6)-(11), the dependent variables are the inverse hyperbolic sine transformation of the number of new technology adoption contracts made in a given year. ln(Spill-Sales) and ln(Input-MA) are additional controls defined in Equations (4.4) and (4.5). In all specifications, we control for region-sector fixed effects and the initial dependent variable at the start of the sample period. Standard errors are two-way clustered at both region and conglomerate level and are reported in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C9: Local Productivity Spillovers from Technology Adoption: Robustness - Spillover Defined at the Broader Level

Dep. Var.	Log sales					Log revenue TFP				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Panel A: Never-Adopter Sample									
Spill	3.54**	3.51**	4.12**	3.36*	3.83**	5.60*	5.37**	5.99*	5.48*	5.24*
	(1.69)	(1.61)	(1.78)	(1.73)	(1.63)	(2.80)	(2.38)	(3.13)	(2.81)	(2.68)
ln(Spill-Sales)			–0.02		–0.02			–0.02		–0.01
			(0.02)		(0.01)			(0.02)		(0.02)
ln(Input-MA)				–0.03*	–0.01				–0.03**	–0.02
				(0.01)	(0.01)				(0.01)	(0.01)
Adj. R^2	0.18	0.21	0.18	0.18	0.22	0.43	0.42	0.44	0.44	0.42
# clusters (region)	38	38	38	38	38	30	27	30	30	27
# clusters (conglomerate)	636	630	636	636	630	324	275	324	324	275
N	1079	1073	1079	1079	1073	344	292	344	344	292
	Panel B: Full Sample									
Spill	4.08***	3.36*	4.31***	3.80**	3.23*	5.28*	3.90	5.31*	5.10*	3.58
	(1.40)	(1.74)	(1.55)	(1.43)	(1.81)	(2.71)	(2.88)	(2.86)	(2.65)	(2.78)
$\mathbb{1}[Adopt]$	0.31*	0.24	0.31*	0.30	0.23	0.16	0.13	0.16	0.16	0.12
	(0.18)	(0.22)	(0.18)	(0.18)	(0.22)	(0.11)	(0.11)	(0.11)	(0.11)	(0.11)
ln(Spill-Sales)			–0.01		–0.01			–0.00		0.01
			(0.01)		(0.01)			(0.02)		(0.01)
ln(Input-MA)				–0.04***	–0.03**				–0.04***	–0.04***
				(0.01)	(0.01)				(0.01)	(0.01)
Adj. R^2	0.19	0.23	0.19	0.19	0.24	0.37	0.42	0.37	0.38	0.43
# clusters (region)	39	39	39	39	39	34	31	34	34	31
# clusters (conglomerate)	702	697	702	702	697	381	338	381	381	338
N	1264	1259	1264	1264	1259	431	387	431	431	387
Region-Sector FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Conglomerate FE		✓			✓		✓			✓

Notes. This table reports the OLS estimates of Equation (4.3). We aggregate regions to 39 regions and construct the spillover measure similar to Equation (4.2) at this broader level. We lag the adoption status of firms by four years. In Panel A, we use the subsample that include only firms that did not adopt any technology during the sample period. In Panel B, we use the full sample of adopters and non-adopters and control for adoption status. The dependent variables are log sales in columns (1)-(5) and revenue TFP in columns (6)-(10). We estimate revenue TFP based on Wooldridge (2009). The additional controls ln(Spill-Sales) and ln(Input-MA) are additional controls defined in Equations (4.4) and (4.5). In all specifications, we control for region-sector fixed effects and for the initial dependent variable at the start of the sample period. Standard errors are two-way clustered at the regiona level defined more broadly and conglomerate level and are reported in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C10: Complementarity in Firms' Technology Adoption Decisions

Dep. Var.	$\mathbb{1}$[New Contract]					*asinh*(# New Contract)				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Spill	0.42***	0.46***	0.42***	0.42***	0.47***	0.49**	0.55*	0.49**	0.49**	0.55*
	(0.15)	(0.16)	(0.15)	(0.15)	(0.16)	(0.20)	(0.28)	(0.21)	(0.20)	(0.29)
ln(Spill-Sales)			–0.00		–0.00			–0.00		–0.00
			(0.00)		(0.00)			(0.00)		(0.00)
ln(Input-MA)				–0.00	0.00				–0.00	–0.00
				(0.00)	(0.00)				(0.00)	(0.00)
Region-Sector FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Conglomerate FE		✓			✓		✓			✓
Adj. R^2	0.02	0.07	0.02	0.02	0.07	0.09	0.19	0.09	0.09	0.19
# cluster (region)	42	42	42	42	42	42	42	42	42	42
# cluster (conglomerate)	1414	1413	1414	1414	1413	1414	1413	1414	1414	1413
N	2689	2688	2689	2689	2688	2689	2688	2689	2689	2688

Notes. This table reports the OLS estimates of Equation (4.3). We aggregate regions to 39 regions and construct the spillover measure similar to Equation (4.2) at this broader level. When we construct the spillover measure defined in Equation (4.2), we lag firms' adoption status by four years. In columns (1)-(5), the dependent variables are a dummy variable of whether a firm makes a new technology adoption contracts made in a given year. In columns (6)-(11), the dependent variables are the inverse hyperbolic sine transformation of the number of new technology adoption contracts made in a given year. ln(Spill-Sales) and ln(Input-MA) are additional controls defined in Equations (4.4) and (4.5). In all specifications, we control for region-sector fixed effects and the initial dependent variable at the start of the sample period. Standard errors are two-way clustered at the regiona level defined more broadly and conglomerate level and are reported in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C11: Local Productivity Spillovers from Technology Adoption: Robustness - Alternative Dependent Variables: Log Employment and Labor Productivity

Dep. Var.	Log employment					Log labor productivity				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Panel A: Never-Adopter Sample									
Spill	4.39***	3.79**	4.94***	4.23***	4.07**	5.55***	5.41***	5.81***	5.34***	5.11**
	(1.54)	(1.64)	(1.70)	(1.50)	(1.76)	(1.84)	(1.62)	(2.08)	(1.78)	(1.92)
ln(Spill-Sales)			–0.02		–0.02			–0.02		–0.01
			(0.01)		(0.01)			(0.02)		(0.02)
ln(Input-MA)				–0.03	–0.02				–0.04**	–0.03
				(0.02)	(0.02)				(0.02)	(0.02)
Adj. R^2	0.18	0.22	0.19	0.19	0.22	0.44	0.42	0.44	0.44	0.42
# clusters (region)	42	39	42	42	39	41	36	41	41	36
# clusters (conglomerate)	351	312	351	351	312	324	275	324	324	275
N	375	335	375	375	335	344	292	344	344	292
	Panel B: Full Sample									
Spill	4.23***	3.93***	4.45***	3.86***	3.72**	4.75***	3.99**	4.72***	4.45***	3.44*
	(1.18)	(1.43)	(1.31)	(1.19)	(1.52)	(1.63)	(1.90)	(1.73)	(1.58)	(1.82)
$\mathbb{1}[Adopt]$	0.32**	0.26	0.32**	0.31**	0.25	0.15*	0.14	0.15*	0.14	0.12
	(0.15)	(0.20)	(0.15)	(0.15)	(0.19)	(0.09)	(0.10)	(0.09)	(0.09)	(0.10)
ln(Spill-Sales)			–0.01		–0.01			0.00		0.00
			(0.01)		(0.01)			(0.02)		(0.02)
ln(Input-MA)				–0.05***	–0.04*				–0.05***	–0.05**
				(0.02)	(0.02)				(0.02)	(0.02)
Adj. R^2	0.19	0.24	0.19	0.19	0.24	0.37	0.43	0.37	0.38	0.43
# clusters (region)	54	54	54	54	54	45	41	45	45	41
# clusters (conglomerate)	411	375	411	411	375	381	338	381	381	338
N	466	430	466	466	430	431	387	431	431	387
Region-Sector FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Conglomerate FE		✓			✓		✓			✓

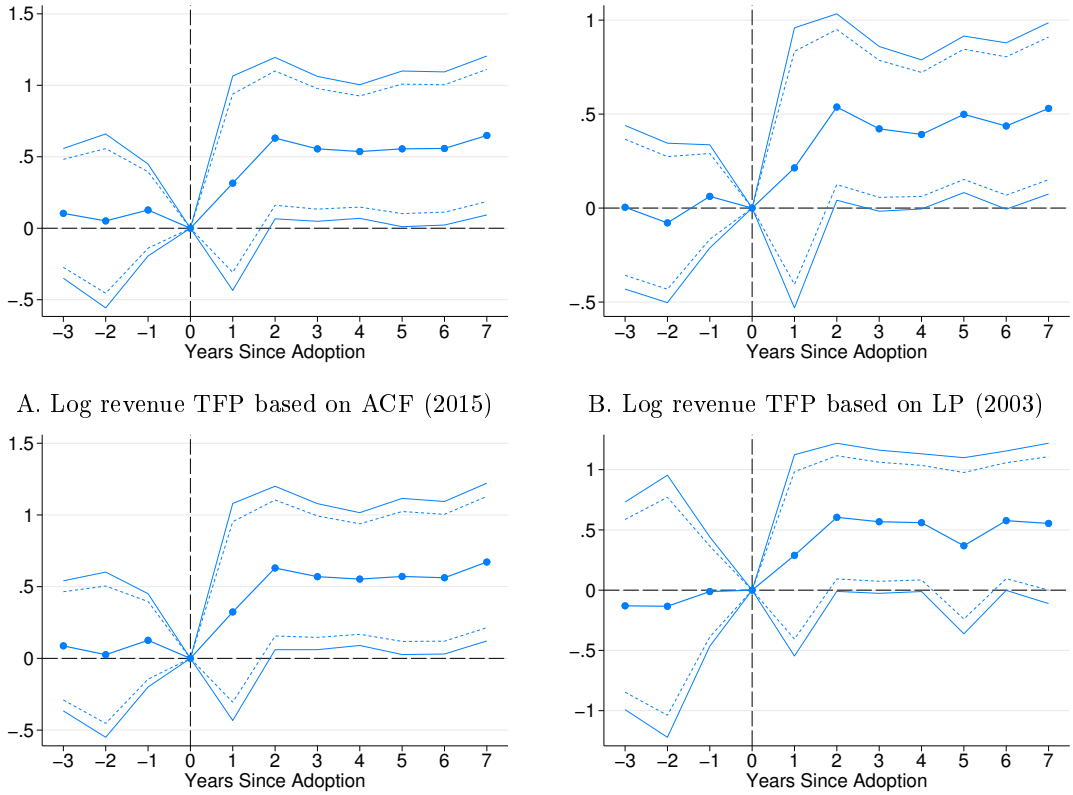
Notes. This table reports the OLS estimates of Equation (4.3). When we construct the spillover measure defined in Equation (4.2), we lag firms' adoption status by four years. In Panel A, we use the subsample that include only firms that did not adopt any technology during the sample period. In Panel B, we use the full sample of adopters and non-adopters and control for adopters' adoption status. The dependent variables are log employment in columns (1)-(5) and labor productivity in (6)-(10). Labor productivity is defined as value added per worker. ln(Spill-Sales) and ln(Input-MA) are additional controls defined in Equations (4.4) and (4.5). In all specifications, we control for region-sector fixed effects and the initial dependent variable at the start of the sample period. Standard errors are two-way clustered at both region and conglomerate level and are reported in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C12: Local Productivity Spillover from Technology Adoption: Robustness - Alternative Dependent Variables: Log Fixed Assets and Assets

Dep. Var.	Log fixed assets					Log assets				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Panel A: Never-Adopter Sample									
Spill	4.55**	5.39***	5.73***	4.51**	6.49***	3.88**	4.08***	4.64**	3.81**	4.70***
	(2.10)	(1.86)	(2.08)	(2.10)	(1.83)	(1.62)	(1.51)	(1.75)	(1.61)	(1.64)
ln(Spill-Sales)			–0.04***		–0.04**			–0.03**		–0.03*
			(0.01)		(0.02)			(0.01)		(0.01)
ln(Input-MA)				–0.01	0.00				–0.01	–0.00
				(0.02)	(0.03)				(0.02)	(0.02)
Adj. R^2	0.12	0.18	0.13	0.12	0.19	0.10	0.17	0.11	0.10	0.17
# clusters (region)	53	53	53	53	53	53	53	53	53	53
# clusters (conglomerate)	631	625	631	631	625	635	629	635	635	629
N	1072	1066	1072	1072	1066	1078	1072	1078	1078	1072
	Panel B: Full Sample									
Spill	3.05**	4.13***	3.68***	2.93**	4.63***	2.88**	3.27***	3.26**	2.69**	3.39**
	(1.41)	(1.18)	(1.36)	(1.41)	(1.18)	(1.20)	(1.21)	(1.31)	(1.19)	(1.29)
$\mathbb{1}[Adopt]$	0.50***	0.39**	0.50***	0.49***	0.39**	0.38***	0.34**	0.38***	0.37***	0.33**
	(0.13)	(0.17)	(0.13)	(0.13)	(0.17)	(0.12)	(0.15)	(0.12)	(0.12)	(0.15)
ln(Spill-Sales)			–0.03**		–0.03*			–0.02		–0.01
			(0.01)		(0.01)			(0.01)		(0.01)
ln(Input-MA)				–0.02	–0.01				–0.02	–0.02
				(0.02)	(0.02)				(0.01)	(0.02)
Adj. R^2	0.15	0.22	0.16	0.15	0.23	0.15	0.20	0.15	0.15	0.20
# clusters (region)	54	54	54	54	54	54	54	54	54	54
# clusters (conglomerate)	696	691	696	696	691	701	696	701	701	696
N	1254	1249	1254	1254	1249	1263	1258	1263	1263	1258
Region-Sector FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Conglomerate FE		✓			✓		✓			✓

Notes. This table reports the OLS estimates of Equation (4.3). When we construct the spillover measure defined in Equation (4.2), we lag firms' adoption status by four years. In Panel A, we use the subsample that include only firms that did not adopt any technology until the end of the sample period. In Panel B, we use the full sample including both adopters and non-adopters and control for adopters' adoption status. Dependent variables are log fixed assets in columns (1)-(5) and assets in columns (6)-(10). ln(Spill-Sales) and ln(Input-MA) are additional controls defined in Equations (4.4) and (4.5). In all specifications, we control for region-sector fixed effects and the initial dependent variable at the start of the sample period. Standard errors are two-way clustered at both region and conglomerate level and are reported in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

C.2 Additional Figures



A. Log revenue TFP based on ACF (2015)

B. Log revenue TFP based on LP (2003)

C. Log revenue TFP based on OLS

D. Log labor productivity

Figure C1. Robustness Checks for Direct Productivity Gains of Technology Adoption: Winners vs. Losers Research Design - Alternative TFP Measures

Notes. This figure illustrates the estimated β_{τ}^{diff} in Equation (4.1) based on winners vs. losers research design. The dependent variables are log revenue TFP and labor productivity. In Panels A, B, and C, we estimate revenue TFPs based on Ackerberg et al. (2015), Levinsohn and Petrin (2003), and OLS, respectively. In Panel D, labor productivity is defined as value-added per worker. We normalize β_{0}^{diff} to zero. All specifications control for event time dummies, and firm, pair, and calendar year fixed effects. The figure reports 90 and 95 percent confidence intervals based on standard errors two-way clustered at the levels of pairs and firms.

C.3 Ruling out Alternative Hypotheses: Winners vs. Losers Research Design

Empirical Evidence on Winners' Exports We provide empirical evidence that winners were more likely to become an exporter and exported more than losers. This evidence supports that our findings based on winners vs. losers research design were driven by productivity shocks rather than domestic demand shocks or markups, because domestic demand shocks or markups are of less concern in international markets.

We merge the pairs of winners and losers with KIS-VALUE that covers firms' exports data after 1980. Because KIS-VALUE coverage is smaller than our firm balance sheet data, some pairs were dropped while merging with KIS-VALUE. 23 out of 34 pairs could be merged with KIS-VALUE.

We pool the sample of matched firms' exports observed 7 or 8 years after the cancellation occurred, which we label as 7-year and 8-year samples, respectively. Then using these 7-year and 8-year samples, we estimate the following pooled OLS regression model:

$$y_{ip,t(p)+\tau} = \beta^{export} \times \mathbb{1}[Adopt_{ip,t(p)}] + \delta_{p\tau} + \epsilon_{ip,t(p)+\tau}, \tag{C.1}$$

where i denotes firm, p pair, and $t(p)$ year in which the event happened for pair p. τ denotes years after the event. $\mathbb{1}[Adopt_{ip,t(p)}]$ is a dummy variable which equals 1 if firm i adopted technology at the time of the event. Dependent variables are $\mathbb{1}[\text{Export}_{ip\tau}]$, $asinh(\text{Export}_{ip\tau})$, and $\ln(\text{Export}_{ip\tau} + 1)$. $\mathbb{1}[\text{Export}_{ip\tau}]$ is a dummy variable of firms' adoption status. $asinh(\text{Export}_{ip\tau})$ is the inverse hyperbolic sine transformation of exports. We use the inverse hyperbolic sine transformation to deal with zero exports, so $asinh(\text{Export}_{ip\tau})$ captures both intensive and extensive margins of exports . $\ln(\text{Export}_{ip\tau} + 1)$ is log one plus exports. $\delta_{p\tau}$ is pair and τ specific fixed effects. $\epsilon_{ip\tau}$ is the error term. We cluster standard errors at the pair level.

Because we are controlling for $\delta_{p\tau}$, β^{export} is identified by variation within pair. If $\mathbb{1}[Adopt_{ip,t(p)}]$ is uncorrelated with the error term, the estimates admit causal interpretation. Because the sample period of KIS-VALUE begins in 1980, we cannot check pre-trends of exports as in Equation (4.1). Although we cannot check pre-trends for exports, the fact we do not find pre-trends in sales or revenue TFP measures supports that $\mathbb{1}[Adopt_{ip,t(p)}]$ is uncorrelated with the error term.

Table C13 reports the results. In column (1), the dependent variable is a dummy variable of firms' adoption status. We pool the 7-year and 8-year samples. The estimated coefficient is positive and statistically significant. We find that the adoption increased firms' probability of exporting by 29 percentage points. In columns (2) and (3), we only use the 7-year and 8-year samples, respectively. The estimates remain statistically significant and similar to those in column (1). In columns (4)-(6), the dependent variable is $asinh(\text{Export}_{ip,t(p)+\tau})$. The coefficients are statistically significant and positive, and their magnitude implies that the adoption increased a 0.55 standard deviation of $asinh(\text{Export}_{ip,t(p)+\tau})$. In columns (7)-(9), the dependent variable is $\ln(\text{Export}_{ip,t(p)+\tau} + 1)$. The magnitude of the estimates is similar to those in columns (4)-(6).

Given the small number of clusters, we report the p-values based on the wild cluster bootstrap-t method of Cameron et al. (2008) in the bracket (p-val (CGM)). Using the wild cluster bootstrap-t, the estimates remain statistically significant across all specifications.

Firm-to-Firm Input Sourcing. Another alternative hypothesis is that cancellations could be systematically related to demand shocks if foreign firms only purchased inputs from South Korean firms that adopted advanced foreign technologies. This could have happened if adoption of foreign

Table C13: Technology Adoption Increased Firms' Exports: Winners vs. Losers Research Design

Dep. Var.	$\mathbb{1}$[Export]			*asinh*(Export)			ln(Export + 1)		
Years after the event (τ)	$\tau = 7, 8$	$\tau = 7$	$\tau = 8$	$\tau = 7, 8$	$\tau = 7$	$\tau = 8$	$\tau = 7, 8$	$\tau = 7$	$\tau = 8$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Adopt	0.29**	0.26*	0.32**	5.25**	4.75*	5.79**	5.05**	4.57*	5.56**
	(0.13)	(0.13)	(0.14)	(2.40)	(2.49)	(2.60)	(2.31)	(2.41)	(2.50)
p-val (CGM)	[0.06]	[0.04]	[0.01]	[0.04]	[0.08]	[0.04]	[0.04]	[0.08]	[0.04]
Pair-τ FE	✓			✓			✓		
Pair FE		✓	✓		✓	✓		✓	✓
Adj. R^2	0.20	0.24	0.14	0.24	0.24	0.21	0.24	0.24	0.22
# cluster (pair)	23	23	22	23	23	22	23	23	22
N	90	46	44	90	46	44	90	46	44

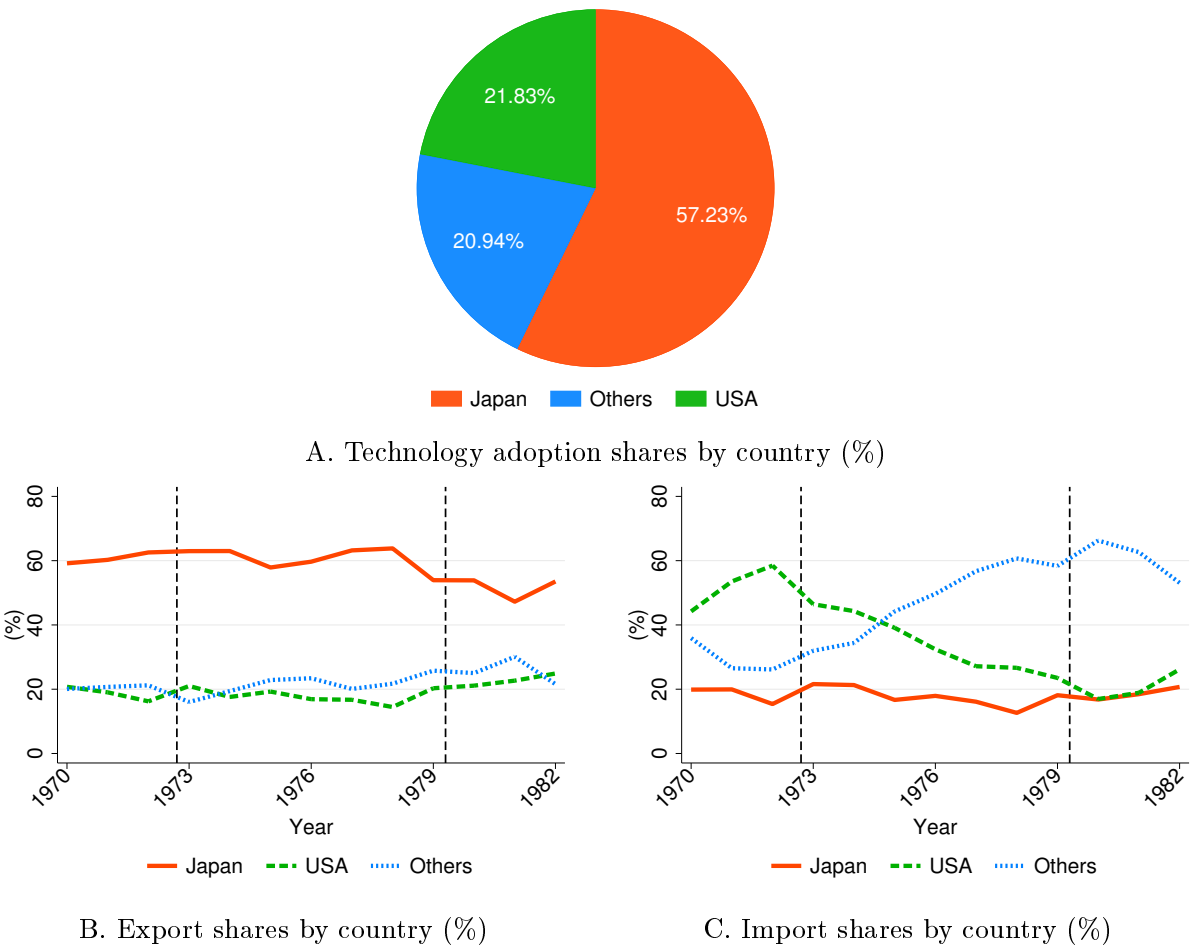
Notes. This table reports the estimates of $\mathbb{1}[\text{Adopt}_{ip,t(p)}]$ in Equation (C.1). Robust standard errors in parenthesis are two-way clustered at the region and firm levels. P-values based on the wild cluster bootstrap-t method of Cameron et al. (2008) are reported in the bracket (p-val (CGM)). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

technology had worked as a signal for potential foreign buyers or foreign firms required customized inputs produced with specific foreign technologies. Under this scenario, increases in sales may reflect both increases in productivity and demands from foreign firms. Without detailed information on firm-to-firm trade, we cannot completely rule out this hypothesis. However, we present empirical evidence that is inconsistent with this hypothesis at the aggregate level.

Panels A, B, and C of Figure C2 are technology adoption, export, and import shares by country for heavy manufacturing sectors. Most of technologies came from Japan and the US. About 57% and 22% contracts were made with Japanese and the US firms, respectively. If technology contracts were systematically related to demand shocks through foreign firms' input sourcing, we would expect aggregate exports of Japan or the US to increase relative more than aggregate exports to the rest of the world. However, we do not find such patterns. Both aggregate exports shares of Japan and the US are stable during the 1970s.

Decreases in Sales of Losers. Another alternative hypothesis is that the results are driven by decreases in sales of losers rather than increases in sales of winners. For example, suppose losers incurred some costs for preparation of adopting a new technology. This loss could have persistently decreases losers' sales after the cancellations. In such case, it is possible that these relative differences identified by the event study were driven by these decreases in sales of losers rather than increases in productivity of winners.

We present empirical evidence that is inconsistent with this hypothesis. Figure C3 plots the mean of log sales of both winners and losers for event horizons, where we normalize the mean to be zero



A. Technology adoption shares by country (%)

B. Export shares by country (%)

C. Import shares by country (%)

Figure C2. Technology Adoption, Export, and Import Shares by Country of the Heavy Manufacturing Sectors

for both groups. The trend diverged only after the cancellation happened. For 1 and 2 years after the cancellation, we do observe temporary drops of the average log sales of losers, which is consistent with this alternative hypothesis. However, after 3 years since the cancellation, the average log sales of losers returned to the original trend.

C.4 Full-Sample Event Study Design

In this subsection, we compare our estimates from the winners vs. losers research design to estimates based on the following two-way fixed effects event-study specification that uses the full-sample:

$$y_{it} = \sum_{\tau=-3}^{\tau=7} \beta_\tau \times \mathbb{1}[\text{Adopt}_{it}^{\tau}] + \mathbf{X}_{it}'\boldsymbol{\gamma} + \delta_i + \delta_t + \epsilon_{it}, \tag{C.2}$$

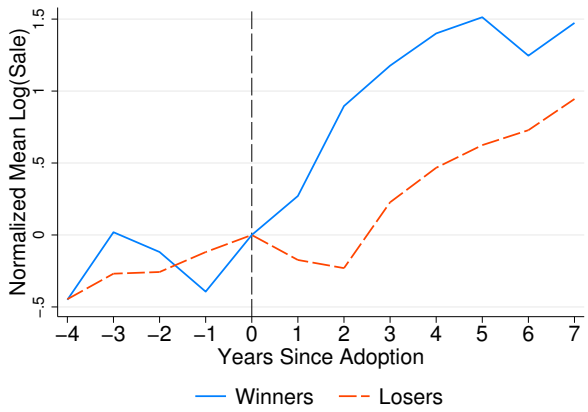


Figure C3. Mean of Log Sales of Winners and Losers

Notes. This figure plots the mean of log sales of winners and losers for the event horizons. The mean of both groups is normalized to be zero at the time of the event.

where $\mathbb{1}[\text{Adopt}_{it}^{\tau}]$ are event-study variables defined as $\mathbb{1}[\text{Adopt}_{it}^{\tau}] := \mathbb{1}[t - \tau = t(i)]$ and $t(i)$ is year in which firm i adopted technology from foreign firms for the first time. Dependent variables y_{ipt} are log sales, log revenue TFP estimated, and labor productivity defined as value added per worker. δ_i and δ_t are firm and calendar year fixed effects. ϵ_{it} is the error term. We additionally control for observables $\mathbf{X}_{it}$ depending on specifications. We two-way cluster standard errors at the region and firm levels.

Unlike the specification in Equation (4.1) based on the winners vs. losers research design, the specification in Equation (C.2) uses the full sample. However, the specification in Equation (C.2) can be problematic for two reasons. First, technology adoption decisions are endogenous, which can make $\mathbb{1}[\text{Adopt}_{it}^{\tau}]$ be correlated with the error term. This endogeneity problem will result in biased estimates for the true impact of technology adoption. Second, because this specification uses pre-treated firms as control groups, it is less robust to problems related to staggered diff-in-diffs design.

Table C14 reports the results. In columns (1)-(4), the dependent variable is log sales. Across different specifications, we find positive correlation between technology adoption and log sales. Also, there are no pre-trends. However, the magnitude of the estimated coefficients is smaller than those from the winners vs. losers research design. The magnitude becomes smaller, and the coefficients are less precisely estimated once we control additional fixed effects in columns (2)-(4). We observe a similar pattern in columns (5)-(12), where we use log revenue TFP and log labor productivity as dependent variables. For log revenue TFP and log labor productivity, we also find that the magnitude of the estimated coefficients is smaller than those from the winners vs. losers research design.

Suppose the identifying assumption of the winners vs. losers research design holds. Then, the estimates from the naive event study design are downward biased. One potential scenario for this bias is that the government selectively approved technology adoption contracts or provided subsidies for the adoption based on political connections rather than productivity. If less productive firms that are more politically connected were targeted by the government, this might result in the downward bias of the estimates in Equation (C.2). However, the winners vs. losers research design can deal

with this bias induced by the subsidies or political connections. From the fact that both winners and losers got approvals from the government, we can indirectly infer that the two groups had a similar level of political favors. Although the misallocation effects are not the focus of this paper, with these potential misallocation effects of the subsidies, the welfare effects of our quantitative analysis should be interpreted as the upper bound.

C.5 Cross-Sector Spillover

This section provides additional empirical results on cross-sector spillover effects. We augment Equation (4.3) with the cross-spillover measures as follows:

$$\triangle y_{injt} = \beta^S \triangle \text{Spill}_{inj(t-4)} + \sum_{g \neq j} \beta^S_{gj} \triangle \text{Spill}_{ing(t-4)} + \gamma y_{injt_0} + \mathbf{X}'_{injt_0}\boldsymbol{\beta} + \triangle\delta_{njt} + \triangle\epsilon_{injt}, \tag{C.3}$$

where β^S_{gj} captures the cross-sector spillover effect from sectors g to j.

A problem of Equation (4.3) is that there are too many parameters to be estimated given the data. There are $|\mathcal{J}| \times (|\mathcal{J}| - 1)$ of cross-sector spillover parameters. Following Ellison et al. (2010), we parametrize β^S_{gj} using the input-output tables of 1970:

$$\beta^S_{gj} = \beta^S_{for}\gamma^g_j + \beta^S_{back}\gamma^j_g$$

where γ^g_j represents shares of sector g intermediate inputs used by sector j obtained from the input-output table. β^S_{for} and β^S_{back} capture spillover effects through forward and backward linkages, respectively. After substituting the above expression, we can derive the following regression model:

$$\triangle y_{injt} = \beta^S \triangle \text{Spill}_{inj(t-4)} + \beta^S_{for}\left(\sum_{g \neq j} \gamma^g_j \triangle \text{Spill}_{ing(t-4)}\right) + \beta^S_{back}\left(\sum_{g \neq j} \gamma^j_g \triangle \text{Spill}_{ing(t-4)}\right) + \gamma y_{injt_0} + \mathbf{X}'_{injt_0}\boldsymbol{\beta} + \triangle\delta_{njt} + \triangle\epsilon_{injt}. \tag{C.4}$$

The cross-sector spillover is governed by only two parameters β^S_{for} and β^S_{back} in Equation (C.4).

Table C15 reports the OLS estimates for β^S, β^S_{for}, and β^S_{back}. In Panels A and B, we separately control the forward and backward linkage spillovers, respectively. In Panel C, we jointly control them. Across different specifications, we do not find statistically significant results for the cross-sector spillovers. The statistically insignificant results may come from the fact that our sector classification is defined at the broad level.

C.6 Matching Algorithm

This section describes the matching algorithm used for matching a loser to a winner for Section 4.1. Let $\mathbf{X} \in \mathcal{R}_k$ denotes the k-dimensional observable variables. The matching proceeds in two steps.

1. Pick two subsets of variables $\mathbf{X}^e \in \mathbf{X}$ that are exactly matched and $\mathbf{X}^d \in \mathbf{X}$ that are distance matched.

2. For each loser f, pick an adopter g such that
   - both firms have the same values of the variables of $\mathbf{X}^e$ with a loser f, then

Table C14: Event Study Estimates of Direct Productivity Gains to Adopters: Standard Two-Way Fixed Effects Event-Study Design

Dep. Var.	Log sales				Log revenue TFP				Log labor productivity			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
3 years before event	–0.10	–0.07	–0.08	–0.06	–0.04	–0.05	–0.04	–0.03	0.04	0.01	0.03	0.06
	(0.06)	(0.06)	(0.06)	(0.07)	(0.06)	(0.07)	(0.07)	(0.09)	(0.07)	(0.07)	(0.06)	(0.09)
2 years before event	–0.02	–0.00	–0.00	0.02	0.02	0.02	0.03	0.04	0.05	0.04	0.05	0.06
	(0.05)	(0.05)	(0.05)	(0.06)	(0.06)	(0.06)	(0.07)	(0.07)	(0.07)	(0.07)	(0.08)	(0.08)
1 year before event	–0.01	0.00	–0.01	0.01	0.06	0.06	0.06*	0.06	0.06	0.05	0.04	0.07
	(0.07)	(0.06)	(0.06)	(0.06)	(0.04)	(0.03)	(0.04)	(0.04)	(0.05)	(0.05)	(0.05)	(0.06)
Year of event												
1 year after event	0.05	0.03	0.04	0.10*	0.05	0.04	0.04	0.06	0.11	0.10	0.10	0.14*
	(0.06)	(0.06)	(0.05)	(0.05)	(0.07)	(0.07)	(0.07)	(0.07)	(0.08)	(0.07)	(0.07)	(0.08)
2 years after event	0.18**	0.15**	0.13*	0.18**	0.15*	0.13	0.14*	0.15	0.19**	0.18**	0.18**	0.18**
	(0.07)	(0.06)	(0.07)	(0.08)	(0.08)	(0.08)	(0.08)	(0.10)	(0.08)	(0.08)	(0.08)	(0.09)
3 years after event	0.20**	0.17**	0.11	0.16*	0.18**	0.14*	0.16*	0.13	0.25***	0.23***	0.25***	0.23**
	(0.08)	(0.08)	(0.08)	(0.09)	(0.08)	(0.08)	(0.08)	(0.10)	(0.08)	(0.08)	(0.09)	(0.10)
4 years after event	0.21**	0.15*	0.08	0.15	0.15	0.08	0.07	0.07	0.19*	0.15	0.15	0.15
	(0.08)	(0.08)	(0.08)	(0.10)	(0.10)	(0.10)	(0.10)	(0.12)	(0.10)	(0.10)	(0.10)	(0.11)
5 years after event	0.27**	0.21*	0.14	0.16	0.20	0.13	0.12	0.07	0.18	0.14	0.16	0.11
	(0.11)	(0.11)	(0.11)	(0.13)	(0.13)	(0.12)	(0.13)	(0.15)	(0.12)	(0.12)	(0.13)	(0.15)
6 years after event	0.31***	0.26***	0.16*	0.21**	0.23**	0.16*	0.13	0.12	0.20	0.16	0.15	0.14
	(0.10)	(0.09)	(0.09)	(0.09)	(0.11)	(0.09)	(0.11)	(0.11)	(0.12)	(0.11)	(0.12)	(0.12)
7 years after event	0.32***	0.25***	0.16	0.20**	0.24**	0.14	0.15	0.11	0.21	0.16	0.20	0.15
	(0.10)	(0.08)	(0.09)	(0.08)	(0.11)	(0.09)	(0.10)	(0.10)	(0.15)	(0.13)	(0.15)	(0.14)
Firm FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓				✓				✓			
Sector-Year FE		✓	✓			✓	✓			✓	✓	
Region-Year FE			✓				✓				✓	
Region-Sector-Year FE				✓				✓				✓
Adj. R^2	0.83	0.83	0.83	0.83	0.87	0.87	0.86	0.86	0.53	0.54	0.52	0.50
# clusters (region)	59	59	59	59	59	59	58	58	59	59	58	58
# clusters (firm)	3366	3366	3365	3323	2163	2163	2147	2105	2170	2170	2154	2112
N	15955	15955	15915	15639	9216	9216	9136	8923	9242	9242	9162	8950

Notes. This table reports the estimated event study coefficients $\mathbb{1}[\text{Adopt}_{it}^{\tau}]$ in Equation (C.2). $\mathbb{1}[\text{Adopt}_{it}^{0}]$ is normalized to zero. The dependent variables are log sales, log revenue TFP, and log labor productivity defined as value added divided by employment. Value added is obtained as sales multiplied by the value added shares obtained from input-output tables corresponding to each year. We estimate log revenue TFP based on Wooldridge (2009). Robust standard errors in parenthesis are two-way clustered at the region and firm levels. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C15: Cross-Sector Local Productivity Spillovers from Technology Adoption

Dep. Var.	Log sales					Log revenue TFP				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Panel A: Forward Linkage Spillovers									
Spill	4.19**	3.64**	4.68**	4.10**	3.94**	5.98***	5.67***	6.26***	5.79***	5.44***
	(1.64)	(1.65)	(1.75)	(1.59)	(1.73)	(1.95)	(1.71)	(2.13)	(1.90)	(1.89)
Forward Spill (β^S_{for})	−1.71	−1.52	−1.17	−1.36	−0.93	−1.54	−2.86***	−1.23	−1.81	−2.89***
	(1.29)	(1.55)	(1.45)	(1.20)	(1.55)	(2.23)	(0.58)	(2.22)	(2.09)	(0.65)
Adj. R^2	0.18	0.22	0.19	0.18	0.22	0.43	0.42	0.44	0.44	0.42
	Panel B: Backward Linkage Spillovers									
Spill	4.10**	3.57**	4.60**	4.02**	3.87**	5.88***	5.50***	6.17***	5.66***	5.28***
	(1.69)	(1.70)	(1.82)	(1.65)	(1.78)	(1.88)	(1.64)	(2.10)	(1.82)	(1.86)
Backward Spill (β^S_{back})	−7.47	−8.35	−5.56	−7.00	−6.87	−3.47	−8.45	−3.20	−4.63	−9.28
	(6.03)	(5.78)	(7.26)	(6.14)	(6.29)	(7.78)	(5.41)	(8.33)	(7.28)	(5.50)
Adj. R^2	0.18	0.22	0.19	0.19	0.22	0.43	0.42	0.44	0.44	0.42
	Panel C: Forward & Backward Linkage Spillovers									
Spill	4.11**	3.56**	4.61**	4.01**	3.85**	5.94***	5.60***	6.21***	5.72***	5.36***
	(1.73)	(1.71)	(1.85)	(1.68)	(1.80)	(2.03)	(1.76)	(2.22)	(1.97)	(1.94)
Forward Spill (β^S_{for})	−0.35	0.32	−0.05	0.21	0.98	−1.16	−1.69	−0.75	−1.13	−1.43
	(2.65)	(1.99)	(2.79)	(2.54)	(2.03)	(3.22)	(1.87)	(3.02)	(3.19)	(1.79)
Backward Spill (β^S_{back})	−6.58	−9.23	−5.42	−7.54	−9.52	−1.58	−5.37	−1.97	−2.77	−6.69
	(11.38)	(7.78)	(12.60)	(11.25)	(8.42)	(10.77)	(8.27)	(11.06)	(10.37)	(8.26)
Adj. R^2	0.18	0.22	0.19	0.18	0.22	0.43	0.42	0.43	0.44	0.42
Region-Sector FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Conglomerate FE		✓			✓		✓			✓
ln(Spill-Sales)			✓		✓			✓		✓
ln(Input-MA)				✓	✓				✓	✓
# clusters (region)	53	53	53	53	53	41	36	41	41	36
# clusters (conglomerate)	636	630	636	636	630	324	275	324	324	275
N	1079	1073	1079	1079	1073	344	292	344	344	292

Notes. This table reports the OLS estimates of Equation (C.4). When we construct the spillover measure defined in Equation (4.2), we lag the adoption status of firms by four years. We use the subsample that include only firms that did not adopt any technology during the sample period. The dependent variables are log sales in columns (1)-(5) and revenue TFP in columns (6)-(10). We estimate revenue TFP based on Wooldridge (2009). The additional controls ln(Spill-Sales) and ln(Input-MA) are additional controls defined in Equations (4.4) and (4.5). In all specifications, we control for region-sector fixed effects and for the initial dependent variable at the start of the sample period. Standard errors are two-way clustered at both region and conglomerate level and are reported in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C16: Cross-Sector Local Productivity Spillovers from Technology Adoption

Dep. Var.	Log sales					Log revenue TFP				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Panel A: Forward Linkage Spillovers									
Spill	0.49**	0.46***	0.49***	0.49**	0.46***	0.46**	0.42**	0.46**	0.45**	0.41**
	(0.18)	(0.16)	(0.18)	(0.19)	(0.16)	(0.17)	(0.16)	(0.17)	(0.18)	(0.16)
Forward Spill (β^S_{for})	–0.14	–0.51*	–0.14	–0.14	–0.50*	–0.20	–0.84	–0.19	–0.20	–0.85
	(0.17)	(0.29)	(0.16)	(0.17)	(0.28)	(0.20)	(0.52)	(0.18)	(0.20)	(0.51)
Adj. R^2	0.02	0.07	0.02	0.02	0.07	0.09	0.19	0.09	0.09	0.19
	Panel B: Backward Linkage Spillovers									
Spill	0.49**	0.46***	0.49***	0.48**	0.46***	0.45**	0.41**	0.46**	0.45**	0.41**
	(0.18)	(0.16)	(0.18)	(0.19)	(0.15)	(0.17)	(0.16)	(0.17)	(0.18)	(0.16)
Backward Spill (β^S_{back})	–0.54	–1.43*	–0.54	–0.55	–1.41*	–0.69	–2.08*	–0.69	–0.71	–2.09*
	(0.70)	(0.80)	(0.67)	(0.70)	(0.79)	(0.81)	(1.15)	(0.78)	(0.81)	(1.14)
Adj. R^2	0.02	0.07	0.02	0.02	0.07	0.09	0.19	0.09	0.09	0.19
	Panel C: Forward & Backward Linkage Spillovers									
Spill	0.49**	0.46***	0.49***	0.48**	0.46***	0.45**	0.42**	0.46**	0.45**	0.41**
	(0.18)	(0.16)	(0.18)	(0.19)	(0.16)	(0.17)	(0.16)	(0.17)	(0.18)	(0.16)
Forward Spill (β^S_{for})	–0.02	–0.38	–0.02	–0.02	–0.37	–0.06	–0.80	–0.06	–0.07	–0.80
	(0.25)	(0.41)	(0.26)	(0.26)	(0.41)	(0.28)	(0.72)	(0.28)	(0.28)	(0.72)
Backward Spill (β^S_{back})	–0.49	–0.54	–0.49	–0.50	–0.54	–0.54	–0.19	–0.54	–0.56	–0.19
	(1.11)	(1.25)	(1.11)	(1.11)	(1.25)	(1.25)	(1.51)	(1.26)	(1.26)	(1.52)
Adj. R^2	0.02	0.07	0.02	0.02	0.07	0.09	0.19	0.09	0.09	0.19
Region-Sector FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Conglomerate FE		✓			✓		✓			✓
ln(Spill-Sales)			✓		✓			✓		✓
ln(Input-MA)				✓	✓				✓	✓
# clusters (region)	61	61	61	61	61	61	61	61	61	61
# clusters (conglomerate)	1414	1413	1414	1414	1413	1414	1413	1414	1414	1413
N	2689	2688	2689	2689	2688	2689	2688	2689	2689	2688

Notes. This table reports the OLS estimates of Equation (C.4). When we construct the spillover measure defined in Equation (4.2), we lag the adoption status of firms by four years. In columns (1)-(5), the dependent variables are a dummy variable of whether a firm makes a new technology adoption contracts made in a given year. In columns (6)-(11), the dependent variables are the inverse hyperbolic sine transformation of the number of new technology adoption contracts made in a given year. The additional controls ln(Spill-Sales) and ln(Input-MA) are additional controls defined in Equations (4.4) and (4.5). In all specifications, we control for region-sector fixed effects. Standard errors are two-way clustered at both region and conglomerate level and are reported in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

- minimize the Mahalanobis distance with loser f in terms of $\mathbf{X}^d$:

$$\text{adopter}_g \in \arg\min_{g' \in \mathcal{F}} \{((\mathbf{X}_f^d - \mathbf{X}_g^d)' \mathbf{S}^{-1} (\mathbf{X}_f^d - \mathbf{X}_g^d)\},$$

  where $\mathcal{F}$ is a set of firms, $\mathbf{S}$ is the sample covariance of $\mathbf{X}^d$, and $\mathbf{X}_f^d$ and $\mathbf{X}_g^d$ represent the variables of firms f and g that are distance matched, respectively.

While we implement this matching algorithm, we pick regions and sectors as $\mathbf{X}^e$, and log assets as $\mathbf{X}^d$. Because we are exactly matching on regions and sectors, our matching procedure absorbs out any region-sector level common shocks, costs of production, and market size. By distance matching on log assets, we can compare winners and losers with similar size.

C.7 Production Function Estimation

In this section, we discuss the procedure we use to estimate revenue TFP measures. We obtain the revenue TFP measures as the residuals after estimating production using the methodologies in Wooldridge (2009), Levinsohn and Petrin (2003), Ackerberg et al. (2015), and OLS. We estimate the following the Cobb-Douglas value added production function as follows:

$$\log VA_{it} = \alpha_L \log L_{it} + \alpha_K \log K_{it} + u_{it}, \tag{C.5}$$

where VA_{it} is value added; L_{it} is employment; K_{it} are fixed assets; and α_L and α_K are Cobb-Douglas labor and capital shares.

When we use the methodologies developed by Wooldridge (2009), Levinsohn and Petrin (2003), and Ackerberg et al. (2015), we use material inputs as a proxy variable. However, information on material inputs is not available for our main firm-balance sheet data digitized from the Annual Reports of Korean Companies. Therefore, we estimate the production function separately for each sector using alternative firm-level data. We used KIS-VALUE from 1980 to 1990. The Act on External Audits of Joint Stock Corporations, which was introduced in 1981, required South Korean firms whose assets were above 3 billion Korean Won to report their balance sheet data. That data is the source for KIS-VALUE. The coverage of our dataset is larger than KIS-VALUE. Also, because we observe sales but not value added, we calculate value added as sales times the value added shares from the input-output tables of corresponding years. Using these estimated coefficients from KIS-VALUE, we obtain revenue TFP for the sample period from 1970 to 1982.

Appendix D Quantification

D.1 Additional Figures

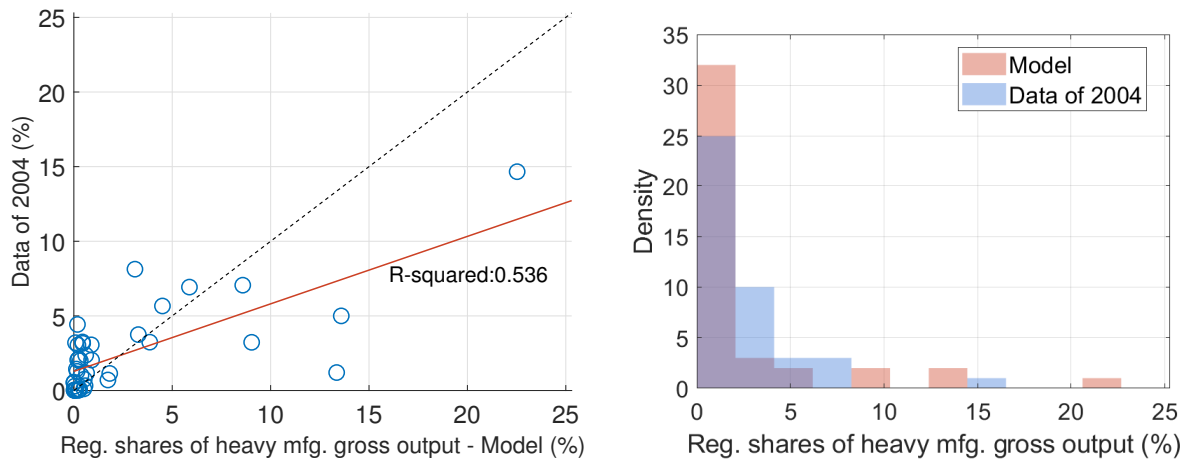
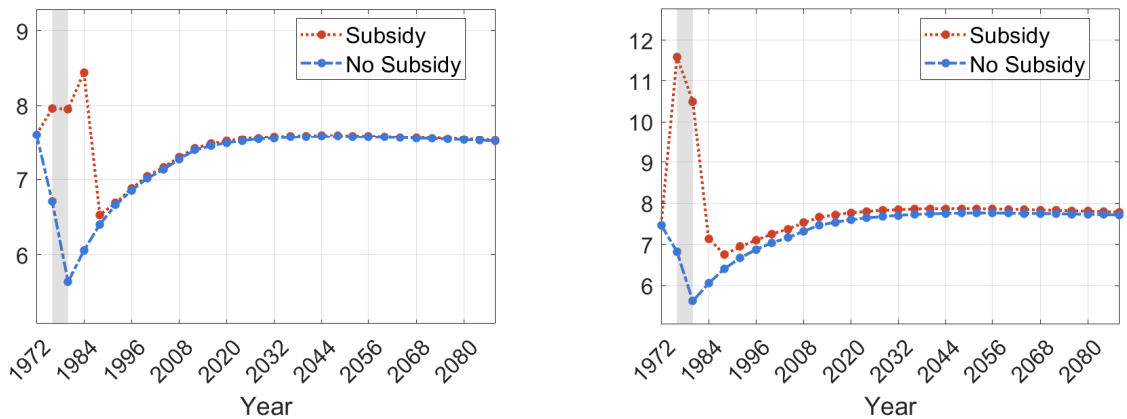


Figure D1. Non-targeted Moments: Spatial Distribution of the Heavy Manufacturing's Gross Output

Notes. This figure compares regional shares of the heavy manufacturing sector obtained from the data in 2004 and those calculated from the model of the corresponding model period. To calculate the regional shares of the data, we use the Mining and Manufacturing Survey that covers the universe of establishments with more than 5 employees. X and y-axes of Panel A are regional shares computed from the model and the data counterpart, respectively. The red solid line of Panel A is the linear fit. Panel B plots the histogram of the regional shares of the data and the model.



Heavy mfg. share of GDP (%)

A. Low η B. Low δ

Figure D2. Comparative Statistics of δ and η

Notes. This figure plots the comparative statistics of δ and η. In Panel A, we set η to be 1.05. In Panel B, we set δ to be 1. The red dotted line and the blue dashed lines plot the outcomes of the baseline and counterfactual economies.

A. Heavy mfg. share of GDP (%)

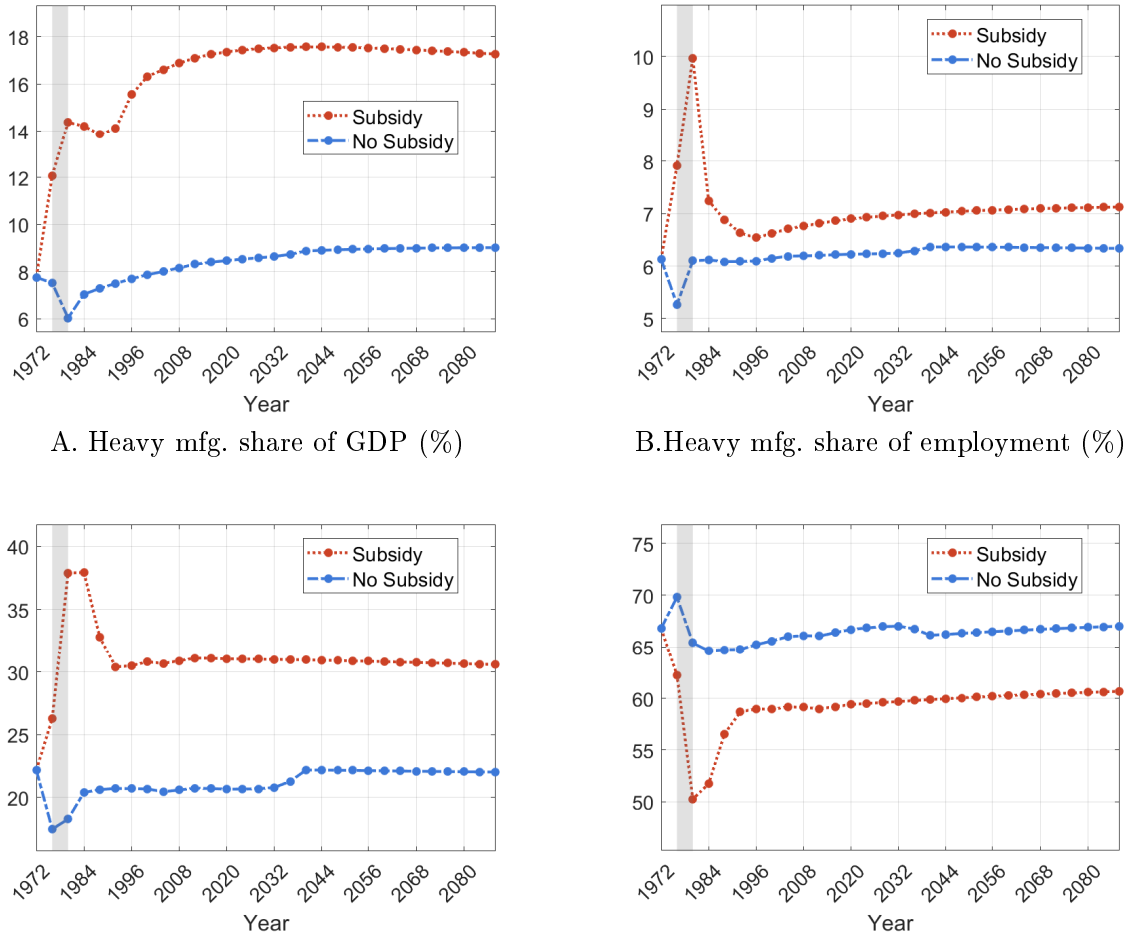
B.Heavy mfg. share of employment (%)

C. Heavy mfg. share of export (%)

D. Light mfg. share of (%)

Figure D3. The Effects of the Temporary Subsidies When there is No Roundabout Production Structure

Notes. This figure plots counterfactual results without a roundabout production structure. Panels A, B, C, and D report the results for the heavy manufacturing sector employment, GDP, and export shares, and the light manufacturing sector export shares, respectively. The red dotted line plots the outcomes of the baseline economy and the blue dotted line plots the outcomes of the counterfactual economy.



A. Heavy mfg. share of GDP (%)

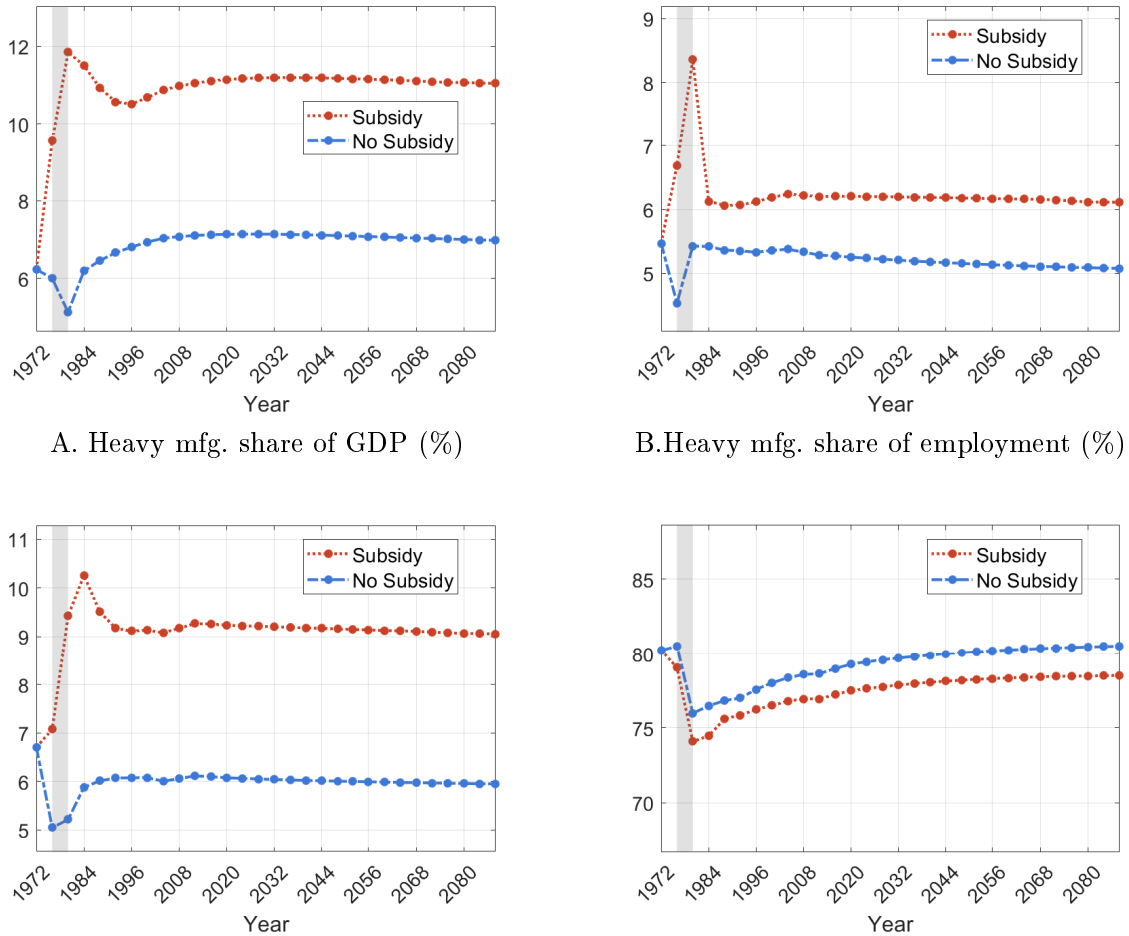
B.Heavy mfg. share of employment (%)

C. Heavy mfg. share of export (%)

D. Light mfg. share of (%)

Figure D4. The Effects of the Temporary Subsidies with Higher Migration Costs

Notes. This figure plots counterfactual results with a 10% higher level of migration costs than the calibrated value in the baseline economy. The red line plots the outcome of the baseline economy and the blue line plots the outcome of the counterfactual economy.



A. Heavy mfg. share of GDP (%)

B.Heavy mfg. share of employment (%)

C. Heavy mfg. share of export (%)

D. Light mfg. share of (%)

Figure D5. The Effects of the Temporary Subsidies When Foreign Market Size is Smaller

Notes. This figure plots counterfactual results with a lower level of foreign market size than the calibrated values in the baseline economy. The red line plots the outcome of the baseline economy and the blue line plots the outcome of the counterfactual economy.

D.2 Calibration Procedure

Data Inputs. The quantitative exercises requires the following data inputs:

- Aggregate data
  1. Initial conditions:
     - Initial shares of adopters in the previous period: $\{\lambda^T_{njt_0}\}_{n\in\mathcal{N},j\in\mathcal{J}^T,t_0=1968}$
     - Initial population distribution: $\{L^{Data}_{nt_0}\}_{n\in\mathcal{N},t_0=1968}$
  2. Sectoral gross output of each region: $\{GO^{Data}_{njt}\}_{n\in\mathcal{N},j\in\mathcal{J},t\in\{1972,1976,1980\}}$
  3. Regional population: $\{L^{Data}_{nt}\}_{n\in\mathcal{N},t\in\{1972,1976,1980\}}$
  4. Sectoral export shares at the national level: $\{EX^{Data}_{jt}/GO^{Data}_{jt}\}_{j\in\mathcal{J},t\in\{1972,1976,1980\}}$ where EX^{Data}_{jt} and GO^{Data}_{jt} are sector j's exports and gross output at the national level
  5. Sectoral import shares at the national level: $\{IM^{Data}_{jt}/E^{Data}_{jt}\}_{j\in\mathcal{J},t\in\{1972,1976,1980\}}$ where IM^{Data}_{jt} and E^{Data}_{jt} are imports and total expenditure on sector j goods at the national level
  6. Import and export tariffs: $\{t^{im}_{jt}\}_{j\in\mathcal{J},t\in\{1972,1976,1980\}}$ and $\{t^{ex}_{jt}\}_{j\in\mathcal{J},t\in\{1972,1976,1980\}}$
- Micro moments
  1. Identifying moment $\hat{\beta}^{pol}$ (Equation (6.4))
  2. Median of light and heavy mfg. shares of exports in 1972 across regions
  3. Median of heavy mfg. shares of adopters in 1972 and 1982 across regions
  4. Percent of zero adoption regions in 1972 and 1982

Algorithm. Taking the values of $\mathbf{\Theta}^E$ and data inputs as given, we obtain the values of $\mathbf{\Theta}^M$, $\{\bar{s}\}_{t\in\{1976,1980\}}$, and $\mathbf{\Psi}_t$ using the following calibration algorithm:

1. Guess parameters.
2. Guess fundamentals $\{c_{fj}, D_{fj}\}_{j\in\mathcal{J}}$, $\{V_{nt}\}_{n\in\mathcal{N}}$, and $\{\phi^{min}_{nj}\}_{n\in\mathcal{N},j\in\mathcal{J}}$
3. Given parameters $\{\mathbf{\Theta}^M, \bar{s}_t\}$, we solve the model and update the fundamentals $\mathbf{\Psi}_t$ for each period. Then, we fit region- and sector level aggregate outcomes to the data counterparts. This step corresponds to the constraints of Equation (6.2). The dimension of fundamentals is $|\{1972, 1976, 1980\}| \times (|\mathcal{N}| \times |\mathcal{J}| + 2 \times |\mathcal{J}^x| + |\mathcal{N}|)$, where $|\{1972, 1976, 1980\}|$ is the number of years when the model is exactly fitted to the region and sector data, $|\mathcal{N}| \times |\mathcal{J}|$ are the number of ϕ^{min}_{nj}, $|\mathcal{J}^x|$ is the number of D^f_j and c^f_j, and $|\mathcal{N}|$ is the number of V_n. For $t = 1$, we take the initial conditions from the data inputs as given. For $t = 2, 3$, we compute the initial conditions from the model outcomes in the previous period.
   (a) Update new $\{D^{f'}_{jt}\}$ using the following equation:

$$\underbrace{\frac{\mathrm{EX}^{Data}_{jt}}{\mathrm{GO}^{Data}_{jt}}}_{\text{Data}} = \underbrace{\frac{\sum_{n\in\mathcal{N}} \left(\frac{\sigma}{\sigma-1}\frac{c_{njt}t^{ex}_{jt}\tau^x_{nj}}{\bar{\phi}^{avg,x}_{njt}}\right)^{1-\sigma} D^{f'}_{jt}}{\sum_{n\in\mathcal{N}} \left(\frac{\sigma}{\sigma-1}\frac{c_{njt}}{\bar{\phi}^{avg}_{njt}}\right)^{1-\sigma} \left(\sum_{m\in\mathcal{N}} \tau_{nmj} P^{\sigma-1}_{mjt} E_{mjt}\right) + \left(\frac{\sigma}{\sigma-1}\frac{c_{njt}t^{ex}_{jt}\tau^x_{nj}}{\bar{\phi}^{avg,x}_{njt}}\right)^{1-\sigma} D^{f'}_{jt}}}_{\text{Model}}$$

(b) Update new $\{c'_{fj}\}$ using the following formula:

$$\underbrace{\frac{\mathrm{IM}_{jt}^{Data}}{\mathrm{E}_{jt}^{Data}}}_{\text{Data}} = \underbrace{\frac{\sum_{n\in\mathcal{N}} \left(\tau_{nj}^{x} t_{jt}^{im} c_{jt}^{f'} / P_{njt}\right)^{1-\sigma} E_{njt}}{\sum_{n\in\mathcal{N}} E_{njt}}}_{\text{Model}}$$

(c) Update new $\{V'_{nt}\}$ until the population outcome of the model fits the actual distribution of population:

$$\underbrace{L_{nt}^{Data}}_{\text{Data}} = \underbrace{\sum_{m\in\mathcal{N}} \frac{\left(V'_{nt} \frac{(1-\tau_t^w+\bar{\pi}_t^h) w_{nt}}{P_{nt}} d_{mn}\right)^{\nu}}{\sum_{n'=1}^{N} \left(V'_{n't} \frac{(1-\tau_t^w+\bar{\pi}_t^h) w_{n't}}{P_{n't}} d_{mn'}\right)^{\nu}} L_{mt-1}}_{\text{Model}} .$$

Only relative levels of $\{V'_{nt}\}$ are identified from the above equation, so we normalize the value of the amenity of the first region to be 1 for each period: $V'_{1t} = 1, \forall t$.

(d) Update new $\{\phi_{nj}^{min'}\}$ until shares of regional gross output are exactly fitted to the data counterparts:

$$\underbrace{\frac{GO_{njt}^{Data}}{\sum_{m\in\mathcal{N}} \sum_{k\in\mathcal{J}} GO_{mkt}^{Data}}}_{\text{Data}}$$

$$= \underbrace{\frac{\left(\frac{\sigma}{\sigma-1} \frac{c_{njt}}{\bar{\phi}_{njt}^{avg}}\right)^{1-\sigma} \left(\sum_{m\in\mathcal{N}} \tau_{nmj} P_{mjt}^{\sigma-1} E_{mjt}\right) + \left(\frac{\sigma}{\sigma-1} \frac{c_{njt} t_{jt}^{ex} \tau_{nj}^{x}}{\bar{\phi}_{njt}^{avg,x}}\right)^{1-\sigma} D_{jt}^{f'}}{\sum_{n'\in\mathcal{N}} \sum_{k'\in\mathcal{J}} \left(\frac{\sigma}{\sigma-1} \frac{c_{n'k't}}{\bar{\phi}_{n'k't}^{avg}}\right)^{1-\sigma} \left(\sum_{m\in\mathcal{N}} \tau_{n'mk'} P_{mk't}^{\sigma-1} E_{mk't}\right) + \left(\frac{\sigma}{\sigma-1} \frac{c_{n'k't} t_{k't}^{ex} \tau_{n'k'}^{x}}{\bar{\phi}_{n'k't}^{avg,x}}\right)^{1-\sigma} D_{k't}^{f'}}}_{\text{Model}},$$

The above equations only identify the relative levels of $\{\phi_{njt}^{min'}\}$, so we normalize the Pareto lower bound parameter of the first region and sector pair to 1 for each period.

4. After updating the geographic fundamentals, given values of parameters and subsidies, we evaluate the following objective function:

$$(m(\{\mathbf{\Theta}^M, \mathbf{s}_t\}) - \bar{m}^{Data})' \mathbf{W} (m(\{\mathbf{\Theta}^M, \mathbf{s}_t\}) - \bar{m}^{Data}),$$

where $m(\Theta)$ is the moments from the model, $\bar{m}^{Data}$ is the data counterparts, and $\mathbf{W}$ is the weighting matrix. We use the identity matrix for the weighting matrix.

5. For each value of $\{\mathbf{\Theta}^M, \mathbf{s}_t\}$, we iterate steps 2, 3, and 4 and find the values of $\{\hat{\mathbf{\Theta}}^M, \hat{\mathbf{s}}_t\}$ that minimize the objective function in the step 4.

D.3 Construction of Data Inputs

In this section, we describe how we constructed data inputs for the calibration procedure. We aggregate 10 manufacturing sectors into light and heavy manufacturing sectors.

D.3.1 Aggregate Data

Initial Shares of Adopters in 1968. While our firm balance sheet data covers from 1970 to 1982, technology adoption contracts cover from 1966 to 1985. We do not directly observe firm balance sheet data in 1968. Therefore, we use the information on the start year of firms to construct a set of firms that were operating in 1968. Then, we merge this set of firms with our data about their adoption activities and construct shares of adopters in the heavy manufacturing sector for each region.[48]

Regional Population Distributions in 1968, 1972, 1976, and 1980. The regional population data comes from the Population and Housing Census, the 2% random sample of the total population. The survey was conducted in 1966, 1970, 1975, and 1980. For the years not covered by this Census survey, we impute population using the geometric average of the two observed samples. For example, the population share of region n in 1973 is imputed as Pop. share$_{n,1973}$ = (Pop. share$_{n,1970}$)$^{\frac{3}{5}}$ × (Pop. share$_{n,1975}$)$^{\frac{2}{5}}$. From these imputed values, we obtain regional population in 1968, 1972, 1976, and 1980. The regional population distribution in 1968 is the initial condition that is taken as given in the model when solving for $t = 1$, whereas the regional population distributions in 1972, 1976, and 1980 are fitted by the regional population distributions of the model at $t = 1, 2, 3$, which are the endogenous outcomes of the model.

Regional and Sectoral Level Gross Output in 1972, 1976, and 1980. We compute gross output at the regional and sectoral level by harmonizing firm-level data and data from input-output tables following di Giovanni et al. (2020). Using firm-level data, we calculate a share of firm sales in region n and sector j and then multiply this share by the gross output of sector j at the national level. Specifically, we calculate

$$GO_{njt}^{Data} = \left(\frac{\sum\limits_{i \in nj} Sale_{it}}{\sum\limits_{m \in \mathcal{N}} \sum\limits_{k \in \mathcal{J}} \sum\limits_{i \in mk} Sale_{it}} \right) \times GO_{jt}^{IO},$$

where GO_{jt}^{IO} is sector j's gross output from the input-output tables. By doing so, we preserve the spatial distribution of firm sales but ensures that the total sum of sales across firms is consistent with the national input-output tables for each year.

Aggregate Export and Import Shares in 1972, 1976, and 1980. Both aggregate export and import shares are obtained from the national input-output tables. We calculate aggregate export share as $EX_{jt}^{Data}/GO_{jt}^{Data}$, where EX_{jt}^{Data} is sector j's exports of the input-output tables. In the model, we treat the service sector as a non-tradable sector, so we assume that exports and imports of the service sector are zero. We calculate aggregate sectoral import share is calculated as $IM_{jt}^{Data}/E_{jt}^{Data}$, where IM_{jt}^{Data} represent imports of sector j and E_{jt}^{Data} represent expenditures of sector j. We calculate E_{jt}^{Data} as follows:

$$E_{jt}^{Data} = \alpha_j \sum_{k \in \mathcal{J}} \left(\gamma_k^L \frac{\sigma - 1}{\sigma} GO_{kt}^{IO} \right) + \sum_{k \in \mathcal{J}} \gamma_k^j \frac{\sigma - 1}{\sigma} GO_{kt}^{IO},$$

where GO_{jt}^{IO} is sector j's gross output from the input-output table in year t.

[48] Given the facts that we cannot observe entry and exit of firms in 1968 and 1969 and we construct the shares based on the firms that operated in 1970 and these firms' start year.

Export and Import Tariffs Data in 1972, 1976, and 1980. We use data on export and import tariffs data are not used for the reduced-form empirical analysis but only for the quantitative exercises and not for the reduced-form empirical analysis. We obtain the data on export tariffs from Magee (1986).[49] The original dataset's industry code is in four-digit 1972 SIC codes. We first convert those codes into four-digit 1987 SIC codes and then into ISIC Revision 3 codes.[50]

We digitize import tariff data from Luedde-Neurath (1986) for 1974, 1976, 1978, 1980, and 1982, which are in the Customs Cooperation Council Nomenclature (CCCN). We convert CCCN to ISIC Revision 3 and then average the results across four-digit ISIC codes. For missing years, we impute values using the geometric average. We assume that the tariff level in 1972 was the same as that in 1974.

We aggregate trade tariffs up to four sectors for each year by taking the average across sectors. We do not use the weighted average, where the weight is given by import values. The weighted average gives zero weight to sectors with zero import values, which can underestimate the magnitude of the tariffs.

D.3.2 Micro moments

We compute shares of adopters for each year using our dataset. After computing these shares across regions and years, we compute the median for 1972 and 1980. Using this information, we compute shares of regions with zero values. We also obtain shares of exporters. However, because of many missing data points on exports, we take the three-year moving averages of shares of exports for each region and sector. We count firms with missing information on exports as non-exporters. Section D.4 describes how we calculate the identifying moment in more detail.

D.4 The Identifying Moment for Subsidy

We formally describe the identifying moment of the subsidy level using the following proposition.

Proposition D.1. *(Identifying Moment for Subsidies) Suppose a subsidy plan is given by Equation (6.3). Assume that (a) exogenous firm productivity follows the unbounded Pareto distribution ($\kappa \to \infty$), (b) goods are freely tradable ($\tau_{nmj} = 1$ and $\tau^x_{nj} = 1$), and (c) $j \in \mathcal{J}^T$ are symmetric. Consider the following regression model for $j \in \mathcal{J}^T$, $n \in \mathcal{N}$:*

$$\ln \lambda^T_{njt} - \theta\delta\lambda^T_{njt-1} = \beta^{pol} \times D^{pol}_{jt} + \delta_{nt} + \epsilon_{njt},$$

where D^{pol}_{jt} is a dummy variable of whether $j \in \mathcal{J}^{pol}$, and δ_{nt} are time-varying regional fixed effects. Then, when $\mathbb{E}[\ln \phi^{min}_{njt} | D^{pol}_{jt}] = 0$ holds,

$$\hat{\beta}^{pol} \xrightarrow{p} \beta^{pol} = \frac{\theta}{\sigma - 1}\left[\ln\left(\left(\frac{\eta}{1-\bar{s}}\right)^{\sigma-1} - 1\right) - \ln(\eta^{\sigma-1} - 1)\right],$$

and $\hat{\beta}^{po}$ uniquely identifies $\bar{s}$ for given values of η, δ, σ, and θ.

Proof. Suppose that a subsidy plan of the government is given as follows:

$$s_{njt} = \begin{cases} \bar{s} & \text{if} \quad t \in \{2,3\}, \quad \forall n \in \mathcal{N}, \quad \forall j \in \mathcal{J}^T \cap \mathcal{J}^{pol} \\ 0 & \text{otherwise.} \end{cases}$$

[49] We download the United States export tariff data from https://cid.econ.ucdavis.edu/ust.html.

[50] The concordance between 1972 SIC and 1987 SIC is obtained from https://www.nber.com.

Under the assumption that goods are freely traded, sectoral price index and real wage are equalized across regions, that is, $P_{njt} = P_{jt}$, $\forall n \in \mathcal{N}, \forall j \in \mathcal{J}$. Also, because of the symmetry assumption for $j \in \mathcal{J}^T$, $P_{jt} = P_{j't}$, $D^f_{jt} = D^f_{j't}$, and $F^T_j = F^T_{j'}$ hold for all $j, j' \in \mathcal{J}^T$. Under the symmetry and free trade assumptions, $\gamma^k_j = \gamma^k_{j'}$ and $\gamma^L_j = \gamma^L_{j'}$ hold for all for all $j, j' \in \mathcal{J}^T$, which in turn make $P_{jt} = P_{j't}$ holds. These two assumptions in turn imply that firms in sectors where technology adoption is available have the same market size.

From Equations (5.6) and (5.7), taking log, we can derive the following relationship:

$$\ln \lambda^T_{njt} = \theta\delta\lambda^T_{njt-1} + \underbrace{\frac{\theta}{\sigma-1}\ln\left(\left(\frac{\eta}{1-s_{njt}}\right)^{\sigma-1} - 1\right)}_{\beta D^{pol}_{jt}}$$
$$\underbrace{-\theta\ln\left(\frac{\mu c_{njt}(\sigma c_{njt}F^T_j)^{\frac{1}{\sigma-1}}}{\left(\sum_{m\in\mathcal{N}} P^{\sigma-1}_{jt}E_{mjt} + D^f_{jt}\right)^{\frac{1}{\sigma-1}}}\right)}_{=\delta_{nt}} + \underbrace{\theta\ln\phi^{min}_{njt}}_{=\epsilon_{njt}}, \quad \text{(D.1)}$$

where the second, third, and fourth terms can be mapped to the policy dummy variable D^{pol}_{jt} which equals one if sector j was targeted by the government in period t, region fixed effects δ_{nt}, and the error term ϵ_{njt}. Variation in the third term of the RHS across regions comes from wages w_{nt}, because $c_{njt} = (w_{nt}/\alpha^L_j)^{\alpha^L_j}\prod_{k\in\mathcal{J}}(P_{nkt}/\alpha^k_j)^{\alpha^k_j}$. This mapping gives us the following regression model:

$$\ln \lambda^T_{njt} - \theta\delta\lambda^T_{njt-1} = \beta D^{pol}_{jt} + \delta_{nt} + \epsilon_{njt}.$$

The condition for the estimates to be unbiased is $\mathbb{E}[\epsilon_{njt}|D^{pol}_{jt}]$. Under the model structure, this is equivalent to $\mathbb{E}[\ln\phi^{min}_{njt}|D^{pol}_{jt}]$ (Equation (D.1)). When this condition is satisfied,

$$\hat{\beta} \xrightarrow{p} \beta = \frac{\theta}{\sigma-1}\left[\ln\left(\left(\frac{\eta}{1-\bar{s}}\right)^{\sigma-1} - 1\right) - \ln(\eta^{\sigma-1} - 1)\right].$$

Given the values of θ, σ, and η, the RHS of the above equation has one-to-one relationship with $\bar{s}$. Therefore, $\bar{s}$ is uniquely identified.

□

The proposition shows that sudden increases in shares of adopters in 1980 captured by β^{pol} are informative about subsidies when they are uncorrelated with exogenous natural advantages; that is, when $\mathbb{E}[\ln\phi^{min}_{njt}|D^{pol}_{jt}] = 0$ holds. This proposition motivates our approach. Since the simplifying assumptions of Proposition D.1 do not hold exactly in either the model or the data, we identify the subsidy level by indirect inference.

Equation (6.4) differs from the regression model in Proposition D.1 in two ways. First, because the heavy manufacturing sector is the only sector where technology adoption is available in our quantitative exercises, we cannot control for δ_{nt}, and D^{pol}_t cannot be separately identified from time fixed effects. More specifically, we run this regression for shares of adopters in the heavy manufacturing sector in 42 regions for 1972 and 1980, so we use 84 samples in total. Note that we assumed that (i)

technology adoption is available only for heavy manufacturing firms and (ii) that common subsidies are provided across regions, and (iii) that we aggregate heavy manufacturing sectors into one sector when we took the model to the data, so we cannot control for region, sector, or time fixed effects. Ideally, a richer model that incorporates multiple heavy manufacturing sectors or more information on subsidy schedules across regions will allow us to control for additional fixed effects.

Second, we control for previous shares rather than subtract them from current shares in dependent variables. This is because the PPML is not defined for dependent variables with negative values and subtracting the previous shares from the current shares with zero values generates observations with dependent variables that take negative values. The estimated coefficients for β^{pol} and β_1 are 0.65 and 5.62, and statistically significant at the 1% level. The value of the estimated coefficient for β_1 (5.62) that corresponds to $\theta \times \delta$ in the model is consistent with the externally calibrated values $4.77 = 1.06 \times 4.5 = \theta \times \delta$. The estimation procedure and results.

Note that the assumptions of Proposition D.1 are not satisfied in the data or the model and we run the regression only for the heavy manufacturing sector. However, we can show that $\hat{\beta}^{pol}$ is still informative for $\bar{s}$. Under the unbounded Pareto distributional assumption, we can derive the following relationship from the model for the heavy manufacturing sector without any additional assumptions:

$$\begin{aligned}
&\ln \lambda^T_{n,heavy,t} - \theta\delta\lambda^T_{n,heavy,t-1} \\
&\quad = \underbrace{\frac{\theta}{\sigma-1}\ln\left(\left(\frac{\eta}{1-s_{n,heavy,t}}\right)^{\sigma-1}-1\right) \underbrace{-\theta\ln\left(\frac{\mu c_{n,heavy,t}(\sigma c_{n,heavy,t}F^T_{heavy})^{\frac{1}{\sigma-1}}}{\left(\sum_{m\in\mathcal{N}} P^{\sigma-1}_{heavy,t}E_{m,heavy,t}+D^f_{heavy,t}\right)^{\frac{1}{\sigma-1}}}\right)}_{=\mathbf{GE}_{n,heavy,t}(\mathbf{\Psi}_t,\mathbf{s}_t)}}_{=D^{pol}_t} \\
&\qquad + \underbrace{\theta\ln\phi^{min}_{n,heavy,t}}_{=\epsilon_{n,heavy,t}}.
\end{aligned}$$

Because the heavy manufacturing sector is the only sector that adopted technology and the only sector the government targeted, we can not identify D^{pol}_t separately from additional time fixed effects.

D^{pol}_t captures both the subsidies in the second term of the right hand side and the general equilibrium effects in the third term of the right hand side ($\mathbf{GE}_{n,heavy,t}(\mathbf{\Psi}_t, \mathbf{s}_t)$). $\mathbf{GE}_{n,heavy,t}(\mathbf{\Psi}_t, \mathbf{s}_t)$ depend all other regions' geographic fundamentals $\mathbf{\Psi}_t$ and subsidies $\mathbf{s}_t$. $\mathbf{GE}_{n,heavy,t}(\mathbf{\Psi}_t, \mathbf{s}_t)$ is a function of own exogenous natural advantage in the error term and therefore is correlated with the error term. This leads to the endogeneity problem of the regression model above. In Proposition D.1, we could absorb out these general equilibrium effects using region fixed effects by imposing the additional assumptions. However, that is not the case in the regression model above.

However, although $\hat{\beta}^{pol}$ is biased, it is still informative for $\bar{s}$. For given values of $\bar{s}$ and other structural parameters, we back out geographic fundamentals by exactly fitting region- and sector-level data. From these obtained geographic fundamentals, we can compute the error term and the general equilibrium effects. Therefore, our indirect inference for fitting $\hat{\beta}^{pol}$ can be thought of fitting the joint effects of both the subsidies in the second term and $\mathbf{GE}_{n,heavy,t}(\mathbf{\Psi}_t, \mathbf{s}_t)$.

D.5 Gravity Equation of Migration Flows

The data on migration shares comes from the 1995 Population and Housing Census, which is the closest to the sample period of our dataset among the accessible population census data. Because of data availability, regions are aggregated into 35 groups. μ^{1995}_{nm1990} is obtained as the total number of migrants moving from region n to region m from 1990 to 1995 divided by the total population of region n in 1990. When we compute the total population and the number of migrants, we restrict our sample age to 20 to 55. We also exclude outward migration flows from Jeju Island and inward migration flows to Jeju Island.

We estimate the above equation using OLS and PPML. The results are reported in Table D1. The estimated coefficient is around -1.3. The magnitude of the estimate implies that a 1 percent increase in distance decreases the share of outward migration by 1.3%.

Table D1: Gravity Equation of Migration Shares

Dep. Var.	Migration Shares from 1990 to 1995	
	OLS	PPML
	(1)	(2)
$LogDist_{mn}$	–1.30***	–1.39***
	(0.06)	(0.03)
Adj. R^2	0.88	.
# clusters (origin)	35	35
# clusters (destination)	35	35
N	1210	1225

Notes. This table reports the gravity estimates of Equation (6.1). The dependent variable is the log of the share of migration from region m to region n from 1990 to 1995. In column (1), we estimate the model using OLS. In column (2), we estimate the model using the Poisson pseudo-maximum likelihood estimation (Silva and Tenreyro, 2006). Clustered errors are two-way clustered at the origin and destination levels. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.