Technology Adoption and Late Industrialization*

Jaedo Choi
University of Michigan, Ann Arbor

Younghun Shim
University of Chicago

JOB MARKET PAPER

October, 2021

Click here for the latest version

Abstract

We study how the adoption of foreign technology and its local spillovers contributed to late industrialization in developing countries during the postwar period. Using novel historical firm-level data for South Korea, we provide causal evidence of direct productivity gains to adopters, and local productivity spillovers of the adoption. Based on these empirical findings, we develop a dynamic spatial model with firms' technology adoption decisions and local productivity spillovers. The spillovers induce dynamic complementarity in firms' technology adoption decisions. Due to dynamic 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 the model to the micro data and econometric estimates, and evaluate the impact of temporary adoption subsidies provided by the Korean government in the 1970s. Had no adoption subsidies been provided, Korea would have converged to an alternative less-industrialized steady state in which the heavy manufacturing GDP share and the aggregate welfare would be 15% and 10% lower than the steady state with successful industrialization. Thus, temporary subsidies to technology adoption had permanent effects.

Keywords: Technology Adoption, Industrialization, Knowledge Spillover, Path Dependence
JEL Codes: O14, O33, O53, R12

*We are extremely grateful to Andrei A. Levchenko, Dominick Bartelme, Jagadeesh Sivadasan, and Sebastian Sotelo for their guidance and invaluable suggestions and encouragement. We also thank Maria Aristizabal-Ramírez, Barthélémy Bonadio, Javier Cravino, Alan Deardorff, Kyle Handley, Yongjun Kim, Nishaad Rao, Linda Tesar, and seminar participants at MWIT, ETSG, and the University of Michigan for helpful comments. We would like to gratefully acknowledge research support from Kwanjeong Educational Foundation. `E-mail: jaedohi@umich.edu`.

1 Introduction

Large cross-country total factor productivity (TFP) differences suggest that technology is fundamental to economic development.[1] Thus, it has been argued by both economists and policy-makers that adoption of advanced technology used by rich countries 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 new knowledge from adopted foreign technologies can be spread to other local firms.[2]

In the postwar period, there have been divergent patterns of industrialization among developing countries. A subset of developing countries such as South Korea, Taiwan, and Turkey transformed their economies from agricultural- to a manufacturing-based in the postwar period, while many others remained stagnant. These developing countries achieved industrialization by adopting foreign technology rather than their own innovation.[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 casual look at the rapid industrialization of these latecomers provides suggestive evidence on the 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 on firms' adoption activities in these late industrializing countries. The key data challenge is that technology adoption is typically not observed directly, and must be inferred from other equilibrium outcomes.

This paper answers the following question: how do the adoption of foreign technology and its local spillover contribute to late industrialization? We study transition of South Korea (henceforth Korea) toward heavy manufacturing sectors in the 1970s. Korea is known for the most successful and rapid industrialization among the latecomers.[5] This paper makes three contributions. First, we overcome the empirical challenge of the existing literature by constructing a novel historical data set that covers the universe of technology adoption contracts between Korean and foreign firms. Thus, we can measure firm-level technology adoption directly. Most of adopted technologies during this time was know-how on operation of plants and capital equipment.

Second, we provide reduced-form empirical evidence on the firm-level effects of technology adop-

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

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

[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 (1962) 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 adoption of foreign technology.

[5]See Lucas (1993).

tion. We develop causal estimates of the direct productivity gains using a "winners vs. losers" research design following Greenstone et al. (2010). An empirical challenge for identifying the direct productivity gains is that firms make adoption decisions endogenously, which leads to the standard selection problem. We deal with this selection problem by comparing winners (the treated) who successfully adopted technology to losers (the control) who received the approval from the government to pursue foreign technology and made an actual contract with a foreign firm but failed to adopt technology in the end because of exogenous contract cancellations by foreign contractors that are not related to the circumstances surrounding the domestic firm. 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 these losers form a valid counterfactual for matched winners conditional on matched observables. We collect these cancellation episodes from historical contract documents. Our estimates imply that technology adoption increased adopters' sales and revenue total factor productivity (TFP) by 40-50%.

We also provide empirical evidence on the local productivity spillovers. The key identification challenge when estimating the spillovers is omitted spatially correlated shocks that 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, where the median land area of our geographic unit of analysis is about Manhattan-sized. Within each region-sector, we construct a spillover measure for each firm as the weighted average of local adopters of the same sector, where the weight is given by the inverse of distance to other firms. This measure varies at firm-level within region depending on firms' geographical proximity to adopters. We then regress non-adopters' sales and productivity on this spillover measure while controlling for time-varying region-sector fixed effects. Because we are controlling these fixed effects, our results are driven by variation in distances to adopters of the same sector within Manhattan-sized regions rather than driven by variation across regions and sectors, so the usual regional or sectoral unobservables are not a concern in our empirical analysis. We find non-adopters' sales and productivity grew faster if more neighboring firms adopted foreign technology. Our estimates indicate that a 1% increase in a local share of adopters leads to a 4-5% increase in non-adopters' sales and revenue TFP.

Our third contribution is that we build a dynamic spatial general equilibrium model with heterogeneous firms that make technology adoption decisions and local productivity spillovers. We use the model to evaluate the general equilibrium effect of the Korean government policy of subsidizing technology adoption by domestic firms. Firms' adoption decisions and the spillover endogenously shape regional- and national-level comparative advantage and export patterns. 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 a 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, a share of adopters becomes a time-varying state variable. The spillover generates a dynamic complementarity in firms' adoption decisions. A higher share of adopters in the previous period leads to higher gains from adoption and, in turn, induces more firms to adopt technology in the current period. Because this spillover is not internalized by adopters, the amount of adoption is suboptimal, which justifies policy interventions that promote adoption.

In a simplified model, we show analytically that the 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, the dynamic complementarity induces more firms to adopt technology, which in turn magnifies the strength of the complementarity in subsequent periods and vice versa. By changing initial conditions, 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 the micro and regional data. There are four important objects of the model related to technology adoption: two parameters that govern the strength of the direct productivity gains and the spillover, the adoption subsidies, and the fixed adoption cost. The first two parameters are related to benefits from technology adoption, and the latter two are related to the costs of adoption. The model delivers structural equations that can be mapped to our reduced-form regression specifications. Thus, we can use the empirical estimates to pin down the two parameters related to the direct productivity gains and the spillover. The 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 in which subsidies were available, relative to the initial period when the subsidies were not provided. We formally show that this moment pins down the input subsidy rate and estimate the subsidy rate by fitting this moment. The intuition behind this approach is that given information on the direct and spillover gains from adoption, pinned down by our reduced-form estimates, the relative increases in shares of adopters are attributable to a reduction in adoption costs induced by the subsidies. Finally, a fixed adoption cost is identified 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 Korea would have evolved differently had the government not provided the subsidies. Consistent with the analytical results, had no subsidies been provided, Korea would have converged to an alternative less-industrialized steady state. In this less-industrialized steady state of the counterfactual economy, when compared to the steady state of the baseline economy where the subsidies were provided, heavy manufacturing GDP, export, and employment shares would have been 10%, 20%, and 2%

permanently lower. Also, the aggregate welfare would have been 10% lower. These aggregate changes 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 impacts of industrial technology adoption on firm performance in developing countries (see, among many others, Atkin et al., 2017; Juhász, 2018; Giorcelli and Li, 2021; Juhász et al., 2020; Hardy and McCasland, 2021). We provide new empirical evidence on the direct productivity gains to the adopters, which is consistent with previous studies that have found positive impacts of technology adoption on firms' performance.

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, 2019). Relative to previous papers that 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 Korea.

Third, we contribute to the quantitative literature on multiple equilibria and big push. According to theoretical big push literature that dates to Rosenstein-Rodan (1943); Hirschman (1958), underdevelopment results from complementarity and coordination failures (see, among others, Murphy et al., 1989; Rodríguez-Clare, 1996; Ciccone, 2002). However, little is known about quantitative implication of multiple equilibria and coordination failure. We contribute to this literature by quantifying permanent effects of the temporary subsidies by the Korean government. Crouzet et al. (2020) and Buera et al. (2021) study complementarity in firms' technology adoption decisions, caused by network externalities and higher intermediate intensities of the adoption goods, respectively. Unlike these papers, we study local productivity spillovers of technology adoption. In terms of the modeling framework, our paper is most closely related to Allen and Donaldson (2020) who study the role of history in determining the spatial distribution of economic activity. Similar to Allen and Donaldson (2020), amounts of technology adoption are also determined by history in our model. Our quantitative results are also consistent with the macroeconomic literature on technology adoption (Parente and Prescott, 1994; Comin and Hobijn, 2010) that finds large aggregate consequences of technology adoption in developing countries.

Finally, this paper contributes to the trade literature on the evolution of comparative advantage. Aggregate data show that comparative advantage evolves (Hausmann and Klinger, 2007; Hanson et al., 2015; Levchenko and Zhang, 2016; Schetter, 2019; Atkin et al., 2021), but there has been limited understanding of driving forces of the evolution of comparative advantage.[6] Using detailed micro

[6] For theoretical works, see, among others, Krugman (1987) and Matsuyama (1992) for learning-by-doing and Buera and Oberfield (2020) and Cai et al. (2021) for knowledge diffusion.

data, Pellegrina and Sotelo (2021) document that knowledge diffusion through migration shaped the comparative advantage of Brazil and Arkolakis et al. (2019) study the role of knowledge diffusion through immigration on the technology frontier of the US in the nineteenth century. We contribute to this literature by quantifying how technology adoption shaped Korea's comparative advantage in heavy manufacturing sectors.

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

2 Data

We construct our main data set by merging firm-level balance sheet data with data of firm-level technology contracts between domestic and foreign firms. These two data sets are matched based on firms' names. The final data set only covers manufacturing sectors. Firms are classified into 10 manufacturing sectors, 4 of which are heavy manufacturing. The sample period of the constructed data set is between 1970 to 1982. The data cover the universe of technology adoption contracts during this period. Firm balance sheet information is representative at the national level. On average, the data set covers 75% of sectoral gross output from the input-output (IO) tables and 66% of the gross national output.[7] We describe our data construction procedure in Appendix Section A in more detail.

Firm-Level Technology Adoption Contracts. Firm-level technology adoption data were hand-collected and digitized from official documents on domestic firms' technology contracts with foreign firms from *National Archives of Korea* and from Korea Industrial Technology Association (1988).[8] The law required domestic firms to submit related documents when making a technology contract with foreign firms.[9] Most of the adopted technologies were know-how on installation or operation of capital equipment or turnkey plants.[10] In this period, there were 2,679 contracts made by 887 unique firms in total. Of these 2,679 contracts, 84% were in the heavy manufacturing sectors.

[7] The ratio between the total sum of firm sales of the data and the gross output from the IO tables is reported in Appendix Figure A2. Also, see Appendix Table A2 for descriptive statistics of the data.

[8] Korea Industrial Technology Association (1988) only recorded names of domestic and foreign firms, whereas the contract documents from the *National Archives of Korea* provides institutional details of each contract.

[9] 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 authorities.

[10] Specifically, about 74% of technology adoption contracts provided know-how, 21.2% granted licenses, and 4% permitted the use of trade-marks. For example, Appendix Figure A1 is one page of the contract document between *Kolon* (Korean) and *Mitsui Toatsu* (Japanese). The contract shows that *Mitsui Toatsu* had to provide technical assistance and blue prints to *Kolon*.

Balance Sheet Data. We obtain firm balance sheet data by digitizing the *Annual Report of Korean Companies*.[11] The firm balance sheet data has information on sales, assets, fixed assets, exports, and addresses of locations of establishments for the sample period between 1970 and 1982.[12] Employment is available only after 1972. Using information of addresses of plants and factories, we can map firms' adoption activities to their location of production.[13]

3 Historical Background on Late Industrialization in South Korea

In late 1972, the 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 the adoption of foreign industrial technology.[14] Adoption of foreign technologies and imported capital equipment related to these adopted technologies were the main channel of technology transfer from foreign developed economies in Korea.[15]

One feature of the HCI Drive is that the timing of the policy and selection of the targeted sectors were driven by a political shock rather than economic conditions (Lane, 2019).[16] After the Vietnamese war, President Nixon changed the US diplomatic policy toward the East Asian allies. In Nixon Doctrine (1969), he declared that the East Asian allies, including South Korea, should take the primary responsibility for their self-defense rather than relying on the US armies. He further planned the complete withdrawal of the US armies out of South Korea. However, at the same time, the military tension between South and North Korea was rising. Because South Korea was heavily relying on the US armies, Nixon Doctrine raised a threat for the national defense of South Korea. In late 1972, in order to modernize its military forces and achieve self-reliant defense against North Korea, President Park of South Korea announced the HCI Drive for promoting the heavy and chemical manufacturing

[11] *Annual Reports of Korean Companies* were published by the *Korea Productivity Center*. It covers firms with employment larger than 50.

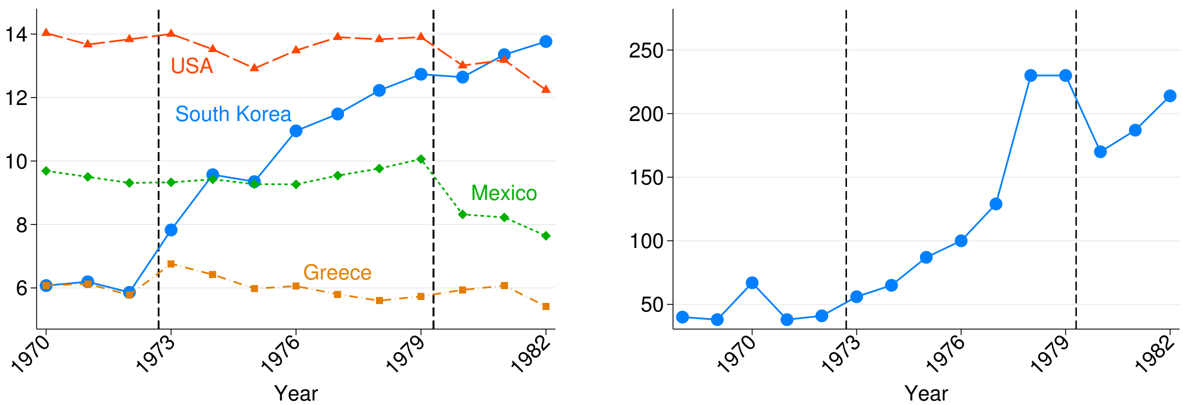
[12] All monetary values are converted to 2015 US dollars.

[13] We convert addresses into the 2010 administrative divisions of Korea.

[14] For example, *Hyundai Motors*, the largest automotive company in Korea, did not have its own models until 1972. It merely re-assembled the existing car model developed by *Ford* and most of the automobile parts were imported. It was only after 1972 that *Hyundai Motors* could start to produce its own models, which was possible because of technology adoption. In 1974, *Hyundai Motors* hired George Turnbull who was 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* for importing new capital equipment and construction of new turnkey plants related to adopted technologies.

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

[16] In Appendix Section B.2, we provide empirical evidence that supports this historical narrative of the political shock using an event-study specification. Also, see Choi and Levchenko (2021) and Kim et al. (2021) for furthur background on the HCI Drive.



A. Heavy mfg. GDP shares (%)

B. # of new foreign technology adoption contracts made by Korean heavy mfg. firms

Figure 1. Late Industrialization and Technology Adoption in Korea

Notes. Panel A illustrates heavy manufacturing GDP shares. Panel B illustrates the number of new technology adoption contracts with foreign firms in the heavy manufacturing sectors between 1968 and 1982 in Korea. The two dotted vertical lines represent the start and the end of the Korean government policy that subsidized technology adoption between 1973 and 1979. Cross-country heavy manufacturing GDP shares are obtained from OECD STAN Database and OECD National Accounts Statistics.

sectors that are related to arms industry.[17] 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 Korea and other selected economies.[18] While at the beginning of the period of our analysis, Korea's heavy manufacturing share was only 6%, during the sample period is achieved a remarkable takeoff, surpassing Mexico by the mid-1970s, and the US by 1982. The right panel plots the yearly number of new adoption contracts between Korean and foreign firms. Our novel data set reveals that with the GDP share of the heavy manufacturing sectors rapidly increasing from 6 to 14% between 1970 and 1982, the yearly number of new adoption contracts between Korean and foreign firms quadrupled. This sudden and rapid increase in amounts of the adoption coincided with the Korean government policy that provided temporary subsidies for technology adoption between 1973 and 1979.[19] Even after the policy ended in 1979, the Korean economy remained specialized in the heavy manufacturing sectors.

[17]At the same time, President Park declared martial law and amended the country's constitution into an authoritarian constitution, called Yushin constitution, which extended his terms of office for the president.

[18]Cross-country heavy manufacturing GDP shares are obtained from OECD STAN Database and OECD National Accounts Statistics.

[19]In Appendix Figure B1, we additionally report the heavy manufacturing employment and export shares, and the measure of revealed comparative advantage (Balassa, 1965). Consistent with Figure 1, the employment share increased from 4% to 9%, the export share increased from 13.5% to 35%, and the revealed comparative advantage measure rose

4 Reduced-Form Evidence on Technology Adoption

In this section, we empirically examine how technology adoption benefited Korean firms. We provide econometric evidence on direct productivity gains to adopters and the local productivity spillovers to non-adopters.

4.1 Direct Productivity Gains to Adopters

Empirical Strategy: "Winners vs. Losers" Design. When estimating the direct productivity gains of the adoption on adopters' performance, one of the key econometric challenges is that firms endogenously make adoption decisions. 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 such ideal random assignment, we implement a "winners vs. losers" research design motivated by 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 (i) initially made actual contracts with foreign firms that got approved by the government, but (ii) these contracts were not implemented in the end because of exogenous cancellations by foreign firms that are not related to a domestic contracting firm (loser). Examples of these exogenous cancellations by foreign firms include cancellations due to changes in foreign firms' management teams or bankruptcy of foreign firms.[20] When contracts are canceled after approvals from the government, domestic firms had to report the related documents on why contracts were canceled.[21] We collect cancellation episodes by reading these thousands of historical documents from the archives.

After identifying losers, we match each loser with an adopter using the exact Mahalanobis matching algorithm.[22] The matching proceeds in two steps. First, we exactly match on region and sector. By doing so, we can absorb shocks that are common at the region-sector level.[23] Second, within region-sector, we pick 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.[24] We match losers and winners with replacement, so one winner can be matched to multiple losers in a given year if they are in the same sector and region. The matching procedure gives us 34 pairs with 57 unique firms.

from 0.2 to 0.65.

[20]We exclude cancellations by domestic firms. Examples of these cancellations include domestic firms' sudden decreases in cash flow.

[21]See Appendix Figure D4.

[22]See Appendix Section D.3 for more detail on the matching procedure.

[23]For example, the matched pair will be subject to common sectoral shocks and have a similar level of market size or the local cost of production.

[24]The Mahalanobis distance is defined as $\sqrt{(x-\mu)S^{-1}(x-\mu)}$ where x is a set of observables, and μ and S are the mean and the covariance matrix of these observables.

All the matched pairs were heavy manufacturing firms.[25]

Using the matched pairs of winners and losers, we estimate the following event study specification which is a generalized triple difference design where a matched winner adopts in different periods and a loser is the control group: for firm i of pair p in period t,

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

where i denotes firm, p pair, and t time. D_{pt}^{k} are event-study variables defined as $D_{pt}^{k} := \mathbb{1}[t - \tau = t(p)]$ where $t(p)$ is event year of pair p. $\mathbb{1}[Adopt_{it}]$ is a dummy variable of adoption status. δ_i, δ_p, and δ_t are firm, pair, and year fixed effects; and ϵ_{ipt} is an error term. Dependent variables y_{ipt} are log sales, log revenue total factor productivity (TFP) estimated based on Wooldridge (2009), and labor productivity defined as value-added per worker.[26] We two-way cluster standard errors at both firm and pair level because of possible appearance of the same firm that introduces mechanical correlation across residuals.

Identifying Assumption. The 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 orthogonal to domestic firms' unobservables.

Our matching procedure makes the first condition likely to hold. It ensures that losers and winners are well-balanced in terms of observable covariates. Also, because we are comparing winners and losers, both of which wanted to adopt technology, we are indirectly controlling underlying unobservables that made these firms self-select into the adoption. Finally, unobservable political favors or subsidies provided during the the periods in which subsidies were available could affect firms' adoption decisions. However, because our definition of losers requires contracts to be approved by the government, we expect that both winners and losers had a similar level of political favor from the government at the time of their 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 will be violated only by unobservable shocks that (i) affect losers' performance after the event and (ii) that are correlated with foreign firms' cancellations, but (iii) that does not affect losers' performance before the event. One example would be that losers' negative shock at the time of the event made them be matched with bad foreign contractors that experienced changes in manager teams or went into bankruptcy. Although we cannot completely rule out this scenario, we believe that shocks that satisfy the above conditions are very unlikely. These losers are the ones that self-select into the adoption. Negative shocks that

[25] Therefore, our empirical estimates should be interpreted as the direct productivity gains to adopters of heavy manufacturing sectors.

[26] Appendix Section D.4 describes our revenue TFP estimation procedure in more detail.

hit losers at the time of the event would cause them not to adopt technology rather than adopt from bad foreign firms. Also, in many cases, the documents related to these cancellations reveal that these losers contacted foreign firms for several periods and tried to make a deal after the cancellation.[27] If negative shocks were driving our results, these losers might not have made further contacts with foreign contractors.

Balance. To assess covariate balance between two groups, we report descriptive statistics of the matched pairs and covariate balance test results. Descriptive statistics (Appendix Table D1) 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.[28] In Appendix Table D2, we report the covariate balance test results where we estimate a linear probability model on the effects of pre-event firm observables on an actual adoption status. Across all specifications, none of the estimated coefficients on 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 that cancellations by foreign firms were exogenous shocks to domestic firms.

Baseline Results. Table 1 and Figure 2 report the estimated diff-in-diffs coefficients in Equation (4.1). There were no pre-trends and winners' sales, revenue TFP, and labor productivity started increasing only after actual adoptions occurred. On average, the adoption increased sales, revenue TFP, and labor productivity by 51%, 46%, and 64%, respectively.[29]

Nearest Neighbors Matching. We supplement our analysis using "winners vs. losers" research design with nearest neighbors matching. We match a firm that adopt foreign advanced technology for the first time to a non-adopter that had a similar path of pre-event outcomes before the event using the whole adoption sample. For each adopter and each potential control firm within the same sector, we compute the Mahalanobis distance of an outcome and growth of an outcome for four years prior to one year to the event. Then, we pick a non-adopter that minimizes the Mahalanobis distance with an adopter. We restrict our sample to be heavy manufacturing firms, consistent with 34 pairs

[27] See the right panel of Appendix Figure D4.

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

[29] This average is calculated by estimating the following pooled diff-in-diffs model:

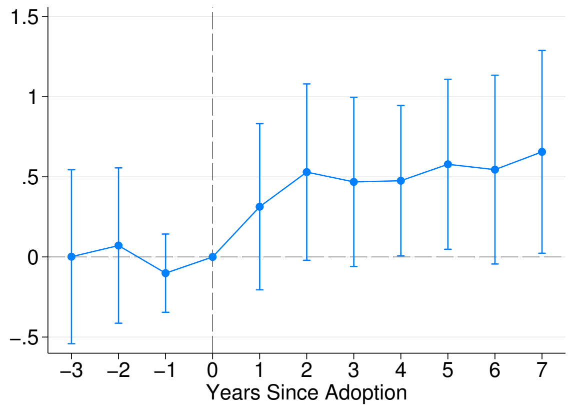
$$y_{ipt} = \sum_{k=\underline{T}}^{\bar{T}} \beta_k \times D^{\tau}_{pt} + \beta^{diff} \times Post_{pt} \times \mathbb{1}[Adopt_{it}] + \delta_i + \delta_p + \epsilon_{ipt},$$

where $Post_{it}$ is an indicator for period t after the event and other variables are defined similarly to Equation (4.1). We report the average as the estimated coefficient of β^{diff}. The estimation results are reported in Appendix Table D3.

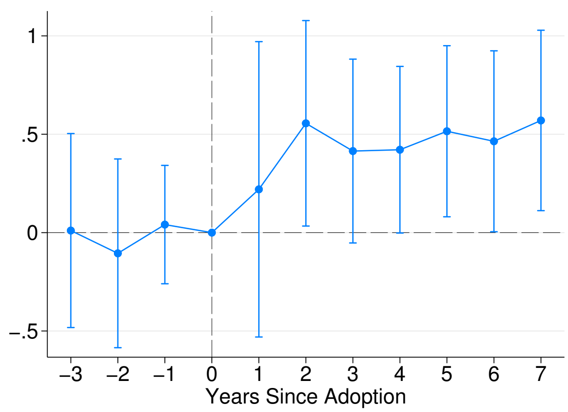
Table 1: Direct Productivity Gains from Technology Adoption: "Winners vs. Losers" Research Design and Nearest Neighbors Matching - Event Study Estimates

Research Design	"Winners vs. losers"			Nearest neighbor		
	log sales	log revenue TFP	log labor productivity	log sales	log revenue TFP	log labor productivity
	(1)	(2)	(3)	(4)	(5)	(6)
3 years before event	0.00	0.01	–0.09	0.07	0.07	0.03
	(0.27)	(0.24)	(0.41)	(0.06)	(0.06)	(0.09)
2 years before event	0.07	–0.11	–0.36	0.01	0.01	0.02
	(0.24)	(0.24)	(0.46)	(0.05)	(0.06)	(0.07)
1 year before event	–0.10	0.04	–0.02	–0.03	–0.01	–0.03
	(0.12)	(0.15)	(0.23)	(0.04)	(0.04)	(0.05)
Year of event						
1 year after event	0.31	0.22	0.28	0.08*	0.05	0.07
	(0.25)	(0.37)	(0.41)	(0.04)	(0.04)	(0.06)
2 years after event	0.53*	0.56**	0.64**	0.19***	0.14**	0.08
	(0.27)	(0.26)	(0.30)	(0.06)	(0.05)	(0.07)
3 years after event	0.47*	0.41*	0.62**	0.25***	0.21***	0.14**
	(0.26)	(0.23)	(0.29)	(0.07)	(0.07)	(0.07)
4 years after event	0.48**	0.42*	0.62**	0.31***	0.26***	0.15**
	(0.23)	(0.21)	(0.27)	(0.08)	(0.07)	(0.07)
5 years after event	0.58**	0.52**	0.43	0.39***	0.34***	0.24***
	(0.26)	(0.21)	(0.36)	(0.09)	(0.08)	(0.08)
6 years after event	0.54*	0.46**	0.55*	0.42***	0.37***	0.24***
	(0.29)	(0.23)	(0.28)	(0.12)	(0.10)	(0.08)
7 years after event	0.66**	0.57**	0.56*	0.42***	0.34***	0.20***
	(0.31)	(0.23)	(0.32)	(0.10)	(0.08)	(0.07)
Firm FE	Y	Y	Y	Y	Y	Y
Pair FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Adj. R^2	0.88	0.86	0.61	0.87	0.89	0.78
# cluster (pair)	34	34	34	151	151	151
# cluster (firm)	57	57	57	177	177	177
N	951	827	835	4824	4124	4137

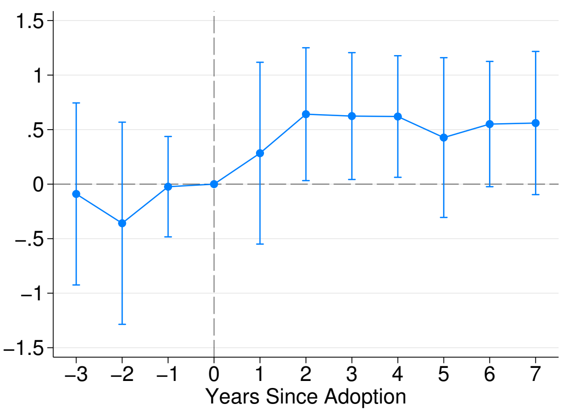
Notes. The table reports the estimated event study coefficients β_τ^{diff} in Equation (4.1). Columns (1)-(3) present the baseline event study estimates based on the "winners vs. losers" research design. Columns (4)-(6) present the event-study estimates using the nearest neighbors matching. β_0^{diff} is normalized to be 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. Log revenue TFP is estimated based on Wooldridge (2009). Across all specification, event time dummies, firm fixed effects, pair fixed effects, and calendar year fixed effects are controlled. Robust standard errors in parenthesis are two-way clustered at pair and firm levels. * p<0.1, ** p<0.05, *** p<0.01.



A. log sales



B. log revenue TFP



C. log labor productivity

Figure 2. Direct Productivity Gains from Technology Adoption: "Winners vs. Losers" Design

Notes. The 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, revenue TFP, and labor productivity. Revenue TFP is estimated based on Wooldridge (2009). Labor productivity is defined as value added per worker. β_0^{diff} is normalized to be zero. All specifications control for event time dummies and firm, pair, and calendar year fixed effects. Error bars represent 95 percent confidence intervals based on standard errors, two way clustered at pair and firm level.

of "winners vs. losers" research design. These selected nearest neighbors serve as a counterfactual for adopters.[30] The results are reported in columns (4)-(6) of Table 1.[31] Reassuringly there are no pre-trends, and the estimates have qualitatively similar patterns to the baseline results obtained from the "winners vs. losers" research design.

Other Robustness Checks. For robustness check, we also run the same regressions using different revenue TFP measures.[32] The results are reported in Appendix Figure D1 and Tables D4. Regardless of using different measures, the estimated event study shows no pre-trends, and the estimated coefficients were within a standard error of the estimates in Panel B of Figure 2.

We also run the regression using inputs for production as dependent variables: log employment and fixed assets. The results are reported in Appendix Figure D3. We do not find that winners increased usage of inputs relative to losers after the event. This shows that winners' increases in sales or revenue TFP measures are not input-driven and supports our interpretation of increases in sales or revenue TFP measures as the direct productivity gains to adopters.

4.2 Local Productivity Spillovers

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

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

where $nj/\{i\}$ is a set of sector j firms in region n excluding firm i; $dist_{ik}$ is distance between firm i and k; and $\mathbb{1}[\text{Adopt}_{k(t-h)}]$ is a dummy variable of firm k's adoption status lagged by h years. By lagging the variable, we allow for new knowledge from adopted technologies to take some time to be diffused locally. When constructing the spillover measure for firm i, we exclude firm i to rule out mechanical correlation. For our baseline specification, we take $h = 4$ and conduct robustness checks for different values of h. Within the same region and sector, each firm has different values of the spillover depending on its location and distances to adopters, which is the main variation we utilize for our empirical analysis.

The spillover measure can be interpreted more structurally as a probability of firm i's manager meeting other managers who work in firms that adopted technologies. Each manager is endowed

[30] Although we employ nearest neighbors matching, we still have 151 pairs, which is a much smaller number than the total number of contracts (around 2,600). This is because of the following reasons. First, we are restricting the treated group to be firms that adopted for a first time. Second, we require observations to have sales information one and four years before the event. Third, there were firms that made multiple contracts in a given year, but we do not consider intensive margins of technology adoption.

[31] Appendix Figure D2 plots the estimated coefficients and confirm no pre-trends.

[32] We use different revenue TFP estimates based on Ackerberg et al. (2015), Levinsohn and Petrin (2003), and OLS.

with a unit of time and can randomly meet at most one manager from other firms. A probability of meeting firm k's manager that is given by the inverse of the distance between firms i and k. The inverse of the distance proxies for spatial frictions that impede local interaction between managers of two firms.[33] The spillover measure captures that knowledge spillovers are highly localized and quickly decays with distance, as supported by the recent empirical evidence on knowledge spillovers.

Using this spillover measure, we consider the following fixed effect regression model:

$$y_{injt} = \beta^S \text{Spill}_{injt} + \delta_i + \delta_{njt} + \epsilon_{injt}, \tag{4.3}$$

where i denotes firm, j sector, n region, and t time. δ_i are time-invariant firm fixed effects and δ_{njt} are time-varying region-sector fixed effects. We use log sales and revenue TFP as dependent variables (y_{injt}). δ_i and δ_{njt} absorb time-invariant firm factors and time-varying region-sector level shocks, respectively.

To difference out firm fixed effects, we estimate Equation (4.3) in long-difference:

$$\triangle y_{injt} = \beta^S \triangle \text{Spill}_{injt} + \gamma y_{injt_0} + \mathbf{X}'_{injt_0}\boldsymbol{\beta} + \triangle \delta_{njt} + \triangle \epsilon_{injt}, \tag{4.4}$$

where $\triangle$ is a time difference operator. All specifications include initial dependent variable y_{injt_0}. The baseline sample includes firms that were operating before 1973 and after 1979 and did not adopt foreign technologies between these periods. $\mathbf{X}_{injt_0}$ are firm controls measured at the initial sample period, which allows for heterogeneous trends that depend on firm observables. Standard errors are two-way clustered at regional and ownership levels. In Korea, there are large conglomerate groups known as *Chaebols* that own multiple firms across different sectors and regions. By clustering at ownership level, we allow 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, where each set covers the period between 1973 and 1979 when the temporary subsidies were provided. Because we cluster firms at both regional and ownership levels, this is innocuous. We add dummies for each set of differences and interaction terms between these dummies and δ_{njt}.

Identifying Assumption. The identifying assumption for the estimates to admit a causal interpretation is that distance to adopters within region-sector (Spill_{injt}) is uncorrelated with the error

[33] By taking the weighted average, we are implicitly assuming that the spillover measure is invariant to the total number of firms. As far as we know, there is no consensus on the functional form of knowledge spillovers. 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 (Jovanovic and Rob, 1989; Lucas and Moll, 2014; Buera and Oberfield, 2020; Perla et al., 2021). Given that time of managers is limited in the real world, this theoretical interpretation seems to be more natural than an alternative story where a manager can interact with all firms in a local area. Under this alternative story, the spillover varies depending on the total number of firms. Second, the literature on externalities has commonly used averages to capture agglomeration forces, such as share of skilled labor (Moretti, 2004) and population density (Ciccone and Hall, 1996).

term ϵ_{injt} conditional on δ_{njt}, δ_i, and other controls. There are two identification concerns highlighted by Manski (1993). First, neighborhood shocks within region-sector level that are correlated across firms can affect both firm i's outcomes and adoption decisions of firm i's neighboring firms, leading to spurious correlation. Second, firm size of adopters tends to be larger than non-adopters, and there can be other effects of being close to these large-sized firms, leading 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, which is much finer than many other previous studies in the literature. Identifying variation comes purely from distances to adopters of the same sector within region but not from variation across regions or sectors. This is equivalent to comparing outcomes of two non-adopters of the same sector within Manhattan-sized regions, one of which is closer to adopters than the other. Variation of Spill_{injt} 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.[34] Because we are controlling for δ_{njt} and differencing out δ_i, neighboring firms' adoption decisions based on time-varying region-sector factors do not bias our estimates. Only firms' adoption or entry decisions based on time-varying firm-specific factors that are spatially correlated at neighborhood level would bias our estimates.[35] Exploiting spatial variation at a fine level mitigates this potential spatial correlation at neighborhood level within region-sector.

We deal with the second concern by isolating variations of being close to adopters from being close to big-sized firms through controlling other potential channels of local spatial interactions between firms. We control for the average sales of local firms inversely weighted by the distances similar to Equation (4.2):

$$\ln\left(\text{Spill-Sales}_{injt}\right) = \ln\left(\sum_{k\in nj/\{i\}}\left\{\frac{(1/dist_{ik})\text{Sales}_{kt}}{\sum_{k'\in nj/\{i\}}(1/dist_{ik'})}\right\}\right). \tag{4.5}$$

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

$$\ln\left(\text{Input-MA}_{injt}\right) = \ln\left(\sum_{j'}\sum_{k\in nj'/\{i\}}\gamma_j^{j'}(1/dist_{ik})\text{Sales}_{kt}\right), \tag{4.6}$$

[34] These new firm that enter between the start and the end of the sample period affect the spillover measure of the continuing firms, but they are not included in the sample, because we are restricting the sample to be firms who were operating at the start of the period.

[35] For example, infrastructure improvement at the neighborhood level that affects both firms' outcomes and adoption decisions would bias our estimates.

where $\gamma_j^{j'}$ is a share of sector j' intermediate inputs used by sector j.[36] The market access measure proxies differential market size due to localized input sourcing.

Baseline Empirical Results. Table 2 reports the OLS estimates for β^S. Column (1) of Panel A is our baseline estimate.[37] The estimated coefficient is statistically significantly positive. One standard deviation increase in the spillover of adoption contributes to 14.5% increases in sales.[38] 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 distanced. With this interpretation, in this hypothetical region, a 1% increase of a local share of adopters leads to a 4.39% increase in non-adopters' sales. In columns (2), (3), and (4), we additionally control for ownership fixed effects, ln(Spill-Sales), and ln(Input-MA) respectively.[39] In column (5), we jointly control for all additional 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.[40] In columns (6)-(10), we use log revenue TFP as a dependent variable.[41] The estimates for log revenue TFP are about 20% larger than the estimates for log sales.

In Panel B, we run the same regression for the full sample, including both adopters and non-adopters. For the full sample, we control a dummy variable of 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.

Robustness. We provide a battery of robustness checks. Instead of using the spillover measure with a 4-year lag, we use the spillover measure with 3-year or 5-year lags. The results are reported in Appendix Tables D5 and D6. When constructing the market access measure Input-MA, we try

[36] Because we do not have information on commodity or service sector firms, we sum j' only across manufacturing sectors. Instead of using $(1/dist_{ik})$, we also consider alternative weights $(1/dist_{ik}^{\alpha})$ for Input-MA$_{injt}$ where $\alpha = 1.1$ which is the value of the average coefficient based on a meta-analysis performed by Head and Mayer (2014).

[37] The magnitude of the estimated coefficients is consistent with the existing 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. The elasticity of the adoption spillover is calculated as follows. The mean level of a 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 percentile 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. The reason why we calculate the elasticity above the 90th percentile is that shares of adopters are highly skewed, where the 75th, 90th, 95th, and 99th percentiles are 0, 0.03, 0.06, and 0.18, respectively.

[38] This is calculated as $14.5 = 100 \times 0.033 \times 4.39$, where 0.033 is one standard deviation of the adoption spillover.

[39] When controlling for ownership fixed effects, we categorize independent firms into one single group.

[40] One potential explanation for the null results of ln(Spill-Sales) and ln(Input-MA) is that knowledge spillovers decay more quickly with distances than other spatial interactions captured by the other two controls and other two spatial interactions operate at broader spatial scale than knowledge spillovers. Then, conditional on δ_{njt}, ln(Spill-Sales) and ln(Input-MA) will not have significant results.

[41] For revenue TFP, there is a smaller number of samples because employment is only available after 1972. We only use one set of differences between 1972 and 1980.

Table 2: Local Productivity Spillovers of Technology Adoption

Dep. Var.	log sales					log revenue TFP				
	(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)	53	53	53	53	53	41	36	41	41	36
# clusters (Ownership)	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	4.23***	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(int-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 (Ownership)	702	697	702	702	697	381	338	381	381	338
N	1264	1259	1264	1264	1259	431	387	431	431	387
Region-Sector FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Ownership FE	N	Y	N	N	Y	N	Y	N	N	Y

Notes. The table reports OLS estimates of Equation (4.4). When constructing the spillover measure defined in Equation (4.2), we lag firms' adoption status by four years. In Panel A, we use the never-adopter sample, which only includes firms that did not adopt any technology until the end of the sample period. In Panel B, we use the full sample, which includes both adopters and non-adopters, and additionally control adoption status. Dependent variables are log sales and revenue TFP in columns (1)-(5) and (6)-(10), respectively. revenue TFP is estimated based on Wooldridge (2009). ln(Spill-Sales) and ln(int-MA) are additional controls defined in Equations (4.5) and (4.6). In all specifications, we control for region-sector fixed effects and initial dependent variable at the start of the sample period. Standard errors are two-way clustered at both region and ownership levels and are reported in parenthesis. * p<0.1, ** p<0.05, *** p<0.01.

alternative weights $1/dist^{1.1}$ where 1.1 is the average of the gravity coefficient reported by Head and Mayer (2014). The results are reported in Appendix Table D9.

Instead of using log sales or revenue TFP, we also use different dependent variables: log fixed assets, assets, and employment, labor productivity. The results are reported in Appendix Tables D7 and D8. The estimated coefficients are statistically significant and positive for different dependent variables except for employment.[42]

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 with J sectors.[43] Home consists of N regions. 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. The time is discrete and indexed by $t \in \{1, 2, \ldots\}$.

There are two types of goods: intermediate goods and final goods. Intermediate goods are produced by intermediate goods producers, indexed by subscript i. There is a fixed mass of firms M_{nj} in each region-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, a firm additionally incurs a fixed export cost (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 a fixed adoption cost. In each region, there is a competitive labor market.

5.2 Firms

Production. Each intermediate variety is produced by intermediate goods producers, which we call firms. Firms are heterogeneous in productivity. The output of firm i with productivity z_{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 L_{it} is labor inputs; M_i^k is sector k intermediate inputs; and γ_j^k are Cobb-Douglas shares. A unit cost of an input bundle is $c_{njt} = (w_{nt}/\alpha_j^L)^{\alpha_j^L} \prod_{k \in \mathcal{J}} (P_{nkt}/\alpha_j^k)^{\alpha_j^k}$, where w_{nt} is wage and P_{njt} is a price of intermediate inputs.

In each region and sector, a perfectly-competitive final good producer produces non-tradable local sectoral aggregate goods used for final consumption and for intermediate inputs. A final goods

[42] These results are consistent with the estimates from the "winners vs. losers" research design where we also find that the adoption increased firms' sales, fixed assets, and productivity measures except for employment.

[43] The small open economy set-up of our model is similar to that of Bartelme et al. (2021).

producer aggregates all available varieties from all regions and Foreign using a constant elasticity of substitution (CES) aggregator:

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

where Q_{njt} is the quantity produced of local aggregate sectoral goods; Ω_{mj} is available sector j varieties in region m; q_{it} is the quantity demanded of an intermediate variety ω produced by domestic firm i; and q_{it}^f is the quantity demanded from Foreign firm. We assume that the available set of Foreign varieties Ω_j^f is exogenously given to Home and is the same across regions.[44]

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 a price of Home variety; and p_{it}^f is a FOB price of an imported variety from Foreign. Under a small open economy, Home takes foreign firms' FOB price as given and c_{jt}^f is exogenous to Home.

Technology Adoption and Export. In each period, a firm makes two static decisions: (i) whether to adopt advanced technology and (ii) whether to export. Both adoption and exporting requires fixed adoption costs F_j^T and F_j^x in units of input bundles.[45] Once a firm decides to adopt technology and pays a fixed adoption cost, a firm can increase its productivity (Yeaple, 2005; Lileeva and Trefler, 2010; Bustos, 2011).[46]

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 which increases in a share of adopters in the previous period λ_{njt-1}^T;

[44] Because there is no fixed export costs for internal trade, each region faces the same set of available varieties.

[45] The fact that both adoption and export costs are fixed costs make firms' decisions static. If they are sunk costs, the firm's problem will be dynamic.

[46] F_j^T is a reduced form parameter that incorporates direct payment to foreign sources, costs of the installation of a new structure or capital equipment related to a newly adopted technology, and any barriers of 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, the political surroundings in the 1970s affected F_j^T. Due to the Cold War, the US government wanted the Korean economy to be self-sustaining and promoted Korea's economic growth. Therefore, it did not block transfers of technology to Korean firms within the free world (Vogel, 1991, p.8).

and ϕ_{it} is exogenous productivity.[47] 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 Korean economy within 10 years rather than long-run outcomes which have been studied more frequently in the standard trade literature.[48] When making adoption decisions, adopters internalize direct productivity gain η, but does not internalize the spillover $f(\lambda^T_{njt-1})$. This externality makes the social returns to adoption larger private returns, which leads to insufficient amounts of adoption than the socially optimum level. For sectors where technology adoption is not available $j \notin \mathcal{J}^T$, a firm's productivity only consists of the 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 a 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.

We briefly describe two possible sets of microfoundations based on historical case studies of Korea in the 1970s. Complete derivations of the microfoundations and related historical cases are described in Appendix C.4. In the first set of microfoundations, we consider local diffusion of new engineering knowledge from foreign technologies that were adopted in the previous period.[49] We present a model based on Desmet and Rossi-Hansberg (2014) where firms choose amounts of innovation that determine their productivity after incurring innovation costs. Local diffusion of new ideas decreases innovation costs (Romer, 1990) and amounts of diffusion increase in a local share of adopters in the previous period. Therefore, a higher local share of adopters in the previous period increases local firms' amounts of innovation by decreasing the innovation costs, which in turn increases the overall productivity of local firms.

In the second set of microfoundations, we consider learning externality and labor mobility across firms.[50] We introduce local engineers who are immobile across regions (unlike production workers)

[47] This specification implicitly assumes that the direct productivity gains are hicks-neutral, which is supported by empirical evidence. We run the event study regression in Equation (4.1) where a dependent variable is log capital per worker, reported in Appendix Figure D3. We find no changes in log capital per worker after adoption.

[48] Allowing the spillover to operate with a lag makes an economy have a deterministic static equilibrium 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 the other equilibrium in a different period depending on agents' self-fulfilling belief. Kline and Moretti (2014) and Allen and Donaldson (2020) also allowed agglomeration to operate with some lags. See Krugman (1991) and Matsuyama (1991) for further discussion on self-fulfilling beliefs. Second, because there is a unique static equilibrium for each period, the model can be easily mapped to the cross-sectional data. Multiple static equilibria models in general suffers from identification issues due to multiplicity (Jovanovic, 1989).

[49] The microfoundation based on local diffusion of knowledge is supported by the historical evidence which show that new ideas and tacit knowledge of adopted technologies were transmitted to local capital goods producers through reverse-engineering of capital equipment related to adopted technologies.

[50] This microfoundation is based on the historical episodes that technical personnel of adopters moved frequently to other firms and their movement played an important role in diffusing their knowledge on adopted technologies. This is

and live only for two periods.[51] An engineer only works in her adulthood (in her second period). Learning externality arises from that if an adult engineer was matched with a firm that adopted technologies in the previous period, her child learns about new knowledge from adopted technologies and obtains a higher level of engineering skills next period. Knowledge is transmitted to a child only when her parent is working at adopters. The local engineer labor market is frictional in that a local engineer and a local firm are randomly matched, and profits are split based on Nash bargaining (Acemoglu, 1996). Due to this random matching, the probability of being matched with more skilled engineers is equivalent to a local share of adopters in the previous period. Firms' productivity increases once they are matched with more skilled engineers. Therefore, a local adopter share in the previous period increases the overall productivity of local firms.

ϕ_{it} is drawn from a distribution $G_{njt}(\phi)$ that vary 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; Melitz and Redding, 2015):

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

which is parametrized by ϕ_{njt}^{max}, ϕ_{njt}^{min}, and θ. We 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.[52]

Adoption Subsidy. The adoption subsidies in Section 3 are modeled as input subsidies, which is based on the institutional details that the Korean government provided subsidies to big-sized adopters for purchases of intermediate inputs and new capital equipment related to adopted technologies. Adopters are subject to input subsidies $0 < s_{njt} < 1$ potentially varying across regions, sectors, and periods. Therefore, firm i's unit cost of production $\tilde{c}_{it}$ is $\frac{c_{njt}}{\phi_{it}f(\lambda_{njt-1}^T)}$ or $\frac{1-s_{njt}}{\eta} \times \frac{c_{njt}}{\phi_{it}f(\lambda_{njt-1}^T)}$ depending on whether firm i adopts technology or not, respectively. Adopters enjoy a lower unit cost of production than non-adopters because of both higher productivity (η) and input subsidies (s_{njt}).

further supported by higher aggregate labor mobility rates in Korea when compared to Japan and the US in the 1970s (Kim and Topel, 1995). Also, both learning externalities and knowledge spillovers through labor mobility have been widely studied in the literature. For example, see Lucas (1988) for learning externality of human capital; see Serafinelli (2019) for empirical evidence on the effects of labor mobility across firms on knowledge diffusion.

[51] The assumption that local engineers cannot move across regions is just for simplicity. As long as there are some mobility costs of engineers across regions, the learning externality will be local.

[52] 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 adopting technology regardless of the values of F_j^T. Helpman et al. (2008) also uses truncated Pareto distributional assumption to rationalize zero trade flows across countries.

The government imposes a labor tax τ_t^w to finance these subsidies.[53] We assume that the labor tax rate is constant across regions, so the after-tax wage in region n is $(1-\tau_t^w)w_{nt}$. 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 and Foreign are $q_{inmjt} = (\tilde{p}_{it})^{-\sigma} P_{mjt}^{\sigma-1} E_{mjt}$ and $q_{injt}^x = (\tilde{p}_{it})^{-\sigma} D_{jt}^f$ when firm i charges price $\tilde{p}_{it}$, respectively. A firm optimally charges a constant mark-up $\mu = \sigma/(\sigma-1)$ over its marginal cost. Thus, price charged by firm i in region n of sector j selling to region m is $p_{inmjt} = \mu\tau_{nmj}\tilde{c}_{it}$ and an export price of firm i is $p_{injt}^x = \mu\tau_{nj}^x\tilde{c}_{it}$.

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

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

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_{jt}^f is exogenous Foreign market. $\pi^m(T_{it};\phi_{it})$ 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 conditional on adoption status.

Adoption and Export Cutoff Productivities. Firms participate in adoption and exporting when gains from them are larger than fixed costs. These net gains from adoption and exporting are higher when a firm is more productive. Therefore, with fixed costs, firms' adoption and export decisions are characterized by cutoff productivities, where only firms with productivity above these cutoffs participate in adoption and exporting. We assume that fixed adoption costs are sufficiently higher than fixed export costs so that adopters always export to Foreign.

The export cutoff $\bar{\phi}_{njt}^x$ is defined at where operating profits in Foreign are equal to fixed export

[53] The assumption that the government finances its adoption subsidies through labor tax is based on the labor market policies and the pro-business attitude of the authoritarian Korean government in the 1970s. The government restricted firms' nominal wage growth to be 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).

costs, $\pi^x(T_{it}=1;\bar{\phi}^x_{njt}) = c_{njt}F^x_j$:

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

The adoption cutoff $\bar{\phi}^T_{njt}$ is defined at where profits conditional adopting and not adopting are equalized $\pi(T_{it}=1, x_{it}=1;\bar{\phi}^T_{njt}) = \pi(T_{it}=0, x_{it}=1;\bar{\phi}^T_{njt})$:

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

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

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

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

Dynamic Complementarity. The spillover generates dynamic complementarity in firms' adoption decisions: a higher share of adopters in the previous period increases gains from adoption in the current period. The dynamic complementarity operates through two channels. The first channel comes from complementarity between market size and productivity increase from adoption.[54] Because stronger spillover increases the productivity of one region relative to other regions, firms in this more productive region will have a larger market size and gains from adoption become larger due to scale effects. This complementarity further incentivizes firms in this more productive region to adopt technology in the current period, which in turn magnifies the spillover force in subsequent periods.

The second channel comes from lowered 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 makes local sectoral aggregate goods cheaper, which in turn makes fixed adoption costs lower.[55] These lowered 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.

[54] See Bustos (2011) and Lileeva and Trefler (2010).
[55] See Matsuyama (1995) and Buera et al. (2021).

5.3 Households.

Households make migration and consumption decisions. 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. In each period, we assume that each household owns w_{nt} shares of a fund that collects profits from all firms across regions and sectors and redistributes back to households (Chaney, 2008). We normalize the total population of Home country to be 1, $L_t = \sum_{n\in\mathcal{N}} L_{nt} = 1$, where L_{nt} is population in region n.

Preferences. Households have Cobb-Douglas preferences over final consumption baskets:

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

where C_{njt} is the consumption of local sector j aggregate goods in region n at period t and α_j is the final good consumption shares. Households are subject to their budget constraints each period: $\sum_{j\in\mathcal{J}} P_{njt}C_{njt} = (1-\tau_t^w+\bar{\pi}_t)w_{nt}$, 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 wage $(1-\tau_t^w)w_{nt}$ and dividends income $\bar{\pi}_t^h w_{nt}$.[56] We denote the ideal price index for households in region n by $P_{nt} = \prod_{j=1}^{J} P_{njt}^{\alpha_j}$.

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

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

where V_{nt} is an exogenous amenity in region n; d_{mn} is a moving cost 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). Under the distributional assumption, a share of households moving to region n from region m in period t is given by

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

The shape parameter ν is migration elasticity that governs the responsiveness of migration flows to

[56] Under the given portfolio structure, dividend per share is proportional to w_{nt}. Specifically, $\bar{\pi}_t^h = \frac{\sum_{n\in\mathcal{N}}\sum_{j\in\mathcal{J}}\int_{\omega\in\Omega_{nj}}\pi(\omega)d\omega}{\sum_{n\in\mathcal{N}} w_{nt}L_{nt}}$.

real income changes of destination.[57] 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. At the of 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:

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

5.4 Static Equilibrium

Define the average productivity inclusive of subsidies for all firms as follows:

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

$\bar{\phi}_{njt}^{avg}$ and capture how sector j firms in region n can produce at relatively lower cost on average. $\bar{\phi}_{njt}^{avg}$ increases in amounts of adoption (lower $\bar{\phi}_{njt}^T$), higher subsidy (higher s_{njt}), or higher spillovers (higher λ_{njt-1}^T). The expression for the average productivity inclusive of subsidies for exporters can be defined similarly, but the difference with $\bar{\phi}_{njt}^{avg}$ is that the lower bound of the distribution is replaced with $\bar{\phi}_{njt}^x$ instead of ϕ_{njt}^{min} because of selection effects induced by fixed exporting costs.

Regional variables in this economy can be expressed as a function of $\bar{\phi}_{njt}^{avg}$ and $\bar{\phi}_{njt}^{avg,x}$. Price index is expressed as:

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

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

[57] 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.

are expressed as[58]:

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

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}.

Market Clearing. In each period, the economy satisfies the following market clearing conditions. 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:

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

where $r(\phi_{it})$ is firm i's revenues.[59] The left hand and the right-hand sides are the total government tax revenues and spending.

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

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

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

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

[58]Note that $\sum_{m \in \mathcal{N}} \pi_{nmjt} + \pi^f_{njt} = 1$ holds.

[59]Firm i's revenues in Home and Foreign are proportional to operating profits in Home and Foreign: $r(\phi_{it})^d = \sigma \pi^d(T^*_{it}; \phi_{it})$ and $r(\phi_{it})^x = \sigma \pi^x(T^*_{it}; \phi_{it})$, where t^*_{it} is firm i's optimal adoption decisions. The total revenue $r(\phi_{it})$ is the sum of $r(\phi_{it})^x$ and $r(\phi_{it})^d$.

Goods market and labor market clearing conditions imply that trade is balanced.

5.5 Dynamic Equilibrium

Laws of Motion of Dynamic State Variables. 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.

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

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

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 Allocation)*** *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).*

The equilibrium conditions (i)-(vi) determine the static equilibrium allocation each period and the equilibrium conditions (vii) and (viii) determine the laws of motion for the dynamic state variables $\{\lambda^T_{njt}, L_{nt}\}$.

5.6 Analytical Results: Multiple Steady States

In this subsection, we analytically 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, firms' exogenous productivity follows unbounded Pareto distribution with a normalized location parameter, and firm mass is normalized to be one. Assumptions 1(iv) and (v) are non-trivial. 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 the adoption does not operate in this environment, and the dynamic complementarity only comes from fixed adoption costs in units of final goods.[60] Assumption (v) is a sufficient condition that guarantees a unique static equilibrium each period.[61]

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[62]:

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

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

The equilibrium share is characterized by the cutoff productivity which is determined at where marginal adopters' net gains from the adoption is equal to zero. Similarly, the time-invariant steady state share of adopters satisfies $\lambda^{T*} = \lambda^{T*}(\lambda^{T*}; \eta, \delta)$ and is determined at $\lambda^{T*} = \lambda_t^{T*} = \lambda_{t-1}^{T*}$.

Given any initial share of adopters $\lambda_{t_0}^T$, this economy has a unique deterministic equilibrium path to the steady state due to Assumption 1(v). Because a static equilibrium is unique each period, there exists a unique sequence of static equilibrium which forms a unique deterministic dynamic

[60] With only one region, each firm's increase in productivity due to the spillover is exactly canceled out by competition forces due to other firms' increases in productivity under the CES demand structure. Therefore, the overall increase in productivity through the spillover does not change the relative market size of each firm, nullifying the complementarity between market size and gains from the adoption. However, because fixed adoption costs are in units of final goods, the spillover from the previous period lowers fixed adoption costs today, which further incentivizes more firms to adopt technology today 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-1}^T, because overall productivity increase induced by the spillover increases labor demand, which in turn increases wages and the total fixed adoption cost $w_t F^T$. As long as final goods are used for fixed adoption costs, the model generates dynamic complementarity. This is formally stated in Appendix Section C.2.3.

[61] When a fixed adoption cost is in units of final goods, and $\sigma \leq 2$, multiple static equilibria can arise each period regardless of the existence of the spillover because firms do not take aggregate price index into account when making adoption decisions. When there is a sufficiently large share of adopters, the aggregate price index becomes lower and this, in turn, decreases the adoption costs and vice versa. This degree of responsiveness of the price index to a share of adopters decreases in the elasticity of substitution σ. When σ is sufficiently low, two static 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 these multiple static equilibria. These multiple static equilibria are studied by Matsuyama (1995) and Buera et al. (2021). Our model differs from their models because our multiple long-run steady states are generated due to 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).

[62] λ_t^{T*} is bounded by 1. $\lambda_t^{T*}(\lambda_{t-1}^T; \eta, \delta) = 1$ when $A(\hat{\lambda}_t^T(\lambda_{t-1}^T; \eta, \delta))^{2-\sigma} f(\lambda_{t-1}^T) \frac{\eta^{\sigma-1}-1}{\sigma F^T} \geq 1$.

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 increases λ_t^{T*} by increasing marginal adopter's net gains.[63] δ increases λ_t^{T*} by magnifying the dynamic complementarity. Most importantly, we show that in this economy, multiple steady can arise due to the dynamic complementarity. If multiple steady states exist, depending on the initial condition $\lambda_{t_0}^T$, the economy may converge to different steady states that are Pareto-ranked based on the steady state share of adopters. These results are summarized in Proposition 1.

Proposition 1. *Under Assumption 1,*

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

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

(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 $\eta \in [\underline{\eta},\bar{\eta}]$;

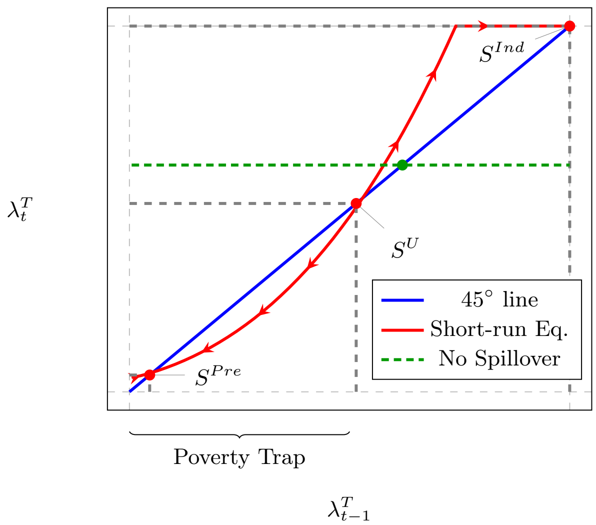
and (v) (Pareto-Ranked) If there exist multiple steady states, these steady states can be Pareto-ranked based on the equilibrium share of adopters.

The case of multiple steady states is illustrated in Panel A of Figure 3 where there are three different steady state with two basins of attraction.[64] The red locus is defined by Equation (5.21), where each point of 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 in the next period $t+1$, given λ_t^{T*}, λ_{t+1}^{T*} is determined, and so on. Therefore, the equilibrium moves along the red locus as time passes. The steady-state is determined at 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 call the pre-industrialized, the unstable, and the industrialized steady states, respectively.

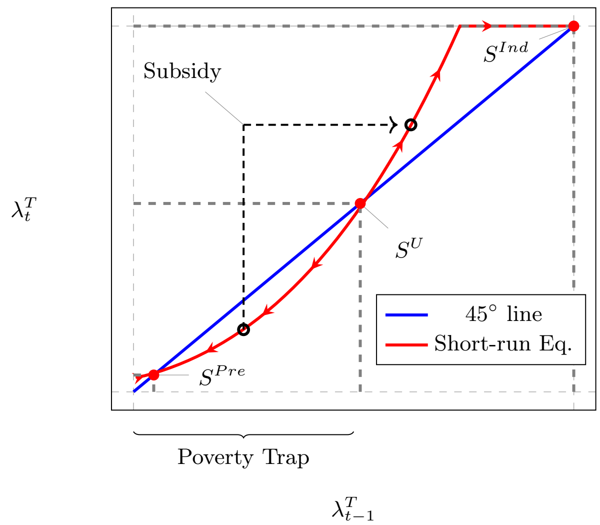
Because the adoption increases firms' productivity, these steady states can be Pareto-ranked depending on the steady state share of adopters. At S^{Ind}, all firms adopt technology, and at S^{Pre} there is a smaller share of adopters than the other two, so S^{Ind} Pareto-dominates the other two

[63] 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 the adoption internalized by the adopter. The term $A(\lambda_t^T)^{2-\sigma} = A(\lambda_t^T)^{1-\sigma} \times A(\lambda_t^T)$ captures the two composite general equilibrium effects of the direct productivity gains that work in the opposite directions on marginal adopter's incentives for the adoption. First, $A(\lambda_t^T)^{1-\sigma}$ captures competition effects, which decreases in λ_t^T. As more firms adopt technology (an increase in λ_t^T), competitors' productivity increases, which in turn intensifies competition across firms and decreases a firm's incentives for adoption due to increased competition (or decreased market size). The second general equilibrium effect is that more adoption decreases the price of final goods and, therefore a fixed adoption cost (Assumption 1(iv)). Assumption 1(v) ensures that the competition force dominates this second general equilibrium effect.

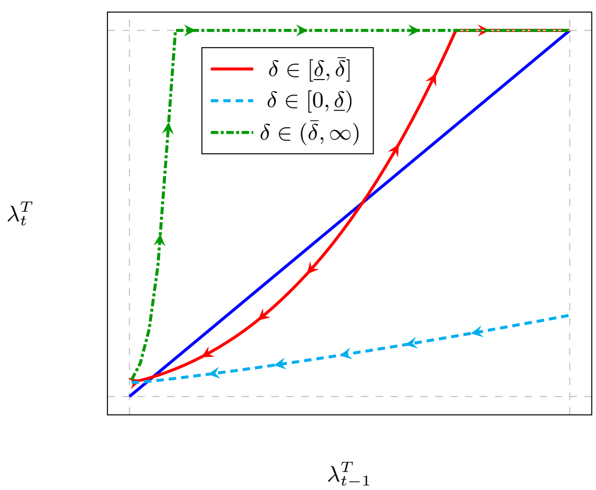
[64] In this economy, there are at most three multiple steady states because of the functional form assumption imposed on the spillover: $f(\lambda_{t-1}^T) = exp(\delta\lambda_{t-1}^T)$. The imposed spillover functional form makes λ_t^T to be strictly convex in λ_{t-1}^T so that the red locus in Figure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



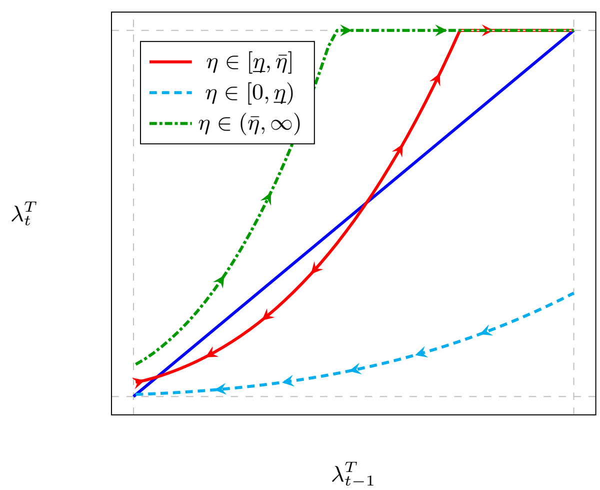
Panel A. Multiple Steady States and Nonlinearity



Panel B. The Role of Adoption Subsidies



Panel C. Comparative Statistics: δ



Panel D. Comparative Statistics: η

Figure 3. Multiple Steady States and Comparative Statistics

Notes. Panel A illustrates the case when multiple steady states arise and the role of nonlinearity of the short-run equilibrium condition. Panel B illustrates the role of adoption sudsidies. Panels C and D illustrate the cmpoarative statistics of δ and η.

steady states and S^{Pre} is Pareto-dominated by the other two.[65] 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 makes the locus intersect with the 45-degree line multiple period t 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 pinned down by the green dashed horizontal line and the 45-degree line.[66]

For the initial conditions given by $\lambda^T_{njt_0} \geq S^U$, the economy converges to S^{Pre}, and for $\lambda^T_{njt_0} \geq S^U$, the economy converges to S^{Ind}. Because firms are not internalizing the spillover, if an economy is locked into the region $\lambda^T_{njt_0} \geq S^U$, 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).

Permanent Effects of Temporary Subsidies and the Multiple Steady States. When the multiple steady states exist, temporary subsidies for technology adoption in the initial period can have permanent effects by moving an economy that is initially in the poverty trap to a new transition path that converges to the 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 Korea's pattern of industrialization toward heavy manufacturing sectors and the temporary policy between 1973 and 1979.

In this model, only the multiple steady states can rationalize the permanent effects of temporary subsidies.[67] When there is only a unique steady state, subsidies temporarily shift the short-run equilibrium curve while subsidies are provided, but the shifted curve moves back to the original position after the end of the subsidies. An economy converges to the original steady state.[68]

Comparative Statistics. Proposition 1(iv) implies that the multiple steady states arise only for medium ranges of $\delta \in [\underline{\delta}, \bar{\delta}]$ and $\eta \in [\underline{\eta}, \bar{\eta}]$, that is when the spillover or the 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 a sufficient degree of nonlinearity of the short-run equilibrium locus, making the locus intersect with the 45-degree line only once. This is graphically illustrated in Panels C and D of Figure 3.

[65] The fact that all firms adopt technology at S^{Ind} is the artifact of that λ^T_t is strictly convex in λ^T_{t-1} which comes from the imposed functional form assumption of the spillover.

[66] 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.

[67] Even if there exists a unique steady state, there is room for policy interventions because of the externality. However, with a unique steady state, these policy interventions should be permanently implemented each period to have permanent effects. This point is graphically illustrated in Appendix Figure C1.

[68] Similarly, Kline and Moretti (2014) studied the Tennessee Valley Authority program in the US and did not detect nonlinearities in the agglomeration elasticity, so they concluded that the program had limited indirect gains through agglomeration.

The comparative statistics can give one potential explanation on why the Korean economy experienced remarkable transformation toward heavy manufacturing sectors but not other developing countries, although these developing countries benchmarked policies of Korea. Both η and δ can be country-specific and depends on various features of each country.[69] Unlike other developing countries, Korea could have been a special case where its values of η and δ were in a range that generated the existence of multiple steady states.

6 Taking the Model to the Data

In our quantitative exercises, we aggregate sectors into four: commodity, light manufacturing, heavy manufacturing, and service sectors. Technology adoption is only available in the heavy manufacturing sector, and the service sector is non-tradable across regions and countries. We also aggregate the data to 42 regions.[70] One period in the model corresponds to 4 years in the data.

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:

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

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

6.1 Externally Calibrated Parameters

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

[69] η and δ are generally related to the absorptive capacity of new technology and degree of barriers to knowledge diffusion, respectively. Countries with lower amounts of skilled labor endowments and higher language barriers may have lower values of η and δ. When compared to other developing countries, Korea had higher amounts of skilled labor endowment and level of cultural proximity (Rodrik, 1995), which can make Korea have higher values of η and δ.

[70] We aggregate up to 42 regions so that each region has at least two firms in each sector based on the administrative divisions in the 1970s and electoral districts.

can derive the following regression model:

$$\log Sales_{it} = (\sigma-1)\log(\eta)T_{it} + \underbrace{\delta\lambda^T_{njt} + \log\Big(\sum_{m\in\mathcal{N}} \tau_{nmj} P^{\sigma-1}_{mjt} E_{mjt} + \tau^x_{nj} D^f_{jt}\Big)}_{\text{Absorbed out by exactly matching on region and sector}} + (\sigma-1)\log\phi_{it}, \quad (6.1)$$

which can be mapped to our "winners vs. losers" specification (Equation (4.1)). By exactly matching on region and sector, we can absorb out the spillover, a unit cost of production, and the market size that are common across firms within region-sector.[71] Exogenous cancellations by foreign firms can be interpreted as a shock to the fixed adoption cost F^T_j in our model framework that is orthogonal to firms' productivity $\log\phi_{it}$. We calibrate η using the average estimates of $\{\beta^{diff}_\tau\}^7_{\tau=1}$ in Equation (4.1) and set $\eta = exp(0.51)/(\sigma-1)$.[72]

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

$$\log Sales_{it} = (\sigma-1)\delta\lambda^T_{njt} + \log\Big(\sum_{m\in\mathcal{N}} \tau_{nmj} P^{\sigma-1}_{mjt} E_{mjt} + \tau^x_{nj} D^f_j\Big) + (\sigma-1)\log\phi_{it}. \quad (6.2)$$

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

Spatial Mobility $\{\nu, d_{mn}\}$. We parametrize migration costs as a function of the geographic distance: $d_{nm} = dist^\zeta_{nm}\epsilon^d_{nm}$ where $dist_{nm}$ is the geographical distance between region n and m and ϵ^d_{nm} is a residual that is not explained by the distance. ν is set to be 2, which is the estimate from Peters (2019). Using Equation (5.10), we can derive a gravity equation for migration flows. ζ is externally calibrated by estimating this gravity equation. Using migration flows between 1990 and 1995, we run the following regression model:

$$\log\mu_{nm,1990-1995} = -\nu\zeta\log dist_{nm} + \delta_m + \delta_n + \epsilon^d_{mn}, \quad (6.3)$$

[71]More precisely, we allow for variations in the spillover in our reduced-form regression model whereas the spillover is assumed to be common across firms within region in our model. We can consistently estimate $(\sigma-1)\log(\eta)$ when the cancellations are exogenous to both $\ln\phi_{it}$ and the residual of the spillover measure net of pair-common effects $(Spill_{injt} - D^\tau_{pt})$.

[72]0.51 is based on the pooled diff-in-diffs estimates reported in column (1) of Appendix Table D3.

where $\mu_{nm,1990-1995}$ is a share of migrants from region n to region m; and δ_n and δ_m are region fixed effects.[73] To address attenuation bias arising from statistical zeros in the gravity models, we estimate the above equation via pseudo poisson maximum likelihood (PPML) (Silva and Tenreyro, 2006). Under the assumed value for ν, we obtain the value for ζ from the estimated coefficients. The gravity estimate implies that $\hat{\zeta} = 1.39/\nu$.[74]

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 take the distance elasticity estimate from Monte et al. (2018) and set $\xi = 1.29/(\sigma - 1)$.[75]

For international trade costs, we assume that to export or import from Foreign, a firm has to ship its products to the nearest port and then additionally pay both variable and fixed international trade costs at the nearest port. Under this assumption, international trade costs can be parametrized as $\tau^x = \tilde{\tau}^x \times (dist^{port}_n)^{\xi}$.$\tilde{\tau}^x$ are variable costs that are incurred at the port. We set $\tilde{\tau}^x = 1.7$ following Anderson and Van Wincoop (2004). $(dist^{port}_n)^{\xi}$ is a variable cost associated with shipping between region n to the nearest port, 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.[76]

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).[77] We set M_{nj} to be proportional to the GDP share of each region-sector and set $\sum_{n \in \mathcal{N}} M_{nj} = 1$ following Chaney (2008).[78] 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 in 1972.

[73] The estimation procedure is described in detail in Appendix Section E.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 the data availability, regions are aggregated up to 35 regions. $\mu_{nm,1990-1995}$ is obtained as the total number of migrants moving from region n to region m between 1990 and 1995 divided by the total population of region n in 1990. When computing the total number of population and migrants, we restrict our sample age between 20 and 55.

[74] We find statistically significant results at 1% where we two-way cluster errors at origin and destination levels. The OLS estimates of Equation (6.3) is 1.30 which is similar to 1.39 obtained from PPML. See Appendix Table E2 for the detailed gravity estimates of migration flows.

[75] Using internal trade flows within the US, Monte et al. (2018) estimate a distance elasticity of -1.29. We calibrate ξ to be $1.29/(\sigma - 1)$.

[76] When computing the nearest distance to ports, we use seven main ports in Korea: Busan, Incheon, Gunsan, Guje, Pohang, Ulsan, and Yeosu.

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

[78] Chaney (2008) assumes that the total mass of firm of each country is proportional to its total value added in the multi-country setting.

6.2 Internally Calibrated Parameters

$\mathbf{\Theta}^M = \{F_j^x, F_j^T, \kappa\}$, $\{s_{njt}\}$, and $\mathbf{\Psi}_t = \{\phi_{njt}^{min}, V_{nt}, D_{jt}^f, c_{jt}^f\}$ are calibrated using the method of moments. Our calibration procedure requires moments from firm-level micro data and a set of cross-sectional aggregate variables in 1972, 1976, and 1980 which cover the period when the subsidies were provided between 1973 and 1979: region-sector level gross output $\{R_{njt}\}$, regional population distribution $\{L_{nt}\}$, aggregate export and import shares, initial shares of adopters $\{\lambda_{nj-1}^T\}$ and initial population distribution $\{L_{n,-1}\}$ that are taken as given when solving the model for $t = 1$.[79] Appendix Section E.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 $\bar{s}$ by solving the following constrained minimization problem:

$$\{\hat{\mathbf{\Theta}}^M, \hat{\bar{s}}\} \equiv \arg\min_{\mathbf{\Theta}^M, \bar{s}} \{L(\mathbf{\Theta}^M, \bar{s})\} = \underbrace{(\bar{m} - m(\mathbf{\Theta}^M, \bar{s}, \mathbf{\Psi}_t))'\mathbf{W}(\bar{m} - m(\mathbf{\Theta}^M, \bar{s}, \mathbf{\Psi}_t))}_{\text{Micro moments}}$$

$$\text{subject to} \qquad \underbrace{\mathbf{C}(\mathbf{\Theta}^M, \bar{s}, \mathbf{\Psi}_t) = \mathbf{C}}_{\text{Aggregate data}}, \tag{6.4}$$

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

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, corresponding to 1976 and 1980 in the data. Given the lack of information on the distribution of subsidies across regions, we assume that the government provides the same subsidy level $\bar{s}$ across regions in $t = 2, 3$:

$$s_{njt} = \begin{cases} \bar{s} & \text{if} \quad t \in \{2, 3\}, \quad \forall n \in \mathcal{N}, \quad \forall j \in \mathcal{J}^T \\ 0 & \text{otherwise.} \end{cases} \tag{6.5}$$

Despite the lack of data on subsidies, with the above parametrization on subsidies, we can identify $\bar{s}$ using the model structure and the reduced-form estimates that measure direct and spillover benefits from the adoption. The intuition is that given information on the benefits from the adoption (direct productivity gains and the spillover), increases in shares of adopters in 1976 or 1980 relative to 1972 are attributable to the policy. The following proposition summarizes this result.

[79] The time interval in the model is 4 years, so 1972, 1976, and 1980 correspond to $t = 1$, $t = 2$, and $t = 3$ of the model periods.

Proposition 2. *(Identifying Moment for Subsidies) Assume (a) unbounded Pareto distributed exogenous firm productivity ($\kappa \to \infty$), (b) free trade ($\tau_{nmj} = 1$ and $\tau^x_{nj} = 1$), and (c) free mobility of labor ($\theta \to \infty$ and $d_{nm} = 1$), and (d) amenity is equalized across regions ($V_{nt} = \bar{V}, \forall n \in \mathcal{N}$). Consider the following regression model for sectors in which the adoption is available:*

$$\ln \lambda^T_{njt} - \theta\delta\lambda^T_{njt-1} = \alpha + \beta^{policy} \times D^{policy}_t + \epsilon_{njt},$$

D^{policy}_t is a dummy variable for periods when subsidies are provided; and $\mathbf{X}_{nj1}$ is region-sector level controls measured at $t = 1$. Then, given values of η, δ, σ, and θ, when $\mathbb{E}[\ln \phi^{min}_{njt} | D^{policy}_t] = 0$ holds,

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

and the estimated coefficient of the policy dummy variable uniquely identifies $\bar{s}$.

Proposition shows that a sudden increase in shares of adopters in 1980 captured by β^{policy} is informative on subsidies, when subsidies are exogenous to region-sector exogenous productivity: $\mathbb{E}[\ln \phi^{min}_{njt} | D^{policy}_t] = 0$. The proposition motivates our approach. Since both in the model and in the data, the simplifying assumptions of Proposition 2 do not hold, we pin down the subsidy level by indirect inference. In particular, using both actual and model-generated data, we estimate the following regression for the heavy manufacturing sector in 1972 and 1980 via PPML to incorporate zeros[80]:

$$\ln \lambda^T_{njt} = \alpha + \beta^{policy} \times D^{policy}_t + \beta_1 \lambda^T_{njt-1} + \epsilon_{njt}. \tag{6.6}$$

Then, we use the estimated β^{policy} from the actual data as the identifying moment for $\bar{s}$ and fit the estimated β^{policy} from the model-generated data to this moment (Nakamura and Steinsson, 2018). Unlike the regression model in Proposition 2, we control previous shares rather than subtract them from current shares in a dependent variable.[81] The estimated coefficients for β^{policy} and β_1 are 0.65 and 5.62, and statistically significant at the 1% level.[82]

Objective Function: Micro Moments, $\{\boldsymbol{\Theta}^M, \bar{s}\}$. We pin down $\boldsymbol{\Theta}^M = \{F^T_j, F^x_j, \kappa\}$ 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 in 1972 and 1980. We identify κ using

[80] More specifically, we run this regression for shares of adopters in the heavy manufacturing sector in 42 regions for years 1972 and 1980, so we use 84 samples in total. Note that we assumed that (i) technology adoption is only available for heavy manufacturing firms and (ii) common subsidies are provided across regions, and (iii) we aggregated heavy manufacturing sectors into one sector when taking the model to the data, we cannot control 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 additional fixed effects.

[81] This is because PPML is not defined for dependent variables with negative values and substracting the previous shares from the current shares with zero values generates observations with dependent variables that take negative values.

[82] 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 results are reported in Appendix Table E2.

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 bound becomes narrower, the cutoff adoption productivity becomes larger than the Pareto upper bound $\kappa\phi_{njt}^{min}$ for some regions, resulting in zero adoption in these regions. Fixed export costs of the light and heavy manufacturing sectors F_j^x are calibrated to match the median shares of exporters across regions in 1972.[83] Because we do not have detailed data on commodity sector firms, we set a fixed export cost of commodity sectors to be the same as that of the light manufacturing sector.

Constraints: Aggregate Data, $\mathbf{\Psi}_t$. The constraints in Equation (6.4) identify geographic fundamentals $\mathbf{\Psi}_t$. We impose the constraints in a way that shares of gross output at region-sector level, 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 region-sector level data in 1972, 1976, and 1980. The number of the constraints is the same with the dimension of the geographic fundamentals.[84] Therefore, for any given parameters $\mathbf{\Theta}^M$ and subsidy rate $\bar{s}$, geographic fundamentals are exactly identified by these constraints and there exist a set of geographic fundamentals that rationalize the data. Because geographic fundamentals are exactly identified, we can identify average productivity $\bar{\phi}_{nj}^{avg}$, D_{jt}^f, and c_{jt}^f, following the model-inversion logic from (Allen and Arkolakis, 2014).[85] However, we cannot identify how much portion of $\bar{\phi}_{nj}^{avg}$ is attributable to natural advantage ϕ_{nj}^{min}, shares of adopters λ_{nj}^T, or subsidies $\bar{s}$ from aggregate data alone. To isolate ϕ_{nj}^{min}, λ_{nj}^T, and $\bar{s}$ separately from $\bar{\phi}_{nj}^{avg}$, we need information on the fixed adoption cost F_j^T and subsidy rate $\bar{s}$ from the micro moments.

6.3 Calibration Results and Model Fit

Table 3 presents the summary of our calibration strategy and the values of the externally and internally calibrated parameters. The estimated adoption cost is 8 times larger than the fixed export cost. The estimated subsidy rate is 0.14, which indicates that adopters are subsidized with 14% of input expenditures. Table 4 reports the model fit. The data moments are well-approximated in the model. The model also matches a variety of non-targeted moments.

[83] Our firm balance sheet data has information on exports. However, many of observations were missing. Given that export data is 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.

[84] 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 where the model is exactly fitted to the region-sector data; $|\mathcal{N}| \times |\mathcal{J}|$ are the number of ϕ_{nj}^{min}; $|\mathcal{J}^x|$ is the number of D_j^f and c_j^f; and $|\mathcal{N}|$ is the number of V_n.

[85] By fitting the input-output tables, we can only identify relative productivity differences across regions and sectors, but cannot identify aggregate shifters of productivity. Therefore, while fitting gross output shares at region-sector level, we normalize ϕ_{njt}^{min} of one region-sector to be 1 for each period. This is of less concern because our interest is the comparison between the baseline economy to the counterfactual economy, which differences out the common aggregate components.

Table 3: Calibration Strategy

Parameters			Identification / Moments
	Description	Value	
	External calibration		
	Structural parameters		
η	Direct productivity gains	1.3	"Winners vs. Losers", Table 2 col. 1 $(\sigma-1)\log(\eta)=0.51$
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 (2019)
ζ	Migration cost, $d_{mn}=(dist_{nm})^{\zeta}$	0.78	Gravity estimates
ξ	Internal trade cost, $\tau_{nmj}=(dist_{nm})^{\xi}$	0.43	Monte et al. (2018)
$\bar{\tau}^x$	International trade costs, $\tau^x_{nj}=\bar{\tau}^x(dist^{port}_{nm})^{\xi}$	1.7	Anderson and Van Wincoop (2004)
α_j	Preferences		IO table, 1972
γ^k_j	Production		IO table, 1972
M_{nj}	Exogenous firm mass		Value-added, 1972 (Chaney, 2008)
	Internal calibration: Method of moment		
	Structural parameters		
F^T_j	Fixed adoption cost	0.20	Share of adopters, heavy mfg.
F^x_j	Fixed export cost, commodity & light mfg.	0.20	Share of exporters, light mfg.
F^x_j	Fixed export cost, heavy mfg.	1.65	Share of exporters, heavy mfg.
κ	Pareto upper bound	6.53	# of regions with zero adoption
	Geographical fundamentals		
ϕ^{min}_{nj}	Natural advantage (Pareto lower bound)		Region-sector sales dist., 1972, 1976, 1980
D^f_j	Foreign market size		Sectoral export intensity, 1972, 1976, 1980
c^f_j	Foreign imported input price		Sectoral import intensity, 1972, 1976, 1980
V_{nt}	Amenity		Pop. dist., 1972, 1976, 1980
	Subsidy		
$\bar{s}$	Subsidy rate	0.11	Identifying moment $\hat{\beta}^{policy}$, Equation (6.6)

Notes. This table reports calibrated objects of the model, their values, and their identifying moments. The calibration procedure is described in Appendix Section E.2 in more detail.

Table 4: Model Fit

Moment	Model	Data
Identifying moment $\hat{\beta}^{policy}$, Equation (6.6)	0.65	0.83
med. exporter share in 1972, light mfg.	0.22	0.21
med. exporter share in 1972, heavy mfg.	0.14	0.18
med. adopter share in 1972	0.06	0.07
med. adopter share in 1982	0.12	0.18
Share of zero adoption regions in 1972	0.59	0.53
Share of zero adoption regions in 1982	0.83	0.93

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

7 The Aggregate and Regional Effects of the Temporary Adoption Subsidy

In this section, we ask how the aggregate and regional pattern of industrialization in Korea would have evolved differently if the temporary subsidies had not been provided. We compare the baseline economy with the subsidies to the counterfactual economy without the subsidies. Unlike the simplified model where there are at most three steady states in Section 5.6, the full quantitative model potentially admits a larger number of steady states. Which steady state will be reached in the long-run is of a 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. The red dotted and blue dashed lines plot the equilibrium path when subsidies were provided and not provided, respectively. The solid green line depicts the actual path of the Korean economy observed in the data, which is computed from the input-output tables. In Panels A, B, C, and D, we compare the heavy manufacturing sector shares in GDP, employment, and export, and the light manufacturing sector shares in export. Had not been temporary adoption subsidies provided, Korea's pattern of industrialization and comparative advantage would have evolved differently. Heavy manufacturing GDP, employment, and export shares would have decreased by 15%, 3%, and 22.5% permanently.

Our calibration strategy only fits the cross-section data in 1972, 1976, and 1980 and does not fit the evolution of variables after 1980. Therefore, our model does not explain evolution of the employment or export shares after 1980 well. However, our model fits the evolution of GDP share of the heavy manufacturing sector quite well even after 1980 (Panel A of Figure 4), which is a non-targeted moment.

Panel A of Figure 5 reports the average productivity changes across regions in the steady state,

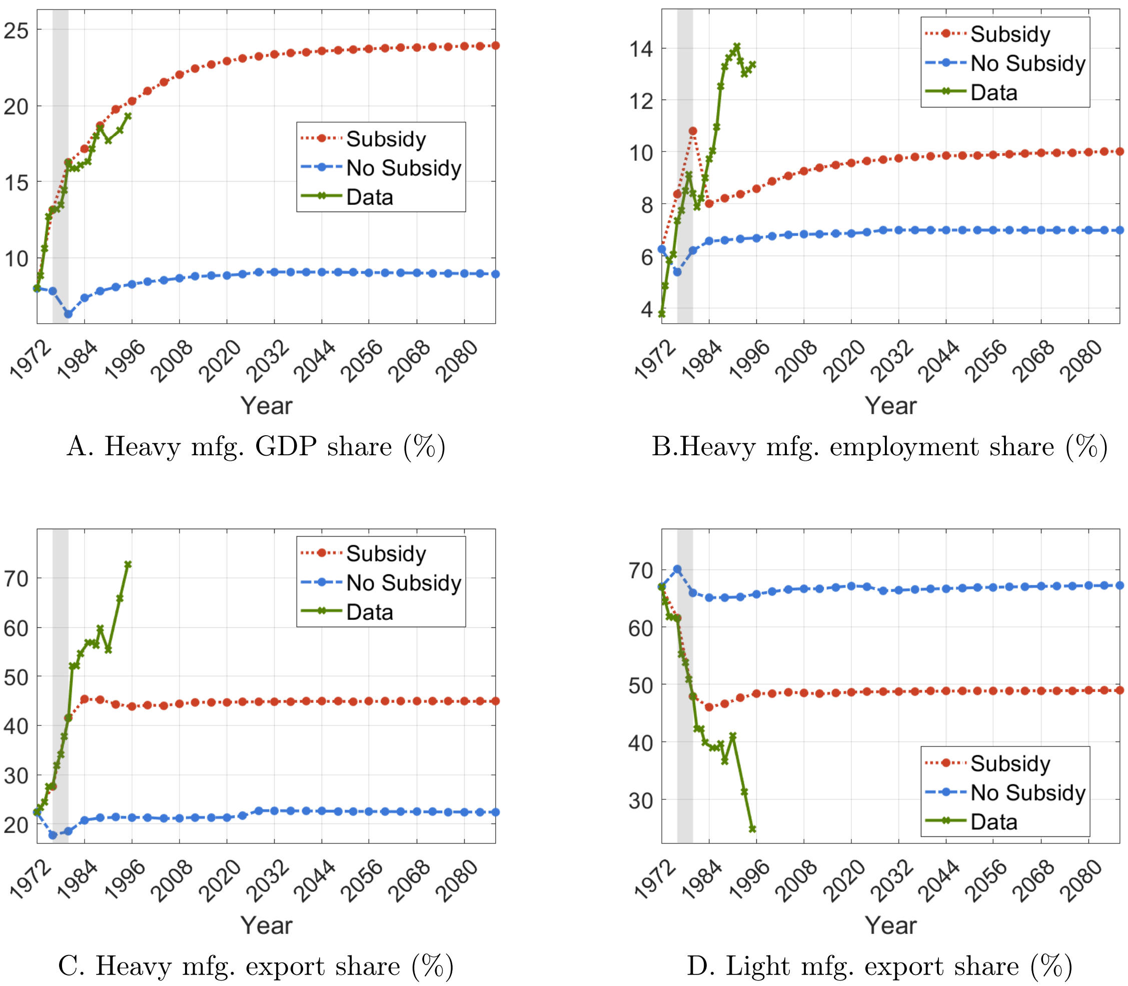


Figure 4. Counterfactual results. Industrialization pattern in South Korea when the temporary subsidies had not been provided.

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

where the average productivity is defined as $M_{nj}[\int_{\omega\in\Omega_{nj}} z_{it}(\omega)^{\sigma-1}d\omega]^{1/(\sigma-1)}$.[86] X- and Y-axis report the average heavy manufacturing sector productivity of each region under the baseline and counterfactual economies, respectively. Regions that experienced productivity increases due to subsidy-induced technology adoption are colored red. The figure shows that the aggregate industrialization

[86]Because M_{nj} and ϕ_{njt}^{min} are not separately identifiable under the fixed entry, $M_{nj}[\int_{\omega\in\Omega_{njt}} z(\omega)^{\sigma-1}d\omega]^{1/(\sigma-1)}$ can be considered as the average productivity when $M_{nj} = 1, \forall n, j$.

pattern of the baseline economy was driven by large productivity increases of a few regions due to subsidy-induced technology adoption.

Panels B and C of Figure 5 plot the regional and aggregate welfare gains in the baseline economy over the counterfactual economy. On average, the regional and aggregate welfare is 10% permanently higher once the economies reach steady states. Large productivity increases of of a few regions and their specialization into the heavy manufacturing sector led to increases in welfare across all regions through trade linkages. Also, note at the beginning of the implementation of the subsidies, the aggregate welfare first decreases in the short run compared to the counterfactual economy because calibrated subsidies are not optimally designed.[87]

Roundabout Production. We find that a roundabout production structure plays an important role in generating multiple steady states and permanent effects of subsidies. A roundabout production structure amplifies the impact of subsidies through cost and demand linkages (Krugman and Venables, 1995). Because of these linkages, complementarity between firm-scale and gains from technology adoption causes more firms to adopt technology. We do the same exercise with a new production structure where labor is the only factor of production and there are no intermediate inputs. The results are reported in Appendix Figure E2. While holding other parameters, subsidies, and geographic fundamentals constant, we find that subsidies do not generate multiple steady states.

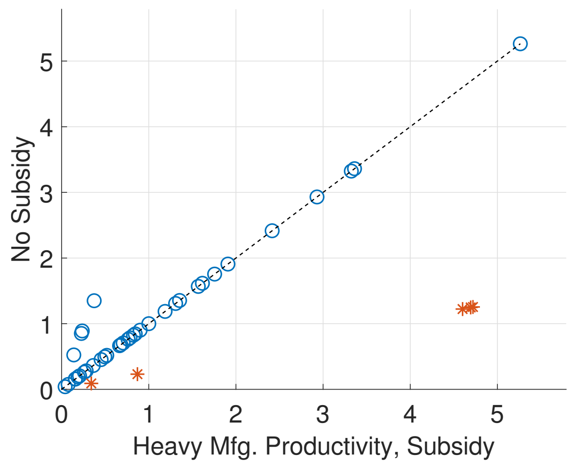
Role of Geography: Foreign Market Size and Migration Costs. We examine how geography interacts with the effects of the temporary adoption subsidies. While comparing the baseline and counterfactual economies, we change geographical features of the Korean economy and examine how its long-term effects differ from the baseline comparison results in Table 4. We specifically consider how the foreign market size and migration costs interact with the policy. In the 1960s and 1970s, Korea experienced large expansion of exports.[88] Also, there were dramatic increases in migration flows from rural to urban areas during industrialization, which is a common feature during industrialization.[89] We quantitatively find big impacts of the joint implications between Foreign market size and lower migration costs and temporary subsidies.

We examine how complementarity between foreign market size and productivity gains from tech-

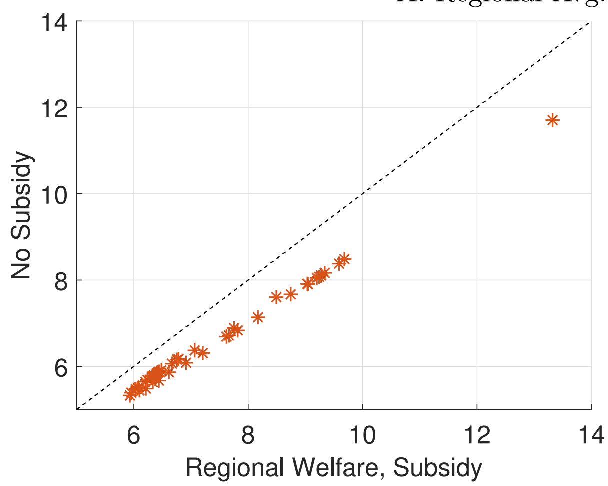
[87]Analyzing the optimal subsidy in this economy is out of the focus of this paper. For the optimal policy, see Bartelme et al. (2020), Fajgelbaum and Gaubert (2020) and Lashkaripour and Lugovskyy (2020) in the static setting.

[88]Dramatic rapid increases in Korea's exports were outcomes of both 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 for tariffs on imported inputs for exports (Connolly and Yi, 2015). Korea also devalued its overvalued currency in 1964, which boosted its exports (Irwin, 2021). Also, the US demands of foreign imports increased dramatically between 1960 to 1980. During these periods, shares of US imports to the total gross national product rose from 6 to 22%.

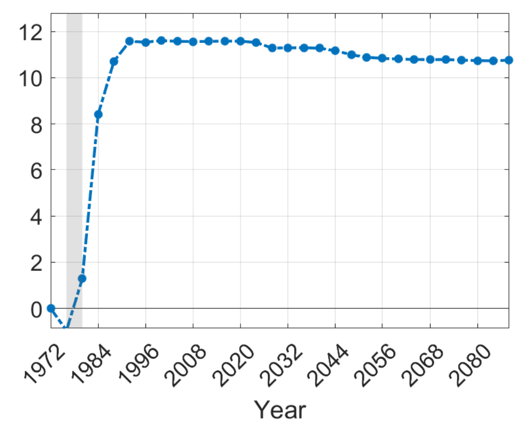
[89]According to World Development Indicators (World Bank), Korea's shares of rural population decreased from 60 to 40% between 1970 and 1982. Annual shares of migrants to the total population increased from 12.6% in 1970 to 21.9% in 1982. Many developing countries similarly underwent rapid transitions from rural to urban during industrialization in the twentieth century. See Table 1 of Lucas (2004).



A. Regional Avg. Productivity Gain



B. Regional Welfare Gain (%)



C. Aggregate Welfare Gain (%)

Figure 5. Counterfactual Results. Aggregate Welfare and Regional Specialization Pattern

Notes. This figure plots the counterfactual results. Panels A reports the ratio of the aggregate welfare of the baseline economy to that of the counterfactual economy, where the aggregate welfare is defined in Equation (5.13). Panel B reports the regionalization specialization pattern where the X-axis plots each region's overall productivity defined as $M_{nj}[\int_{\omega\in\Omega_{njt}} z(\omega)^{(\sigma-1)}]^{1/(\sigma-1)}$.

nology adoption affect firms' technology adoption decisions while the subsidies are provided and its long-term consequences. We decreased the Foreign market size so that export shares in the heavy manufacturing sector in 1972 was 6.6% which is the level in 1966 (instead of the original 22% in 1972). The results are reported in Appendix Figure E4. The gap of the heavy manufacturing GDP shares between two steady states is about 5%, which is 10% smaller than the counterfactual results (15%) under the 1972 Foreign market size level. These quantitative results provide suggestive evidence that Korea's export-promotion policies might have played an important role for shaping Korea's economic development jointly with subsidies.

We next examine joint implications between migration costs and firms' adoption decisions. We set 10% higher migration costs. Because of higher migration costs, people will be moving less toward regions with higher productivity induced by the adoption, which increases wages and costs of production. Due to the complementarity between firm scale and gains from the adoption, fewer firms will adopt technology with higher migration costs. These results are reported in Appendix Figure E3. The gap of heavy manufacturing GDP shares between two steady states is around 9%, which is 6% smaller than the counterfactual results (15%) under the baseline calibrated values.

Comparative Statistics. We conduct the comparative statistics of δ and η to examine how the particularly chosen parameters drive these multiple steady states results. In Appendix Figure E1, 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 a large impact of technology adoption on late industrialization in Korea both empirically and quantitatively. Our finding confirms a widely held belief by economists that technology adoption can foster the economic development of developing countries. We empirically find that technology adoption not only directly benefited adopters but also had large local spillover effects. Based on these empirical findings, we build a dynamic spatial model to conduct the counterfactual analysis on the Korean government temporary subsidies for technology adoption in the heavy manufacturing sectors. Using the quantitative model calibrated to the micro data, we show that the temporary adoption subsidies can have a permanently large impact on the economy by moving an economy to a new transition path that converges to an alternative more-industrialized steady state.

We believe that our empirical findings are important for two reasons. First, they highlight that externalities technologies diffuse slowly to developing countries and why appropriate policy intervention 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 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 uncertainty about future technology, and forward-looking technology adoption decisions by agents. Incorporating more realistic assumptions on agents' beliefs in the economy and how these beliefs interact with history 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 w28059*, 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," *Mimeo*, 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?," *Mimeo*, 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, R. 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.

Balassa, Bela, "Trade Liberalisation and "Revealed" Comparative Advantage," *The Manchester School*, 1965, *33* (2), 99–123.

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 w26193*, 2020.

— , **Ting Lan, and Andrei Levchenko**, "Specialization, Market Access and Real Income," *NBER Working Paper w28274*, 2021.

Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts, "Does Management Matter? Evidence from India," *Quarterly Journal of Economics*, 2013, *128* (1).

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.

Buera, Francisco J, Hugo Hopenhayn, Yongseok Shin, and Nicholas Trachter, "Big Push in Distorted Economies," *NBER Working Paper w28561*, 2021.

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 Maria Santacreu, "Knowledge Diffusion, Trade, and Innovation across Countries and Sectors," *American Economic Journal: Macroeconomics*, 2021.

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 w29263*, 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.

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," *Mimeo*, 2020.

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

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 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.

Enos, J. L. and W. H. 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.

Feenstra, Robert C., Robert E Lipsey, Haiyan Deng, Alyson C Ma, and Hengyong Mo, "World Trade Flows: 1962-2000," *NBER Working Paper w11040*, 2005.

Gerschenkron, Alexander, *Economic Backwardness in Historical Perspective*, Cambridge, MA: Harvard University Press, 1962.

Giorcelli, Michela and Bo Li, "Technology Transfer and Early Industrial Development: Evidence from the Sino-Soviet Alliance," *Mimeo*, 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.

Hanson, Gordon H., Nelson Lind, and Marc-Andreas Muendler, "The Dynamics of Comparative Advantage," *NBER Working Paper w21753*, 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.

Head, Keith and Thierry Mayer, *Gravity Equations: Workhorse, Toolkit, and Cookbook*, Vol. 4, Elsevier B.V., 2014.

Helpman, Elhanan, Marc 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 w29299*, 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 w27503*, 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.

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.

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 w29252*, 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," *Mimeo*, 2019.

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

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

Lucas, Robert E., "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.

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.

— **and Stephen J. Redding**, "New Trade Models, New Welfare Implications," *American Economic Review*, 2015, *105* (3), 1105–1146.

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," *NBER Working Paper w12610*, 2019.

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

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 w28421*, 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," *Mimeo*, 2019.

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

Rodrik, D., "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, P. N., "Problems of Industrialisation of Eastern and South-Eastern Europe," *The Economic Journal*, 1943, *53*, 202–211.

Schetter, Ulrich, "A Structural Ranking of Economic Complexity," *Mimeo*, 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.

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

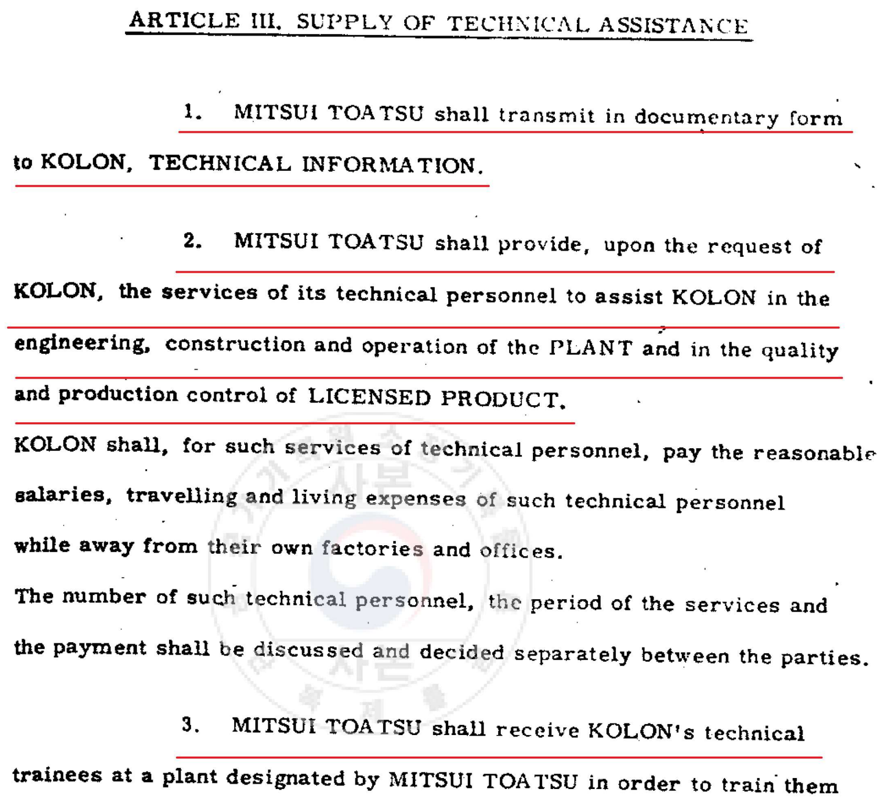
Wooldridge, Jeffrey M., "On Estimating Firm-Level Production Functions Using Proxy Variables to Control for Unobservables," *Economics Letters*, 2009, *104* (3), 112–114.

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

APPENDIX

Appendix A Appendix: 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. Our technology adoption data mainly comes from historical contract documents from *National Archives of Korea*. The *Foreign Capital Act* strictly regulated domestic firms' transactions with foreign firms, including technology adoption contracts. Because of the law, when domestic firms wanted to make technology adoption contracts with foreign firms, they first had to get approvals from the government and after the contract, they had to submit documents related to newly made contracts to the government authorities. In these documents, domestic firms had to submit (i) their plans on what to do with newly adopted technologies and (ii) copies of actual contract documents.

After Chung-Hee Park came to power through a military coup, he created the *Economic Planning Board* in 1961 to promote economic development and to design better economic policies. After he became elected as a president of the civilian government in 1963, the EPB was at the center of the economic policy-making process of Korea until President Park was assassinated in 1979. Starting from 1961, the EPB had monthly meetings until the mid 1980s. In each meeting, they examined newly made transactions between domestic and foreign firms. Documents that were examined in each meeting were collected and preserved by *National Archives of Korea*. These are our main source of data. Figure A1 is one example of these documents. This is one page from the actual contract document between *Kolon* (Korean) and *Mitsui Toatsu* (Japanese).

Contents of Technology Adoption Documents.

A.2 Firm Balance Sheet Data.

We match firm balance sheet data obtained from *Annual Report of Korean Companies* between 1970 and 1982, and technology adoption contracts obtained from the archival contract documents. These reports are made by *Korea Productivity Center*. The data set covers firm with employment more than 50. Balance sheet information includes sales, assets, fixed assets, and employment. Employment starts from 1972. We convert all monetary values into 2015 US dollars.

It has detailed information on address of the location of production. We convert addresses into the 2010 administrative divisions of Korea upto town level.[90] Then, using the distances between towns, we calculate distances between firms within district.

Sector Groupings. We classify firms into 10 manufacturing sectors. Table A1 reports the classification. This classification is similar to Choi and Levchenko (2021) who use the same firm balance sheet data. The numbers inside the parenthesis are ISIC Rev. 3.1 (ISIC) code. We use this ISIC codes to map our firm data to other trade or tariff data sets.

A.3 Identify Changes of Firms' Names.

We match technology adoption and firm balance sheet data sets using firms' names and information on start year. One of the key challenges when merging two data sets based on firms' names is that many firms changed their names during the sample period. We track each firm's name using the following steps.

- Step 1: If firms reported their changes of names in *Annual Report of Korean Companies*.
- Step 2: Information on firms' history from their websites
- Step 3: If companies' website is not available, search for firms' names from *https://www.jobkorea.co.kr* or *https://www.saramin.co.kr*.[91]
  - Only when reported start year information in *Annual Report of Korean Companies* coincides with search results from the above two websites, we identify as the same firm.
  - In the case where a firm became merged with other firms, we count such firm as an exit.
- Step 4: Search for newspapers. Newspapers in the 70s sometimes had an article related to changes of firms' names.

A.4 Coverage.

Figure A2 reports the average coverage of micro 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 in corresponding years.

[90] We classifies firms' location of production into *Li* and *Dong* levels.

[91] These are two largest job posting websites in Korea.

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 Non-Metallic Mineral Products	Glass and glass products (261) Non-metallic mineral products n.e.c. (269)

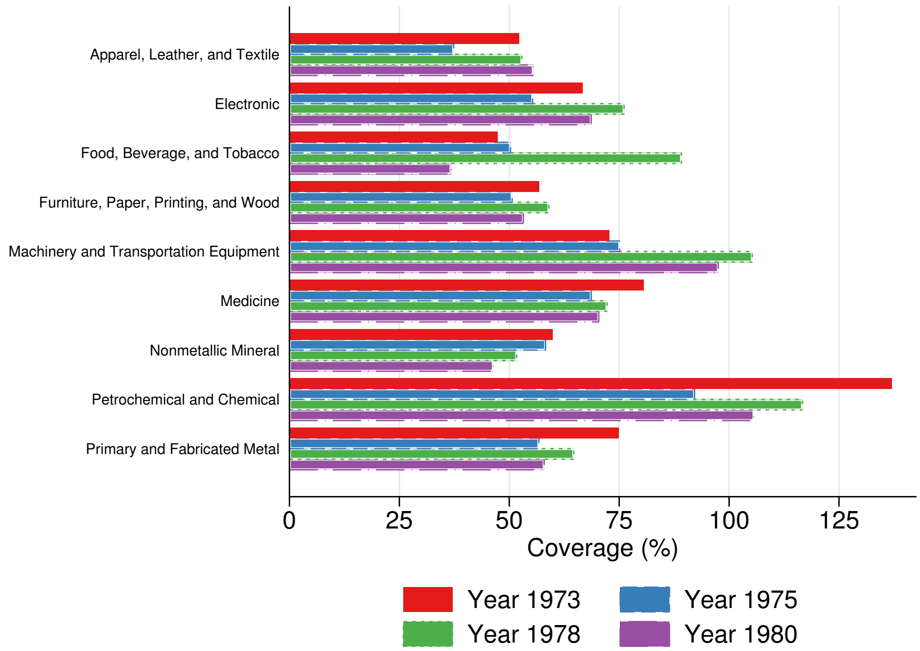


Figure A2. Coverage

When computing this coverage, for some observations with missing information on sales, we impute using information on assets. 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 micro data set covers gross output from the input output table about 70%. However, there are some heterogeneity across sectors. "Machinery and Transportation Equipment" or "Petrochemical and Chemical" have higher coverage rate, whereas "Food, Beverage, and Tobacco" or "Apparel, Leather, and Textile" are relatively less covered than other sectors.

A.5 Descriptive Statistics.

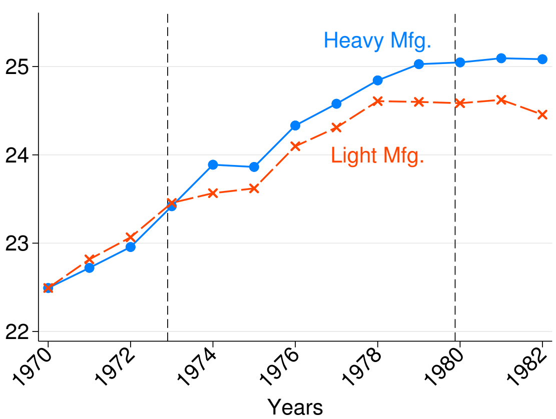
Table A2 reports the descriptive statistics of the constructed data set. The table reports firm balance sheet variables including log of sales, assets, fixed assets, and employment, and variables related to firms' adoption activities. In columns (1), (2), and (3), we include all, heavy, and light manufacturing firms. $\mathbb{1}[\text{Adopt}]$ is a dummy variable which equals one if a firm is in a contract relationship with any foreign firms. From the adoption contract data, we can observe when firms made adoption contracts and associated contract years. The dummy variable equals one if a domestic firm is in the middle of contract years. $\mathbb{1}[\text{Adopt}]$ is a dummy

Table A2: Descriptive Statistics.

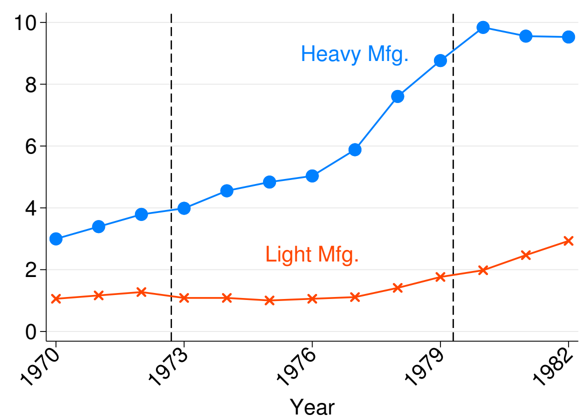
	All mfg. firms (1)	Heavy mfg. firms (2)	Light mfg. firms (3)
	Firm Balance Sheet		
ln(Sales)	15.65	15.54	15.75
	(1.925)	(1.938)	(1.910)
ln(Assets)	15.14	15.10	15.18
	(1.766)	(1.764)	(1.767)
ln(Fixed Assets)	13.96	13.94	13.98
	(1.966)	(1.933)	(1.992)
ln(Emp)	5.166	5.028	5.285
	(1.321)	(1.319)	(1.311)
	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. In column (1), descriptive statistics of all manufacturing firms are reported. In columns (2) and (3), descriptive statistics of heavy and light manufacturing firms are reported. 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.

variable which equals one if a firm ever adopted foreign technology during the sample period. Consistent with the historical narrative, adoption activities are concentrated among heavy manufacturing firms. Between 1970 and 1982, 13% of heavy manufacturing firms on average adopted technology at least once, whereas it was only 4.2% for light manufacturing firms.



A. log sum of sales



B. Shares of adopters

Figure A3. Evolution of Size of Manufacturing Sectors and Shares of Adopters

Notes. Panels A and B of this figure plots evolution of size of manufacturing sectors and shares of adopters, respectively. Size of each sector is measured as log of the total sum of firms' sales in each sector, which is normalized by the level 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.

Appendix B Appendix: Historical Background

B.1 Additional Aggregate Facts on Late Industrialization in South Korea

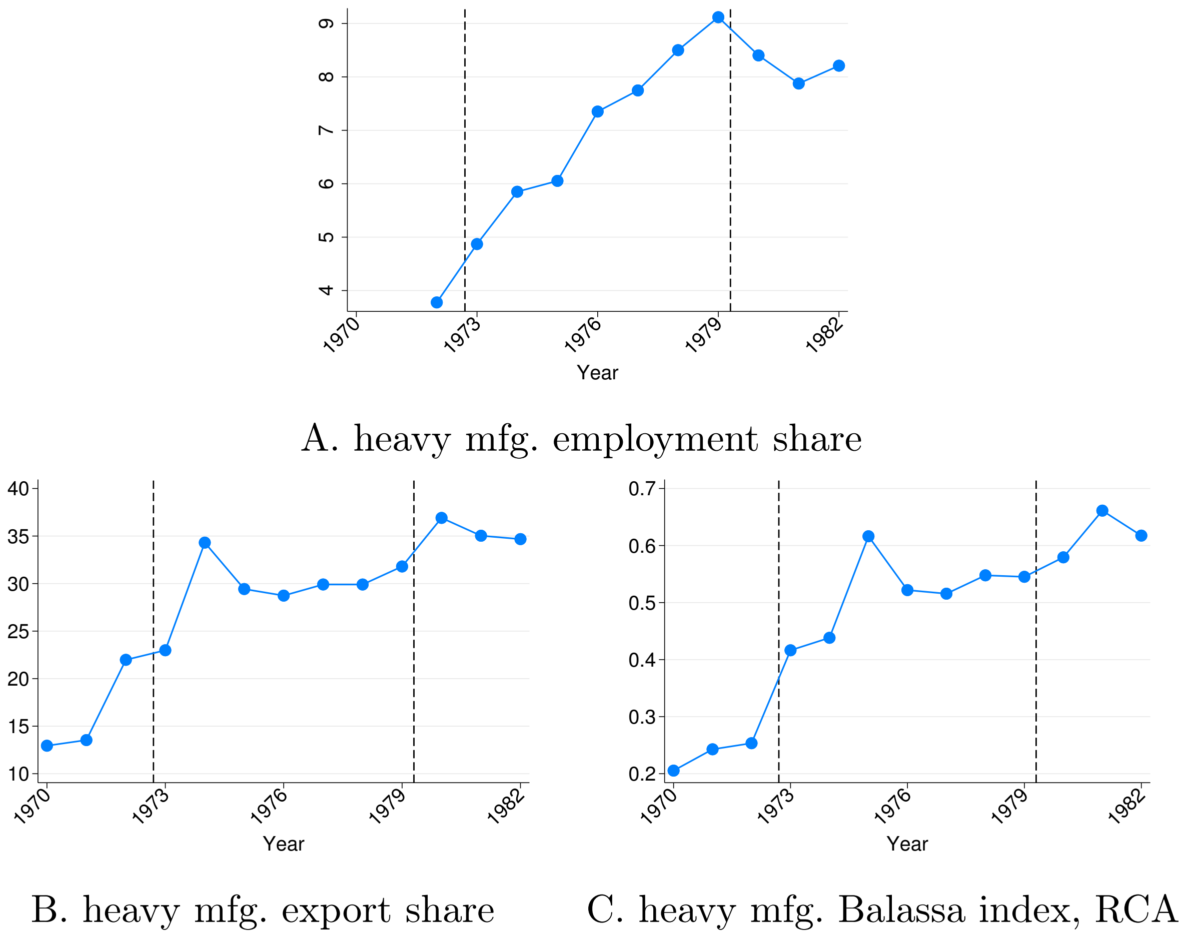


Figure B1. Aggregate Patterns of Late Industrialization in South Korea in the 1970s

Notes. The figure illustrates aggregate changes of the Korean economy during the 1970s. Panels A, B, and C reports heavy manufacturing employment shares, export shares, and Balassa revealed comparative advantage index.

In this section, we provide additional aggregate patterns on late industrialization in South Korea in the 1970s. In Panels A, B, and C of Figure B1, we report the heavy manufacturing employment share, the heavy manufacturing export share, and Balassa's revealed comparative advantage index, which is defined as:

$$RCA_{heavy,t} = \frac{EX^{KOR}_{heavy,t} / \sum_{j \in \mathcal{J}} EX^{KOR}_{jt}}{EX^{RoW}_{heavy,t} / \sum_{j \in \mathcal{J}} EX^{RoW}_{jt}},$$

where EX^{c}_{jt} is the sector j exports of country c. The $RCA_{heavy,t}$ measures specialization pattern in the heavy manufacturing sectors of Korea to that of the rest of the world. The employment shares are constructed based on OECD Stan database, and the export shares and the revealed comparative advantage index are computed based on the trade data from Feenstra et al. (2005).[92] Consistent with the heavy manufacturing GDP shares in Figure 1, the employment shares increased from 4 to 8% 1972 and 1982. Korea's sectoral employment data of OECD Stan database starts from 1972, so we could not compute the shares for 1970 and 1971. The export shares increased from 13.7 to 35% between 1970 and 1982.

[92] The trade data was downloaded from https://cid.econ.ucdavis.edu/nberus.html.

B.2 Firm Level Evidence on Techonlogy Adoption and the Korean Government Policy

In this section, we provide empirical evidence on the impact of the Korean government policy on firms' technology adoption decisions. We run the following event study specification for the sample of heavy manufacturing firms:

$$100 \times \mathbb{1}[Adopt_{it}] = \sum_{\tau=-3}^{9} \beta^{\tau} D_t^{\tau} + \delta_i + \epsilon_{it}, \tag{B.1}$$

where i denotes firm and t time. $\mathbb{1}[Adopt_{it}]$ is a dummy variable of firms' adoption status. We multiply the dummy variable by 100 for the ease of interpretation. D_t^{τ} is the event study variables defined as $D_t^{\tau} := \mathbb{1}[t - \tau = 1973]$. δ_i and δ_t are firm and time fixed effects. ϵ_{it} is the error term. The key variable of interests are $\{\beta^{\tau}\}_{\tau=-3}^{9}$. These coefficients detect how firm sales or adoption decisions were affected by the policy. $\{\beta^{\tau}\}_{\tau=-2}^{0}$ are pre-trends. If confounding factors were driving the implementation of the policy, it may show up in the pre-trends. Standard errors are clustered at regional level.

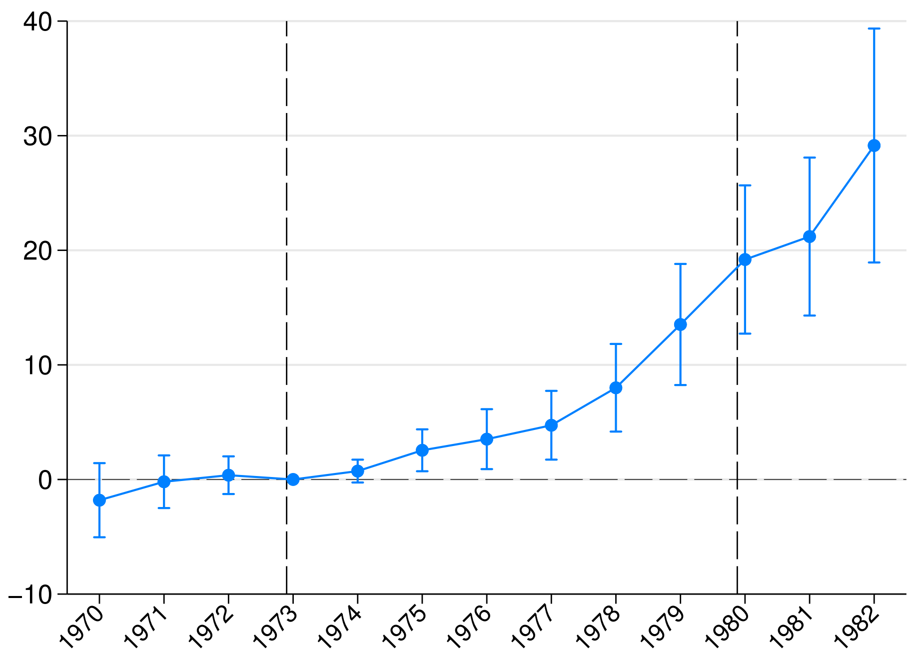


Figure B2. The Impact of the HCI Drive on Firms' Technology Adoption Decisions

Notes. The figure illustrates the estimated β_{τ} in Equation (B.1). β_0^{diff} is normalized to be zero. All specifications control for firm and calendar year fixed effects. The two dashed vertical lines denote the start and the end of the HCI Drive. Error bars represent 95 percent confidence intervals based on standard errors, clustered at regional level.

Figure B2 illustrates the estimated coefficients with 95 percent confidence intervals. Because we are normalizing β_0^{diff} to be zero, the estimated coefficients capture how firms' adoption activities increased relative to the level in 1973. There are no pre-trends. Firms' overall activities before 1973 were not statistically distinguishable from those in 1973. However, only after 1973, more firms started adopting foreign technology. This sudden rapid increases after 1973 support the historical narrative of the sudden launch of the HCI Drive driven by the political shock. The estimated coefficients imply that the probability of adopting foreign technology in 1980 increased 20% relative to 1973.

Appendix C Appendix: Model

C.1 Closed-Form Expressions for Regional Variables

In this section, we derive closed-form expressions for regional variables. Given the firms' optimal adoption and export decisions and the bounded Pareto distributional assumption, regional-level variables summed across firms within region-sector can be expressed as a function of shares of adopters and exporters.

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

$$P_{njt}^{1-\sigma} = \sum_{m\in\mathcal{N}} M_{mj}\Bigg\{ \int_{\phi_{mjt}^{min}}^{\bar{\phi}_{mjt}^{T}} \left(\frac{\sigma}{\sigma-1}\frac{\tau_{mnj}c_{mjt}}{f(\lambda_{mjt-1}^{T})\phi}\right)^{1-\sigma} dG_{mjt}(\phi) + \int_{\bar{\phi}_{mjt}^{T}}^{\phi_{mjt}^{max}} \left(\frac{\sigma}{\sigma-1}\frac{\tau_{mnj}(1-s_{mjt})c_{mjt}}{f(\lambda_{mjt-1}^{T})\phi}\right)^{1-\sigma} dG_{mjt}(\phi)\Bigg\} + (\tau_{nj}^{x}c_{jt}^{f}).$$

The above 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:

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

where

$$\begin{aligned}\bar{\phi}^{avg}_{njt} &= \bar{\phi}^{avg}(\lambda^T_{njt-1}, \lambda^T_{njt}, s_{njt}, \phi^{min}_{njt}) \\ &= \frac{\theta f(\lambda^T_{njt-1})(\phi^{min}_{njt})^{\sigma-1}}{\tilde{\theta}(1-\kappa^{-\theta})}\Bigg\{\bigg(\Big(\frac{\eta}{1-s_{njt}}\Big)^{\sigma-1}-1\bigg)(\tilde{\lambda}^T_{njt})^{\frac{\tilde{\theta}}{\theta}} + \bigg(1-\Big(\frac{\eta}{1-s_{njt}}\Big)^{\sigma-1}\kappa^{-\tilde{\theta}}\bigg)\Bigg\},\end{aligned} \tag{C.2}$$

$\tilde{\lambda}^T_{njt} = (1-\kappa^{-\theta})\lambda^T_{njt} + \kappa^{-\theta}$, and $\tilde{\theta} = \theta - (\sigma-1)$.[93] A price index depends on the three terms: unit cost, average productivity inclusive of subsidies $\bar{\phi}^{avg}_{njt}$, and consumer foreign market access $(\tau^x_{nj}c^f_{jt})^{1-\sigma}$. $\bar{\phi}^{avg}_{njt}$ captures how region n can produce sector j intermediate varieties at cheaper cost relative to other regions. Region n can produce at cheaper costs if it has technological advantages (λ^T_{njt}, λ^T_{njt-1}, ϕ^{min}_{njt}) or higher subsidies (s_{njt}). Holding other variables constant, a price index is lower when (i) other neighboring regions have lower unit costs (either lower τ_{nmj} or c_{mjt}), (ii) neighboring regions have higher productivity or obtain more subsidies (higher $\bar{\phi}^{avg}_{njt}$), or (iii) a price of imported inputs is lower (lower τ^x_{nj} or c^f_{jt}).

The average productivity inclusive of subsidy (Equation (C.2)) increases in a share of adopters in the previous period λ^T_{njt-1}, a share of adopters in the current period λ^T_{njt}, subsidies s_{njt}, and a natural advantage captured by the Pareto lower bound ϕ^{min}_{njt}. A share of adopters in $t-1$ increases average productivity directly through the spillover and indirectly by inducing more firms to adopt technology in period t (Equation (5.7)). A current share of adopters increases the average productivity through direct productivity gains. Subsidies increase the average productivity directly by lowering the cost of production of 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^d_{njt} and the total value of export R^x_{njt}: $R_{njt} = R^d_{njt} + R^x_{njt}$.

Regional exports can be written as

$$R^x_{njt} = M_{nj}\Bigg[\int_{\bar{\phi}^T_{njt}}^{\bar{\phi}^{max}_{njt}}\bigg(\frac{\sigma}{\sigma-1}\frac{\tau^x_{nj}(1-s_{njt})c_{njt}}{\eta f(\lambda^T_{njt-1})\phi}\bigg)^{1-\sigma}dG_{njt}(\phi) + \int_{\bar{\phi}^x_{njt}}^{\bar{\phi}^T_{njt}}\bigg(\frac{\sigma}{\sigma-1}\frac{\tau^x_{nj}c_{njt}}{f(\lambda^T_{njt-1})\phi}\bigg)^{1-\sigma}dG_{njt}(\phi)\Bigg]D^f_{jt}, \tag{C.3}$$

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

The first term inside the bracket can be expressed as:

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

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

[93] When $\lambda^T_{njt} \to 0$, the average productivity becomes $\bar{\phi}^{avg}_{njt} = \frac{\theta}{\tilde{\theta}(1-\kappa^{-\theta})}f(\lambda^T_{njt-1})(\phi^{min}_{njt})^{\sigma-1}(1-\kappa^{-\tilde{\theta}})$. When $\lambda^T_{njt} \to 1$, the average productivity becomes $\bar{\phi}^{avg}_{njt} = \frac{\theta}{\tilde{\theta}(1-\kappa^{-\theta})}f(\lambda^T_{njt-1})(\phi^{min}_{njt})^{\sigma-1}\Big(\frac{\eta}{1-s_{njt}}\Big)^{\sigma-1}(1-\kappa^{-\tilde{\theta}})$.

can be re-expressed as:

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

where $\tilde{\lambda}^x_{njt} = (1-\kappa^{-\theta})\lambda^x_{njt} + \kappa^{-\theta}$. The last equality comes from that $\lambda^x_{njt} = 1 - G_{njt}(\bar{\phi}^x_{njt})$. Using Equations (C.3), (C.4), and (C.5), and $M^x_{njt} = M_{nj} \times \lambda^x_{njt}$, we can derive exports of Equation (5.16) in the main text.

Using the algebra above, the total gross output can be expressed as

$$
R^d_{njt} = M_{nj}(\mu c_{njt})^{1-\sigma} \times (\bar{\phi}^{avg}_{njt})^{\sigma-1} \times \underbrace{\sum_{m\in\mathcal{N}} \tau^{1-\sigma}_{nmj} P^{\sigma-1}_{mjt} E_{mjt}}_{\substack{\text{Firm domestic}\\ \text{market access}}}. \tag{C.6}
$$

Similarly, regional exports can be expressed as

$$
R^x_{njt} = M^x_{njt}(\mu c_{njt})^{1-\sigma} \times \underbrace{(\bar{\phi}^{avg,x}_{njt})^{\sigma-1}}_{\substack{\text{Exporters'}\\ \text{average productivity}\\ \text{inclusive of subsidy}}} \times \underbrace{(\tau^x_{nj})^{1-\sigma} D^f_{jt}}_{\substack{\text{Firm foreign}\\ \text{market access}}}, \tag{C.7}
$$

where

$$
\begin{aligned}
\bar{\phi}^{avg,x}_{njt} &= \bar{\phi}^{avg,x}(\lambda^T_{njt-1}, \lambda^T_{njt}, \lambda^x_{njt}, s_{njt}, \phi^{min}_{njt}) \\
&= \frac{\theta f(\lambda^T_{njt-1})(\phi^{min}_{njt})^{\sigma-1}}{\tilde{\theta}(1-\kappa^{-\theta})}\frac{(\tilde{\lambda}^x_{njt})^{\frac{\tilde{\theta}}{\theta}}}{\lambda^x_{njt}} \\
&\qquad \times \left\{\left(\left(\frac{\eta}{1-s_{njt}}\right)^{\sigma-1} - 1\right)\left(\frac{\tilde{\lambda}^T_{njt}}{\tilde{\lambda}^x_{njt}}\right)^{\frac{\tilde{\theta}}{\theta}} + \left(1 - \left(\frac{\eta}{1-s_{njt}}\right)^{\sigma-1}\kappa^{-\tilde{\theta}}(\tilde{\lambda}^x_{njt})^{-\frac{\tilde{\theta}}{\theta}}\right)\right\},
\end{aligned}
\tag{C.8}
$$

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

Both the total domestic sales and exports (i) increase in the average productivities inclusive ob subsidies, (ii) increase in degree of access to markets, (iii) increase in subsidies, and (iv) decrease in the cost of production. One difference between $\bar{\phi}^{avg,x}_{njt}$ (Equation (C.8)) and $\bar{\phi}^{avg}_{njt}$ (Equation (C.2)) is that $\bar{\phi}^{avg,x}_{njt}$ additionally depends on a share of exporters λ^x_{njt}. λ^x_{njt} captures selection induced by a fixed export cost. Because of a fixed export cost, 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 a share of exporters λ^x_{njt} because a larger share of exporters implies that less productive firms participate in exporting, which in turn leads to a weaker selection effect 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 $\bar{\phi}^{avg}_{njt}$.

C.2 Analytical Results on Multiple Steady States

C.2.1 Derivation of the Equilibrium Share of Adopters.

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

$$(\bar{\phi}_t^T)^{\sigma-1} = \frac{\sigma P_t F^T}{(\eta^{\sigma-1}-1)(\mu w_t)^{1-\sigma} f(\lambda_{t-1}^T)^{\sigma-1} P_t^{\sigma} Q_t}, \tag{C.9}$$

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

$$(\lambda_t^T)^{-\frac{1}{\theta}} = \bar{\phi}_t^T \tag{C.10}$$

First, we show that

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

and

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

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 $z(\omega) = \eta(\omega)f(\lambda_{t-1}^T)\phi(\omega)$ for adopters and $z(\omega) = f(\lambda_{t-1}^T)\phi(\omega)$ 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 monopolistic competition assumption into the above equation, we obtain that $Q_t = [\int z(\omega)^{\sigma-1}d\omega]^{\frac{1}{\sigma-1}}$. Using the Pareto distributional assumption and the cutoff property, we can further derive that

$$Q_t = \left[\frac{\theta}{\tilde{\theta}}\Big((\eta^{\sigma-1}-1)(\bar{\phi}_t^T)^{\theta-(\sigma-1)}+1\Big)\right]^{\frac{1}{\sigma-1}} f(\lambda_{t-1}^T) = \underbrace{\left[\frac{\theta}{\tilde{\theta}}\Big((\eta^{\sigma-1}-1)(\lambda_t^T)^{1-\frac{\sigma-1}{\theta}}+1\Big)\right]^{\frac{1}{\sigma-1}}}_{=A(\lambda_t^T)} \times f(\lambda_{t-1}^T), \tag{C.11}$$

where the second equality is derived from Equation (C.10). Similarly, using that $P_t = [\mu w_t \int z(\omega)^{\sigma-1}d\omega]^{\frac{1}{1-\sigma}}$, we can derive that

$$\frac{w_t}{P_t} = \frac{w_t}{[\int(\mu w_t/z_{it}(\omega))^{1-\sigma}]^{\frac{1}{1-\sigma}}} = \frac{\sigma-1}{\sigma}\left[\frac{\theta}{\tilde{\theta}}\Big((\eta^{\sigma-1}-1)(\lambda_t^T)^{1-\frac{\sigma-1}{\theta}}+1\Big)\right]^{\frac{1}{\sigma-1}} f(\lambda_{t-1}^T), \tag{C.12}$$

where the second equality is also derived from Equation (C.10).

Substituting Equations (C.10), (C.11), and (C.12) into Equation (C.9), we can obtain that

$$\lambda_t^T = \left(\frac{(\eta^{\sigma-1}-1)}{\sigma F^T} \times A(\lambda_t^T)^{2-\sigma} \times f(\lambda_{t-1}^T)\right)^{\frac{\theta}{\sigma-1}}. \tag{C.13}$$

Let $\hat{\lambda}_t^T$ be the solution of Equation (C.13). Because the equilibrium share is bounded by 1, the equilibrium share is defined as follows:

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

C.2.2 Proofs of Proposition 1: Multiple Steady States

Proof of Proposition 1(i). The equilibrium is defined by 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 the above equation. If the obtained λ_t^T from the above equation is above 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

$$G(\lambda_t^T; \eta, \delta, \lambda_{t-1}^T) = A(\lambda_t^T)^{2-\sigma} \times f(\lambda_{t-1}^T)\frac{(\eta^{\sigma-1}-1)}{\sigma F^T} - (\lambda_t^T)^{\frac{\sigma-1}{\theta}} \tag{C.14}$$

where

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

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 (C.14) with respect to λ_t^T, we obtain that

$$\frac{\partial G}{\partial \lambda_t^T} = \left(\frac{2-\sigma}{\sigma-1}\right)A(\lambda_t^T)^{3-2\sigma}(\eta^{\sigma-1}-1)\frac{\theta-(\sigma-1)}{\theta}(\lambda_t^T)^{-\frac{\sigma-1}{\theta}}f(\lambda_{t-1}^T)\frac{(\eta^{\sigma-1}-1)}{\sigma F^T} - \frac{\sigma-1}{\theta}(\lambda_t^T)^{\frac{-\theta+(\sigma-1)}{\theta}} < 0, \tag{C.15}$$

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

Taking the derivative of Equation (C.14) with respect to λ_{t-1}^T, we obtain that

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

Applying the implicit function theorem and using the signs of Equations (C.15) and (C.16), we obtain that

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

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

Taking the derivative of Equation (C.14) with respect to η, we obtain that

$$\begin{aligned}\frac{\partial G}{\partial \eta} &= \left(\frac{2-\sigma}{\sigma-1}\right)A(\lambda_t^T)^{3-2\sigma}f(\lambda_{t-1}^T)\frac{\theta}{\theta-(\sigma-1)}(\lambda_t^T)^{\frac{\theta-(\sigma-1)}{\theta}}(\sigma-1)\eta^{\sigma-2}\frac{(\eta^{\sigma-1}-1)}{\sigma F^T} \\ &\quad + A(\lambda_t^T)^{2-\sigma}f(\lambda_{t-1}^T)\frac{(\sigma-1)\eta^{\sigma-2}}{\sigma F^T} \\ &= A(\lambda_t^T)^{3-2\sigma}f(\lambda_{t-1}^T)\frac{(\sigma-1)\eta^{\sigma-2}}{\sigma F^T}\frac{\theta}{\theta-(\sigma-1)}\left[\frac{1}{\sigma-1}(\eta^{\sigma-1}-1)(\lambda_t^T)^{\frac{\theta}{\theta-(\sigma-1)}}+1\right] > 0.\end{aligned} \tag{C.17}$$

Taking the derivative of Equation (C.14) with respect to δ, we obtain that

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

Applying the implicit function theorem and using the signs of Equations (C.15), (C.18), and (C.17),

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

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 again show this by applying the implicit function theorem and doing some tedious algebra. Applying the implicit function theorem,

$$\begin{aligned}\frac{\partial^2 \lambda_t^T}{\partial(\lambda_{t-1}^T)^2} &= -\frac{1}{(\partial G/\partial \lambda_t^T)^3}\\ &\times\left[\frac{\partial G}{\partial \lambda_{t-1}^T}\times\left(\frac{\partial G}{\partial \lambda_t^T}\right)^2 - \left(\frac{\partial^2 G}{\partial \lambda_t^T\partial \lambda_{t-1}^T}+\frac{\partial^2 G}{\partial \lambda_{t-1}^T\partial \lambda_t^T}\right)\times\frac{\partial G}{\partial \lambda_{t-1}^T}\times\frac{\partial G}{\partial \lambda_t^T}+\frac{\partial^2 G}{\partial(\lambda_t^T)^2}\times\left(\frac{\partial G}{\partial \lambda_{t-1}^T}\right)^2\right].\end{aligned} \tag{C.19}$$

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

$$\frac{\partial^2 G}{\partial(\lambda_{t-1}^T)^2} = A(\lambda_t^T)^{2-\sigma}\frac{(\eta^{\sigma-1}-1)}{\sigma F^T}exp(\delta\lambda_{t-1}^T)\delta^2 > 0. \tag{C.20}$$

$$\frac{\partial^2 G}{\partial \lambda_t^T\partial \lambda_{t-1}^T} = \frac{\partial^2 G}{\partial \lambda_{t-1}^T\partial \lambda_t^T} = \frac{2-\sigma}{\sigma-1}A(\lambda_t^T)^{3-2\sigma}\times\left[\frac{\theta-(\sigma-1)}{\theta}(\eta^{\sigma-1}-1)(\lambda_t^T)^{-\frac{\sigma-1}{\theta}}\right]exp(\delta\lambda_{t-1}^T)\lambda_{t-1}^T < 0. \tag{C.21}$$

$$\begin{aligned}\frac{\partial^2 G}{\partial(\lambda_t^T)^2} &= \frac{(2-\sigma)(3-\sigma)}{(\sigma-1)^2}A(\lambda_t^T)^{2-2\sigma)}\left[\frac{\theta-(\sigma-1)}{\theta}(\lambda_t^T)^{-\frac{\sigma-1}{\theta}}(\eta^{\sigma-1}-1)\right]^2 exp(\delta\lambda_{t-1}^T)\frac{(\eta^{\sigma-1}-1)}{\sigma F^T}\\ &+\frac{\sigma-2}{\theta}A(\lambda_t^T)^{3-2\sigma}(\eta^{\sigma-1}-1)\frac{\theta-(\sigma-1)}{\theta}(\lambda_t^T)^{-\frac{\sigma-1}{\theta}-1}exp(\delta\lambda_{t-1}^T)\frac{(\eta^{\sigma-1}-1)}{\sigma F^T}\\ &+\frac{\sigma-1}{\theta}\frac{\theta-(\sigma-1)}{\theta}(\lambda_t^T)^{-\frac{\theta-(\sigma-1)}{\theta}-1} > 0.\end{aligned} \tag{C.22}$$

Substituting Equations (C.15), (C.16), (C.20), (C.21), and (C.22) into Equation (C.19), we obtain that $\frac{\partial^2 \lambda_t^T}{\partial(\lambda_{t-1}^T)^2} > 0$, 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 at most twice. [94] 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

[94]The intercept is always positive because of the unbounded Pareto distributional assumption which guarantees a

Equation (C.14) meet at $\lambda^T_{t-1} = 1$. Also, by the same logic, there exits $\bar{\delta}$ and $\bar{\eta}$ such that the 45 degree line is tangent to Equation (C.14). The two lines meet at least twice for $\delta \in [\underline{\delta}, \bar{\delta}]$ and $\eta \in [\underline{\eta}, \bar{\eta}]$.

□

Proof of Proposition 1(iv). The welfare of household is $\frac{w_t + \Pi_t}{P_t}$ where Π_t is the aggregate profits summed across all firms in the economy.[95] This can be expressed as $\frac{w_t}{P_t} + \frac{\Pi_t}{P_t}$. Using Equations (C.11) and (C.12) and that

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

we can derive that the welfare can be expressed as $f(\lambda^T_{t-1})A(\lambda^T_t)$. 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.

□

positive share of adopters at $\lambda^T_{t-1} = 0$.

[95] Note that $L_t = 1$.

C.2.3 Source of Dynamic Externality

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

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

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

which is similar to Equation (C.9) but $P_t F^T$ is replaced with $w_t F^T$. $\frac{w_t}{P_t}$ and Q_t are defined analogously to Equations (C.11) and (C.12) regardless of that fixed adoption costs are in units of labor. Substituting Equations (C.11) and (C.12) into the above cutoff, we can derive that

$$\lambda_t^T = \Big(\frac{(\eta^{\sigma-1}-1)}{\sigma F^T} \times \mu \times A(\lambda_t^T)^{1-\sigma}\Big)^{\frac{\theta}{\sigma-1}}. \tag{C.23}$$

The above expression differs from the expression of Equation (C.13) in that μ replaces $f(\lambda_{t-1}^T)$.

The equilibrium share of adopters in Equation (C.23) 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. Suppose there was a higher share of adopters in the previous period. Overall productivity of all firms in the current period increases, which in turn increases overall demand for labor. As labor demands increase equilibrium wage, costs of the adoption become higher. In the equlibrium, increases in costs of the adoption exactly cancel out increases in overall productivity, which in turn makes the equilibrium share of adopters be not affected by λ_{t-1}^T (Equation C.23).

C.2.4 Multiple Steady States and Permanent Effects of Temporary Subsidies

We show that temporary subsidies cannot have permanent effects when multiple-steady states do not exist in the simplified model in Section 5.6. 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 to generate multiple steady state and there is a unique steady-state. For simplicity, we assume that the economy is at the original steady states before t_0.

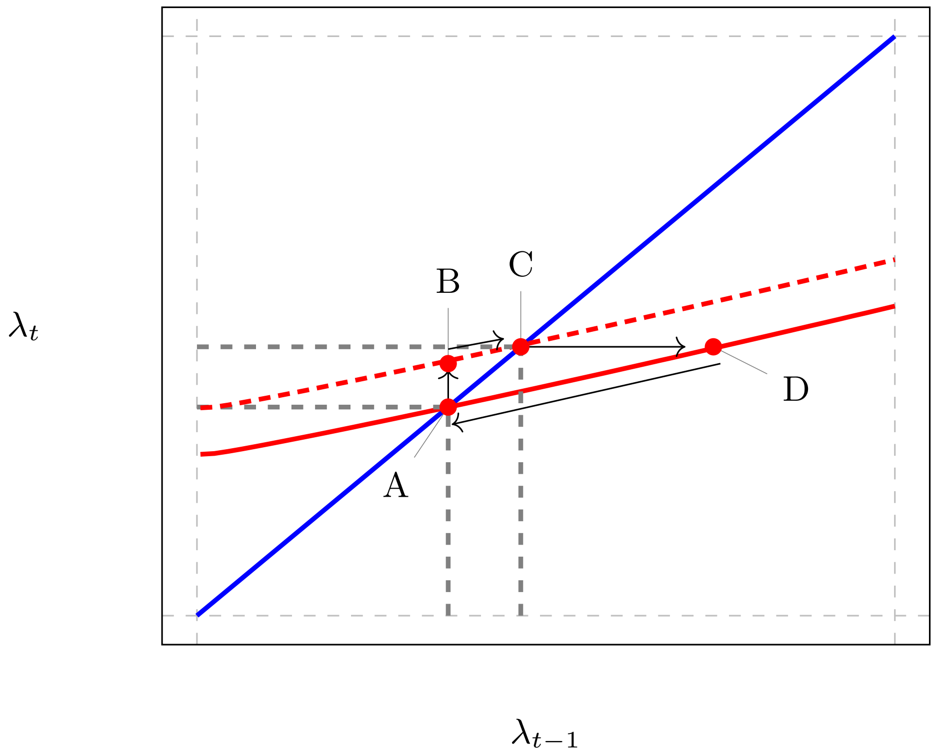


Figure C1. When multiple steady states do not exist, temporary adoption subsidies cannot have permanent effects.

Notes. This figure illustrates that when multiple steady states do not exist, temporary adoption subsidies cannot have permanent effects. The solid red locus is the short-run equilibrium curve and the dashed red locus is the short-run equilibrium curve when adoption subsidies are provided permanently.

Figure C1 graphically illustrates that temporary subsidies only have temporary effects when there exists a unique steady state. The red solid locus is the original short-run equilibrium curve without any subsidies. 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 with the subsidy rate $\bar{s}$. Then, between t_0 and t_1, it converges to the new steady state. 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 an economy jumps right to C and start converging to the original steady state.

Even if there is a unique steady state, because of the externality, there is a room for policy interventions. However, these policy interventions have to be provided permanently to have permanent effects. For example, the new steady state in Figure C1 can be welfare-improving over the original steady state and this new steady state can be sustained when $\bar{s}$ is provided every period, which is similar to the static setting with externality. However, these permanent policies are inconsistent with the industrialization pattern in South Korea where the adoption subsidies were provided between 1973 and 1979.

C.3 Proof of Proposition 2: Identifying Moment for Subsidies

Proof of Proposition 2. Under the assumption that $V_{nt} = \bar{V}, \forall n$, free mobility of labor, and free trade, sectoral price index and real wage are equalized across regions, that is, $P_{njt} = P_{jt}, \forall n, j$ and $w_{nt}/P_{nt} = \bar{W}, \forall n$. This in turn implies that wage is equalized across regions $w_{nt} = w_t, \forall n$ and firms face the same market size.

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

$$\ln \lambda_{njt}^{T} = \theta\delta\lambda_{njt-1}^{T} + \underbrace{\frac{\theta}{\sigma-1}\left[\left(\frac{\eta}{1-s_t}\right)^{\sigma-1} - 1\right]}_{\beta D_t^{policy}} \underbrace{-\theta \ln \left[\frac{\mu c_{jt}(\sigma w_t F_j^T)^{\frac{1}{\sigma-1}}}{\left(\sum_{m\in\mathcal{N}} P_{jt}^{\sigma-1} E_{mjt} + D_{jt}^{f}\right)^{\frac{1}{\sigma-1}}}\right]}_{=\alpha} + \underbrace{\theta \ln \phi_{njt}^{min}}_{=\epsilon_{njt}}, \quad \text{(C.24)}$$

where the second, third, and fourth terms can be mapped to the policy dummy variable D_t^{policy}, the constant α, and the error term ϵ_{njt}. This mapping gives us the following regression model:

$$\ln \lambda_{njt}^{T} - \theta\delta\lambda_{njt-1}^{T} = \alpha + \beta D_t^{policy} + \epsilon_{njt}.$$

The condition for the estimates to be unbiased is $\mathbb{E}[\epsilon_{njt}|D_t^{policy}]$, which is equivalent to $\mathbb{E}[\ln \phi_{njt}^{min}|D_t^{policy}]$ (Equation (C.24)). When this condition is satisfied,

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

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

□

C.4 Possible Microfoundations for Adoption Spillovers

C.4.1 Local Diffusion of Knowledge

Set up. Consider a closed economy with one sector and N regions. For notational convenience, we omit a subscript j that denotes for sectors. Firms are facing downward sloping-demand curve and monopolistically competitive. Goods are freely tradable across regions. A firm receives exogenous productivity $\tilde{\phi}_{it}$, which is independent and identically distributed across firms. Given this exogenous productivity, in each period, firms choose whether to adopt advanced foreign technology T_{it} and a level of innovation a_{it} as in Desmet and Rossi-Hansberg (2014).

Firms' Profit Maximization. Given $\tilde{\phi}_{it}$, a firm optimally chooses whether to adopt technology T_{it} or not and a level of innovation a_{it}:

$$\max_{T_{it}\in\{0,1\},a_{it}\in[0,\infty)} \left\{ \frac{1}{\sigma}\left(\frac{\sigma}{\sigma-1}\frac{w_{nt}}{\tilde{\eta}^{T_{it}} a_{it}^{\gamma_1}\tilde{\phi}_{it}}\right)^{1-\sigma} P_t^{\sigma-1}E_t - w_{nt}T_{it}F^T - w_{nt}a_{it}^{\alpha_1} g(\lambda_{nt-1}^T)B_t \right\}, \tag{C.25}$$

where $T_{it} \in \{0,1\}$ is a dummy variable of adoption decision; $\tilde{\eta}$ is the direct productivity increase from adoption; w_{nt} is a local wage; $P_t^{\sigma-1}E_t$ 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, $B_t = b_1 P_t^{\sigma-1}E_t$ with a constant term b_1 .

The positive externality of adoption comes 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 that a larger adopter share in the previous period decreases the cost of innovation in the current period. This cost specification captures local diffusion of a non-rivalrous component of ideas. With more firms adopting advanced technologies, other local firms are more likely to learn new ideas from these adopters and can use these newly learned ideas for their own innovation. $g(\lambda_{nt-1}^T)$ captures this local diffusion of ideas in a reduced form. We assume that $\gamma_1(\sigma-1) - \alpha_1 + 1 < 0$ holds.[96]

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

$$\gamma_1(\sigma-1)a_{it}^{\gamma_1(\sigma-1)}\frac{1}{\sigma}\left(\frac{\sigma}{\sigma-1}\frac{w_{nt}}{\tilde{\eta}^{T_{it}}\tilde{\phi}_{it}}\right)^{1-\sigma} - b_1 w_{nt}\alpha_1 a_{it}^{\alpha_1-1} g(\lambda_{nt-1}^T) = 0,$$

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.[97] Note that $\frac{\delta_{-1}}{\alpha_1-1-\gamma_1(\sigma-1)} > 0$ and $\frac{1-\sigma}{\alpha_1-1-\gamma_1(\sigma-1)} > 0$ hold. This implies that the optimal amounts of innovation is increasing in the previous adopter share λ_{nt-1}^T, technology adoption status T_{it}, and exogeonus productivity $\tilde{\phi}_{it}$. Substituting the optimal

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

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

a_{it}^* into Equation (C.25), a firm's maximization problem can be rewritten as:

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

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

Historical Case Study. Historical evidence on local diffusion of nonrivalrous components of technology comes from the case study of the Pohang Iron and Steel Company Ltd (POSCO) in Enos and Park (1988, Chapter 7). POSCO was the nation's first integrated steel mill, which started operating in 1973.[98] Given Korea's lack of technology, the imported technology played a big role for POSCO at the initial stages. The government heavily subsidized POSCO for adoption of technology and installation of imported capital equipment associated with the imported technologies. Some of the technicians quitting POSCO got employed in capital good producing firms located near POSCO. In these capital good producing firms, they helped these firms producing capital equipment that was used in POSCO, such as water treatment, dust collection, and the large magnetic crane. In the early 1970s, these capital equipment was all imported, but it started to be produced by local suppliers because of knowledge spillover from technicians who worked in POSCO.

C.4.2 Learning Externality and Labor Mobility in an Imperfect Labor Market

Set up. Consider a closed economy with one sector and N regions. For notational convenience, we omit a subscript j that denotes for sectors. Firms are facing downward sloping-demand curve and monopolistically competitive. Goods are freely tradable across regions. A firm receives exogenous productivity $\tilde{\phi}_{it}$ each period.

In each region, there is the unit measure of engineers and firms in each region. Managers lives two periods, child and adult. They only consume and work in their adulthood. Managers cannot move locations. Once engineers become adults in the second period, they give birth to a child. Managers who worked in firms which adopted technologies passes their knowledge to their children. This learning from parents increases engineering skills of children when they become adults in the second period, which increases engineering skills by $\gamma_1 > 1$. If parents did not work in firms with foreign technology, their children's engineering skills are 1. We assume that engineering skills of newly born children are 1 and $\gamma_1 > 1$ if parents worked in non-adopters and adopters respectively.

Following Acemoglu (1996), we assume that engineers and firms are randomly matched one-to-one. The surplus from match, profits generated by operating firms, is divided by the constant shares between engineers and firms based on Nash bargaining. Managers take the proportion of $\tilde{\beta}$. Once engineers and firms are randomly matched within region, they jointly maximize profits.

A firm makes adoption decision before the matching happens, so a firm has to make an adoption decision based on its expected profit. A firm's overall productivity depends on (1)its exogenous productivity $\tilde{\phi}_{it}$, (2) engineering skills of matched engineers, and (3) adoption decision.

[98] POSCO is located in the one of the targeted regions.

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

A firm's maximization problem can be written as

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

where λ_{nt-1} is a local adopter share 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 labor; γ_1 is engineering skills of engineers whose parents worked in firms that adopted foreign technology; and $\tilde{\eta}$ is the direct productivity gain from adoption. Doing some algebra, the above maximization problem can be rewritten as

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

where

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

Note that $\tilde{f}(\lambda^T_{nt-1})$ increases in a local adopter share 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 were high in Korea (Kim and Topel, 1995). The average duration of a job of manufacturing sectors in Korea was around 4 years, which was less than half of the average of the US (9 years). Consistent with this aggregate statistics from Kim and Topel (1995), Enos and Park (1988, p. 166) provides one particular case study on diffusion of knowledge through labor mobility in the machinery sector. *Daewoo Heavy Industries Ltd* (henceforth *Daewoo*) built the first diesel engine plant in Korea after adopting technology from *MAN* in West Germany. However, one year after *Daewoo* started operating the plant, *Hyundai Heavy Indstries* (henceforth *Hyundai*) adopted technology from *Perkins* in the US and started producing diesel engine. When newly starting business, *Hyundai* hired skilled engineers from *Daewoo* who already acquired technological know-how and capabilities within *Daewoo* by offering higher salaries. Because of this incident, *Daewoo* lost its skilled worker by 33%. Both aggregate statistics on labor mobility and one historical case study support one potential channel of knowledge diffusion through labor mobility.

Appendix D Appendix: Reduced-Form

D.1 Additional Tables

Table D1: Descriptive Statistics: "Winners vs. Losers" Design Samples Between the Year of the Cancellation to 5 Years Prior to the Cancellation

	Winner				Loser				t-Statistics
	Mean	Med.	SD	N	Mean	Med.	SD	N	(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. The table reports the descriptive statistics of the "winners vs. losers" design samples between the year of the cancellation to 5 years prior to the cancellation. Column (9) reports t-statistics of the mean difference between winners and losers is reported with its p-value in bracket. Sales, fixed assets, and assets are measured in 2015 US dollars. Standard errors are two-way clustered by pair and firm and reported in parenthesis. The number of pairs and firms are 33 and 55. All monetary values are in 2015 US dollars.

Table D2: Covariate Balance Test: "Winners vs. Losers" Design Samples Between the Year of the Cancellation to 5 Years Prior to 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	Y	Y	Y	Y
Pair FE	N	Y	N	Y

Notes. The table reports the covariate balance tests of the "winners vs. losers" design samples between the year of the cancellation to 5 years prior to the cancellation. The dependent variable is a dummy variable which equals 1 if a firm actually 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 bracket. Standard errors are two-way clustered by pair and firm and reported in parenthesis. The number of pairs and firms are 33 and 55.

Table D3: Direct Productivity Gains from Technology Adoption: "Winners vs. Losers" Research Design and Nearest Neighbors Matching - Pooled diff-in-diffs Estimates.

Research Design	"Winners vs. losers"			Nearest neighbor		
Dep. Var.	log sales	log revenue TFP	log labor productivity	log sales	log revenue TFP TFP	log labor productivity
	(1)	(2)	(3)	(4)	(5)	(6)
$Post_{it} \times \mathbb{1}[Adopt_{it}]$	0.51***	0.46***	0.64**	0.26***	0.21***	0.15***
	(0.17)	(0.16)	(0.25)	(0.08)	(0.06)	(0.05)
Adj. R^2	0.88	0.87	0.62	0.87	0.89	0.78
# cluster (pair)	34	34	34	151	151	151
# cluster (firm)	57	57	57	177	177	177
N	951	827	835	4824	4124	4137

Notes. The table reports the pooled diff-in-diffs estimator. Columns (1)-(3) present the baseline pooled diff-in-diffs estimates based on the "winners vs. losers" research design. Columns (4)-(6) present the estimates using the nearest neighbors matching. 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. Log revenue TFP is estimated based on Wooldridge (2009). Across all specification, event time dummies, firm fixed effects, pair fixed effects, and calendar year fixed effects are controlled. Robust standard errors in parenthesis are two-way clustered at pair and firm levels. * p<0.1, ** p<0.05, *** p<0.01.

Table D4: Direct Productivity Gains from Technology Adoption: "Winners vs. Losers" Research Design and Nearest Neighbors Matching - Robustness, Alternative TFP Measures

Research Design	"Winners vs. losers"			Nearest neighbor		
	log revenue TFP					
	ACF (2015)	LP (2003)	OLS	ACF (2015)	LP (2003)	OLS
	(1)	(2)	(3)	(4)	(5)	(6)
3 years before event	0.06	0.04	0.00	0.06	0.08	0.05
	(0.30)	(0.24)	(0.29)	(0.07)	(0.06)	(0.07)
2 years before event	–0.18	–0.08	–0.19	0.02	0.01	0.02
	(0.34)	(0.24)	(0.34)	(0.07)	(0.06)	(0.07)
1 year before event	0.10	0.06	0.08	–0.01	–0.01	–0.01
	(0.19)	(0.15)	(0.19)	(0.05)	(0.04)	(0.05)
Year of event						
1 year after event	0.37	0.23	0.33	0.05	0.05	0.05
	(0.38)	(0.37)	(0.39)	(0.05)	(0.04)	(0.05)
2 years after event	0.71**	0.56**	0.67**	0.10	0.14**	0.10
	(0.30)	(0.26)	(0.29)	(0.06)	(0.05)	(0.06)
3 years after event	0.66**	0.43*	0.63**	0.16**	0.20***	0.17**
	(0.28)	(0.23)	(0.27)	(0.07)	(0.07)	(0.07)
4 years after event	0.67**	0.45**	0.63**	0.19**	0.25***	0.19***
	(0.25)	(0.21)	(0.24)	(0.07)	(0.07)	(0.07)
5 years after event	0.64**	0.52**	0.57*	0.25***	0.33***	0.26***
	(0.29)	(0.23)	(0.29)	(0.08)	(0.08)	(0.08)
6 years after event	0.59**	0.46*	0.56**	0.26***	0.36***	0.27***
	(0.29)	(0.24)	(0.27)	(0.09)	(0.10)	(0.09)
7 years after event	0.69**	0.58**	0.67**	0.20***	0.32***	0.21***
	(0.29)	(0.23)	(0.28)	(0.08)	(0.08)	(0.08)
Firm FE	Y	Y	Y	Y	Y	Y
Pair FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Adj. R^2	0.94	0.90	0.60	0.71	0.94	0.77
# cluster (pair)	34	34	34	151	151	151
# cluster (firm)	57	57	57	177	177	177
N	827	827	827	4124	4124	4124

Notes. The table reports the estimated event study coefficients β_τ^{diff} in Equation (4.1). Columns (1)-(3) present the baseline event study estimates based on the "winners vs. losers" research design. Columns (4)-(6) present the event-study estimates using the nearest neighbors matching. β_0^{diff} is normalized to be zero. The dependent variables are log revenue TFP. In columns (1) and (4), log revenue TFP is estimated based on Ackerberg et al. (2015). In columns (2) and (5), log revenue TFP is estimated based on Levinsohn and Petrin (2003). In columns (3) and (6), log revenue TFP based on OLS. The detailed procedure of production function estimation is described in Appendix Section D.4. Across all specification, event time dummies, firm fixed effects, pair fixed effects, and calendar year fixed effects are controlled. Robust standard errors in parenthesis are two-way clustered at pair and firm levels. * p<0.1, ** p<0.05, *** p<0.01.

Table D5: Local Spillover of 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 (Ownership)	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(int-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 (Ownership)	702	697	702	702	697	381	338	381	381	338
N	1264	1259	1264	1264	1259	431	387	431	431	387
Region-Sector FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Ownership FE	N	Y	N	N	Y	N	Y	N	N	Y

Notes. The table reports OLS estimates of Equation (4.4). When constructing the spillover measure defined in Equation (4.2), we lag firms' adoption status by 3 years. In Panel A, we use the subsample which only include firms that did not adopt any technology until the end of the sample period. In Panel B, we use the full sample that include both adopters and non-adopters and additionall control adopters' adoption status. Dependent variables are log sales and revenue TFP in columns (1)-(5) and (6)-(10) respectively. Revenue TFP is estimated based on Wooldridge (2009). ln(Spill-Sales) and ln(int-MA) are additional controls defined in Equations (4.5) and (4.6). In all specifications, we control for region-sector fixed effects and initial dependent variable at the start of the sample period. Standard errors are two-way clustered at both region and ownership levels and are reported in parenthesis. * p<0.1, ** p<0.05, *** p<0.01.

Table D6: Local Spillover of 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 (Ownership)	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(int-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 (Ownership)	702	697	702	702	697	381	338	381	381	338
N	1264	1259	1264	1264	1259	431	387	431	431	387
Region-Sector FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Ownership FE	N	Y	N	N	Y	N	Y	N	N	Y

Notes. The table reports OLS estimates of Equation (4.4). When constructing the spillover measure defined in Equation (4.2), we lag firms' adoption status by 5 years. In Panel A, we use the subsample which only include firms that did not adopt any technology until the end of the sample period. In Panel B, we use the full sample that include both adopters and non-adopters and additionall control adopters' adoption status. Dependent variables are log sales and revenue TFP in columns (1)-(5) and (6)-(10) respectively. Revenue TFP is estimated based on Wooldridge (2009). ln(Spill-Sales) and ln(int-MA) are additional controls defined in Equations (4.5) and (4.6). In all specifications, we control for region-sector fixed effects and initial dependent variable at the start of the sample period. Standard errors are two-way clustered at both region and ownership levels and are reported in parenthesis. * p<0.1, ** p<0.05, *** p<0.01.

Table D7: Local Spillover of 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)	53	53	53	53	53	41	36	41	41	36
# clusters (Ownership)	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	4.23***	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(int-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 (Ownership)	702	697	702	702	697	381	338	381	381	338
N	1264	1259	1264	1264	1259	431	387	431	431	387
Region-Sector FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Ownership FE	N	Y	N	N	Y	N	Y	N	N	Y

Notes. The table reports OLS estimates of Equation (4.4). When constructing the spillover measure defined in Equation (4.2), we lag firms' adoption status by 4 years. In Panel A, we use the subsample which only include firms that did not adopt any technology until the end of the sample period. In Panel B, we use the full sample that include both adopters and non-adopters and additionall control adopters' adoption status. Dependent variables are log employment and labor productivity in columns (1)-(5) and (6)-(10) respectively. Labor productivity is defined as value-added per worker. ln(Spill-Sales) and ln(int-MA) are additional controls defined in Equations (4.5) and (4.6). In all specifications, we control for region-sector fixed effects and initial dependent variable at the start of the sample period. Standard errors are two-way clustered at both region and ownership levels and are reported in parenthesis. * p<0.1, ** p<0.05, *** p<0.01.

Table D8: Local Spillover of 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 (Ownership)	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(int-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 (Ownership)	696	691	696	696	691	701	696	701	701	696
N	1254	1249	1254	1254	1249	1263	1258	1263	1263	1258
Region-Sector FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Ownership FE	N	Y	N	N	Y	N	Y	N	N	Y

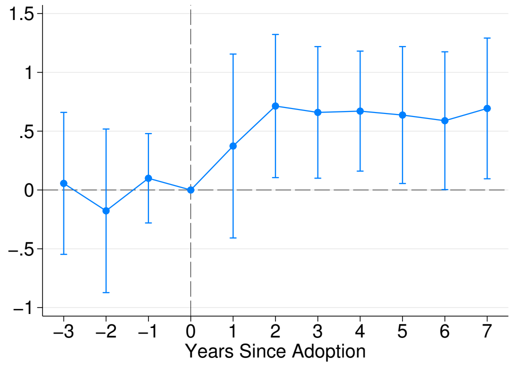
Notes. The table reports OLS estimates of Equation (4.4). When constructing the spillover measure defined in Equation (4.2), we lag firms' adoption status by 4 years. In Panel A, we use the subsample which only include firms that did not adopt any technology until the end of the sample period. In Panel B, we use the full sample that include both adopters and non-adopters and additionall control adopters' adoption status. Dependent variables are log fixed assets and asssets in columns (1)-(5) and (6)-(10) respectively. ln(Spill-Sales) and ln(int-MA) are additional controls defined in Equations (4.5) and (4.6). In all specifications, we control for region-sector fixed effects and initial dependent variable at the start of the sample period. Standard errors are two-way clustered at both region and ownership levels and are reported in parenthesis. * p<0.1, ** p<0.05, *** p<0.01.

Table D9: Local Spillover of Technology Adoption: Robustness - Input Market Access

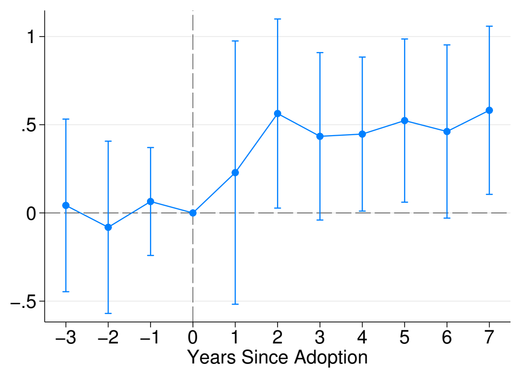
Dep. Var.	log sales		log revenue TFP	
	(1)	(2)	(3)	(4)
	Panel A: Never-Adopter Sample			
Spill	4.15***	3.97**	5.36***	5.32***
	(1.49)	(1.73)	(1.79)	(1.87)
ln(Spill-Sales)	–0.02	–0.01	–0.03**	–0.02
	(0.01)	(0.01)	(0.01)	(0.02)
ln(Input-MA)		–0.02		–0.01
(Weight: $1/dist^{1.1}$)		(0.01)		(0.02)
Adj. R^2	0.19	0.22	0.44	0.42
# clusters (region)	53	53	41	36
# clusters (Ownership)	638	631	326	277
N	1079	1072	346	294
	Panel B: Full Sample			
Spill	3.69***	3.57**	4.51***	3.49*
	(1.18)	(1.48)	(1.57)	(1.91)
$\mathbb{1}[Adopt]$	0.31**	0.25	0.15	0.13
	(0.16)	(0.20)	(0.09)	(0.10)
ln(Spill-Sales)	–0.03**	–0.03*	–0.04***	–0.03**
	(0.01)	(0.01)	(0.01)	(0.02)
ln(Input-MA)		–0.01		0.01
(Weight: $1/dist^{1.1}$)		(0.01)		(0.02)
Adj. R^2	0.19	0.24	0.38	0.43
# clusters (region)	54	54	45	41
# clusters (Ownership)	704	699	382	339
N	1263	1258	432	388
Region-Sector FE	Y	Y	Y	Y
Ownership FE	N	Y	N	Y

Notes. The table reports OLS estimates of Equation (4.4). When constructing the spillover measure defined in Equation (4.2), we lag firms' adoption status by 4 years. In Panel A, we use the subsample which only include firms that did not adopt any technology until the end of the sample period. In Panel B, we use the full sample that include both adopters and non-adopters and additionall control adopters' adoption status. Dependent variables are log sales and revenue TFP in columns (1)-(2) and (3)-(4) respectively. Revenue TFP is estimated based on Wooldridge (2009). ln(Spill-Sales) and ln(int-MA) are additional controls defined in Equations (4.5) and (4.6). In all specifications, we control for region-sector fixed effects and initial dependent variable at the start of the sample period. Standard errors are two-way clustered at both region and ownership levels and are reported in parenthesis. * p<0.1, ** p<0.05, *** p<0.01.

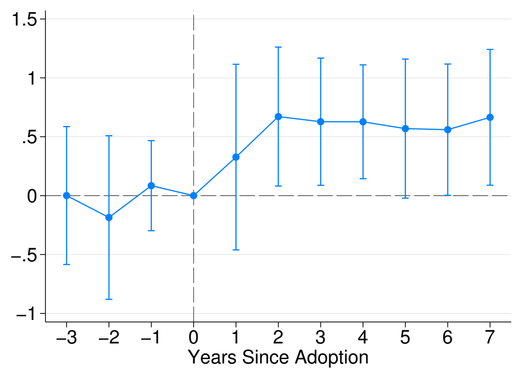
D.2 Additional Figures



A. log revenue TFP-ACF (2015)



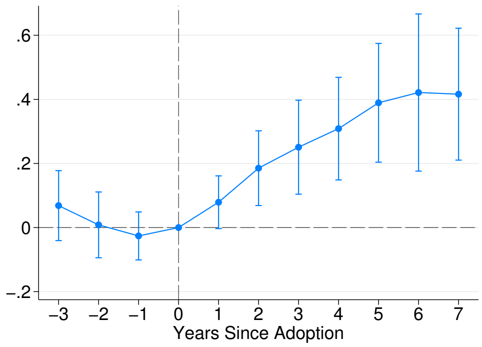
B. log revenue TFP-LP (2003)



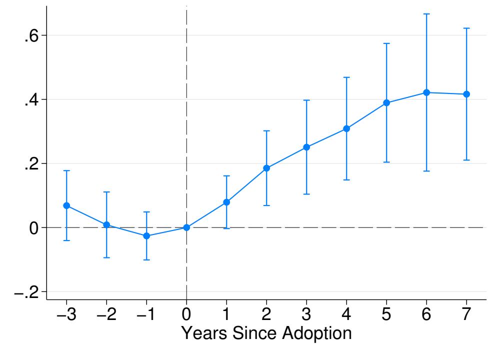
C. log revenue TFP-OLS

Figure D1. Direct Productivity Gains of Technology Adoption: “Winners vs. Losers” Research Design - Robustness, Alternative TFP measures

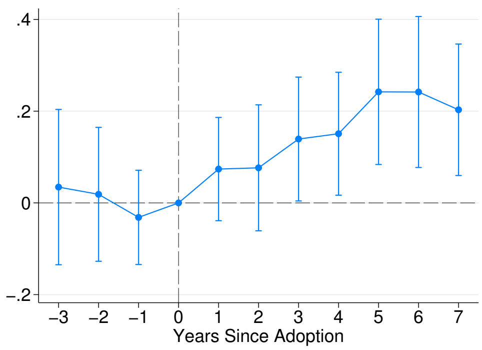
Notes. The figure illustrates the estimated β_{τ}^{diff} in Equation (4.1) based on “winners vs. losers” research design. The dependent variables are log revenue TFP. In Panels A, B, and C, revenue TFPs are estimated based on Ackerberg et al. (2015), Levinsohn and Petrin (2003), and OLS. β_0^{diff} is normalized to be zero. All specifications control for event time dummies and firm, pair, and calendar year fixed effects. The short-dashed red line is the pooled diff-in-diffs estimate of the impact of technology adoption. Error bars represent 95 percent confidence intervals based on standard errors, two way clustered at pair and firm level.



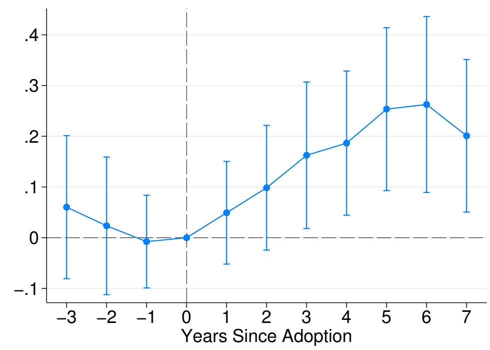
A. log sales



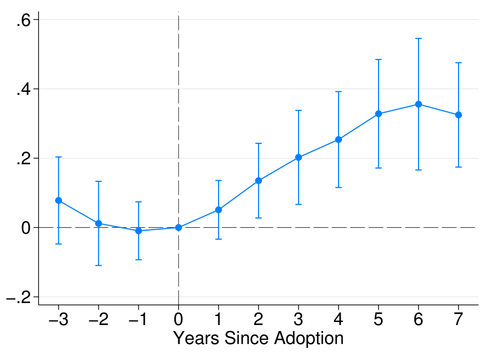
B. log revenue TFP-W (2009)



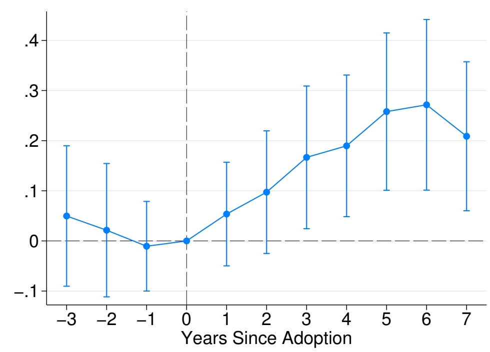
C. log labor productivity



D. log revenue TFP-ACF (2015)



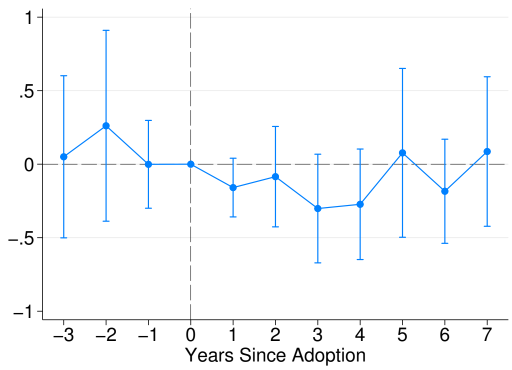
E. log revenue TFP-LP (2003)



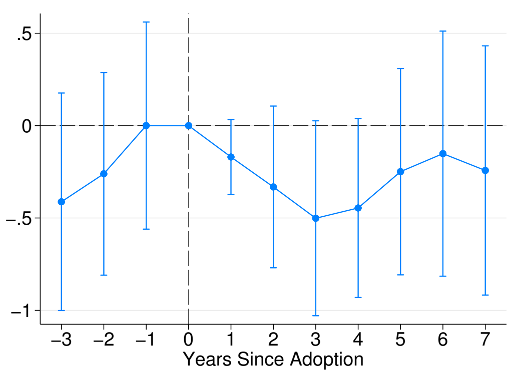
F. log revenue TFP-OLS

Figure D2. Direct Productivity Gains of Technology Adoption: Nearest Neighbors Matching

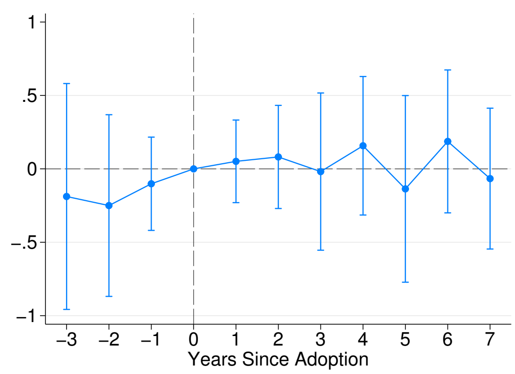
Notes. The figure illustrates the estimated β_{τ}^{diff} in Equation (4.1) based on nearest neighbors matching. In Panels A, B, C, D, E, and F, dependent variables are log sales, log revenue TFP based on Wooldridge (2009), log labor productivity, log revenue TFP based on Ackerberg et al. (2015), log revenue TFP based on Levinsohn and Petrin (2003), and log revenue TFP based on OLS. β_0^{diff} is normalized to be zero. All specifications control for event time dummies and firm, pair, and calendar year fixed effects. The short-dashed red line is the pooled diff-in-diffs estimate of the impact of technology adoption. Error bars represent 95 percent confidence intervals based on standard errors, two way clustered at pair and firm level.



A. log employment



B. log fixed assets



C. log fixed assets per worker

Figure D3. Inputs and Technology Adoption: "Winners vs. Losers" Research Design

Notes. The figure illustrates the estimated β_τ^{diff} in Equation (4.1) when dependent variables are log employment, fixed assets, and fixed assets per worker. β_0^{diff} is normalized to be zero. All specifications control for event time dummies and firm, pair, and calendar year fixed effects. The short-dashed red line is the pooled diff-in-diffs estimate of the impact of technology adoption. Error bars represent 95 percent confidence intervals based on standard errors, two way clustered at pair and firm level.

D.3 Matching Algorithm

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

1. Pick two subsets of variables $\mathbf{X}^e \in \mathbf{X}$ that are matched exactly and $\mathbf{X}^d \in \mathbf{X}$ that are distance matched.
2. For each loser f, pick an adopter g such that
   - have the same values of the variables of X_e with a loser f;
   - minimize the Mahalanobis distance

   $$\text{adopter}_g \in argmin\{((X^d_f - X^d_g)'\mathbf{S}^{-1}(X^d_f - X^d_g)\}$$

   where $\mathbf{S}$ is the sample covariance of $\mathbf{X}^d$ and $\mathbf{X}^d_f$ are the variables of firm f.

D.4 Production Function Estimation

In this section, we discuss the estimation procedure of TFP measures. TFP measures are obtained as the residual after estimating the production using different methodologies: Wooldridge (2009), Levinsohn and Petrin (2003), Ackerberg et al. (2015), and OLS. We estimate Cobb-Douglas value-added production function:

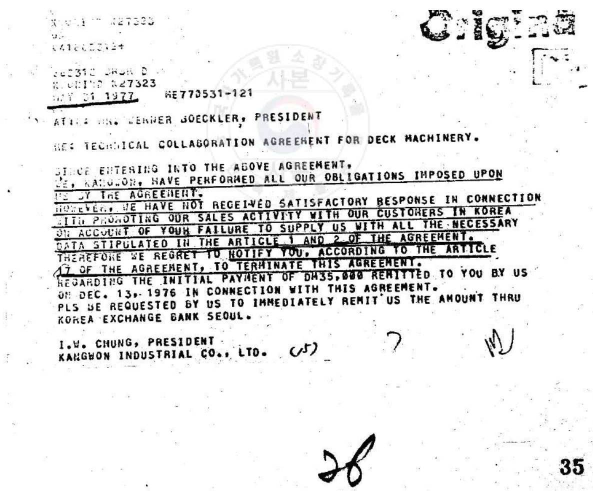
$$\log VA_{it} = \alpha_L \log L_{it} + \alpha_K \log K_{it} + u_{it}. \tag{D.1}$$

When using methodologies developed by Wooldridge (2009), Levinsohn and Petrin (2003), and Ackerberg et al. (2015), we use material inputs as a proxy variable.

The material input data is not available for our data set. Therefore, we estimate the production function separately for each sector using the alternative firm-level data. We use *KIS-VALUE* between 1980 and 1990. The *Act on External Audit of Joint-Stock Corporations*, introduced in 1981, requires the Korean firms whose assets were above 3 billion Korean Won were required to report their balance sheet data, which is the source for *KIS-VALUE*. The coverage of our data set is larger than *KIS-VALUE*. Also, because we only observe sales, when estimating the production function, 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 between 1970 and 1982.

D.5 Example of a Loser

Figure D4 reports an example of a loser. *Kangwon Industrial Co.* (*Kangwon*) and *Broehl Maschinen Fabric GmbH* (*Broehl*) made a contract regarding deck machinery. *Broehl* was a german firm. Although *Kangwon* paid fixed fee in advance, *Broehl* did not sent a blueprint. Panel A is the official english document related to termination of the contract between two firms. Panel B is the official Korean document in which Kangwon reported to the government that the contract failed due to *Broehl*. In the official Korean document, *Kangwon* mentions that it asked *Broehl* several times for fulfillment of the contract.



MAY 31 1977 RE770531-121

ATTN: MR. WERNER GOECKLER, PRESIDENT

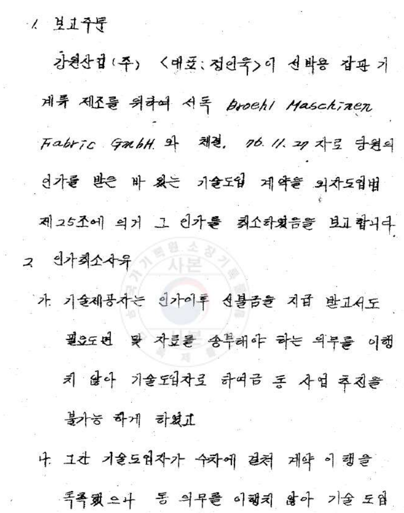
RE: TECHNICAL COLLABORATION AGREEMENT FOR DECK MACHINERY.

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 RESPONSE 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. Official document on termination of the contract between *Kangwon* and *Broehl*



1. 보고주문

강원산업(주) <대표:정인욱>이 선박용 갑판 기계류 제조를 위하여 서독 Broehl Maschinen Fabric GmbH.와 체결, 76. 11. 27자로 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

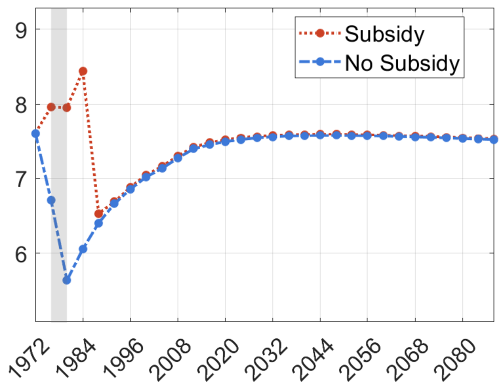
나. 그간 기술도입자가 수차에 걸쳐 계약 이행을 독촉했으나 동 의무를 이행치 않아 기술 도입

B. Official document that *Kangwon* reported to the government

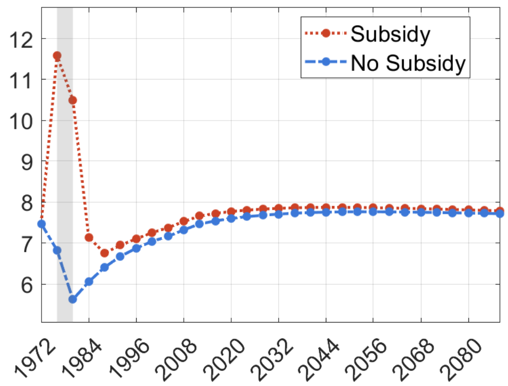
Figure D4. Example a Loser. Contract between *Kangwon Industrial Co.* and *Broehl Maschinen Fabric GmbH*

Appendix E Appendix: Quantification

E.1 Additional Figures



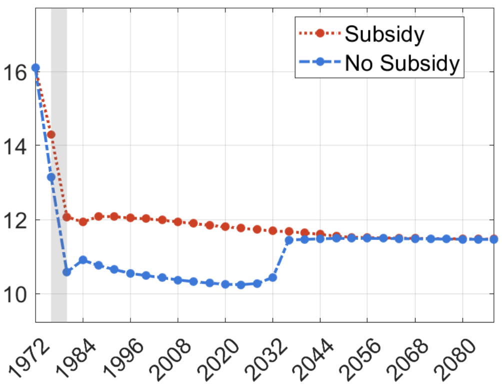
A. Low η



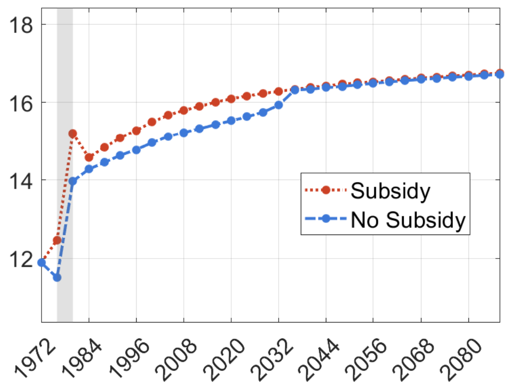
B. Low δ

Figure E1. Heavy Mfg. GDP Shares. Comparative Statistics of δ and η

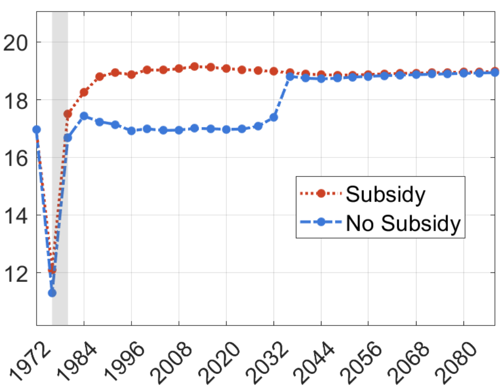
Notes. This figure plots the comparative statistics of δ and η. In Panel A, η is set 1.05. In Panel B, δ is set to be 1. The red dotted line and the blue dashed lines plot outcomes of the baseline and counterfactual economies.



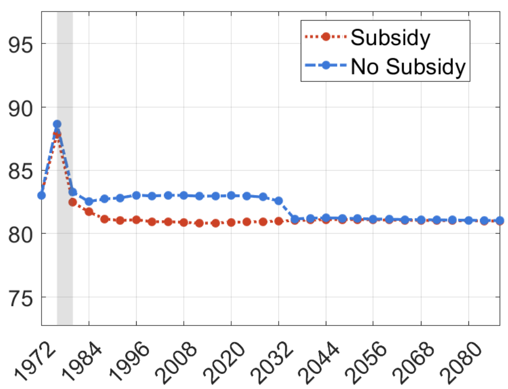
A. Heavy mfg. GDP share (%)



B.Heavy mfg. employment share (%)



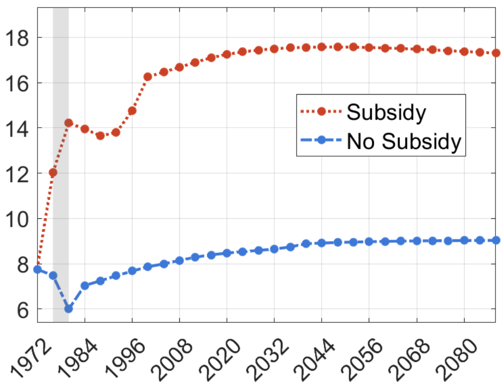
C. Heavy mfg. export share (%)



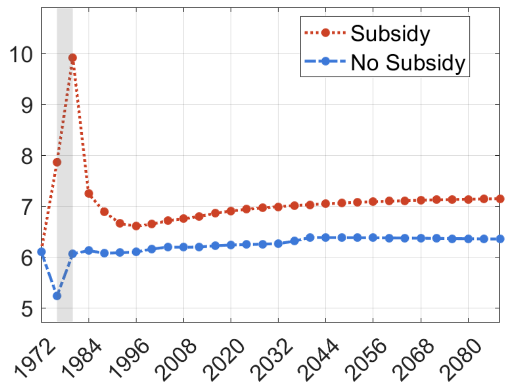
D. Light mfg. export share (%)

Figure E2. The Role of a Roundabout Production Structure and the Temporary Subsidies.

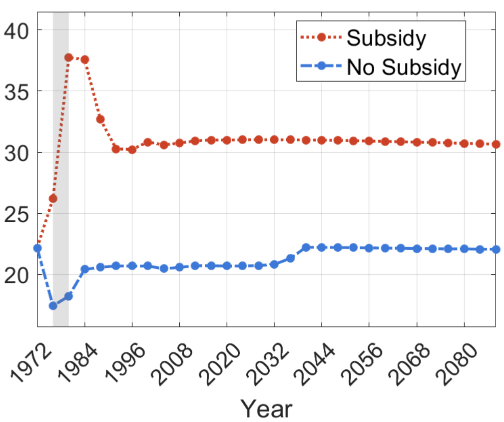
Notes. This figure plots the counterfactual results without a roundabout production structure. Panels A, B, C, and D report the results for the heavy manufacturing sector employment, GDP, and export shares, and the light manufacturing sector export shares respectively. The red dotted and the blue dashed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



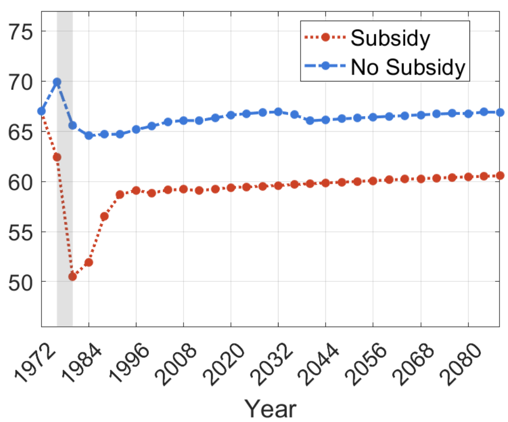
A. Heavy mfg. GDP share (%)



B.Heavy mfg. employment share (%)



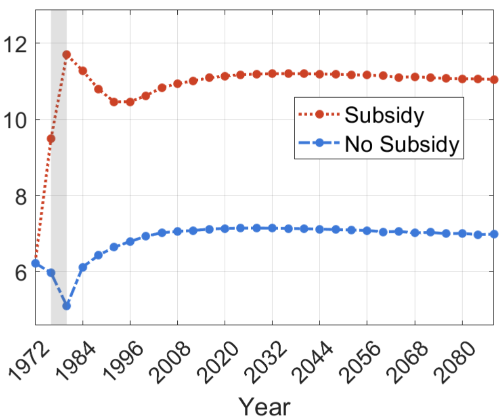
C. Heavy mfg. export share (%)



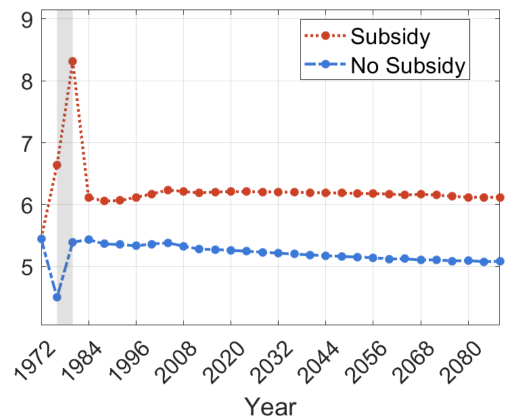
D. Light mfg. export share (%)

Figure E3. The Role of Migration Costs and the Temporary Subsidies.

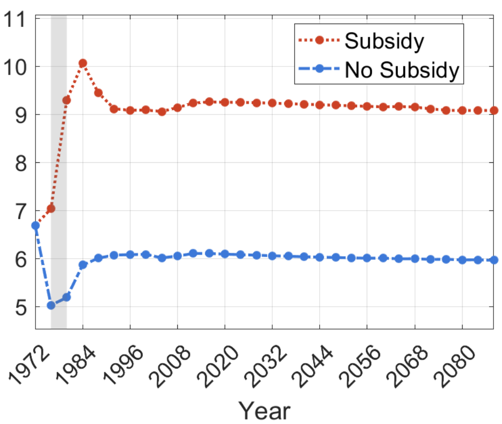
Notes. This figure plots the counterfactual results with a higher level of migration costs than the baseline counterfactual exercises. Panels A, B, C, and D report the results for the heavy manufacturing sector employment, GDP, and export shares, and the light manufacturing sector export shares respectively. The red dotted and the blue dashed lines plot outcomes of the baseline and counterfactual economies.



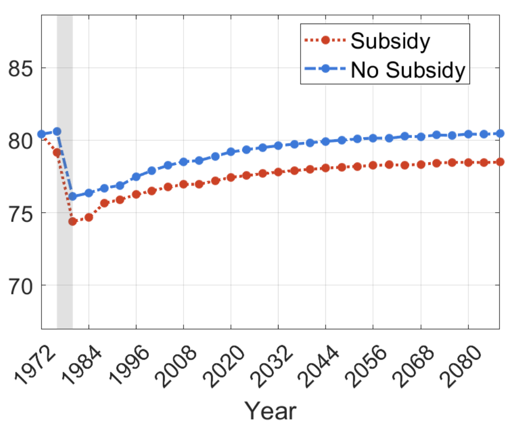
A. Heavy mfg. GDP share (%)



B.Heavy mfg. employment share (%)



C. Heavy mfg. export share (%)



D. Light mfg. export share (%)

Figure E4. The Role of Foreign Market Size and the Temporary Subsidies.

Notes. This figure plots the counterfactual results with a lower level of foreign market size than the baseline counterfactual exercises. 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 and the blue dashed linesplot outcomes of the baseline and counterfactual economies.

E.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 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 national level
  5. Sectoral import shares at 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 the total expenditure on sector j goods at 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 $\hat{\beta}^{policy}$, Equation (6.6)
  2. Median of light & heavy mfg. shares of exports in 1972 across regions
  3. Median of heavy mfg. shares of adopters in 1972 & 1982 across regions
  4. Share of zero adoption regions in 1972 & 1982

Algorithm. Taken 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. We fit region-sector level aggregate outcomes to the data counterparts. This step corresponds to the constraints of Equation (6.4). 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

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

   (b) Update new $\{c'_{fj}\}$ using

$$\underbrace{\frac{\mathrm{IM}^{Data}_{jt}}{\mathrm{E}^{Data}_{jt}}}_{\text{Data}} = \underbrace{\frac{\sum_{n\in\mathcal{N}}\left(\tau^{x}_{nj}t^{im}_{jt}c^{f'}_{jt}/P_{njt}\right)^{1-\sigma}E_{njt}}{\sum_{n\in\mathcal{N}}E_{njt}}}_{\text{Model}}$$

   (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 $\{V_{nt}'\}$ is identified from the above equation, so we normalize the value of the amenity of the first region to be one for each period, $V_{1t}' = 1, \forall t$.

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

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

where

$$\bar{\phi}_{njt}^{avg} = \frac{\theta f(\lambda_{njt-1}^{T})(\phi_{njt}^{min'})^{\sigma-1}}{\tilde{\theta}(1-\kappa^{-\theta})}\left\{\left(\left(\frac{\eta}{1-\bar{s}_{jt}}\right)^{\sigma-1}-1\right)(\tilde{\lambda}_{njt}^{T})^{\frac{\tilde{\theta}}{\theta}}+\left(1-\left(\frac{\eta}{1-\bar{s}_{jt}}\right)^{\sigma-1}\kappa^{-\tilde{\theta}}\right)\right\},$$

$$\begin{aligned}\bar{\phi}_{njt}^{avg,x} = {} & \frac{\theta f(\lambda_{njt-1}^{T})(\phi_{njt}^{min'})^{\sigma-1}}{\tilde{\theta}(1-\kappa^{-\theta})}\frac{(\tilde{\lambda}_{njt}^{x})^{\frac{\tilde{\theta}}{\theta}}}{\lambda_{njt}^{x}} \\ & \times\left\{\left(\left(\frac{\eta}{1-\bar{s}_{jt}}\right)^{\sigma-1}-1\right)\left(\frac{\tilde{\lambda}_{njt}^{T}}{\tilde{\lambda}_{njt}^{x}}\right)^{\frac{\tilde{\theta}}{\theta}}+\left(1-\left(\frac{\eta}{1-\bar{s}_{jt}}\right)^{\sigma-1}\kappa^{-\tilde{\theta}}(\tilde{\lambda}_{njt}^{x})^{-\frac{\tilde{\theta}}{\theta}}\right)\right\},\end{aligned}$$

$\tilde{\lambda}_{njt}^{T} = (1-\kappa^{-\theta})\lambda_{njt}^{T}+\kappa^{-\theta}$, and $\tilde{\lambda}_{njt}^{x} = (1-\kappa^{-\theta})\lambda_{njt}^{x}+\kappa^{-\theta}$. The above equations only identify the relative levels of $\{\phi_{njt}^{min'}\}$, so we normalize the Pareto lower bound parameter of the first region-sector to be 1 for each period, $\phi_{11t}^{min'} = 1, \forall t$.

4. After updating the fundamentals, given parameters, evaluate the following objective function:

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

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 $\{\mathbf{\Theta}^M, \bar{s}_{jt}\}$, we iterate steps 2, 3, and 4 and find $\{\mathbf{\Theta}^M, \bar{s}_{jt}\}$ that minimizes the objective function.

E.3 Construction of Data Inputs for Calibration Procedure

In this section, we describe how we constructed data inputs for the calibration procedure. We aggregate 10 manufacturing to 2 sectors: light and heavy manufacturing sectors.

Aggregate Data

Initial Shares of Adopters in 1968. Our firm-level balance sheet data covers between 1970 and 1982, whereas technology adoption contracts covers between 1966 and 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 to their adoption activities and construct shares of adopters in the heavy manufacturing sector for each region.[99]

Regional Population Distributions in 1968, 1972, 1976, and 1980. The regional population data comes from Population and Housing Census, which is 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 using the two observed samples. For example, the population share of region n in 1973 is imputed as Pop. $\text{share}_{n,1973} = (\text{Pop. share}_{n,1970})^{\frac{3}{5}} \times (\text{Pop. share}_{n,1975})^{\frac{2}{5}}$. From these imputed values, we obtain population distribution 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.

Region-Sector Level Gross Output in 1972, 1976, and 1980. We compute gross output at region-sector level by harmonizing firm-level data and the input-output table. Using the firm-level data, we calculate a share of sales of firms in region n and sector j and then multiply this share with the gross output of sector j at the national-level. Specifically, we calculate

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

where GO_{jt}^{IO} is sector j's gross output from the input output table. 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 table.

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

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

[99] Given that we cannot observe entry and exit of firms in 1968 and 1969 and we construct the shares based on the firms which survived between 1968 and 1970, this constructed shares are likely to overestimate the actual shares of adopters.

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

Export and Import Tariffs Data in 1972, 1976, and 1980. Export and import tariffs data is not used for the reduced-form empirical analysis but only for the quantitative exercises. The export tariffs data is obtained from Magee (1986).[100] The original data set's industry code is in 4-digit 1972 SIC code. It is first converted into 4-digit 1987 SIC codes and then converted into ISIC Revision 3 codes.[101]

Import tariffs data is digitized from Luedde-Neurath (1986) for 1974, 1976, 1978, 1980, and 1982, which are in the Customs Cooperation Council Nomenclature (CCCN). CCCN is converted into ISIC Revision 3, and then it is averaged across 4-digit ISIC codes. For missing years, we impute values using the geometric average. We assume that the tariff level in 1972 is the same as that in 1974.

We aggregate trade tariffs up to the 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 actual magnitude of the tariffs. However, the quantitative results are not affected regardless of weighting or not.

E.3.1 Micro moments

We can compute shares of adopters for each year using our data set. After computing these shares across regions and years, we compute the median for 1972 and 1980. Also, u sing this information, we can compute shares of regions with zero values. Shares of exporters are similarly obtained. However, because of many missing samples on exports, we take the three year moving averages on shares of exports for each region-sector. We count firms with missing information on exports as non-exporters. Section E.4 describes calculating the identifying moment in more detail.

[100]The US export tariff data was downloaded from https://cid.econ.ucdavis.edu/ust.html.
[101]The concordance between 1972 SIC and 1987 SIC is obtained from "www.nber.com."

E.4 Identifying Moment for Subsidy

E.4.1 Calibration Procedure

Using data on shares of adopters of the heavy manufacturing sector across regions in 1972 and 1980, we run the following regression model via PPML:

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

where D^{policy}_t is a dummy variable which equals one in 1980; and $\lambda^T_{n,heavy,t}$ is a heavy manufacturing share of adopters of region n in period t. One period of the model corresponds to 4 years in the data, so $\lambda^T_{n,heavy,t-1}$ is lagged by 4 years. Standard errors are clustered at regional level. The estimated coefficients are reported in Table E1.

Table E1: Identifying Moment for Subsidy

Dep. Var. $\lambda^T_{n,heavy,t}$	
	(1)
D^{policy}_t	0.65^{**}
	(0.25)
$\lambda^T_{n,heavy,t-1}$	5.62^{***}
	(0.80)
# of clusters (region)	42
N	84

Notes. The table reports OLS estimates of Equation (E.1). Standard errors are clustered at region level and are reported in parenthesis. * p<0.1, ** p<0.05, *** p<0.01.

Using model-generated data, we run the same regression model. Following Silva and Tenreyro (2006), we calculate the estimate of β^{policy} by solving the first order condition of log-likelihood of PPML:

$$\sum_{n \in \mathcal{N}} \mathbf{X}_{nt}(\lambda^T_{n,heavy,t} - exp(\mathbf{X}'_{nt}\boldsymbol{\beta})), \tag{E.2}$$

where $\boldsymbol{\beta} = [\alpha, \beta^{Policy}, \beta_1]'$ and $\mathbf{X}_{nt} = [1, D^{policy}_t, \lambda^T_{n,,heavy,t-1}]$.

E.5 Gravity Equation of Migration Flows

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 the data availability, regions are aggregated up to 35 regions. $\mu_{nm,1990-1995}$ is obtained as the total number of migrants moving from region n to region m between 1990 and 1995 divided by the total population of region n in 1990. When computing the total number of population and migrants, we restrict our sample age between 20 and 55. We also exclude both outward migration flows from Jeju island and inward migration flows to Jeju island.

We parametrize migration costs as a function of distance between two regions $dist_{mn}$ and an error term ϵ^d_{mnt} that is orthogonal to distance between two regions: $d_{mn} = (dist_{mn})^{-\zeta}\epsilon^d_{mnt}$. Taking log of Equation (5.10), we can derive the following regression model:

$$\ln \mu_{mnt} = -\nu\zeta \log dist_{mn} + \underbrace{\ln\left(V_{nt}\frac{(1-\bar{\tau}^w_t + \bar{\pi}^h_t)w_{nt}}{P_{nt}}\right)}_{=\delta_n} + \underbrace{\ln\left(\sum_{n'=1}^{N}\left(V_{n't}\frac{(1-\tau^w_t + \bar{\pi}^h_t)w_{n't}}{P_{n't}}d_{mn'}\right)^{\nu}\right)}_{\delta_m} + \epsilon^d_{mnt},$$

which gives Equation (6.3). We estimate the above equation using OLS and PPML. The results are reported in Table E2.

Table E2: Gravity Equation of Migration Shares

Dep. Var.	Migration Shares between 1990 and 1995	
	OLS	PPML
	(1)	(2)
$\log Dist_{mn}$	–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. The table reports the gravity estimates of Equation (6.3). Dependent variable is log of migration share from region m to region n between 1990 and 1995. In column (1), we estimate the model using OLS. In column (2), we estimate the model using the Poisson pseudo likelihood estimation (Silva and Tenreyro, 2006). Clustered errors are two-way clustered at origin and destination levels. * p<0.1, ** p<0.05, *** p<0.01.