

Technology Adoption and Late Industrialization*

Jaedo Choi

University of Michigan, Ann Arbor

Younghun Shim

University of Chicago

JOB MARKET PAPER

November, 2021

[Click here for the latest version](#)

Abstract

We study how the adoption of foreign technology and the local spillovers from such adoption contributed to late industrialization in a developing country during the postwar period. Using novel historical firm-level data for South Korea, we provide 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. Because of this 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 our model to the micro data and econometric estimates. We evaluate the effects of the Korean government policy that temporarily provided adoption subsidies to heavy manufacturing firms in the 1970s. Had no adoption subsidies been provided, Korea would have converged to a less industrialized steady state in which the heavy manufacturing sector's share of GDP would have been 15% lower and aggregate welfare would have been 10% lower than the steady state with successful industrialization. Thus, temporary subsidies for technology adoption had permanent effects.

Keywords: Technology adoption, industrialization, knowledge spillover, path dependence

JEL Codes: O14, O33, O53, R12

*We are grateful to Andrei A. Levchenko, Dominick Bartelme, Jagadeesh Sivadasan, and Sebastian Sotelo for their guidance and invaluable suggestions and encouragement. We are indebted to Javier Cravino, Alan Deardorff, Kyle Handley, and Linda Tesar for their insightful comments. We thank Maria Aristizabal-Ramírez, Barthélémy Bonadio, Ricardo Hausmann, Yongjun Kim, Bunyada (Mos) Laoprapassorn, Kiminori Matsuyama, Nishaad Rao, Ulrich Schetter, and seminar participants at the Harvard Growth Lab, 2021 Midwest International Trade Meetings, 2021 ETSG, 2021 LACEA-LAMES, and University of Michigan for helpful comments and discussions. We also thank the National Archives of Korea for sharing data. We gratefully acknowledge research support from the Kwanjeong Educational Foundation. E-mail: jaedochi@umich.edu.

1 Introduction

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

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

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

This paper makes three contributions. First, we overcome the empirical challenge in the literature by constructing a novel historical data set that covers the universe of technology adoption contracts between Korean and foreign firms. Most of the adopted technology during this period was related to knowledge about how to build and operate plants and capital equipment related to mass production.

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

²See Romer (1990). Recent studies provide empirical evidence about 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, 2021).

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

⁴Building 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 the adoption of foreign technology.

⁵See Lucas (1993).

Using this data set, we can measure firm-level technology adoption directly.

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

We also provide empirical evidence about local productivity spillovers of adoption. The key identification challenge when estimating the spillovers is that spatially correlated shocks affect both firms' performance and their neighbors' adoption decisions ([Manski, 1993](#)). We deal with this challenge by exploiting spatial variation at a fine level of geographic detail. The median land area of our geographic unit of analysis is the size of Manhattan, or almost 23 square miles. Within each region and 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 the firm-level within each region and sector 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 control for these fixed effects, our results are driven by variation in distances to adopters of the same sector within regions instead of being driven by variation across regions and sectors, so the usual regional or sectoral unobservables are not a concern in our empirical analysis. We find that non-adopters' sales and productivity grew faster when more neighboring firms had adopted foreign technology. Our estimates indicate that when the local share of adopters increased by 1%, the sales and revenue TFP of non-adopters increased by 4-5%.

Third, we construct a dynamic spatial general equilibrium model with heterogeneous firms' technology adoption decisions and local productivity spillovers. We use the model to evaluate the general equilibrium effect of the Korean government policy of subsidizing technology adoption by domestic firms. Firms' adoption decisions and the spillover endogenously shape comparative advantage and export patterns at the regional and national levels. Firms can adopt a more productive modern technology after incurring a fixed adoption cost. The spillover operates with a one-period lag, where

the current local productivity increases in the local share of adopters in the previous period. This time lag of the spillover is a source of dynamics in the model. Because of this time lag, the share of adopters becomes a time-varying state variable. The spillover generates a dynamic complementarity in firms' adoption decisions. A higher share of adopters in the previous period leads to higher gains from adoption that in turn induces more firms to adopt technology in the current period. Because adopters do not internalize this spillover, the amount of adoption is suboptimal. This justifies appropriate policy interventions that promote adoption.

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

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

Using the calibrated model, we ask how the pattern of industrialization in Korea would have evolved had the government not provided subsidies. Our results show that if subsidies had not been provided, Korea would have converged to a less industrialized steady state. In the steady state of this counterfactual economy, the heavy manufacturing sector's share of GDP would have been 10% lower, exports would have been 22.5% lower, and employment would have been 2% lower than the steady state of the baseline economy where subsidies had been provided. Also, the aggregate welfare would have been 10% lower. The aggregate differences are driven by a few regions that become more

productive because of subsidy-induced technology adoption.

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

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

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

Finally, this paper contributes to the trade literature on the evolution of comparative advantage. Aggregate data show that comparative advantage evolves ([Hausmann and Klinger, 2007](#); [Hanson et al., 2015](#); [Levchenko and Zhang, 2016](#); [Schetter, 2019](#); [Atkin et al., 2021](#)), but the understanding of what drives the evolution of comparative advantage has been limited so far.⁶ Using detailed micro data, [Pellegrina and Sotelo \(2021\)](#) document how knowledge diffusion through migration shaped the comparative advantage of Brazil, and [Arkolakis et al. \(2019\)](#) study the role immigrants played in diffusing knowledge about new technology in the United States in the nineteenth century. We

⁶For 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.

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 we used for our empirical and quantitative analysis. Section 3 describes the historical background of 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 analysis of 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 on firms’ technology adoption activities. We link these two data sets based on firms’ names. The resulting data set include only firms in the manufacturing sectors. We classified firms into 10 manufacturing sectors, 4 of which are heavy manufacturing. The sample period of the constructed data set is 1970 to 1982. The final data set has 7,223 unique firms of which 49% are heavy manufacturing.

The final data set includes 1,698 contracts made by 628 unique firms.⁷ Of these, 1,361 contracts and 457 firms were in heavy manufacturing sectors. Most of the adopted technologies were related to know-how about how to install or operate capital equipment or turnkey plants.⁸ 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.⁹ We describe our data construction procedure in Appendix Section A in more detail.

Firm-Level Technology Adoption Contracts. We hand-collected and digitized firm-level data on technology adoption from official documents related to domestic firms’ technology contracts with foreign firms from National Archives of Korea and from the [Korea Industrial Technology Association \(1988\)](#). These documents had information about names of domestic and foreign contractors and contract years. The law required domestic firms to submit related documents when they signed technology adoption contracts with foreign firms.¹⁰

⁷There were 1,776 contracts in total in the raw contract data set. However, 78 of them could not be matched with our balance sheet data. This gives us 1,698 contracts.

⁸Specifically, about 74% of technology adoption contracts provided the know-how, 21.2% granted licenses, and 4% permitted the use of trademarks. For example, Appendix Figure A1 is one page of the contract document between Kolon (Korean) and Mitsui Toatsu (Japanese), both of which are chemical manufacturers. The contract shows that Mitsui Toatsu had to provide technical assistance and blueprints to Kolon.

⁹The ratio between the total sum of firm sales of the data and the gross output from the input-output tables is reported in Appendix Figure A2. Also, see Appendix Table A2 for descriptive statistics of the data.

¹⁰Any domestic firms’ transactions with foreign firms, including technology adoption contracts, were strictly regulated under the Foreign Capital Inducement Act, which was first enacted in 1966. According to the law, once a domestic firm got approval from the government for the adoption, it had to report the related information to the Economic Planning Board which played a central role for economic policy-making process in Korea during the sample period.

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

3 Historical Background of Late Industrialization in South Korea

In late 1972, the Korean government launched the Heavy and Chemical Industry (HCI) Drive to modernize and promote heavy manufacturing sectors, including chemicals, electronics, machinery, steel, non-ferrous metal, and transport equipment. One of the main policy instruments was subsidies for adopting foreign industrial technology.¹¹ In the 1970s, the adoption of foreign technologies and imported capital equipment related to those technologies were the main means of technology transfer from foreign developed economies to South Korea.¹²

The timing of the policy that subsidized technology adoption and the selection of the targeted sectors were driven by a political shock rather than economic conditions (Lane, 2019).¹³ After the Vietnam War, President Nixon changed the diplomatic policy of the United States toward its East Asian allies. In the Nixon Doctrine (1969), he declared that the East Asian allies of the United States, including South Korea, should take primary responsibility for their self-defense instead of relying on the US military. He also planned the complete withdrawal of the US military from South Korea. However, at this time, military tension between South and North Korea was rising. Because South Korea was heavily reliant on the US military, the Nixon Doctrine posed a threat for the national defense of South Korea. In late 1972, in order to modernize South Korea’s military forces and achieve self-reliant defense against North Korea, President Park of South Korea announced the drive to promote the heavy and chemical manufacturing sectors that are related to the arms industry.¹⁴ The

¹¹For example, Hyundai Motors, the largest automotive company in Korea, did not have its own models until 1972. It merely reassembled the existing car model developed by Ford and imported most of the automobile parts. Hyundai Motors did not start to produce its own models until 1972, when it became possible because of technology adoption. In 1974, Hyundai Motors hired George Turnbull, the former director at British Leyland as a new vice-president in order to improve its management technology. In 1976, Hyundai Motors adopted engine technology from Perkins Engine, design from Ital Design, and transmission technology from Mitsubishi, which are British, Italian, and Japanese firms, respectively. The government subsidized Hyundai Motors to enable it import new capital equipment and construct new turnkey plants related to the technologies it had adopted. See Choi and Levchenko (2021) for how the Korean government subsidized firms during the 1970s.

¹²Another commonly used means of technology transfer in developing countries is foreign direct investment (FDI) (Keller, 2004). In Korea, however, FDI did not play a big role. The Korean government strictly regulated FDI, and the total value of the technologies and capital equipment domestic firms imported was 22 times greater than that of FDI. Moreover, when compared to other developing countries, Korea had a lower stock of FDI. For example, the value of Korea’s stock of FDI was only 7 percent of the value of Brazil’s stock (Kim, 1997, p.42-43).

¹³In 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 further background on Korea’s HCI Drive.

¹⁴At the same time, President Park declared martial law and amended the country’s constitution into an authoritarian



A. Heavy mfg. share of GDP (%)



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

Figure 1. Late Industrialization and Technology Adoption in Korea

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

government considered Korea's underdeveloped technology in heavy manufacturing sectors as one of the national threats.¹⁵ Given Korea's large technology gap with the world frontier, the government deemed technology adoption to be the most effective way to catch up with the frontier.¹⁶ 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. While at the beginning of the period of our analysis, Korea's heavy manufacturing share was only 6%, it achieved a remarkable takeoff during the sample period, surpassing Mexico by the mid-1970s and the United States by 1982. The right panel plots the yearly number of new adoption contracts between Korean and foreign firms. Our novel data set reveals that the yearly number of contracts between South Korean and foreign firms for new technology quadrupled in the period between 1970 and 1982. This sudden and rapid increase in the rate of adoption coincided with temporary government subsidies for technology adoption in South Korea from 1973 to 1979.¹⁷ Even after the policy ended in 1979, the Korean economy continued to specialize in the heavy

document, called the Yushin constitution, that extended his term of office.

¹⁵“Without rapidly improving our underdeveloped technology, our nation will be unable to secure an independent national defense system ... Inevitably, we will face a decline in our competitiveness of exports goods in international markets and national power, which bodes ill for our chance of a peaceful reunification with North Korea.” (Ministry of Science and Technology, 1972, p. 3)

¹⁶“Considering our nation's current technological state, adopting foreign advanced technologies and continuously adapting them to our needs seem to be the most effective catching-up strategy.” (Ministry of Science and Technology, 1972, p. 4)

¹⁷In Appendix Figure B1, we report heavy manufacturing's share of employment and export and the measure of revealed comparative advantage (Balassa, 1965). Consistent with Figure 1, the employment share increased from 4%

manufacturing sectors.

4 Reduced-Form Evidence on Technology Adoption

In this section, we examine how technology adoption benefited Korean firms. We provide econometric evidence on direct productivity gains for adopters and local productivity spillovers for non-adopters. According to the historical narrative, large-sized Korean firms tend to rely on foreign sources to acquire advanced technologies, whereas small-sized firms on reverse engineering of technologies adopted by neighboring firms or on hiring experienced engineers from local adopters to obtain new technologies.¹⁸ Our econometric evidence on the direct gains and local spillovers capture the former and the latter, respectively.

4.1 Direct Productivity Gains to Adopters

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

We define winners (the treated) as firms that successfully adopted technology from foreign firms. We define losers (the comparison) as non-adopters that made contracts with foreign firms that got approved by the government but were not able to adopt foreign technology because the foreign firm canceled the contract for reasons that had nothing to do with the South Korean firm. Examples include cancellations due to bankruptcy or to changes in the management team of the foreign firm. We exclude cancellations by domestic firms. The reasons for these cancellations include a domestic firm's sudden decreases in cash flow. See Appendix Figure A3 for an example of a cancellation by a loser. When contracts were canceled after approval from the government, domestic firms had to report the related documents on the reason for the cancellation. We collect data on contract cancellations by reading thousands of historical documents from the archives.

After identifying losers, we match each loser with an adopter using the exact Mahalanobis matching algorithm. The matching proceeds in two steps. First, we exactly match on region and sector in order to absorb shocks within regions and sectors, such as market size or local wages. Second,

to 9%, the export share increased from 13.5% to 35%, and the revealed comparative advantage measure rose from 0.2 to 0.65.

¹⁸See [Kim and Kim \(1985\)](#) and [Kim \(1997\)](#). For instance, during the 1970s, there were 15 firms producing black-and-white TV producers. The first four large firms started producing TV after adopting foreign technologies, but the other 11 acquired technologies by hiring experienced engineers from the first four adopters ([Kim, 1997](#), p. 156). See [C.4](#) for historical case studies.

within regions and sectors, we choose a winner that was most similar to a loser in terms of firm size measured by log assets, where the similarity is measured by the Mahalanobis distance. We match losers and winners with replacement, so we can match one winner to multiple losers in a given year if they were in the same sector and region. The matching procedure gives us 34 pairings among 57 unique firms. All the matched pairs consist of heavy manufacturing firms. See Appendix Section D.3 for more detail on the matching procedure.

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

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

where i denotes firm, p pair, and t time. D_{pt}^{τ} are event-study variables defined as $D_{pt}^{\tau} := \mathbb{1}[t - \tau = t(p)]$, where $t(p)$ is event year of pair p . $\mathbb{1}[Adopt_{it}]$ is a dummy variable for adoption status. δ_i , δ_p , and δ_t are firm, pair, and year fixed effects. ϵ_{ipt} is an error term. Dependent variables y_{ipt} are log sales, log revenue TFP estimated based on Wooldridge (2009), and labor productivity defined as value added per worker.¹⁹ Matching with replacement introduces mechanical correlation across residuals, because of the possible appearance of the same firm. Thus, we two-way cluster standard errors at the level of both firms and pairs.

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

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

Because we do not find differential pre-trends between winners and losers (which will be shown below), the second condition of our identifying assumption would be violated only by unobservable shocks that affected losers' performance after the event and were correlated with foreign firms' cancellations, but did not affect losers' performance before the event. One example would be a

¹⁹Appendix Section D.4 describes our revenue TFP estimation procedure in more detail.

negative shock of losers at the time of the event that caused losers to be matched with a bad foreign contractor that experienced a change in its management teams or went into bankruptcy. We can directly test this using firm-to-firm structure of our technology adoption contract data. If our results are driven by matching based on negative shocks, we would expect that characteristics of foreign firms that made contracts winners and losers to be different. Also, we believe that shocks that satisfy the above conditions were very unlikely. These losers are firms that self-select into the adoption. Negative shocks that hit losers at the time of the event would have caused them not to adopt technology rather than adopt technology from bad foreign firms.

Balance. To assess covariate balance between two groups, we report descriptive statistics of the matched pairs and covariate balance test results. The 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.²⁰ In Appendix Table D2, we report the results of covariate balance tests where we estimate a linear probability model of the effects of pre-event firm observables on adoption status. Across all specifications, none of the estimated coefficients of firm observables are statistically different from zero both individually and jointly once we control for pair fixed effects. These results indicate that firm observables cannot predict the cancellations of losers, which supports our identifying assumption that cancellations by foreign firms were exogenous shocks to domestic firms.

We compare two groups of foreign firms that made contracts with winners and losers based on their patenting activities in the US. We obtain data on patenting activities in the US from the United States Patent and Trademark Office (USPTO). We use firms' patenting activities in the US as a proxy for how these firms are close to the world technology frontier. When these foreign firms made contracts, Appendix Table D3 shows that none of the t-statistics of tests that various measures of patent activities of two groups are equal are statistically significant. This rules out an alternative story that negative shocks made losers be matched with bad foreign firms.

Baseline Results. Table 1 and Figure 2 report the estimated coefficients in Equation (4.1). There is no pre-trend. Winners' sales, revenue TFP, and labor productivity did not begin to increase until adoptions occurred. On average, the adoption increased sales by 51%, revenue TFP by 46%, and labor productivity by 64%.²¹

²⁰Both winners and losers were larger than the average of all heavy manufacturing firms. For example, the average log sales of all heavy manufacturing firms was 15.54, but the averages of winners of losers were 17.80 and 18.46, respectively (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.

²¹This average is calculated by estimating the following pooled diff-in-diffs model:

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

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

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

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

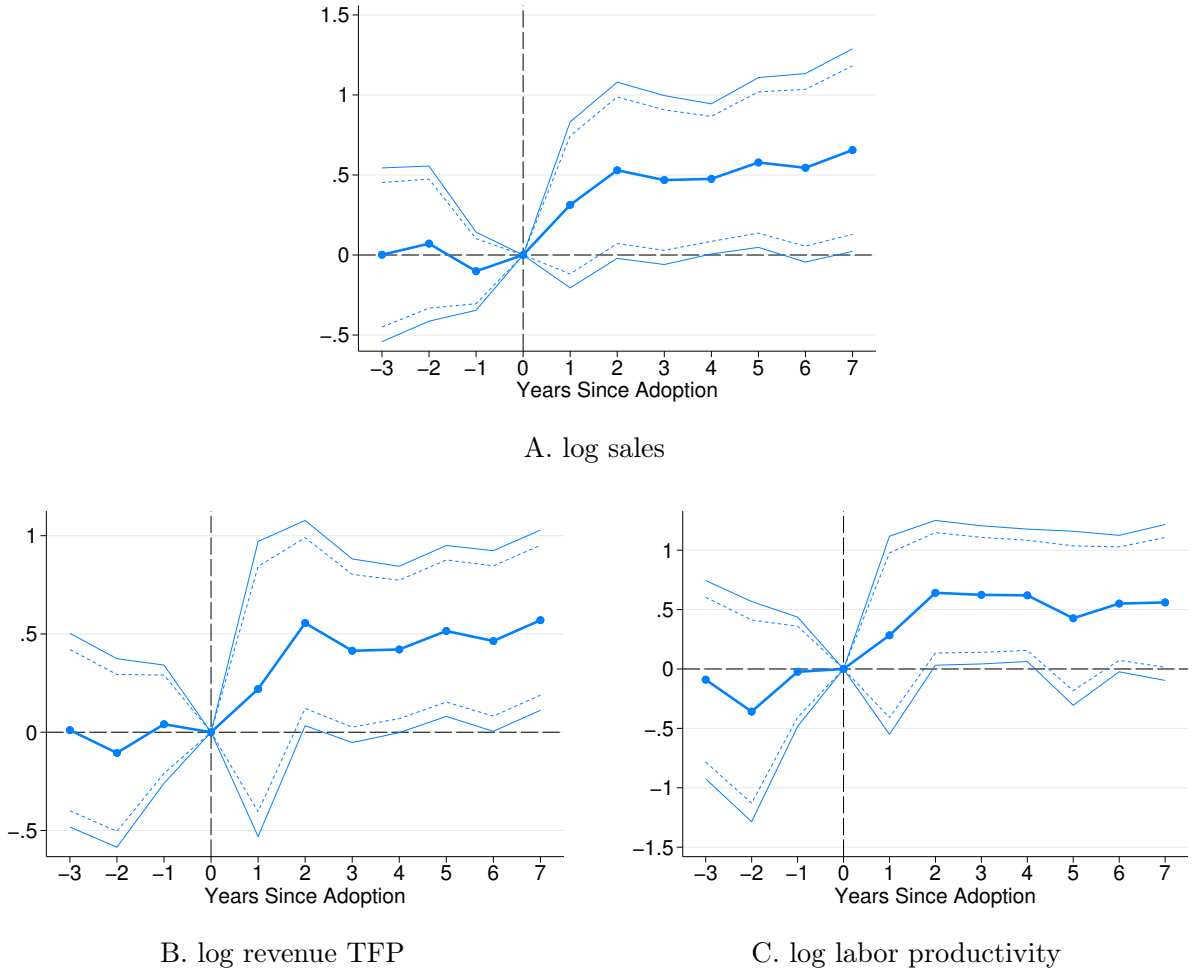


Figure 2. Direct Productivity Gains to Adopters: Winners vs. Losers Design

Notes. This figure illustrates the estimated β_{τ}^{diff} in Equation (4.1) based on winners vs. losers research design. In Panels A, B, and C, the dependent variables are log sales, 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. The plotted coefficients correspond to columns (1)-(3) of Table 1. The figure reports 90 and 95 percent confidence intervals based on standard errors two-way clustered at the levels of pairs and firms.

Robustness. We run the same regressions using different revenue TFP measures to check for robustness.²² The results are reported in Appendix Figure D1 and columns (4)-(6) of Table 1. Even though we use different measures, the estimated event study shows no pre-trend, and the estimated coefficients were within a standard error of the estimates of column (3) of Table 1.

We also run the regression using log employment and fixed assets (inputs for production) as dependent variables. The results are reported in Appendix Figure D2. We do not find that winners increase their use of inputs relatively more than losers do after the event. This shows that winners' increases in sales or revenue TFP measures are not input-driven and supports our interpretation of these increases as direct productivity gains to adopters, respectively.

4.2 Local Productivity Spillovers to Non-Adopters

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

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

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

The spillover measure can be interpreted as the probability that firm i 's manager would meet other managers who worked in firms that had adopted foreign technologies. Each manager is endowed with a unit of time and can randomly meet at most one manager from other firms. The probability that a manager would meet a manager from firm k is given by the inverse of the distance between firms i and k . The inverse of the distance is a proxy for spatial frictions that would have impeded local interaction between managers of two firms.²³ The spillover measure captures the fact that

where Post_{it} is an indicator for period t after the event and other variables are defined as in Equation (4.1). We report the average as the estimated coefficient of β^{diff} . The estimation results are reported in Appendix Table D4.

²²We use revenue TFP estimates based on Akerberg et al. (2015), Levinsohn and Petrin (2003), and OLS.

²³By taking the weighted average, we implicitly assume that the spillover measure is invariant to the total number of firms. As far as we know, there is no consensus about the functional form of knowledge spillovers (Combes and Gobillon, 2015; Gibbons et al., 2015). However, we think the weighted average is more suitable in our setting. First,

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

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

$$y_{injt} = \beta^S \text{Spill}_{inj(t-h)} + \delta_i + \delta_{njt} + \epsilon_{injt}, \quad (4.3)$$

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

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

$$\Delta y_{injt} = \beta^S \Delta \text{Spill}_{inj(t-h)} + \gamma y_{injt_0} + \mathbf{X}'_{injt_0} \boldsymbol{\beta} + \Delta \delta_{njt} + \Delta \epsilon_{injt}, \quad (4.4)$$

where Δ is a time difference operator. All specifications include the 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 the levels of regions and conglomerates. In Korea, large conglomerate groups known as chaebols own multiple firms across sectors and regions. Clustering at the conglomerate level allows for arbitrary correlation of error terms between firms within the same conglomerate group.

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

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

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 and Tonetti, 2014; Perla et al., 2021). Given that managers' time is a limited resource in the real world, this theoretical interpretation seems to be more natural than an alternative scenario where a manager can interact with all firms in the same local area. In this scenario, the spillover varies depending on the total number of adopters rather than the shares. Second, the literature on externalities has commonly used averages to capture agglomeration forces, such as share of skilled labor (Moretti, 2004) and population density (Ciccone and Hall, 1996).

close to large firms can lead to omitted variable bias.

We deal with the first concern by controlling for time-varying region-sector fixed effects at a fine level of geographic detail. The median size of our geographical unit of analysis for the sample is about Manhattan-sized, which is much finer than the unit of analysis in many previous studies. Our identifying variation comes purely from distance to adopters within the same sector and region, but not from variation across regions or sectors. Variation of $\text{Spill}_{inj(t-h)}$ 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.²⁴ Because we control for δ_{njt} and difference out δ_i , neighboring firms' adoption decisions based on time-varying region-sector factors do not bias our estimates. Only adoption or entry decisions based on time-varying firm-specific factors that are spatially correlated at neighborhood level would bias our estimates. For example, infrastructure improvement at the neighborhood level that affected both firms' outcomes and adoption decisions would bias our estimates. Exploiting spatial variation at a fine level mitigates potential spatial correlation at the neighborhood level within region-sector.

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

$$\ln(\text{Spill-Sales}_{inj t}) = \ln \left(\sum_{k \in nj/\{i\}} \left\{ \frac{(1/\text{dist}_{ik})\text{Sales}_{kt}}{\sum_{k' \in nj/\{i\}} (1/\text{dist}_{ik'})} \right\} \right). \quad (4.5)$$

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

$$\ln(\text{Input-MA}_{inj t}) = \ln \left(\sum_{j'} \sum_{k \in nj'/\{i\}} \gamma_j^{j'} (1/\text{dist}_{ik}) \text{Sales}_{kt} \right), \quad (4.6)$$

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

²⁴The firms that enter began production between the start and the end of the sample period affected the spillover measure of firms that were in production at the beginning of the sample period, but we did not include them in the sample because we restrict the sample to firms that were operating at the start of the period.

Table 2: Local Productivity Spillovers from Technology Adoption

Dep. Var.	log sales					log revenue TFP				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
<i>Panel A: Never-Adopter Sample</i>										
Spill	4.39*** (1.54)	3.79** (1.64)	4.94*** (1.70)	4.23*** (1.50)	4.07** (1.76)	5.55*** (1.84)	5.41*** (1.62)	5.81*** (2.08)	5.34*** (1.78)	5.11** (1.92)
ln(Spill-Sales)			-0.02 (0.01)		-0.02 (0.01)			-0.02 (0.02)		-0.01 (0.02)
ln(Input-MA)				-0.03 (0.02)	-0.02 (0.02)				-0.04** (0.02)	-0.03 (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 (conglomerate)	636	630	636	636	630	324	275	324	324	275
N	1079	1073	1079	1079	1073	344	292	344	344	292
<i>Panel B: Full Sample</i>										
Spill	4.23*** (1.18)	3.93*** (1.43)	4.45*** (1.31)	3.86*** (1.19)	3.72** (1.52)	4.75*** (1.63)	3.99** (1.90)	4.72*** (1.73)	4.45*** (1.58)	3.44* (1.82)
$1[Adopt]$	0.32** (0.15)	0.26 (0.20)	0.32** (0.15)	0.31** (0.15)	0.25 (0.19)	0.15* (0.09)	0.14 (0.10)	0.15* (0.09)	0.14 (0.09)	0.12 (0.10)
ln(Spill-Sales)			-0.01 (0.01)		-0.01 (0.01)			0.00 (0.02)		0.00 (0.02)
ln(Input-MA)				-0.05*** (0.02)	-0.04* (0.02)				-0.05*** (0.02)	-0.05** (0.02)
Adj. R^2	0.19	0.24	0.19	0.19	0.24	0.37	0.43	0.37	0.38	0.43
# clusters (region)	54	54	54	54	54	45	41	45	45	41
# clusters (conglomerate)	702	697	702	702	697	381	338	381	381	338
N	1264	1259	1264	1264	1259	431	387	431	431	387
Region-Sector FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Conglomerate FE		✓			✓		✓			✓

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

Baseline Results. Table 2 reports the OLS estimates for β^S . Column (1) of Panel A is our baseline estimate.²⁵ The estimated coefficient is statistically significantly positive. One standard deviation increase in the spillover of adoption contributes to 14.5% increases in sales.²⁶ 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. In this interpretation, a 1% increase in the local share of adopters leads to a 4.39% increase in non-adopters’ sales in the hypothetical region. In columns (2), (3), and (4), we also control for conglomerate fixed effects, $\ln(\text{Spill-Sales})$, and $\ln(\text{Input-MA})$, respectively.²⁷ In column (5), we control for all other variables. The estimates with additional controls all stay within a standard error of the baseline estimate. The estimated coefficients of $\ln(\text{Spill-Sales})$ and $\ln(\text{Input-MA})$ are not statistically significant and do not take positive values.²⁸ In columns (6)-(10), we use log revenue TFP as a dependent variable.²⁹ 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 for a dummy variable for own adoption status. Because they are likely to be correlated with the error term, we do not meaningfully interpret this variable. The estimates based on the full sample are within a standard error of the baseline estimates in column (1) of Panel A.

Robustness. We provide a battery of robustness checks. Instead of using the spillover measure with a four year lag, we use the spillover measure with three year or five year lags. The results are reported in Appendix Tables D5 and D6.

Instead of using $(1/\text{dist}_{ik})$, we also consider alternative weights $(1/\text{dist}_{ik}^\alpha)$ 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). The results are reported in Appendix Table D9.

Instead of using log sales or revenue TFP, we use 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 are positive for different dependent variables except for employment.

²⁵The magnitude of the estimated coefficients is consistent with the estimates in the literature on local knowledge spillover. The estimates in column (1) of Panel A indicate that the elasticity of firms’ sales to the spillover at the mean and the 90th and 95th percentiles is 0.05, 0.13, and 0.26, respectively. We calculate the elasticity of the adoption spillover as follows. The mean level of the local share of adopters is 0.011. An increase of 1% of the mean level (0.00011) increases firms’ sales by 0.05% ($= 100 \times 0.00011 \times 4.39$). The elasticities at the 90th and 95th percentiles are calculated similarly. For example, estimates from Bloom, Nicholas et al. (2013) imply that the elasticity of firms’ sales to their spillover measure based on patents is 0.19–0.26. We calculate elasticity above the 90th percentile because shares of adopters are highly skewed; where the 75th, 90th, 95th, and 99th percentiles were 0, 0.03, 0.06, and 0.18, respectively.

²⁶This is calculated as $14.5 = 100 \times 0.033 \times 4.39$, where 0.033 is one standard deviation of the adoption spillover.

²⁷When we control for the fixed effects of conglomerates, we categorize independent firms into one single group.

²⁸One potential explanation for the null results of $\ln(\text{Spill-Sales})$ and $\ln(\text{Input-MA})$ is that knowledge spillovers decay more quickly with distances than other spatial interactions captured by the other two controls, and the other two spatial interactions operate at a broader spatial scale than knowledge spillovers. Then, conditional on δ_{njt} , $\ln(\text{Spill-Sales})$ and $\ln(\text{Input-MA})$ will not have significant results.

²⁹The number of samples for revenue TFP was smaller because employment data is not available until 1972. We use only one set of differences between 1972 and 1980.

5 Theoretical Framework

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

5.1 Setup

We consider a small open economy Home with N regions and J sectors. We divide the world into Home and Foreign. Home is small because it cannot affect Foreign aggregates. However, its domestic prices are determined by domestic supply and demand conditions, and Home firms face downward sloping demands from Foreign.³⁰ Subscripts $n, m \in \mathcal{N}$ index Home regions, and $j, k \in \mathcal{J}$ sectors, where \mathcal{N} and \mathcal{J} are the sets of Home regions and sectors. Time is discrete and indexed by $t \in \{1, 2, \dots\}$.

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

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

5.2 Firms

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

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

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

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

³⁰The small open economy setup of our model is similar to that of Bartelme et al. (2021).

elasticity of substitution (CES) aggregator:

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

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

The exact CES price index is

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

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

Technology Adoption and Exports. In each period, firms make two static decisions: (i) whether to adopt advanced technology and (ii) whether to export. Both adopting technology and exporting incur adoption and export fixed costs in units of input bundles (F_j^T and F_j^x). The fact that both adoption and export costs are fixed costs make firms' decisions static. This static nature of firms' problems makes the model computable while preserving rich cross-sectional regional heterogeneity, and connecting the model to the data and econometric estimates. Once firms decide to adopt technology and pay fixed adoption costs, they can increase their productivity (Yeaple, 2005; Lileeva and Trefler, 2010; Bustos, 2011).³¹

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

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

where $\eta > 1$ is direct productivity gains from adoption, T_{it} is a binary adoption decision, $f(\lambda_{njt-1}^T)$

³¹ F_j^T is a reduced-form parameter that includes direct payment to foreign sources, the costs of installing a new structure or capital equipment related to a newly adopted technology, and any barriers to adoption. Many previous papers have studied sources of adoption barriers in developing countries (see, among many others, Parente and Prescott, 1994; Banerjee and Dufo, 2005; Acemoglu et al., 2007; Atkin et al., 2017). Also, South Korea's political context in the 1970s might have affected F_j^T . Due to the Cold War, the 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 South Korean firms (Vogel, 1991, p.8).

is a local adoption spillover that increases in the share of adopters in the previous period λ_{njt-1}^T , and ϕ_{it} is exogenous productivity. We allow the spillover to operate with a one-period lag (Allen and Donaldson, 2020), which is more realistic given that our focus is the transformation of the Korean economy within 10 years instead of the long-run outcomes that have been studied more frequently in the trade literature.³² When making adoption decisions, adopters internalize the direct productivity gain η but not the spillover $f(\lambda_{njt-1}^T)$. These externalities mean that social returns to adoption are larger than private returns. This leads to adoption rates that are lower than the socially optimum level. For sectors where technology adoption is not available ($j \notin \mathcal{J}^T$), firms' productivity consists of only exogenous productivity: $z_{it} = \phi_{it}$.

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

$$f(\lambda_{njt-1}^T) = \exp(\delta \lambda_{njt-1}^T),$$

where $\delta > 0$ is the semi-elasticity of firm productivity with respect to a local share of adopters. Under this parametrization, we show that δ can be mapped to the reduced-form spillover estimate in Section 4.2. The spillover can be micro-founded based on (1) local diffusion of new engineering knowledge and (2) learning externalities and labor mobility across firms. These two sets of microfoundations are based on historical case studies of Korea in the 1970s.³³ Complete derivations of the microfoundations and related historical cases are described in Appendix C.4.

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

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

which is parametrized by ϕ_{njt}^{max} , ϕ_{njt}^{min} , and θ . We also assume that the gap between the lower and upper bounds of the distribution is the same across regions, sectors, and periods: $\phi_{njt}^{max} = \kappa \phi_{njt}^{min}$,

³²Allowing the spillover to operate with a lag gives an economy a deterministic static equilibrium in each period (Adserà and Ray, 1998). This is a desirable theoretical property for two reasons. First, we can rule out unrealistic situations where an economy swings from one equilibrium to another in a different period depending on agents' self-fulfilling beliefs. See Krugman (1991) and Matsuyama (1991) for further discussion of self-fulfilling beliefs. Second, because there is a unique static equilibrium for each period, the model can be easily mapped to cross-sectional data. In general, multiple static equilibria models suffer from identification issues due to multiplicity (Jovanovic, 1989). Kline and Moretti (2014) and Allen and Donaldson (2020) also allowed agglomeration to operate with some lags.

³³The historical evidence shows that new ideas and knowledge about adopted technologies were frequently transmitted to local capital goods producers through reverse engineering of capital equipment related to adopted technologies. Also, technical personnel of adopters moved frequently to other firms and their movement played an important role in diffusing knowledge about adopted technologies. It is further supported by higher aggregate labor mobility rates in Korea in the 1970s than those of Japan and the United States (Kim and Topel, 1995). Also, learning externalities and knowledge spillovers through labor mobility have been widely studied in the literature. For example, see Lucas (1988) for learning externalities of human capital; see Serafinelli (2019) for empirical evidence on the effects of labor mobility across firms on knowledge diffusion.

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.³⁴

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

The government imposes a labor tax (τ_t^w) to finance these subsidies.³⁵ We assume that the labor tax rate is constant across regions, so the after-tax wages in region n are $(1 - \tau_t^w)w_{nt}$. The government budget is balanced every period.

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

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

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

³⁴If $\kappa \rightarrow \infty$, the bounded Pareto distribution becomes unbounded Pareto. However, the unbounded Pareto distributional assumption cannot rationalize zeros because as productivity is unbounded, there is always a small share of firms that adopt technology regardless of the values of F_j^T . Helpman et al. (2008) also uses a bounded Pareto distributional assumption to rationalize zero trade flows across countries.

³⁵The assumption that the government finances its adoption subsidies through a labor tax is based on the labor market policies and the pro-business attitude of the authoritarian South Korean government in the 1970s. The government restricted firms' nominal wage growth to below 80% of the sum of inflation and aggregate productivity growth and enacted temporary provisions in 1971 to prohibit labor union activities (Kim and Topel, 1995). Also, see footnote 3 of Itskhoki and Moll (2019).

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 are exogenous foreign demands. $\pi^m(T_{it}; \phi_{it})$ are operating profits conditional on adoption status obtained from region m , and $\pi^d(T_{it}; \phi_{it}) = \sum_{m \in \mathcal{N}} \pi^m(T_{it}; \phi_{it})$ are the sum of all these profits from domestic regions. $\pi^x(T_{it}; \phi_{it})$ are operating profits in foreign markets conditional on adoption status.

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

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

$$\bar{\phi}_{njt}^x = \frac{\mu c_{njt} (\sigma c_{njt} F_j^x)^{\frac{1}{\sigma-1}}}{f(\lambda_{njt-1}^T) \left((\tau_{nj}^x)^{1-\sigma} D_{jt}^f \right)^{\frac{1}{\sigma-1}}}. \quad (5.5)$$

The adoption cutoff $\bar{\phi}_{njt}^T$ is determined at where profits when adopting technology and profits when not adopting are equalized³⁶:

$$\bar{\phi}_{njt}^T = \frac{\mu c_{njt} (\sigma c_{njt} F_j^T)^{\frac{1}{\sigma-1}}}{\left(\left(\frac{\eta}{1-s_{njt}} \right)^{\sigma-1} - 1 \right)^{\frac{1}{\sigma-1}} f(\lambda_{njt-1}^T) \left(\sum_{m \in \mathcal{N}} \tau_{nmj}^{1-\sigma} P_{mjt}^{\sigma-1} E_{mjt} + (\tau_{nj}^x)^{1-\sigma} D_{jt}^f \right)^{\frac{1}{\sigma-1}}}. \quad (5.6)$$

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

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

and a mass of adopters is obtained as $M_{njt}^T = M_{nj} \times \lambda_{njt}^T$. Similarly, a share of exporters is $\lambda_{njt}^x = 1 - G_{njt}(\bar{\phi}_{njt}^x)$ and a mass of exporters is expressed as $M_{njt}^x = M_{nj} \times \lambda_{njt}^x$.

Dynamic Complementarity. Spillovers generate dynamic complementarity in firms' adoption decisions: a higher share of adopters in the previous period increases gains from adoption in the current period. The dynamic complementarity operates in two ways. The first way drives from complementarity between market size and productivity increases from adoption (Bustos, 2011; Lileeva

³⁶ $\bar{\phi}_{njt}^x$ and $\bar{\phi}_{njt}^T$ satisfy $\pi^x(T_{it} = 1; \bar{\phi}_{njt}^x) = c_{njt} F_j^x$ and $\pi(T_{it} = 1, x_{it} = 1; \bar{\phi}_{njt}^T) = \pi(T_{it} = 0, x_{it} = 1; \bar{\phi}_{njt}^T)$, respectively.

and Treffer, 2010). Because stronger spillovers increase the productivity of one region relative to other regions, firms in the more productive region will have a larger market and larger gains from adoption due to scale effects. This complementarity further incentivizes firms in the more productive region to adopt technology in the current period, which in turn magnifies spillover in subsequent periods.

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

5.3 Households.

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

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

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

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 in 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 wages $(1 - \tau_t^w) w_{nt}$ and income from dividends $\bar{\pi}_t^h w_{nt}$.³⁷ We denote the ideal price index for households in region n using $P_{nt} = \prod_{j=1}^J P_{njt}^{\alpha_j}$.

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

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

³⁷Under the given portfolio structure, dividend per share is proportional to w_{nt} . Specifically, $\bar{\pi}_t^h = \left(\sum_{n \in \mathcal{N}} \sum_{j \in \mathcal{J}} \int_{\omega \in \Omega_{nj}} \pi(\omega) d\omega \right) / \left(\sum_{n \in \mathcal{N}} w_{nt} L_{nt} \right)$.

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

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

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

The shape parameter ν is migration elasticity that governs the responsiveness of migration flows to real income changes of destination.³⁸ The population of region n in period t is the sum of all migrants to region n from all other regions in time $t-1$. Therefore, the spatial distribution of population evolves according to the following law of motion:

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

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

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

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

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

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

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

$\bar{\phi}_{njt}^{avg}$ captures the average cost advantage of sector j firms in region n . $\bar{\phi}_{njt}^{avg}$ increases in amounts of adoption that decreases in $\bar{\phi}_{njt}^T$, higher s_{njt} , or higher λ_{njt-1}^T . The expression for the average productivity including subsidies for exporters ($\bar{\phi}_{njt}^{avg,x}$) can be defined similarly, but the lower bound

³⁸Higher 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.

is replaced with $\bar{\phi}_{njt}^x$ instead of ϕ_{njt}^{min} because of selection effects induced by fixed exporting costs.

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

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

Region n 's share of the total sector j expenditure on goods from domestic region m and from Foreign are expressed as³⁹:

$$\pi_{mnjt} = \left(\frac{\tau_{mnj} c_{mjt} / \bar{\phi}_{mjt}^{avg}}{P_{njt}} \right)^{1-\sigma} \quad \text{and} \quad \pi_{njt}^f = \left(\frac{\tau_{nj}^x c_{jnt}^f}{P_{njt}} \right)^{1-\sigma}. \quad (5.15)$$

Regional gross output for domestic expenditures R_{njt}^d and the total value of exports R_{njt}^x are expressed as:

$$R_{njt}^d = M_{nj} \left(\frac{\mu c_{njt}}{\bar{\phi}_{njt}^{avg}} \right)^{1-\sigma} \sum_{m \in \mathcal{N}} \tau_{nmj}^{1-\sigma} P_{mjt}^{\sigma-1} E_{mjt} \quad \text{and} \quad R_{njt}^x = M_{njt}^x \left(\frac{\mu \tau_{nj}^x c_{njt}}{\bar{\phi}_{njt}^{avg,x}} \right)^{1-\sigma} D_{jnt}^f. \quad (5.16)$$

The total regional gross output R_{njt} is the sum of R_{njt}^d and R_{njt}^x .

5.4 Equilibrium

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

$$\Psi_t = \{\phi_{njt}^{min}, V_{nt}, D_{jnt}^f, c_{jnt}^f\} \quad \text{and} \quad \mathbf{s}_t = \{s_{njt}\}.$$

The timing of this model is as follows. Given $\{\lambda_{njt-1}^T, L_{nt-1}\}$, Ψ_t , and \mathbf{s}_t , households make static consumption and migration decisions and firms make static adoption and export decisions in t . These decisions, production, consumption, and wages are determined by the static equilibrium in t , in which households maximize their utility, firms maximize their profits, and market clearing conditions are satisfied. $\{\lambda_{njt}^T, L_{nt}\}$, which are the outcomes of the static equilibrium in t , become the state variables in $t+1$ and so on.

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

³⁹Note that $\sum_{m \in \mathcal{N}} \pi_{mnjt} + \pi_{njt}^f = 1$ holds.

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

$$w_{nt}L_{nt} = \left[\sum_{j \in \mathcal{J}} \gamma_j^L \left(\frac{\sigma-1}{\sigma} R_{njt} + M_{njt}^T c_{njt} F_j^T + M_{njt}^x c_{njt} F_j^x \right) \right], \quad (5.17)$$

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:

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

where $r(\phi_{it})$ is firm i 's revenues.⁴⁰ The left hand side of this equation is total government tax revenue and the right hand side is total government spending.

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

$$E_{njt} = \sum_{k \in \mathcal{J}} \gamma_k^j \left(\frac{\sigma-1}{\sigma} R_{nkt} + M_{nkt}^T c_{nkt} F_k^T + M_{nkt}^x c_{nkt} F_k^x \right) + \alpha_j (1 - \tau_t^w + \bar{\pi}_t^h) w_{nt}L_{nt}. \quad (5.19)$$

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

$$R_{njt} = R_{njt}^x + \sum_{m \in \mathcal{N}} \pi_{nmjt} E_{mjt}. \quad (5.20)$$

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

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

We define the dynamic equilibrium of this economy as follows:

Definition 1. Given initial shares of adopters $\{\lambda_{njt_0}^T\}$ and the path of the geographic fundamentals Ψ_t and subsidies $\{s_{njt}\}$, a dynamic equilibrium is a path of wages $\{w_{nt}\}$, price indices $\{P_{njt}\}$, a set of functions $\{p_{inmjt}(\omega), q_{inmjt}(\omega), p_{injt}^x(\omega), q_{injt}^x(\omega), T_{it}(\omega), x_{it}(\omega)\}$, labor tax $\{\tau_t^w\}$, population $\{L_{nt}\}$, and shares of adopters $\{\lambda_{njt}^T\}$ such that

⁴⁰Firm i 's revenues in Home regions and Foreign are proportional to their operating profits in those settings: $r(\phi_{it})^d = \sigma \pi^d(T_{it}^*; \phi_{it})$ and $r(\phi_{it})^x = \sigma \pi^x(T_{it}^*; \phi_{it})$, where t_{it}^* is firm i 's optimal adoption decisions. Total revenue $r(\phi_{it})$ is the sum of $r(\phi_{it})^x$ and $r(\phi_{it})^d$.

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

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

5.5 Analytical Results: Multiple Steady States

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

Assumption 1. (i) $|\mathcal{N}| = |\mathcal{J}| = 1$ and $\tau_{nj}^x \rightarrow \infty$ (closed economy with one region and one sector); (ii) $M = 1$ (normalization); (iii) $\kappa \rightarrow \infty$ and $\phi_t^{min} = 1$ (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 an unbounded Pareto distribution with a normalized location parameter and firm mass is normalized to be one. Assumption 1(iv) is a source of dynamic complementarity in firms' adoption decisions in this environment. With only one region and the CES demand structure, the complementarity between market size and gains from adoption does not operate in this environment, and the dynamic complementarity comes only from fixed adoption costs in units of final goods.⁴¹ Assumption (v) is a sufficient condition to guarantee a unique static equilibrium each period.⁴²

⁴¹With only one region, each firm's increase in productivity due to spillovers is exactly canceled out by competition due to other firms' increases in productivity under the CES demand structure. Therefore, the overall increase in productivity through spillovers does not change the relative market size of each firm, which nullifies the complementarity between market size and gains from adoption. However, because fixed adoption costs are in units of final goods, the spillover from the previous period lowers fixed adoption costs in the current period, which further incentivizes more firms to adopt technology in the current period and generates dynamic complementarity. At the other extreme, when fixed adoption costs are in units of labor, there is no dynamic complementarity, and the equilibrium share of adopters λ_t^{T*} is not affected by λ_{t-1}^T , because overall productivity increase induced by the spillover increases labor demand, which in turn increases wages and the total fixed adoption cost $w_t F^T$. This is formally stated in Appendix Section C.2.3. Although we assume that fixed adoption costs are only in units of final goods for simplicity, as long as parts of fixed adoption costs are in units of final goods, the model generates dynamic complementarity.

⁴²When a fixed adoption cost is in units of final goods, and $\sigma \leq 2$ holds, multiple static equilibria can arise each period regardless of the existence of the spillover. This is because firms do not take the aggregate price index into account when making adoption decisions. When the share of adopters becomes larger, the aggregate price index becomes lower and this, in turn, decreases fixed adoption costs and vice versa. This degree of responsiveness of the price index to the share of adopters decreases in the values of the elasticity of substitution σ . When σ is sufficiently low, two static equilibria can arise where one has a higher share of adopters and the other has a lower share. By imposing $\sigma > 2$, we

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 equation⁴³:

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

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

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

Given any initial shares of adopters $\lambda_{t_0}^T$, this economy has a unique deterministic equilibrium path to the steady state due to Assumption 1(v). Because static equilibrium is unique each period, there exists a unique sequence of static equilibrium that forms a unique deterministic dynamic path. λ_t^{T*} increases in λ_{t-1}^T due to dynamic complementarity. λ_t^{T*} also increases in two parameters: η and δ . η increases λ_t^{T*} by increasing the net gains for marginal adopters.⁴⁴ δ increases λ_t^{T*} by magnifying dynamic complementarity. Most importantly, we show that multiple steady can arise due to the dynamic complementarity in this economy. When multiple steady states exist, these steady states can be Pareto-ranked based on the steady state share of adopters, and the initial share of adopters $\lambda_{t_0}^T$ determines which steady state is realized in the long-run. These results are summarized in Proposition 1.

Proposition 1. *Under Assumption 1,*

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

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

⁴³ λ_t^{T*} is bounded by 1. $\lambda_t^{T*}(\lambda_{t-1}^T; \eta, \delta) = 1$ when $A(\hat{\lambda}_t^T(\lambda_{t-1}^T; \eta, \delta))^{2-\sigma} f(\lambda_{t-1}^T) \frac{\eta^{\sigma-1}-1}{\sigma F^T} \geq 1$.

⁴⁴ Note that two terms are related to the direct productivity gains governed by η in Equation (5.21): $(\eta^{\sigma-1} - 1)$ and $A(\lambda_t^T)^{2-\sigma}$. The term $(\eta^{\sigma-1} - 1)$ captures marginal adopters' net gains from adoption, which is internalized by their adoption decisions. The term $A(\lambda_t^T)^{2-\sigma} = A(\lambda_t^T)^{1-\sigma} \times A(\lambda_t^T)$ captures two composite general equilibrium effects of direct productivity gains. These two general equilibrium effects work in the opposite directions in the net gains for marginal adopters. First, $A(\lambda_t^T)^{1-\sigma}$ captures competition effects, which decrease in λ_t^T . As more firms adopt technology (increases in λ_t^T), the productivity of competitors increases, which in turn intensifies competition across firms and decreases marginal adopters' net gains due to increased competition (or decreased market size). The second general equilibrium effect is that a higher share of adopters decreases the price of final goods and therefore fixed adoption costs (Assumption 1(iv)). Assumption 1(v) ensures that the first general equilibrium effect dominates this second general equilibrium effect.

rium;

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

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

Because technology adoption increases firms' productivity, these steady states can be Pareto-ranked depending on the steady state share of adopters. At S^{Ind} , all firms adopt technology, and at S^{Pre} there is a smaller share of adopters than the other two, so S^{Ind} Pareto-dominates the other two steady states and S^{Pre} is Pareto-dominated by the other two.⁴⁶ S^U is unstable in that the economy converges to this steady state only when the initial condition is given by the value of S^U . The nonlinearity of the red locus means that it intersects with the 45-degree line multiple times and generates the multiple steady states, where the spillover ($\delta > 0$) generates such nonlinearity. For example, if there is no spillover ($\delta = 0$), there is always a unique steady state illustrated by the intersection of the green dashed horizontal line and the 45-degree line. When there is no spillover, the equilibrium share of adopters is determined regardless of the share of adopters in the previous period, which gives the horizontal line in the graph.

For the initial conditions given by $\lambda_{n_{jt_0}}^T \geq S^U$, the economy converges to S^{Pre} , and for $\lambda_{n_{jt_0}}^T \geq S^U$, the economy converges to S^{Ind} . Because firms do not internalize the spillover, if an economy is locked into the region $\lambda_{n_{jt_0}}^T \geq S^U$, it converges to S^{Pre} , although an economy has the potential to reach S^{Ind} . This region is known as a poverty trap in the literature (Azariadis and Stachurski, 2005).

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

⁴⁶The fact that all firms adopt technology at S^{Ind} is the artifact of the fact that λ_t^T is strictly convex in λ_{t-1}^T , which comes from the imposed functional form assumption of the spillover.



Panel A. Multiple Steady States and Nonlinearity



Panel B. The Role of Adoption Subsidies



Panel C. Comparative Statistics of δ



Panel D. Comparative Statistics of η

Figure 3. Multiple Steady States and Comparative Statistics

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

Multiple Steady States and the Permanent Effects of Temporary Subsidies. When multiple steady states exist, temporary subsidies for technology adoption in the initial period can have permanent effects that move an economy that is initially in the poverty trap to a new transition path that converges to an industrialized steady state. This is illustrated in Panel A of Figure 3. Suppose the initial condition is given as $\lambda_{njt_0}^T < S^U$, so that an economy converges to S^{Pre} . However, if the government implements a one-time policy that subsidizes technology adoption, this can shift the share of adopters above the S^U level, which causes an economy to converge to the industrialized steady state S^{Ind} . This can rationalize South Korea's pattern of industrialization toward heavy manufacturing sectors and the temporary policy between 1973 and 1979.

In this model, only multiple steady states can rationalize the permanent effects of temporary subsidies.⁴⁷ When there is only one steady state, subsidies temporarily shift the short-run equilibrium curve while they are provided, but the curve moves back to the original position after subsidies end and the economy converges to its original steady state.⁴⁸

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

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

6 Taking the Model to the Data

In our quantitative exercises, we aggregate sectors into four categories: commodity, light manufacturing, heavy manufacturing, and service sectors. Given that most of the adoption occurred in the heavy manufacturing sectors, we assume that technology adoption is only available in the heavy manufac-

⁴⁷Even if a unique steady state exists, there is room for policy interventions because of externalities. However, when a unique steady state exists, these policy interventions should be implemented in each period in order to produce permanent effects. This point is graphically illustrated in Appendix Figure C1.

⁴⁸Similarly, Kline and Moretti (2014), who studied the Tennessee Valley Authority program in the United States, did not detect nonlinearities in the agglomeration elasticity, so they concluded that the program had limited indirect gains through agglomeration.

turing sector. The service sector is nontradable across regions and countries. We also aggregate the data to 42 regions.⁴⁹ One period in the model corresponds to 4 years in the data, so the timing of the spillover in the model is consistent with the spillover estimates in Section 4.2.

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

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

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

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

Externally calibrated parameters

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

6.1 Externally Calibrated Parameters

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

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

which can be mapped to our winners vs. losers specification (Equation (4.1)). By exactly matching on region and sector, we absorb out the spillover, the unit cost of production, and the market size that are common across firms within regions and sectors.⁵⁰ Exogenous cancellations by foreign firms can

⁴⁹We 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 on electoral districts.

⁵⁰More precisely, we allow for variation in the spillover in our reduced-form regression model. In contrast, our model assumes that the spillover is common across firms within each region and sector. 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 ($\text{Spill}_{inj(t-h)} - D_{pt}^T$).

be interpreted as a shock to the fixed adoption cost F_j^T in our model framework that is orthogonal to firms' productivity $\log \phi_{it}$. We calibrate η using the average estimates of $\{\beta_{\tau}^{diff}\}_{\tau=1}^7$ in Equation (4.1) and set $\eta = \exp(0.51)/(\sigma - 1)$.⁵¹

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

$$\log Sales_{it} = (\sigma - 1)\delta\lambda_{njt}^T + \log \left(\sum_{m \in \mathcal{N}} \tau_{nmj} P_{mjt}^{\sigma-1} E_{mjt} + \tau_{nj}^x D_j^f \right) + (\sigma - 1) \log \phi_{it}.$$

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

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

$$\log \mu_{nm1990}^{1995} = -\nu\zeta \log \text{dist}_{nm} + \delta_m + \delta_n + \epsilon_{mn}^d, \quad (6.1)$$

where μ_{nm1990}^{1995} represent shares of migrants from region n to region m ; and δ_n and δ_m are region fixed effects.⁵² To address attenuation bias arising from statistical zeros in the gravity models, we estimate the equation using the Poisson pseudo-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 $\hat{\zeta} = 1.39/\nu$.⁵³

⁵¹0.51 is based on the pooled diff-in-diffs estimates reported in column (1) of Appendix Table D4.

⁵²The 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 data availability, regions are aggregated up to 35 regions. μ_{nm1990}^{1995} is obtained as the total number of migrants who moved from region n to region m in the period 1990 to 1995 divided by the total population of region n in 1990. When we compute the total number of population and migrants, we restrict our sample age to 20 to 55.

⁵³We find statistically significant results at 1% when we two-way cluster errors at origin and destination levels. The OLS estimates of Equation (6.1) is 1.30, which is similar to 1.39 obtained from the PPML. See Appendix Table E2 for the detailed gravity estimates of migration flows. These estimated values are consistent with the estimates by recent papers. For example, Pellegrina and Sotelo (2021) find the estimated elasticity of Brazilian migration flows to distance is -1.32.

Variable Trade Costs $\{\tau_{nmj}, \tau_{nj}^x\}$. We parametrize variable internal trade costs as a function of the geographic distance $\tau_{nmj} = (\text{dist}_{nm})^\xi$ and assume that ξ is the same across different sectors. We do not observe internal trade flows, so we borrow the estimates from the literature. We use the distance elasticity estimate from [Monte et al. \(2018\)](#) and set $\xi = 1.29/(\sigma - 1)$.⁵⁴

For international trade costs, we assume that to export or import, firms have to ship their products to the nearest port and then pay both variable and fixed international trade costs at the port. Under this assumption, international trade costs can be parametrized as $\tau^x = \tilde{\tau}^x \times (\text{dist}_n^{\text{port}})^\xi$. $\tilde{\tau}^x$ represent variable costs incurred at the port. We set $\tilde{\tau}^x$ to be 1.7 following [Anderson and Van Wincoop \(2004\)](#). $(\text{dist}_n^{\text{port}})^\xi$ represent variable costs associated with shipping from region n to the nearest port, where $\text{dist}_n^{\text{port}}$ is the distance between region n and the port and ξ is the same parameter of the parametrization of internal trade costs.⁵⁵

The Remaining Parameters $\{\sigma, \theta, M_{nj}, \alpha_j, \gamma_j^L, \gamma_j^k\}$. 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](#)).⁵⁶ We set M_{nj} to be proportional to the GDP share of each region and sector and set $\sum_{n \in \mathcal{N}} M_{nj} = 1$ following [Chaney \(2008\)](#).⁵⁷ The Cobb-Douglas shares of preference (α_j) and production function (γ_j^k and γ_j^L) are taken from the input-output table for 1972.

6.2 Internally Calibrated Parameters

$\Theta^M = \{F_j^x, F_j^T, \kappa\}$, $\mathbf{s}_t = \{s_{njt}\}$, and $\Psi_t = \{\phi_{njt}^{\text{min}}, V_{nt}, D_{jt}^f, c_{jt}^f\}$ 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 periods when the subsidies were provided between 1973 and 1979. The required set of aggregate variables include region-sector level gross output $\{R_{njt}\}$, regional population distribution $\{L_{nt}\}$, aggregate export and import shares, initial shares of adopters $\{\lambda_{nj-1}^T\}$ and initial population distribution $\{L_{n,-1}\}$. $\{\lambda_{nj-1}^T\}$ and $\{L_{n,-1}\}$ are taken as given when solving the model for $t = 1$.⁵⁸ Appendix Section E.2 explains the algorithm of the calibration procedure and how we construct the data inputs in detail.

⁵⁴Using internal trade flows within the United States, [Monte et al. \(2018\)](#) estimate a distance elasticity of -1.29 . We calibrate ξ to be $1.29/(\sigma - 1)$.

⁵⁵When computing the nearest distance to ports, we use seven main ports in South Korea: Busan, Incheon, Gunsan, Guje, Pohang, Ulsan, and Yeosu.

⁵⁶Under the Pareto distributional assumption with shape parameter θ , the distribution of firm sales follows the Pareto distribution with shape parameter $\theta/(\sigma - 1)$. Many previous studies that have estimated θ using firm sales distribution have found that $\theta/(\sigma - 1)$ is close to 1 ([Axtell, 2001](#); [di Giovanni et al., 2011](#); [di Giovanni and Levchenko, 2012, 2013](#)).

⁵⁷[Chaney \(2008\)](#) assumes that the total mass of firm of each country is proportional to its total value added in the multi-country setting.

⁵⁸The 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.

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

$$\begin{aligned} \{\hat{\Theta}^M, \hat{\mathbf{s}}_t\} \equiv \arg \min_{\Theta^M, \mathbf{s}_t} \{L(\Theta^M, \mathbf{s}_t)\} = & \underbrace{(\bar{m} - m(\Theta^M, \mathbf{s}_t, \Psi_t))' \mathbf{W} (\bar{m} - m(\Theta^M, \mathbf{s}_t, \Psi_t))}_{\text{Micro moments}} \\ \text{subject to } & \underbrace{\mathbf{C}(\Theta^M, \mathbf{s}_t, \Psi_t) = \mathbf{C}_t}_{\text{Aggregate data}}, \quad t \in \{1, 2, 3\}, \end{aligned} \quad (6.2)$$

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

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

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

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

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

Proposition 2. (*Identifying Moment for Subsidies*) Suppose a subsidy plan is given by Equation (6.3). Assume that (a) exogenous firm productivity follows the unbounded Pareto distribution ($\kappa \rightarrow \infty$), (b) goods are freely tradable ($\tau_{nmj} = 1$ and $\tau_{nj}^x = 1$), and (c) $j \in \mathcal{J}^T$ are symmetric. Consider the following regression model for $j \in \mathcal{J}^T$, $n \in \mathcal{N}$:

$$\ln \lambda_{njt}^T - \theta \delta \lambda_{njt-1}^T = \beta^{policy} \times D_{jt}^{policy} + \delta_{nt} + \epsilon_{njt},$$

where D_{jt}^{policy} is a dummy variable of whether $j \in \mathcal{J}^{policy}$, and δ_{nt} are time-varying regional fixed effects. Then, when $\mathbb{E}[\ln \phi_{njt}^{min} | D_{jt}^{policy}] = 0$ 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 $\hat{\beta}^{policy}$ uniquely identifies \bar{s} for given values of η , δ , σ , and θ .

The proposition shows that sudden increases in shares of adopters in 1980 captured by β^{policy} are informative about subsidies when they are uncorrelated with exogenous natural advantages; that is, when $\mathbb{E}[\ln \phi_{njt}^{min} | D_{jt}^{policy}] = 0$ holds. This proposition motivates our approach. Since the simplifying assumptions of Proposition 2 do not hold in either the model or the data, we identify the subsidy level by indirect inference. In particular, using both actual and model-generated data, we estimate the following regression for only the heavy manufacturing sector in 1972 and 1980 using the PPML to incorporate zeros:

$$\ln \lambda_{n,heavy,t}^T = \alpha + \beta^{policy} \times D_t^{policy} + \beta_1 \lambda_{n,heavy,t-1}^T + \epsilon_{n,heavy,t}. \quad (6.4)$$

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).

Equation (6.4) differs from the regression model in Proposition 2 in two ways. First, because the heavy manufacturing sector is the only sector where technology adoption is available in our quantitative exercises, we cannot control for δ_{nt} , and D_t^{policy} cannot be separately identified from time fixed effects.⁵⁹ However, we show that $\hat{\beta}^{policy}$ is still informative on \bar{s} in Appendix Section E.4. Second, we control for previous shares rather than subtract them from current shares in dependent variables.⁶⁰ The estimated coefficients for β^{policy} and β_1 are 0.65 and 5.62, and statistically significant at the 1% level.⁶¹

Objective Function: Micro Moments, $\{\Theta^M, \bar{s}\}$. We pin down $\Theta^M = \{F_j^T, F_j^x, \kappa\}$ and subsidy rate \bar{s} using micro-moments. \bar{s} is identified by the identifying moment discussed above. We identify fixed adoption costs F_j^T using the median of shares of adopters in 1972 and 1980. We identify κ using the share of regions with zero adoption in 1972 and 1980. κ rationalizes zero adoption in some regions observed in the data. If κ is sufficiently low, that is, if the gap between the Pareto lower and upper 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

⁵⁹More 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, so we cannot control for region, sector, or time fixed effects. Ideally, a richer model that incorporates multiple heavy manufacturing sectors or more information on subsidy schedules across regions will allow us to control for additional fixed effects.

⁶⁰This is because PPML is not defined for dependent variables with negative values and subtracting the previous shares from the current shares with zero values generates observations with dependent variables that take negative values.

⁶¹The value of the estimated coefficient for β_1 (5.62) that corresponds to $\theta \times \delta$ in the model is consistent with the externally calibrated values $4.77 = 1.06 \times 4.5 = \theta \times \delta$. The estimation procedure and results are reported in Appendix Table E1.

across regions in 1972.⁶² Because we do not have detailed data on commodity sector firms, we set fixed export costs of commodity sectors to be the same as those of the light manufacturing sector.

Constraints: Aggregate Data, Ψ_t . The constraints in Equation (6.2) identify geographic fundamentals Ψ_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 the data in 1972, 1976, and 1980. The number of the constraints is the same with the dimension of the geographic fundamentals.⁶³ Therefore, for any given parameters Θ^M and subsidy rate \bar{s} , geographic fundamentals are exactly identified by these constraints and there exists a set of geographic fundamentals that rationalizes the data.

Because geographic fundamentals are exactly identified, we can identify the average productivity including subsidies $\bar{\phi}_{nj}^{avg}$, following the model-inversion logic from Allen and Arkolakis (2014).⁶⁴ However, we cannot identify what portion of $\bar{\phi}_{nj}^{avg}$ is attributable to natural advantages 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 fixed adoption costs F_j^T and subsidies \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 5.6 times larger than the estimated fixed export cost. The estimated subsidy rate is 0.11, which indicates that adopters are subsidized with 11% of input expenditures. Table 4 reports the model fit. The data moments are well-approximated in the model.

7 The Aggregate and Regional Effects of the Temporary Adoption Subsidy

In this section, we ask how the aggregate and regional patterns of industrialization in Korea would have evolved differently if the temporary subsidies had not been provided. In the baseline economy, the subsidies are provided, whereas in the counterfactual economy, the subsidies are not provided. We compare the baseline and the counterfactual economies. Unlike the simplified model in Section 5.5 where there is a maximum of three steady states, the full quantitative model potentially admits a

⁶²Our firm balance sheet data has information on exports. However, many observations were missing. Given that export data are very noisy, we do not use export information for our reduced-form empirical analysis, but only for computing the moment on shares of exporters for our quantitative analysis.

⁶³The 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 .

⁶⁴By 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	
<i>External calibration</i>			
<u>Structural parameters</u>			
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 (2021)
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, $\tau_{nj}^x = \bar{\tau}^x (dist_{nm}^{port})^\xi$	1.7	Anderson and Van Wincoop (2004)
α_j	Preferences		IO table, 1972
γ_j^k	Production		IO table, 1972
M_{nj}	Exogenous firm mass		Value added, 1972 (Chaney, 2008)
<i>Internal calibration: Method of moment</i>			
<u>Structural parameters</u>			
F_j^T	Fixed adoption cost	0.28	Share of adopters, heavy mfg.
F_j^x	Fixed export cost, commodity & light mfg.	0.06	Share of exporters, light mfg.
F_j^h	Fixed export cost, heavy mfg.	0.05	Share of exporters, heavy mfg.
κ	Pareto upper bound	4.42	# of regions with zero adoption
<u>Geographical fundamentals</u>			
ϕ_{nj}^{min}	atural advantage (Pareto lower bound)		Region-sector sales dist., 1972, 1976, 1980
D_j^f	Foreign market size		Sectoral export intensity, 1972, 1976, 1980
c_j^f	Foreign imported input price		Sectoral import intensity, 1972, 1976, 1980
V_{nt}	Amenity		Pop. dist., 1972, 1976, 1980
<u>Subsidy</u>			
\bar{s}	Subsidy rate	0.11	Identifying moment $\hat{\beta}^{policy}$, Equation (6.4)

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 (6.4)	0.65	0.83
med. shares of exporters in 1972, light mfg.	0.22	0.21
med. shares of exporters in 1972, heavy mfg.	0.14	0.18
med. shares of adopters in 1972	0.06	0.07
med. shares of adopters in 1982	0.12	0.19
Share of zero adoption regions in 1972	0.59	0.53
Share of zero adoption regions in 1982	0.83	0.94

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

larger number of steady states. Which steady state will be reached in the long-run is of computational question, given calibrated values of $\{\Psi_t, \bar{s}, \Theta\}$ that are chosen to match cross-sectional data in 1972, 1976, and 1980 rather than chosen arbitrarily.

Figure 4 reports our main counterfactual results. In Panels A, B, C, and D, we compare the heavy manufacturing sector shares of GDP, employment, and exports, and the light manufacturing sector shares of exports. Had temporary adoption subsidies not been provided, South Korea’s pattern of industrialization and its comparative advantage would have evolved differently. When compared to the steady state of the baseline economy, heavy manufacturing’s share of GDP would have decreased by 15%, its share of employment would have decreased by 3%, and its share of exports would have decreased by 22.5%, and these changes would have been permanent in the steady state of the counterfactual economy.

Our calibration strategy only fits the cross-section data for 1972, 1976, and 1980 and does not fit the evolution of variables after 1980. Thus, our model does not explain the evolution of employment or export shares after 1980 well. However, our model fits the evolution of the heavy manufacturing’s share of GDP 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, where the average productivity is defined as $M_{nj}[\int z_{it}(\phi)^{\sigma-1}dG_{njt}(\phi)]^{1/(\sigma-1)}$. Because M_{nj} and ϕ_{njt}^{min} are not separately identifiable under the fixed entry, $M_{nj}[\int z_{it}(\phi)^{\sigma-1}dG_{njt}(\phi)]^{1/(\sigma-1)}$ can be considered as the average productivity when $M_{nj} = 1, \forall n, j$. The x and y axes 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 pattern of the baseline

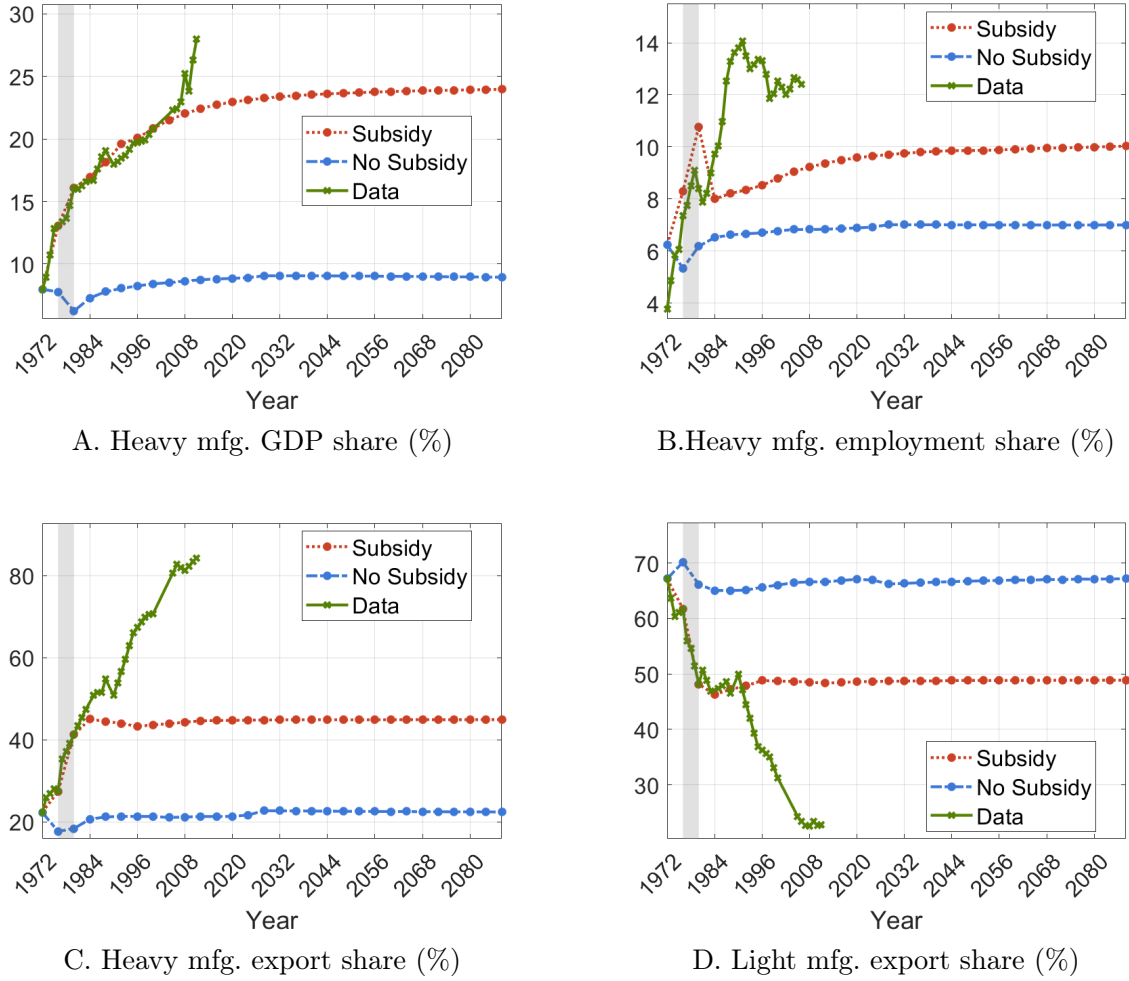


Figure 4. Results of the Counterfactual Analysis

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.

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 of the baseline are 10% permanently higher than the counterfactual once the economies reach steady states. Large productivity increases of a few regions and their specialization into the heavy manufacturing sector led to increases in welfare across all regions through trade linkages. Also, note that the beginning of

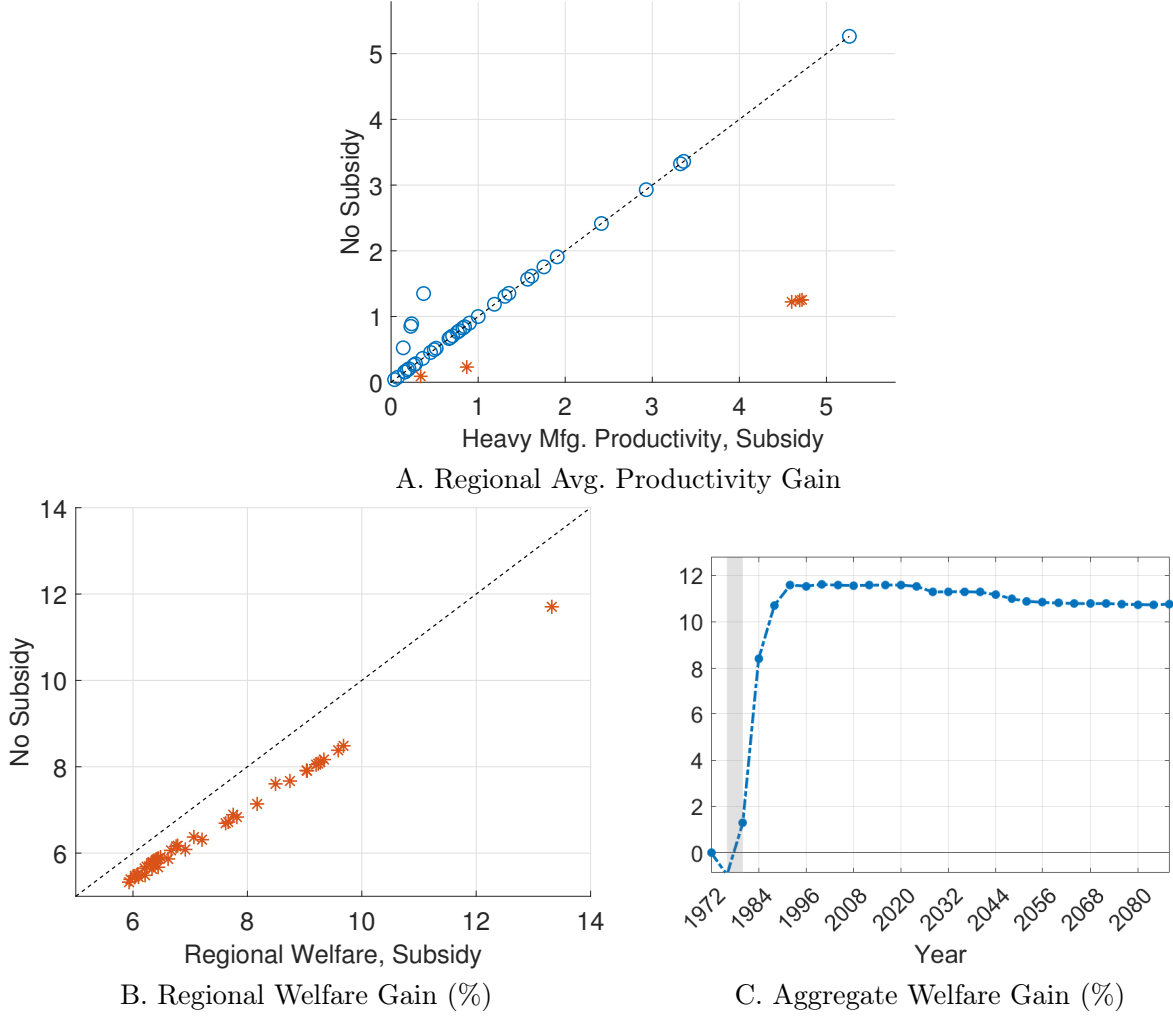


Figure 5. Counterfactual Results. Welfare and Regional Average Productivity

Notes. This figure plots the counterfactual results. Panel A reports the regional average productivity defined as $M_{nj}[\int z_{it}(\phi)^{\sigma-1}dG_{njt}(\phi)]^{1/(\sigma-1)}$. The x and y axes plot each region's average productivity under the baseline and counterfactual economies. Panel B reports regional welfare (Equation (5.12)). The X- and Y-axis plot each region's welfare under the baseline and counterfactual economies. In Panels A and B, each dot is colored red if a corresponding region experienced increases in the average productivity and regional welfare. Panel C reports the ratio of the aggregate welfare (Equation (5.13)) of the baseline economy to that of the counterfactual economy.

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.⁶⁵

Roundabout Production. We find that a roundabout production structure plays an important role in generating permanent effects of subsidies. A roundabout production structure amplifies the

⁶⁵ 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 Lugovsky (2020) in the static setting.

impact of subsidies through cost and demand linkages (Krugman and Venables, 1995).⁶⁶ Because of these linkages, complementarity between firm-scale and gains from technology adoption causes more firms to adopt technology. We do the same counterfactual exercises with a new production structure where labor is the only factor of production, and there are no intermediate inputs. The results are reported in Appendix Figure E2. While holding other parameters, subsidies, and geographic fundamentals constant, we find that both baseline and counterfactual economies converge to the same steady states.

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 with and without the subsidies, we change geographical features of the Korean economy and examine how its long-term effects differ from the main results in Figure 4. We specifically examine the role of Foreign market size and migration costs. We focus on Foreign market size because of large expansion of Korea’s exports in the 1960s and 1970s, and narratives that suggest the export expansion played an important role for Korea’s economic development.⁶⁷ We study migration costs because there were dramatic increases in migration flows from rural to urban areas in Korea during the 1970s, which is a common feature during industrialization.⁶⁸

We examine joint implications between Foreign market size and the subsidies. We decreased 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 main results in Figure 4. These quantitative results provide suggestive evidence that exports might have played an important role in shaping Korea’s economic development jointly with subsidies.

We next examine joint implications between migration costs and the subsidies. 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 adoption, fewer firms will adopt technology with higher migration costs. These results are reported in Appendix Figure E3. The gap of heavy

⁶⁶Heavy manufacturing sectors had disproportionately larger own cost shares of production γ_j^j than other sectors. Own cost shares of production were 0.09, 0.26, 0.46, and 0.13 for commodity, light manufacturing, heavy manufacturing, and service sectors, respectively.

⁶⁷Dramatic 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 tariffs on imported inputs for exports (Connolly and Yi, 2015). Korea also devalued its over-valued currency in 1964, which boosted its exports (Irwin, 2021). Also, the US demands for foreign imports increased dramatically between 1960 to 1980. During that period, shares of US imports in the total gross national product rose from 6 to 22%.

⁶⁸According to the World Development Indicators (World Bank), Korea’s shares of the 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 underwent rapid transitions from rural to urban during industrialization in the twentieth century. See Table 1 of Lucas (2004).

manufacturing GDP shares between two steady states is around 9%, which is 6% smaller than the main results in Figure 4.

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

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

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

References

- Acemoglu, Daron**, “A Microfoundation for Social Increasing Returns in Human Capital Accumulation,” *Quarterly Journal of Economics*, 1996, *111* (3), 779–804.
- , **Pol Antràs**, and **Elhanan Helpman**, “Contracts and Technology Adoption,” *American Economic Review*, 2007, *97* (3), 916–943.
- Akerberg, Daniel A.**, **Kevin Caves**, and **Garth Frazer**, “Identification Properties of Recent Production Function Estimators,” *Econometrica*, 2015, *83* (6), 2411–2451.
- Adserà, Alicia** and **Debraj Ray**, “History and Coordination Failure,” *Journal of Economic Growth*, 1998, *3* (3), 267–276.
- Allen, Treb** and **Costas Arkolakis**, “Trade and the Topography of the Spatial Economy,” *Quarterly Journal of Economics*, 2014, *129* (2014), 1085–1139.
- and **Dave Donaldson**, “The Geography of Path Dependence in the Spatial Economy,” *NBER Working Paper 28059*, 2020.
- Amsden, Alice H.**, *Asia’s Next Giant: South Korea and Late Industrialization*, Oxford University Press, 1989.
- Anderson, James E.** and **Eric Van Wincoop**, “Trade Costs,” *Journal of Economic Literature*, 2004, *42* (3), 691–751.
- Arkolakis, Costas**, **Michael Peters**, and **Sun Kyoung Lee**, “European Immigrants and the United States’ Rise to the Technological Frontier,” *Working Paper*, 2019.
- Arzaghi, Mohammad** and **J. Vernon Henderson**, “Networking off Madison Avenue,” *Review of Economic Studies*, 2008, *75* (4), 1011–1038.
- Atkin, David**, **Arnaud Costinot**, and **Masao Fukui**, “Globalization and the Ladder of Development: Pushed to the Top or Held at the Bottom?,” *Working Paper*, 2021.
- , **Azam Chaudhry**, **Shamyla Chaudry**, **Amit K. Khandelwal**, and **Eric Verhoogen**, “Organizational Barriers to Technology Adoption: Evidence from Soccer-Ball Producers in Pakistan,” *Quarterly Journal of Economics*, 2017, *132* (3), 1101–1164.
- Axtell, Robert L.**, “Zipf Distribution of U.S. Firm Sizes,” *Science*, 2001, *293* (5536), 1818–1820.
- Azariadis, Costas** and **John Stachurski**, “Poverty Traps,” in Phillippe Aghion and Steven Durlauf, eds., *Handbook of Economic Growth*, Vol. 1, Amsterdam: Elsevier, 2005, chapter 5, pp. 295–384.
- 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 26193*, 2020.
- , **Ting Lan**, and **Andrei Levchenko**, “Specialization, Market Access and Real Income,” *NBER Working Paper 28274*, 2021.
- Bloom, Nicholas**, **Mark Schankerman**, and **John Van Reenen**, “Identifying Technology Spillovers and Product Market Rivalry,” *Econometrica*, 2013, *81* (4), 1347–1393.
- Broda, Christian** and **David E. Weinstein**, “Globalization and the Gains from Variety,” *Quarterly Journal of Economics*, 2006, *121* (2), 541–585.
- Buera, Francisco J.** and **Ezra Oberfield**, “The Global Diffusion of Ideas,” *Econometrica*, 2020, *88* (1), 83–114.

- , **Hugo Hopenhayn**, **Yongseok Shin**, and **Nicholas Trachter**, “Big Push in Distorted Economies,” *NBER Working Paper 28561*, 2021.
- 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 29263*, 2021.
- Ciccone, Antonio**, “Input Chains and Industrialization,” *Review of Economic Studies*, 2002, *69* (3), 565–587.
- and **Robert E. Hall**, “Productivity and the Density of Economic Activity,” *American Economic Review*, 1996, *86* (1), 54–70.
- Combes, Pierre Philippe** and **Laurent Gobillon**, “The Empirics of Agglomeration Economies,” in Gilles Duranton, J. Vernon Henderson, and William C. Strange, eds., *Handbook of Regional and Urban Economics*, elsevier ed., Vol. 5, North Holland, Amsterdam: Elsevier, 2015, chapter 5, pp. 247–338.
- Comin, Diego** and **Bart Hobijn**, “An Exploration of Technology Diffusion,” *American Economic Review*, 2010, *100* (5), 2031–2059.
- Connolly, Michelle** and **Kei-Mu Yi**, “How Much of South Korea’s Growth Miracle Can Be Explained by Trade Policy?,” *American Economic Journal: Macroeconomics*, 2015, *7* (4).
- Crouzet, Nicolas**, **Apoorv Gupta**, and **Filippo Mezzanotti**, “Shocks and Technology Adoption: Evidence from Electronic Payment Systems,” *Working Paper*, 2020.
- de Souza, Gustavo**, “The Labor Market Consequences of Appropriate Technology,” *Working Paper*, 2021.
- Desmet, Klaus** and **Esteban Rossi-Hansberg**, “Spatial Development,” *American Economic Review*, 2014, *104* (4), 1211–1243.
- 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, John L.** and **Woo-Hee Park**, *The Adoption and Diffusion of Imported Technology: The Case of Korea*, Routledge, 1988.
- Fajgelbaum, Pablo D.** and **Cecile Gaubert**, “Optimal Spatial Policies, Geography, and Sorting,” *Quarterly Journal of Economics*, 2020, *135* (2), 959–1036.
- Feenstra, Robert C.**, **Robert E. Lipsey**, **Haiyan Deng**, **Alyson C. Ma**, and **Hengyong Mo**, “World Trade Flows: 1962-2000,” *NBER Working Paper 11040*, 2005.

- Gerschenkron, Alexander**, *Economic Backwardness in Historical Perspective*, Cambridge, MA: Harvard University Press, 1962.
- Gibbons, Steve, Henry G. Overman, and Eleonora Patacchini**, “Spatial Methods,” in Gilles Duranton, J. Vernon Henderson, and William C. Strange, eds., *Handbook of Regional and Urban Economics*, gilles dur ed., Vol. 5, North Holland, Amsterdam: Elsevier, 2015, chapter 3, pp. 115–168.
- Giorcelli, Michela and Bo Li**, “Technology Transfer and Early Industrial Development: Evidence from the Sino-Soviet Alliance,” *Working Paper*, 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 J. Melitz, and Yona Rubinstein**, “Estimating Trade Flows: Trading Partners and Trading Volumes,” *Quarterly Journal of Economics*, 2008, 123 (2), 441–487.
- Hirschman, Albert O.**, *The Strategy of Economic Development*, Yale University Press, 1958.
- Irwin, Douglas A.**, “From Hermit Kingdom to Miracle on the Han: Policy Decisions that Transformed South Korea into an Export Powerhouse,” *National Bureau of Economic Research 29299*, 2021.
- Itskhoki, Oleg and Benjamin Moll**, “Optimal Development Policies With Financial Frictions,” *Econometrica*, 2019, 87 (1), 139–173.
- Jaffe, Adam B., Manuel Trajtenberg, and Rebecca Henderson**, “Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations,” *Quarterly Journal of Economics*, 1993, 108 (3), 577–598.
- Jovanovic, Boyan**, “Observable Implications of Models with Multiple Equilibria,” *Econometrica*, 1989, 57 (6), 1431–1437.
- and **Rafael Rob**, “The Growth and Diffusion of Knowledge,” *Review of Economic Studies*, 1989, 56 (4), 569–582.
- Juhász, Réka**, “Temporary Protection and Technology Adoption: Evidence from the Napoleonic Blockade,” *American Economic Review*, 2018, 108 (11), 3339–3376.
- , **Mara P. Squicciarini, and Nico Voigtländer**, “Technology Adoption and Productivity Growth: Evidence from Industrialization in France,” *NBER Working Paper 27503*, 2020.
- Kantor, Shawn and Alexander Whalley**, “Research Proximity and Productivity: Long-Term Evidence from Agriculture,” *Journal of Political Economy*, 2019, 127 (2), 819–854.
- Keller, Wolfgang**, “Geographic Localization of International Technology Diffusion,” *American Economic Review*, 2002, 92 (1), 120–142.
- , “International Technology Diffusion,” *Journal of Economic Literature*, 2004, 42 (3), 752–782.

- Kerr, William R. and Scott Duke Kominers**, “Agglomerative Forces and Cluster Shapes,” *Review of Economics and Statistics*, 2015, 97 (4), 877–899.
- Kim, Dae-Il and Robert H. Topel**, “Labor Markets and Economic Growth: Lessons from Korea’s Industrialization, 1970-1990,” in “Differences and Changes in Wage Structures,” Cambridge: National Bureau of Economic Research, 1995, pp. 227–264.
- Kim, Linsu**, *Imitation to Innovation: The Dynamics of Korea’s Technological Learning*, Boston: Harvard Business School Press, 1997.
- and **Youngbae Kim**, “Innovation in a Newly Industrializing Country: A Multiple Discriminant Analysis,” *Management Science*, 1985, 31 (3), 312–322.
- Kim, Minho, Munseob Lee, and Yongseok Shin**, “The Plant-Level View of an Industrial Policy: The Korean Heavy Industry Drive of 1973,” *NBER Working Paper 29252*, 2021.
- Klenow, Peter J. and Andrés Rodríguez-Clare**, “The Neoclassical Revival in Growth Economics: Has It Gone Too Far?,” *NBER Macroeconomics Annual*, 1997, 12, 73–103.
- Kline, Patrick and Enrico Moretti**, “Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority,” *Quarterly Journal of Economics*, 2014, 129 (1), 275–331.
- Korea Industrial Technology Association**, *Records of Technology Acquisition Contracts: 1962-1988*, Seoul: Korea Industrial Technology Association, [in Korean], 1988.
- Krugman, Paul**, “The Narrow Moving Band, the Dutch Disease, and the Competitive Consequences of Mrs. Thatcher. Notes on Trade in the Presence of Dynamic Scale Economies,” *Journal of Development Economics*, 1987, 27 (1-2), 41–55.
- , “History versus Expectations,” *Quarterly Journal of Economics*, 1991, 106 (2), 652–667.
- and **Anthony J. Venables**, “Globalization and the Inequality of Nations,” *Quarterly Journal of Economics*, 1995, 110 (4), 857–880.
- Lane, Nathan**, “Manufacturing Revolutions Industrial Policy and Networks in South Korea,” *Working Paper*, 2019.
- Lashkaripour, Ahmad and Volodymyr Lugovskyy**, “Scale Economies and the Structure of Trade and Industrial Policy,” *Working Paper*, 2020.
- Levchenko, Andrei A. and Jing Zhang**, “The Evolution of Comparative Advantage: Measurement and Welfare Implications,” *Journal of Monetary Economics*, 2016, 78, 96–111.
- Levinsohn, James and Amil Petrin**, “Estimating Production Functions using Inputs to Control for Unobservables,” *Review of Economic Studies*, 2003, 70 (2), 317–341.
- Lileeva, Alla and Daniel Trefler**, “Improved Access to Foreign Markets Raises Plant-Level Productivity. . .for Some Plants,” *Quarterly Journal of Economics*, 2010, 125 (3), 1051–1099.
- Lucas, Robert E.**, “On the Mechanics of Economic Development,” *Journal of Monetary Economics*, 1988, 22 (1), 3–42.
- , “Making a Miracle,” *Econometrica*, 1993, 61 (2), 251–272.
- , “Life Earnings and Rural-Urban Migration,” *Journal of Political Economy*, 2004, 112 (1), 29–59.
- and **Benjamin Moll**, “Knowledge Growth and the Allocation of Time,” *Journal of Political Economy*, 2014, 122 (1), 1–51.
- Luedde-Neurath, Richard**, *Import Controls and Export-Oriented Development: A Reassessment of the South Korean Case*, Routledge, 1986.
- Magee, Chris**, “U.S. Tariffs at the SIC Level, 1974-1988,” 1986.

- Manski, Charles F.**, “Identification of Endogenous Social Effects: the Reflection Problem,” *Review of Economic Studies*, 1993, 60 (3), 531–542.
- Matsuyama, Kiminori**, “Increasing Returns, Industrialization, and Indeterminacy of Equilibrium,” *Quarterly Journal of Economics*, 1991, 106 (2), 617–650.
- , “Agricultural Productivity, Comparative Advantage, and Economic Growth,” *Journal of Economic Theory*, 1992, 58 (2), 317–334.
- , “Complementarities and Cumulative Processes in Models of Monopolistic Competition,” *Journal of Economic Literature*, 1995, 33 (2), 701–729.
- Melitz, Marc J.**, “The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity,” *Econometrica*, 2003, 71 (6), 1695–1725.
- Ministry of Science and Technology**, *Science and Technology Annual* 1972.
- Monte, Ferdinando, Stephen J. Redding, and Esteban Rossi-Hansberg**, “Commuting, Migration, and Local Employment Elasticities,” *American Economic Review*, 2018, 108 (12), 3855–3890.
- Moretti, Enrico**, “Workers’ Education, Spillovers, and Productivity: Evidence from Plant-Level Production Functions,” 2004.
- , “The Effect of High-Tech Clusters on the Productivity of Top Inventors,” *American Economic Review*, 2021, 111 (10), 3328–3375.
- Murphy, Kevin M., Andrei Shleifer, and Robert W. Vishny**, “Industrialization and the Big Push,” *Journal of Political Economy*, 1989, 97 (5), 1003–1026.
- Nakamura, Emi and Jón Steinsson**, “Identification in Macroeconomics,” *Journal of Economic Perspectives*, 2018, 32 (3), 59–86.
- Parente, Stephen L. and Edward C. Prescott**, “Barriers to Technology Adoption and Development,” *Journal of Political Economy*, 1994, 102 (2), 298–321.
- and —, *Barriers to Riches*, MIT Press, 2002.
- Pellegrina, Heitor and Sebastian Sotelo**, “Migration, Specialization, and Trade: Evidence from the Brazilian March to the West,” *NBER Working Paper 28421*, 2021.
- Perla, Jesse and Christopher Tonetti**, “Equilibrium Imitation and Growth,” *Journal of Political Economy*, 2014, 122 (1).
- , —, and **Michael E. Waugh**, “Equilibrium Technology Diffusion, Trade, and Growth,” *American Economic Review*, 2021, 111 (1), 73–128.
- Peters, Michael**, “Market Size and Spatial Growth - Evidence from Germany’s Post-War Population Expulsions,” *NBER Working Paper 29329*, 2021.
- Rodríguez-Clare, Andrés**, “The Division of Labor and Economic Development,” *Journal of Development Economics*, 1996, 49 (1), 3–32.
- Rodrik, Dani**, “Getting Interventions Right: How South Korea and Taiwan Grew Rich,” *Economic Policy*, 1995, 20, 53–107.
- Romer, Paul M.**, “Endogenous Technological Change,” *Journal of Political Economy*, 1990, 98 (5), S71–S102.
- Rosenstein-Rodan, Paul N.**, “Problems of Industrialisation of Eastern and South-Eastern Europe,” *The Economic Journal*, 1943, 53, 202–211.
- Schetter, Ulrich**, “A Structural Ranking of Economic Complexity,” *Working Paper*, 2019.

- Serafinelli, Michel**, ““Good” Firms, Worker Flows, and Local Productivity,” *Journal of Labor Economics*, 2019, *37* (3), 747–792.
- Silva, J. M. C. Santos and Silvana Tenreyro**, “The Log of Gravity,” *Review of Economics and Statistics*, 2006, *88* (4), 641–658.
- 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 of Technology Adoption 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. President Park was in power for 19 years. He was the chairman of the military junta for 1961 and 1962. In 1963 and 1967, he became elected as a president of the civilian government. In 1971, he became re-elected, and it was supposed to be his last presidency. In 1972, President Park declared martial law and amended the country's constitution into an authoritarian constitution, called the Yushin constitution, which

extended his terms of office for the president indefinitely. After 1961 and until President Park was assassinated in 1979, the EPB was at the center of the economic policy-making process of Korea.

During his presidency, 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 EPB. After the contract, they also had to submit documents related to newly made contracts to the EPB. 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. 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 the National Archives of Korea. Our technology adoption data mainly comes from historical contract documents from the National Archives of Korea.

Figure A1 is one example of these documents, which is one page from the actual contract document between Kolon (Korean) and Mitsui Toatsu (Mitsui) (Japanese). Most of the adopted technologies were know-how on construction/operation of plants and capital equipment related to mass production. For example, Figure A1 specifies that Mitsui had to provide blueprints, send skilled engineers to train Korean workers, and provide training service by inviting Korean engineers to its plants in Japan.

Available Information. From these contracts, we obtained three main information: (1) domestic firms' names, (2) foreign firms' names, (3) contract years. We use the information on (1) domestic firms' names and (3) contract years to construct a dummy variable of firms' adoption status.

A.2 Firm Balance Sheet Data.

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

Balance Sheet Variables. Balance sheet information includes sales, assets, fixed assets, and employment. Employment starts from 1972. We convert all monetary values into 2015 US dollars. The data set covers firm with employment more than 50. Also, the data set has information on firms' start years. We use this start year information to match firms' changes of names.

Location of Production. It has detailed information on the address of the location of production. We convert addresses into the 2010 administrative divisions of Korea up to the town level. (We classifies firms' location of production into Li and Dong levels. Then, using distance between towns, we calculate distance between firms within district.

Sector Groupings. We classify firms into 10 manufacturing sectors. 4 of them are classified as heavy manufacturing sectors, which mainly follows the sector classification by Lane (2019). 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) codes. We use these ISIC codes to map our firm data to other trade or tariff data.

A.3 Other Data Sets

Input-Output Tables. Input-output tables are obtained from Bank of Korea.⁶⁹ Input-output tables are available for 1970, 1973, 1975, 1978, 1980, 1983, and 1985 during the sample period.

OECD Stan Database. We obtain cross-country heavy manufacturing GDP shares in Figure 1 from OECD Stan database which has sectoral GDP information at 2-digit ISIC3 level.⁷⁰

For Mexico, the total GDP was not available for the early 1970s, but the total value of light and heavy manufacturing shares was available. Therefore, from the OECD Stan database, we could calculate heavy manufacturing value added shares out of the total manufacturing, but we could not calculate its shares to the national GDP. We supplement the Mexico sample with shares of manufacturing to the total GDP obtained from World Bank Indicators.⁷¹ The share of heavy manufacturing GDP shares is then obtained as value added shares of heavy manufacturing sectors to the total value added of the manufacturing times GDP shares of manufacturing sectors to the total GDP.

A.4 Criteria for Matching Two Main Data Sets

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

1. Firms should have the same name in a given year.
2. Start years of firms should be prior to observed years of adoption.
 - Even if we observe same names in both data sets, if observed adoption activities happened before start year information in the balance sheet data, we do not match these firms.
3. Firms should be in the same sector.
 - Each contract document has a brief description about technology that firms were adopting.
 - Even if we observe same names in both data sets, if these descriptions do not align with the recorded sector in the balance sheet data, we do not match these firms.

A.5 Tracking Changes of Firms' Names

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 the Annual Reports of Korean Companies.
- Step 2: Information on firms' history from their websites

⁶⁹The data were downloaded from https://ecos.bok.or.kr/EIndex_en.jsp.

⁷⁰The data were downloaded from <https://stats.oecd.org/Index.aspx?DataSetCode=STAN>.

⁷¹The data were downloaded from <https://data.worldbank.org/indicator/NV.IND.MANF.ZS>.

Table A1: Classification of Sectors

Aggregated Industry	Industry
	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)
(i) Chemicals, Petrochemicals, Rubber, & Plastic Products	
	Office, accounting, & computing machinery (30) Electrical machinery and apparatus n.e.c. (31) Radio, television and communication equipment and apparatus (32) Medical, precision, and optical instruments, watches and clocks (33)
<u>Heavy Mfg.</u>	
(ii) Electrical Equipment	
	Basic metals (27) Fabricated metals (28)
(iii) Basic & Fabricated Metals	
	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)
(iv) Machinery & Transport Equipment	
	Food products and beverages (15) Tobacco products (16)
(v) Food, Beverages, & Tobacco	
	Textiles (17) Apparel (18) Leather, luggage, handbags, saddlery, harness, and footwear (19)
(vi) Textiles, Apparel, & Leather	
	Manufacturing n.e.c. (369)
(vii) Manufacturing n.e.c.	
	Wood and of products, cork (20) Paper and paper products (21) Publishing and printing (22) Furniture (361)
(viii) Wood, Paper, Printing, & Furniture	
	pharmaceuticals and medicine chemicals (2423)
(ix) Pharmaceuticals & Medicine Chemicals	
	Glass and glass products (261) non-metallic mineral products n.e.c. (269)
(x) Other Non-Metallic Mineral Products	

- Step 3: Search for firms' names from <https://www.jobkorea.co.kr> or <https://www.saramin.co.kr> if companies' websites are not available.⁷²
 - Only when reported start year information in the Annual Reports 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.6 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. When computing this coverage, for some observations with missing information on sales, we impute using the 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 is some heterogeneity across sectors. "Machinery and Transportation Equipment" or "Petrochemical and Chemical" have higher coverage rates, whereas "Food, Beverage, and Tobacco" or "Apparel, Leather, and Textile" are relatively less covered than other sectors.

A.7 An Example of a Loser

We identify losers from actual contract documents. The Foreign Capital Act also required firms to submit related documents once their contracts failed in the end if these contracts were approved by the EPB. They had to submit official cancellation contract documents between domestic and foreign firms and documents that describe why contracts failed.

Figure A3 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 send a blueprint. Panel A is the official English document related to the termination of the contract between two firms. Panel B is the official Korean document in which Kangwon reported why the contract failed in the end to the EPB. The document says that the contract failed because although Kangwon asked Broehl several times for its fulfillment of the contract after Kangwon made the fixed payment, there were no responses from Broehl.

⁷²These are two largest job posting websites in Korea.

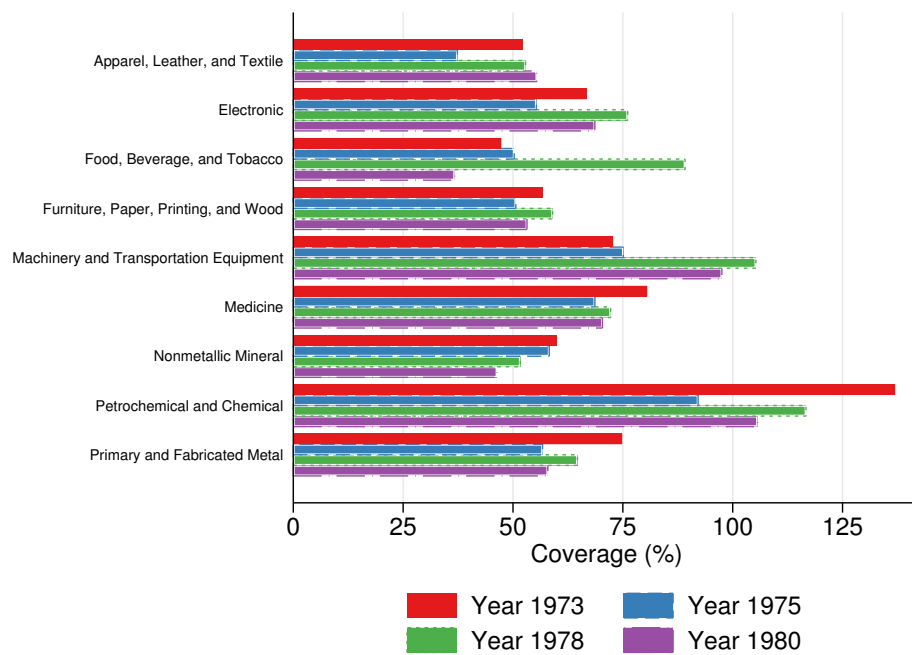
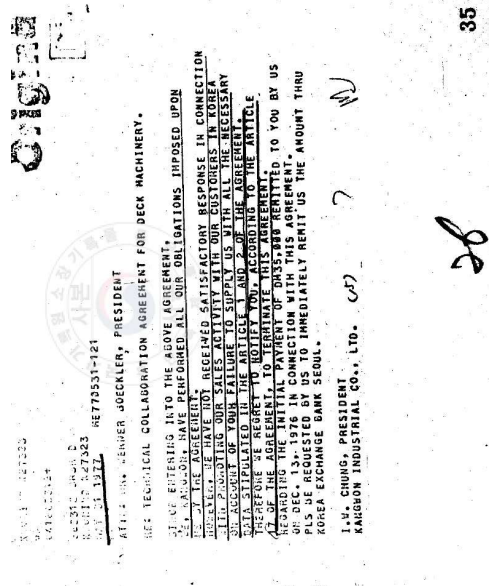
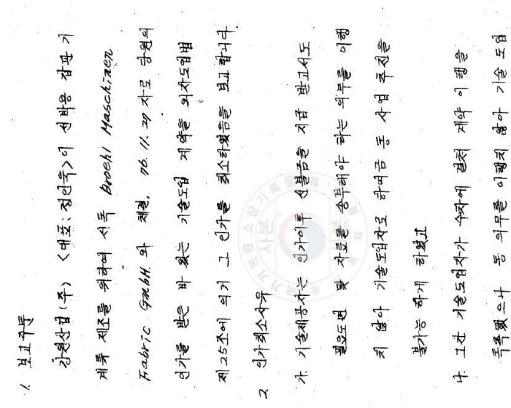


Figure A2. Coverage

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



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



B. Official document that
Kangwon reported to the government
Contract between *Kangwon Industrial Co.* and *Broehl Maschinen Fabric GmbH*

A.8 Descriptive Statistics.

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

In Panel A of Figure A4, we plot evolution of size of heavy and light manufacturing sectors. Size of sector j is measured as follows:

$$\ln \text{Size}_{jt} = \ln \left(\sum_{i \in j} \text{Sale}_{it} \right), \quad j \in \{\text{Light}, \text{Heavy}\}.$$

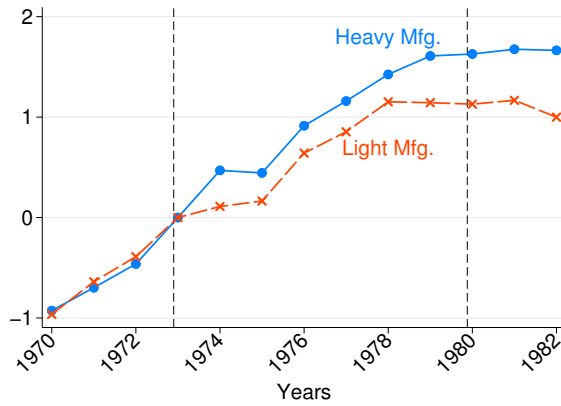
We normalize the size of each sector by their 1973 level so that we can track how heavy and light manufacturing sectors evolved differently after the policy started to be implemented in 1973. In Panel B of Figure A4, we plot shares of adopters in heavy and light manufacturing sectors. The shares are defined as the share of firms who are in contractual relationships with foreign firms to the total number of firms in a given year.

The patterns from the micro data reveal a similar pattern in Figure 1. The total size of heavy manufacturing sectors started increasing faster than that of the light manufacturing sectors after 1973, and this rapid increase coincided with increases in amounts of adoption by heavy manufacturing firms.

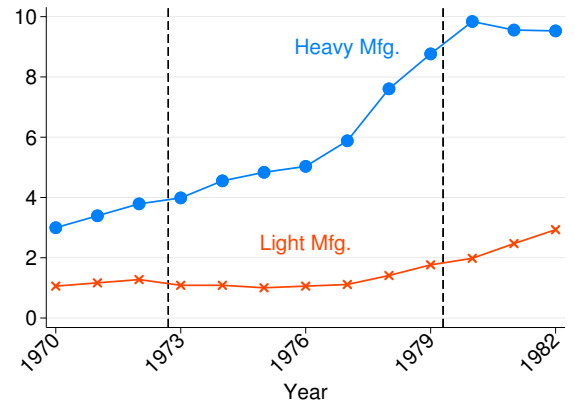
Table A2: Descriptive Statistics.

	All mfg. (1)	Heavy mfg. (2)	Light mfg. (3)
<i>Firm Balance Sheet</i>			
ln(Sales)	15.65 (1.925)	15.54 (1.938)	15.75 (1.910)
ln(Assets)	15.14 (1.766)	15.10 (1.764)	15.18 (1.767)
ln(Fixed Assets)	13.96 (1.966)	13.94 (1.933)	13.98 (1.992)
ln(Emp)	5.166 (1.321)	5.028 (1.319)	5.285 (1.311)
<i>Technology Adoption</i>			
1[Adopt]	0.0587 (0.235)	0.0951 (0.293)	0.0267 (0.161)
1[Ever Adopt]	0.0841 (0.278)	0.132 (0.339)	0.0418 (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. 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. 1[Ever Adopt] is a dummy variable which equals one if a firm ever had technology adoption contracts with foreign firms.



A. log sum of sales



B. shares of adopters

Figure A4. Evolution of Size of Manufacturing Sectors and Shares of Adopters from the Micro Data

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

Appendix B Appendix: Historical Background

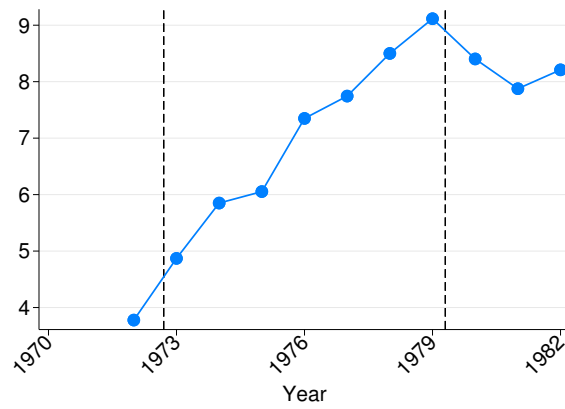
B.1 Additional Aggregate Statistics on Late Industrialization in South Korea

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:

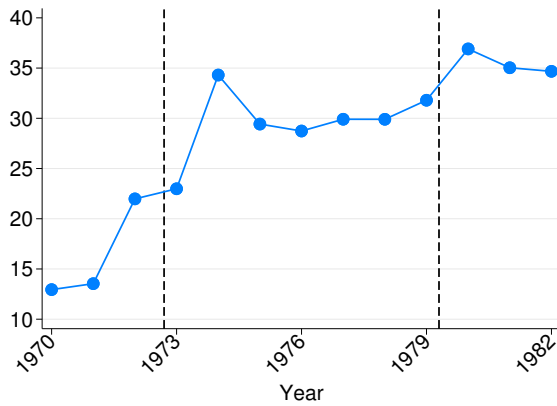
$$RCA_{heavy,t} = \frac{EX_{heavy,t}^{KOR} / \sum_{j \in \mathcal{J}} EX_{jt}^{KOR}}{EX_{heavy,t}^{RoW} / \sum_{j \in \mathcal{J}} EX_{jt}^{RoW}},$$

where EX_{jt}^c 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).⁷³ 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.

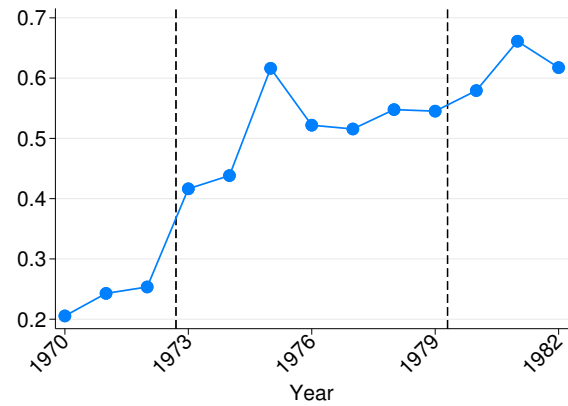
⁷³The trade data was downloaded from <https://cid.econ.ucdavis.edu/nberus.html>.



A. heavy mfg. employment share (%)



B. heavy mfg. export share (%)



C. heavy mfg. Balassa index, RCA

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. The two dotted vertical lines represent the start and the end of the Korean government policy that subsidized technology adoption between 1973 and 1979.

B.2 The Effects of the Policy on Firms' Technology Adoption Decisions

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:

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

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 are time-invariant firm fixed effects. ϵ_{it} are the error terms. Standard errors are clustered at regional level.

The key variable of interests are $\{\beta^{\tau}\}_{\tau=-3}^9$. We normalize β^0 to be zero. Thus, β^{τ} captures how firms' adoption decisions differ relative to the 1973 level. If the policy affected firms' adoption decisions after the policy started to be implemented in 1973, we expect $\{\beta^{\tau}\}_{\tau=1}^9$ to be statistically significantly larger than zero. $\{\beta^{\tau}\}_{\tau=-3}^{-1}$ are pre-trends. If confounding factors were driving the implementation of the policy, it may show up in the pre-trends. If there were no confounding factors, we expect these pre-trends to be statistically indistinguishable from zero. Because we are restring our samples to be heavy manufacturing firms, time fixed effects are not separately identifiable from these event dummies.

Figure B2 illustrates the estimated coefficients with 95 percent confidence intervals. There were no pre-trends. Firms' overall adoption decisions before 1973 were not statistically distinguishable from those in 1973. However, only after 1973, more firms started adopting foreign technologies. This sudden rapid increase after 1973 supports 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.

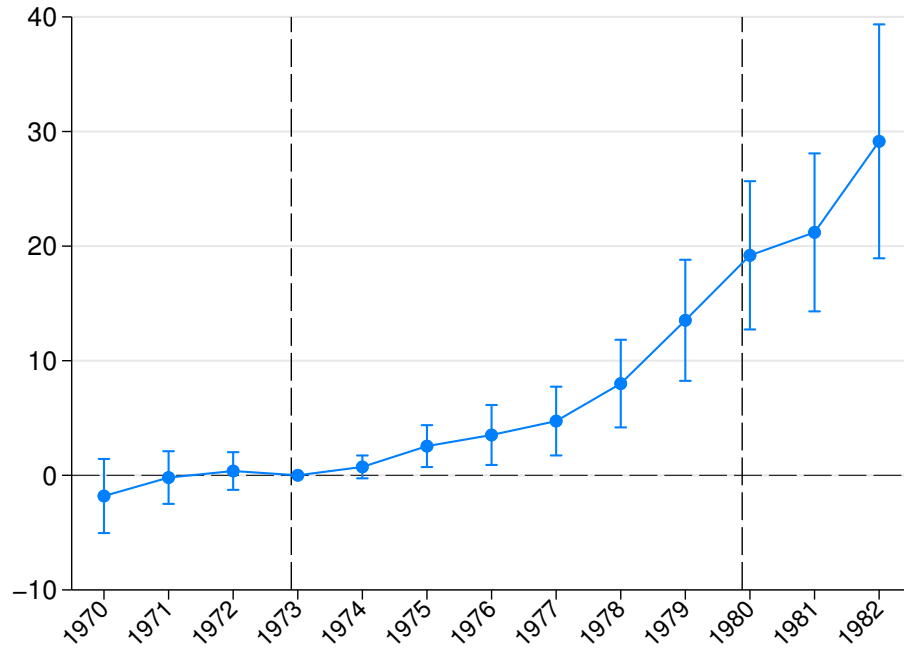


Figure B2. The Impact of the Temporary Subsidies provided by the Korean Government on Firms' Technology Adoption Decisions

Notes. This figure illustrates the estimated β^T in Equation (B.1). β^0 is normalized to zero. All specifications control for firm and calendar year fixed effects. The two dotted vertical lines represent the start and the end of the Korean government policy that subsidized technology adoption between 1973 and 1979. Error bars represent 95 percent confidence intervals based on standard errors clustered at regional level.

Appendix C Appendix: Model

C.1 Closed-Form Expressions for Regional Variables

In this section, we derive closed-form expressions for price index, regional gross output for domestic expenditures, and regional exports. Given 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, shares of exporters, subsidies, and natural advantage.

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

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

The above equation can be rewritten as:

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

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

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

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

where

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

$\tilde{\lambda}_{njt}^T = (1 - \kappa^{-\theta})\lambda_{njt}^T + \kappa^{-\theta}$, and $\tilde{\theta} = \theta - (\sigma - 1)$.⁷⁴ A price index depends on the three terms: unit cost, average productivity including subsidies $\bar{\phi}_{njt}^{avg}$, and consumer foreign market access $(\tau_{nj}^x c_{jt}^f)^{1-\sigma}$. $\bar{\phi}_{njt}^{avg}$ captures how region n can produce sector j intermediate varieties at cheaper cost relative to other regions. Region n can produce at cheaper costs if it has technological advantages ($\lambda_{njt}^T, \lambda_{njt-1}^T, \phi_{njt}^{min}$) 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 $\bar{\phi}_{njt}^{avg}$), or (iii) a price of imported inputs is lower (lower τ_{nj}^x or c_{jt}^f).

The average productivity including subsidy (Equation (C.2)) increases in a share of adopters in the previous period λ_{njt-1}^T , a share of adopters in the current period λ_{njt}^T , subsidies s_{njt} , and a natural advantage captured by the Pareto lower bound ϕ_{njt}^{min} . 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_{njt}^d and the total value of export R_{njt}^x : $R_{njt} = R_{njt}^d + R_{njt}^x$.

⁷⁴When $\lambda_{njt}^T \rightarrow 0$, the average productivity becomes $\bar{\phi}_{njt}^{avg} = \frac{\theta}{\tilde{\theta}(1-\kappa^{-\theta})} f(\lambda_{njt-1}^T)(\phi_{njt}^{min})^{\sigma-1}(1-\kappa^{-\tilde{\theta}})$. When $\lambda_{njt}^T \rightarrow 1$, the average productivity becomes $\bar{\phi}_{njt}^{avg} = \frac{\theta}{\tilde{\theta}(1-\kappa^{-\theta})} f(\lambda_{njt-1}^T)(\phi_{njt}^{min})^{\sigma-1} \left(\frac{\eta}{1-s_{njt}} \right)^{\sigma-1} (1-\kappa^{-\tilde{\theta}})$.

Regional exports can be written as

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

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:

$$\begin{aligned} & \int_{\bar{\phi}_{njt}^T}^{\bar{\phi}_{njt}^{max}} \left(\frac{\sigma}{\sigma-1} \frac{\tau_{nj}^x s_{njt} c_{njt}}{\eta f(\lambda_{njt-1}^T) \phi_{it}} \right)^{1-\sigma} dG_{njt}(\phi_{it}) \\ &= \frac{\theta f(\lambda_{njt-1}^T)^{\sigma-1}}{\tilde{\theta}(1-\kappa^{-\theta})} (\mu c_{njt})^{1-\sigma} \left(\frac{\eta}{1-s_{njt}} \right)^{\sigma-1} (\phi_{njt}^{min})^{-\theta} \left((\bar{\phi}_{njt}^T)^{-\tilde{\theta}} - (\kappa \phi_{njt}^{min})^{-\tilde{\theta}} \right) \\ &= \frac{\theta f(\lambda_{njt-1}^T)^{\sigma-1}}{\tilde{\theta}(1-\kappa^{-\theta})} (\mu c_{njt})^{1-\sigma} \left(\frac{\eta}{1-s_{njt}} \right)^{\sigma-1} (\phi_{njt}^{min})^{\sigma-1} \left((\tilde{\lambda}_{njt}^T)^{\frac{\tilde{\theta}}{\theta}} - \kappa^{-\tilde{\theta}} \right), \end{aligned} \quad (C.4)$$

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

Similarly, the second term can be re-expressed as:

$$\begin{aligned} & \int_{\bar{\phi}_{njt}^x}^{\bar{\phi}_{njt}^T} \left(\frac{\sigma}{\sigma-1} \frac{\tau_{nj}^x c_{njt}}{f(\lambda_{njt-1}^T) \phi_{it}} \right)^{1-\sigma} dG_{njt}(\phi_{it}) \\ &= \frac{\theta f(\lambda_{njt-1}^T)^{\sigma-1}}{\tilde{\theta}(1-\kappa^{-\theta})} (\mu c_{njt})^{1-\sigma} (\phi_{njt}^{min})^{-\theta} \left((\bar{\phi}_{njt}^x)^{-\tilde{\theta}} - (\bar{\phi}_{njt}^T)^{-\tilde{\theta}} \right) \\ &= \frac{\theta f(\lambda_{njt-1}^T)^{\sigma-1}}{\tilde{\theta}(1-\kappa^{-\theta})} (\mu c_{njt})^{1-\sigma} (\phi_{njt}^{min})^{\sigma-1} \left((\tilde{\lambda}_{njt}^x)^{\frac{\tilde{\theta}}{\theta}} - (\tilde{\lambda}_{njt}^T)^{\frac{\tilde{\theta}}{\theta}} \right), \end{aligned} \quad (C.5)$$

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

Using Equations (C.3), (C.4), and (C.5), and $M_{njt}^x = M_{nj} \times \lambda_{njt}^x$, regional exports can be expressed as:

$$R_{njt}^x = M_{njt}^x (\mu c_{njt})^{1-\sigma} \times \underbrace{(\bar{\phi}_{njt}^{avg,x})^{\sigma-1}}_{\text{Exporters' average productivity including subsidy}} \times \underbrace{(\tau_{nj}^x)^{1-\sigma} D_{jt}^f}_{\text{Firm foreign market access}},$$

where

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

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

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

$$R_{njt}^d = M_{nj}(\mu c_{njt})^{1-\sigma} \times \underbrace{(\bar{\phi}_{njt}^{avg})^{\sigma-1}}_{\text{Average productivity including subsidy}} \times \underbrace{\sum_{m \in \mathcal{N}} \tau_{nmj}^{1-\sigma} P_{mjt}^{\sigma-1} E_{mjt}}_{\text{Firm domestic market access}}.$$

and

$$R_{njt}^x = M_{nj}^x(\mu c_{njt})^{1-\sigma} \times \underbrace{(\bar{\phi}_{njt}^{avg,x})^{\sigma-1}}_{\text{Exporters' average productivity including subsidy}} \times \underbrace{(\tau_{nj}^x)^{1-\sigma} D_{jt}^f}_{\text{Firm foreign market access}}.$$

Both the total domestic sales and exports (i) increase in the average productivities inclusive of subsidies, (ii) increase in degree of access to markets, (iii) increase in subsidies, and (iv) decrease in the cost of production.

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

C.2 Analytical Results on Multiple Steady States

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

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

$$(\bar{\phi}_t^T)^{\sigma-1} = \frac{\sigma P_t F^T}{(\eta^{\sigma-1} - 1)(\mu w_t)^{1-\sigma} f(\lambda_{t-1}^T)^{\sigma-1} P_t^\sigma Q_t}, \quad (\text{C.7})$$

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

$$(\lambda_t^T)^{-\frac{1}{\theta}} = \bar{\phi}_t^T \quad (\text{C.8})$$

First, we show that

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

and

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

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

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

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

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

we can derive that

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

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

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

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

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

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

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

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

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.12) with respect to λ_t^T , we obtain that

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

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

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

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

Applying the implicit function theorem and using the signs of Equations (C.13) and (C.14), 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.12) with respect to η , we obtain that

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

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

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

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

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

and

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

This proves Proposition 1(iii). □

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

theorem and doing some tedious algebra. Applying the implicit function theorem,

$$\begin{aligned} \frac{\partial^2 \lambda_t^T}{\partial(\lambda_{t-1}^T)^2} &= -\frac{1}{(\partial G / \partial \lambda_t^T)^3} \\ &\times \left[\frac{\partial G}{\partial \lambda_{t-1}^T} \times \left(\frac{\partial G}{\partial \lambda_t^T} \right)^2 - \left(\frac{\partial^2 G}{\partial \lambda_t^T \partial \lambda_{t-1}^T} + \frac{\partial^2 G}{\partial \lambda_{t-1}^T \partial \lambda_t^T} \right) \times \frac{\partial G}{\partial \lambda_{t-1}^T} \times \frac{\partial G}{\partial \lambda_t^T} + \frac{\partial^2 G}{\partial(\lambda_t^T)^2} \times \left(\frac{\partial G}{\partial \lambda_{t-1}^T} \right)^2 \right]. \end{aligned} \quad (\text{C.17})$$

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

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

$$\begin{aligned} \frac{\partial^2 G}{\partial \lambda_t^T \partial \lambda_{t-1}^T} &= \frac{\partial^2 G}{\partial \lambda_{t-1}^T \partial \lambda_t^T} = \frac{2-\sigma}{\sigma-1} A(\lambda_t^T)^{3-2\sigma} \\ &\times \left[\frac{\theta - (\sigma-1)}{\theta} (\eta^{\sigma-1} - 1) (\lambda_t^T)^{-\frac{\sigma-1}{\theta}} \right] \exp(\delta \lambda_{t-1}^T) \lambda_{t-1}^T < 0. \end{aligned} \quad (\text{C.19})$$

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

Substituting Equations (C.13), (C.14), (C.18), (C.19), and (C.20) into Equation (C.17), 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.⁷⁵ Because $\lambda_t^T(\delta, \eta)$ strictly increases in δ and η , there exists $\underline{\delta}$ and $\underline{\eta}$ such that the 45 degree line and Equation (C.12) meet at $\lambda_{t-1}^T = 1$. Also, by the same logic, there exists $\bar{\delta}$ and $\bar{\eta}$ such that the 45 degree line is tangent to Equation (C.12). The two lines meet at least twice for $\delta \in [\underline{\delta}, \bar{\delta}]$ and $\eta \in [\underline{\eta}, \bar{\eta}]$. □

⁷⁵The intercept is always positive because of the unbounded Pareto distributional assumption which guarantees a positive share of adopters at $\lambda_{t-1}^T = 0$.

Proof of Proposition 1(v). The welfare of household is $\frac{w_t + \Pi_t}{P_t}$ where Π_t is the aggregate profits summed across all firms in the economy.⁷⁶ This can be expressed as $\frac{w_t}{P_t} + \frac{\Pi_t}{P_t}$. Using Equations (C.9) and (C.10) 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 $f(\lambda_{t-1}^T) A(\lambda_t^T)$. The welfare in the steady-state is $f(\lambda^{T*}) A(\lambda^{T*})$ which strictly increases in λ^{T*} . Therefore, the equilibrium with a larger mass of adopters Pareto-dominates the equilibrium with a smaller mass of adopters. \square

C.2.3 Source of Dynamic Externality

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

Suppose fixed adoption costs are in units of labor. The cutoff for adoption is defined as 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.7) 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.9) and (C.10) regardless of that fixed adoption costs are in units of labor. Substituting Equations (C.9) and (C.10) into the above cutoff, we can derive that

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

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

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

⁷⁶Note that $L_t = 1$.

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

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

Figure C1 graphically illustrates that temporary subsidies only have temporary effects when there exists a unique steady state. The solid red 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 when the subsidy \bar{s} is permanently provided. Point C is the steady state of this new short-run equilibrium curve. Therefore, between t_0 and t_1 , it converges to the new steady state C . However, after the end of the temporary subsidies at t_1 , the short-run equilibrium curve moves back to the original short-run equilibrium curve, and an economy jumps right to D and starts converging to the original steady state A .

Even if there is a unique steady state, there is still room for policy interventions due to the externalities. 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 permanently provided each period, which is similar to the static setting with externalities. However, these permanent policies are inconsistent with the industrialization pattern in South Korea, where adoption subsidies were only provided between 1973 and 1979.

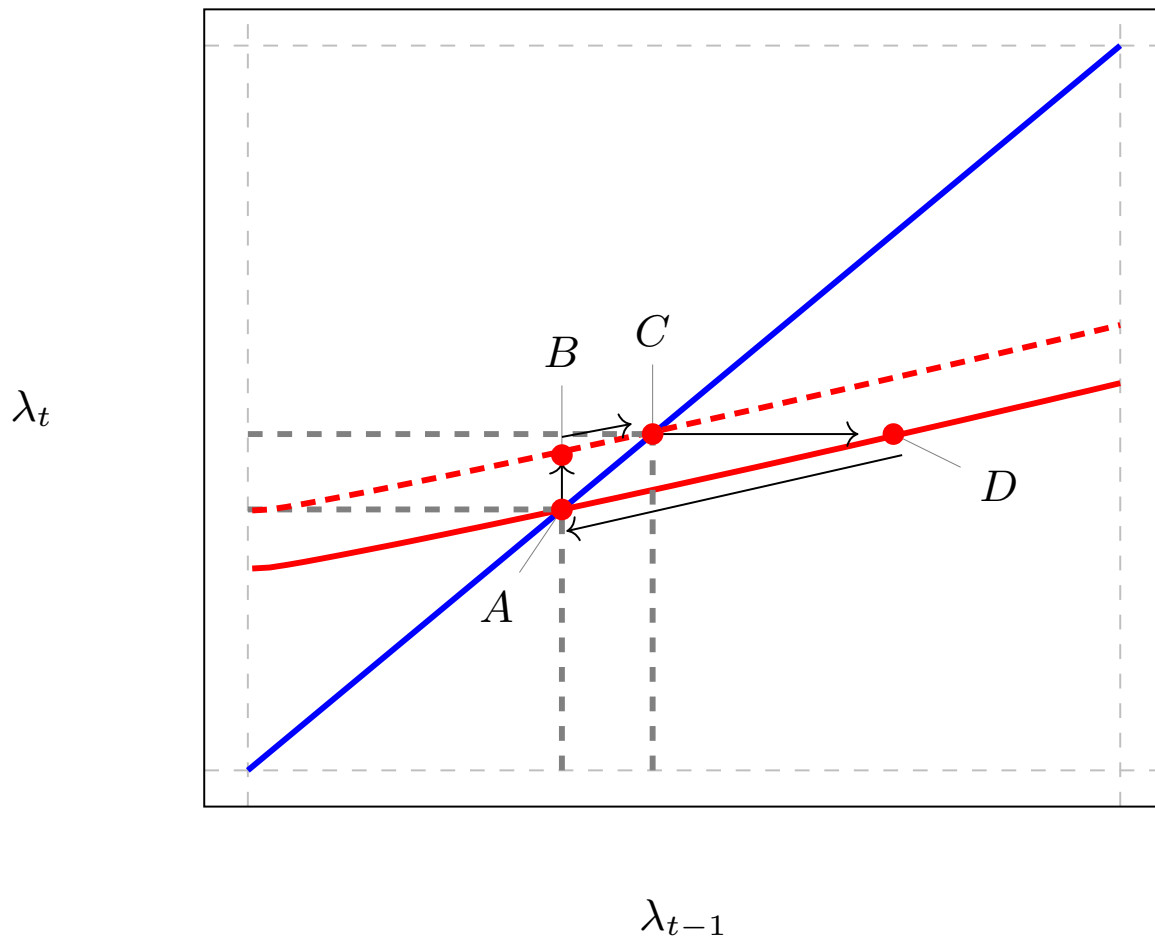


Figure C1. Temporary Subsidies Can Have Permanent Effects Only When Multiple Steady States Exist.

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

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

Proof of Proposition 2. Suppose that a subsidy plan of the government is given as follows:

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

Under the assumption that goods are freely traded, sectoral price index and real wage are equalized across regions, that is, $P_{njt} = P_{jt}$, $\forall n \in \mathcal{N}, \forall j \in \mathcal{J}$. Also, because of the symmetry assumption for $j \in \mathcal{J}^T$, $P_{jt} = P_{j't}$, $D_{jt}^f = D_{j't}^f$, and $F_j^T = F_{j'}^T$ hold for all $j, j' \in \mathcal{J}^T$.⁷⁷ These two assumptions in turn imply that firms in sectors where technology adoption is available face the same market size.

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

$$\begin{aligned} \ln \lambda_{njt}^T = & \underbrace{\theta \delta \lambda_{njt-1}^T + \frac{\theta}{\sigma-1} \ln \left(\left(\frac{\eta}{1-s_{njt}} \right)^{\sigma-1} - 1 \right)}_{\beta D_{jt}^{policy}} \\ & - \underbrace{\theta \ln \left(\frac{\mu c_{njt} (\sigma c_{njt} F_j^T)^{\frac{1}{\sigma-1}}}{\left(\sum_{m \in \mathcal{N}} P_{jt}^{\sigma-1} E_{mjt} + D_{jt}^f \right)^{\frac{1}{\sigma-1}}} \right)}_{=\delta_{nt}} + \underbrace{\theta \ln \phi_{njt}^{min}}_{=\epsilon_{njt}}, \quad (C.22) \end{aligned}$$

where the second, third, and fourth terms can be mapped to the policy dummy variable D_{jt}^{policy} which equals one if sector j was targeted by the government in period t , region fixed effects δ_{nt} , and the error term ϵ_{njt} .⁷⁸ This mapping gives us the following regression model:

$$\ln \lambda_{njt}^T - \theta \delta \lambda_{njt-1}^T = \beta D_{jt}^{policy} + \delta_{nt} + \epsilon_{njt}.$$

The condition for the estimates to be unbiased is $\mathbb{E}[\epsilon_{njt} | D_{jt}^{policy}] = 0$. Under the model structure, this is equivalent to $\mathbb{E}[\ln \phi_{njt}^{min} | D_{jt}^{policy}] = 0$ (Equation (C.22)). When this condition is satisfied,

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

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

⁷⁷ $\gamma_j^k = \gamma_{j'}^k$ and $\gamma_j^L = \gamma_{j'}^L$ for all for all $j, j' \in \mathcal{J}^T$, which leads to $P_{jt} = P_{j't}$ jointly with free trade assumption.

⁷⁸Variation in the third term of the RHS across regions comes from wages w_{nt} . Note that $c_{njt} = (w_{nt}/\alpha_j^L)^{\alpha_j^L} \prod_{k \in \mathcal{J}} (P_{nkt}/\alpha_k^k)^{\alpha_j^k}$.

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 sectors. Each firm faces a CES demand and is monopolistic for its own variety. Goods are freely tradable across regions.

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

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

$$\pi_{it} = \max_{T_{it} \in \{0,1\}, a_{it} \in [0,\infty)} \left\{ \frac{1}{\sigma} \left(\frac{\sigma}{\sigma-1} \frac{w_{nt}}{\tilde{\eta}^{T_{it}} a_{it}^{\gamma_1} \tilde{\phi}_{it}} \right)^{1-\sigma} P_t^{\sigma-1} E_t - T_{it} P_t F^T - w_{nt} a_{it}^{\alpha_1} g(\lambda_{nt-1}^T) B_t \right\}, \quad (\text{C.23})$$

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

The positive externalities of adoption come from $g(\lambda_{nt-1}^T)$ of the cost of innovation. We assume that $\frac{\partial g(\lambda_{nt-1}^T)}{\partial \lambda_{nt-1}^T} < 0$ holds, so that a larger share of adopters 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.⁷⁹

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}^*

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

⁷⁹This parameter restriction guarantees the second-order condition of a firm's maximization problem.

where \bar{C}_{nt}^1 is a collection of constants and variables that are common within region n .⁸⁰ 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 a share of adopters in the previous period λ_{nt-1}^T , technology adoption status T_{it} , and exogenous productivity $\tilde{\phi}_{it}$. Substituting the optimal a_{it}^* into Equation (C.23), a firm's maximization problem can be rewritten as:

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

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

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

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 sectors. Each firm faces a CES demand and is monopolistic for its own variety. Goods are freely tradable across regions.

In each region, there is a unit measure of engineers and firms. Managers live 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 that adopted technologies pass their knowledge to their children. This learning from parents increases the engineering skills of children when they become adults in the second period, which increases engineering skills by $\gamma_1 > 1$. If parents do not work in firms with foreign technology, their children's engineering skills are 1. We assume that the engineering skills of 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-

⁸⁰Specifically, $\bar{C}_{nt}^1 = \left[\frac{\sigma b_1 \alpha_1}{\gamma_1(\sigma-1)} \left(\frac{\sigma}{\sigma-1} \right)^{\sigma-1} \right]^{\frac{1}{\gamma_1(\sigma-1)-\alpha_1+1}} w_{nt}^{\frac{\gamma_1(\sigma-1)-\sigma}{\gamma_1(\sigma-1)-\alpha_1+1}}$.

one. The surplus from a 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 an 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) exogenous productivity $\tilde{\phi}_{it}$ iid drawn each period, (2) engineering skills of matched engineers, and (3) adoption decisions.

Firms' Profits Maximization Problem. Because of the random matching, firms are matched with engineers with higher engineering skills γ_1 with the probability of λ_{nt-1}^T and with engineers with lower skills 1 with the probability of $1 - \lambda_{nt-1}^T$.

A firm's maximization problem can be written as

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

where λ_{nt-1} is a local share of adopters in the previous period; $\tilde{\phi}_{it}$ is exogenous productivity; w_{nt} is a local wage; T_{it} is a binary adoption decision; F^T is a fixed adoption cost in units of final goods; γ_1 is engineering skills of engineers whose parents worked in 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

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

where

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

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

Historical Case Study. In the 1970s, labor mobility across firms was high in Korea (Kim and Topel, 1995). The average duration of a job in the manufacturing sector 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, Chapter 7) provides one historical case study on the diffusion of knowledge through labor mobility in steel industry. The Pohang Iron and Steel Company Ltd. (POSCO) was the nation's first integrated steel mill, which started operating in 1973. Given Korea's lack of technology, the imported technology

played a significant role for POSCO at the initial stages. The government heavily subsidized POSCO for the adoption of technology and installation of imported capital equipment associated with the imported technologies. Some of the technicians quitting POSCO got employed in capital good producing firms located near POSCO. In these capital goods producing firms, the technicians helped them produce capital equipment used in POSCO, such as water treatment, dust collection, and the large magnetic crane. In the early 1970s, this capital equipment was all imported, but it started to be produced by local suppliers because of knowledge spillover from technicians who worked in POSCO.

Enos and Park (1988, p. 166) provides another example. 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 Industries (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 two historical case studies 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	Obs.	Mean	Med.	SD	Obs.	(Col. 1 - Col. 5)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
log sales	17.80	18.21	2.22	133	18.46	18.45	1.78	131	2.36 [0.13]
log employment	7.34	7.60	1.23	109	7.07	7.19	1.54	130	0.23 [0.64]
log fixed assets	17.15	17.10	2.26	162	17.19	17.64	2.26	158	0.01 [0.93]
log assets	18.00	17.99	2.10	162	18.12	18.40	2.08	158	0.07 [0.80]
log value added/emp	9.57	9.70	1.26	102	9.95	9.62	1.35	122	1.55 [0.22]

Notes. 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 with its p-value in brackets. 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	✓	✓	✓	✓
Pair FE		✓		✓

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 adopted technology in the event time. Each cell in columns (1) and (2) reports estimates from a separate bivariate regression. F-statistics of joint significance are reported for multivariate regressions, and their p-values are reported in brackets. Standard errors are two-way clustered by pair and firm and reported in parenthesis. The number of pairs and firms are 33 and 55.

Table D3: Descriptive Statistics of Patenting Activities by Foreign Contractors: Winners vs. Losers Design Samples

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

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

Table D4: Pooled diff-in-diffs Estimates of Direct Productivity Gains to Adopters: Winners vs. Losers Research Design

Dep. Var.	log sales	log labor	log revenue TFP			
		productivity	W. (2009)	ACF (2015)	LP (2003)	OLS
	(1)	(2)	(3)	(4)	(5)	(6)
$Post_{it} \times \mathbb{1}[Adopt_{it}]$	0.51*** (0.17)	0.64** (0.25)	0.46*** (0.16)	0.62*** (0.20)	0.45*** (0.16)	0.60*** (0.20)
Adj. R^2	0.88	0.62	0.87	0.94	0.90	0.61
# cluster (pair)	34	34	34	34	34	34
# cluster (firm)	57	57	57	57	57	57
N	951	835	827	827	827	827

Notes. The table reports the pooled diff-in-diffs estimator based on the winners vs. losers research design. The dependent variables are log sales, log revenue TFP, and log labor productivity defined as value added divided by employment. Value added is obtained as sales multiplied by the value added shares obtained from input-output tables corresponding to each year. In columns (3), (4), (5), and (6), revenue TFPs are estimated based on [Wooldridge \(2009\)](#), [Akerberg et al. \(2015\)](#), [Levinsohn and Petrin \(2003\)](#), and OLS. All specifications control for event time dummies, firm fixed effects, pair fixed effects, and calendar year fixed effects. Robust standard errors in parenthesis are two-way clustered at pair and firm levels. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table D5: Local Productivity Spillovers from Technology Adoption: Robustness - 3 Year Lag

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

Notes. This table reports the OLS estimates of Equation (4.4). When we construct the spillover measure defined in Equation (4.2), we lag the adoption status of firms 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 of both adopters and non-adopters and additional control for 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(Input-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 conglomerate levels and are reported in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table D6: Local Productivity Spillovers from Technology Adoption: Robustness - 5 Year Lag

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

Notes. The table reports the OLS estimates of Equation (4.4). When we construct 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 including both adopters and non-adopters and additional control for 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(Input-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 conglomerate levels and are reported in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table D7: Local Productivity Spillovers from Technology Adoption: Robustness - Alternative Dependent Variables: Log Employment and Labor Productivity

Dep. Var.	log employment					log labor productivity				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
<i>Panel A: Never-Adopter Sample</i>										
Spill	4.39*** (1.54)	3.79** (1.64)	4.94*** (1.70)	4.23*** (1.50)	4.07** (1.76)	5.55*** (1.84)	5.41*** (1.62)	5.81*** (2.08)	5.34*** (1.78)	5.11** (1.92)
ln(Spill-Sales)			-0.02 (0.01)		-0.02 (0.01)			-0.02 (0.02)		-0.01 (0.02)
ln(Input-MA)				-0.03 (0.02)	-0.02 (0.02)				-0.04** (0.02)	-0.03 (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 (conglomerate)	636	630	636	636	630	324	275	324	324	275
N	1079	1073	1079	1079	1073	344	292	344	344	292
<i>Panel B: Full Sample</i>										
Spill	4.23*** (1.18)	3.93*** (1.43)	4.45*** (1.31)	3.86*** (1.19)	3.72** (1.52)	4.75*** (1.63)	3.99** (1.90)	4.72*** (1.73)	4.45*** (1.58)	3.44* (1.82)
1[Adopt]	0.32** (0.15)	0.26 (0.20)	0.32** (0.15)	0.31** (0.15)	0.25 (0.19)	0.15* (0.09)	0.14 (0.10)	0.15* (0.09)	0.14 (0.09)	0.12 (0.10)
ln(Spill-Sales)			-0.01 (0.01)		-0.01 (0.01)			0.00 (0.02)		0.00 (0.02)
ln(Input-MA)				-0.05*** (0.02)	-0.04* (0.02)				-0.05*** (0.02)	-0.05** (0.02)
Adj. R^2	0.19	0.24	0.19	0.19	0.24	0.37	0.43	0.37	0.38	0.43
# clusters (region)	54	54	54	54	54	45	41	45	45	41
# clusters (conglomerate)	702	697	702	702	697	381	338	381	381	338
N	1264	1259	1264	1264	1259	431	387	431	431	387
Region-Sector FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Conglomerate FE		✓			✓		✓			✓

Notes. The table reports the OLS estimates of Equation (4.4). When we construct 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 including both adopters and non-adopters and additional control for 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(Input-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 conglomerate levels and are reported in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table D8: Local Productivity Spillover from Technology Adoption: Robustness - Alternative Dependent Variables: Log Fixed Assets and Assets

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

Notes. The table reports the OLS estimates of Equation (4.4). When we construct 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 including both adopters and non-adopters and additional control for adopters' adoption status. Dependent variables are log fixed assets and assets in columns (1)-(5) and (6)-(10), respectively. ln(Spill-Sales) and ln(Input-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 conglomerate levels and are reported in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table D9: Local Productivity Spillovers from Technology Adoption: Robustness - Input Market Access

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

Notes. The table reports the OLS estimates of Equation (4.4). When we construct 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 including both adopters and non-adopters and additional control for 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(Input-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 conglomerate levels and are reported in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

D.2 Additional Figures

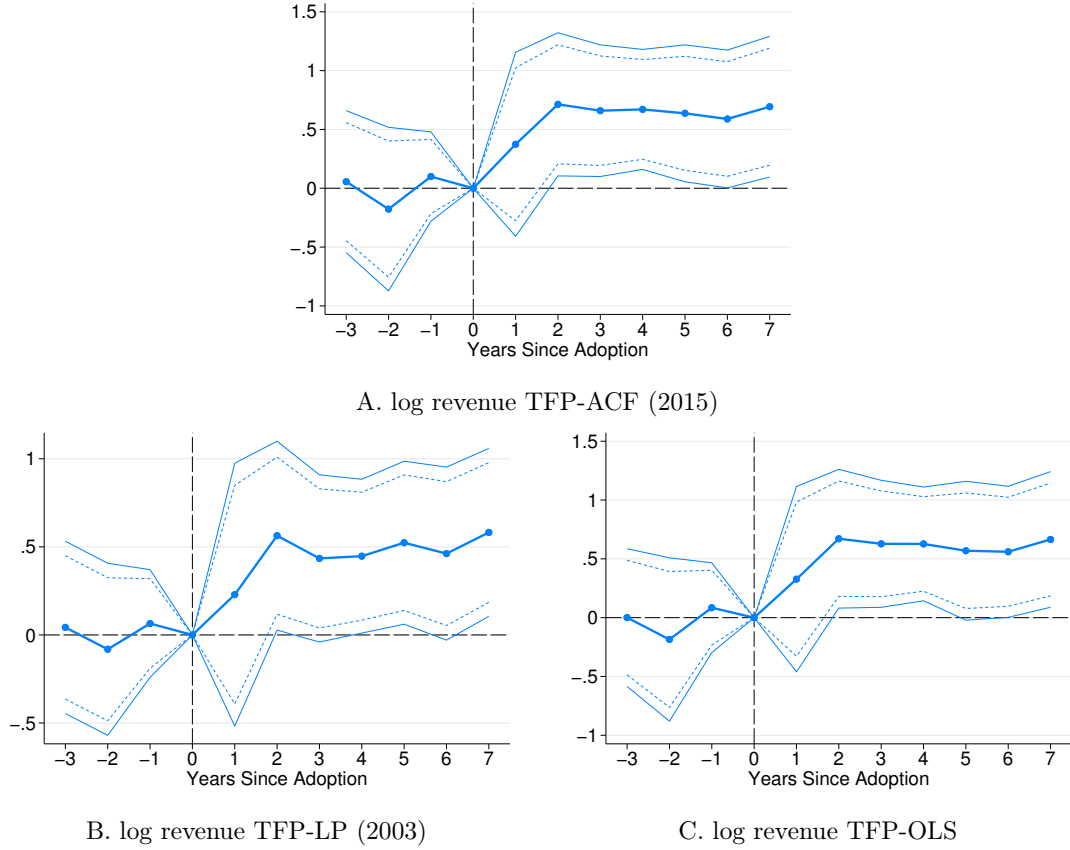


Figure D1. Direct Productivity Gains of Technology Adoption: Winners vs. Losers Research Design - Robustness, Alternative TFP measures

Notes. This figure illustrates the estimated β_{τ}^{diff} in Equation (4.1) based on winners vs. losers research design. The dependent variables are log revenue TFP. In Panels A, B, and C, revenue TFPs are estimated based on [Akerberg 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 figure reports 90 and 95 percent confidence intervals based on standard errors two-way clustered at the levels of pairs and firms.

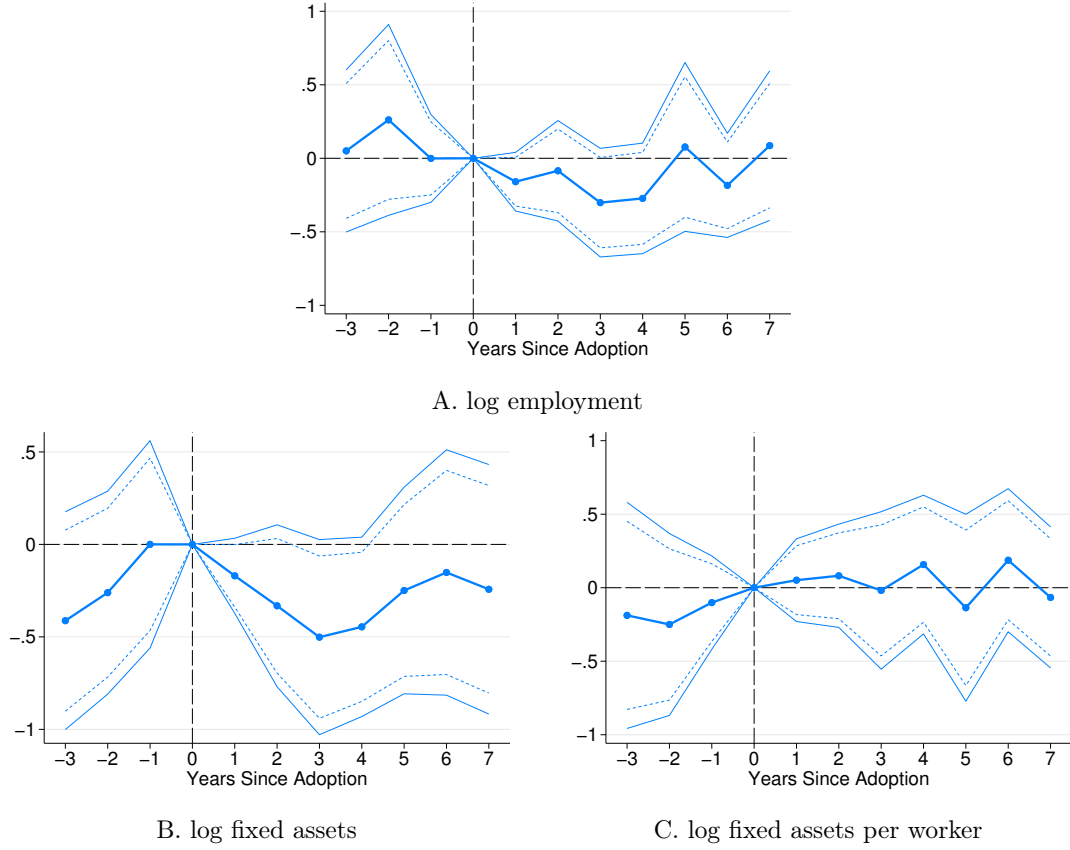


Figure D2. Inputs and Technology Adoption: Winners vs. Losers Research Design

Notes. This 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 zero. All specifications control for event time dummies, and firm, pair, and calendar year fixed effects. The figure reports 90 and 95 percent confidence intervals based on standard errors two-way clustered at the levels of pairs and firms.

D.3 Matching Algorithm

This section describes the matching algorithm used for matching a loser to a winner 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 exactly matched 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 \mathbf{X}^e with a loser f ;
 - minimize the Mahalanobis distance with loser f in terms of \mathbf{X}^d :

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

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

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

D.4 Production Function Estimation

In this section, we discuss the estimation procedure for revenue TFP measures. The revenue TFP measures are obtained as the residuals after estimating the production using different methodologies: [Wooldridge \(2009\)](#), [Levinsohn and Petrin \(2003\)](#), [Akerberg et al. \(2015\)](#), and OLS. We estimate the following Cobb-Douglas value added production function:

$$\log VA_{it} = \alpha_L \log L_{it} + \alpha_K \log K_{it} + u_{it}, \quad (\text{D.1})$$

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

When using methodologies developed by [Wooldridge \(2009\)](#), [Levinsohn and Petrin \(2003\)](#), and [Akerberg et al. \(2015\)](#), we use material inputs as a proxy variable. However, information on material inputs is not available for our main firm-balance sheet data digitized from *Annual Reports of Korean Companies*. Therefore, we estimate the production function separately for each sector using the alternative firm-level data. We used *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 but not value added, we calculate value added as sales times the value added shares from the input-output tables of corresponding years. Using these estimated coefficients from *KIS-VALUE*, we obtain revenue TFP for the sample period between 1970 and 1982.

Appendix E Appendix: Quantification

E.1 Additional Figures

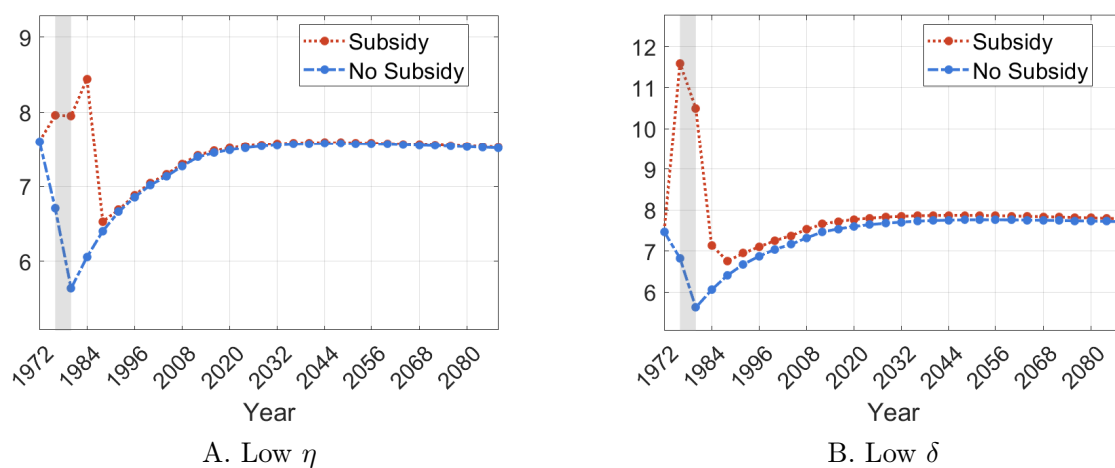
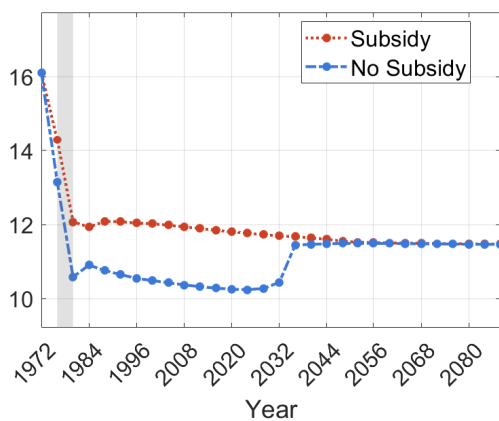
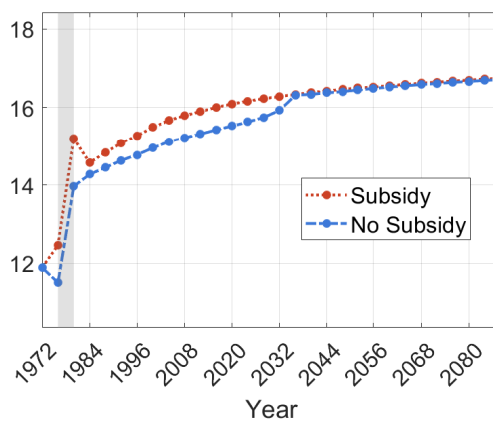


Figure E1. Heavy Mfg. GDP Shares. Comparative Statistics of δ and η

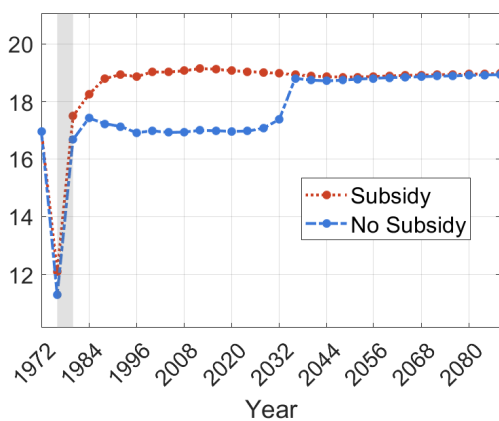
Notes. This figure plots the comparative statistics of δ and η . In Panel A, η is set 1.05. In Panel B, δ is set to be 1. The red dotted line and the blue dashed lines plot outcomes of the baseline and counterfactual economies.



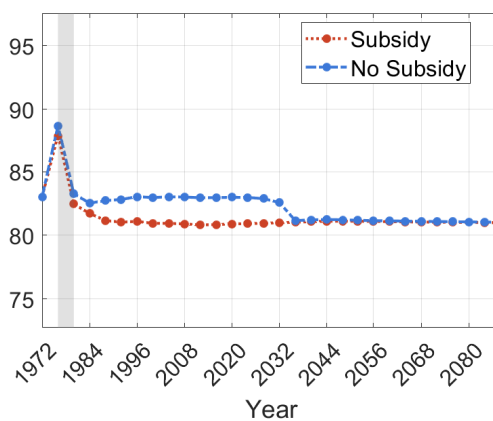
A. Heavy mfg. GDP share (%)



B. Heavy mfg. employment share (%)



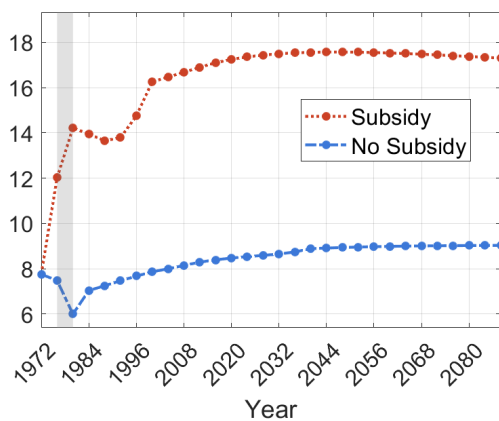
C. Heavy mfg. export share (%)



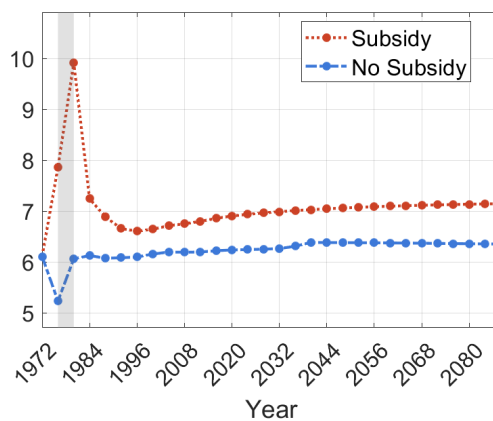
D. Light mfg. export share (%)

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

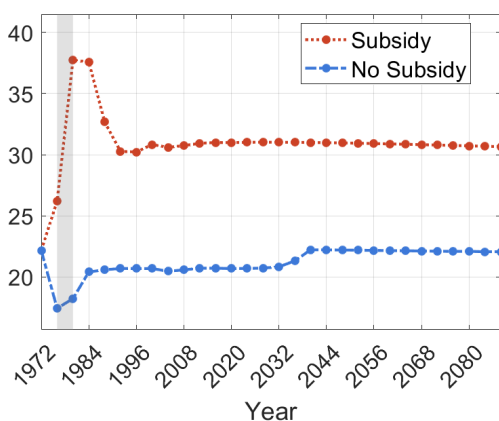
Notes. This figure plots the counterfactual results without a roundabout production structure. Panels A, B, C, and D report the results for the heavy manufacturing sector employment, GDP, and export shares, and the light manufacturing sector export shares, respectively. The red dotted and the blue dashed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



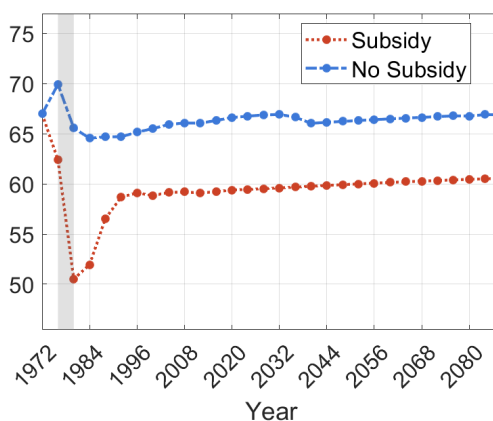
A. Heavy mfg. GDP share (%)



B. Heavy mfg. employment share (%)



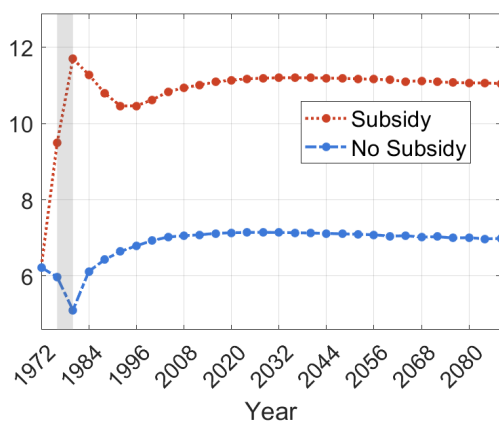
C. Heavy mfg. export share (%)



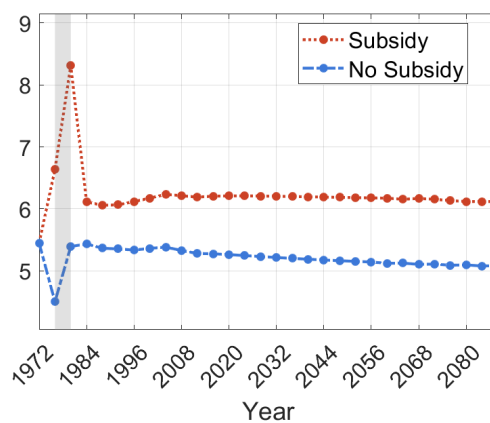
D. Light mfg. export share (%)

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

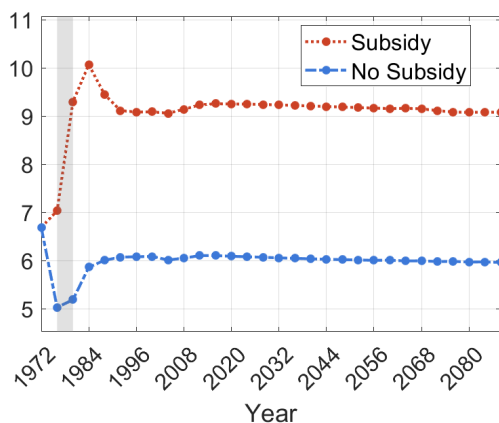
Notes. This figure plots the counterfactual results with a higher level of migration costs than the baseline counterfactual exercises. Panels A, B, C, and D report the results for the heavy manufacturing sector employment, GDP, and export shares, and the light manufacturing sector export shares, respectively.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



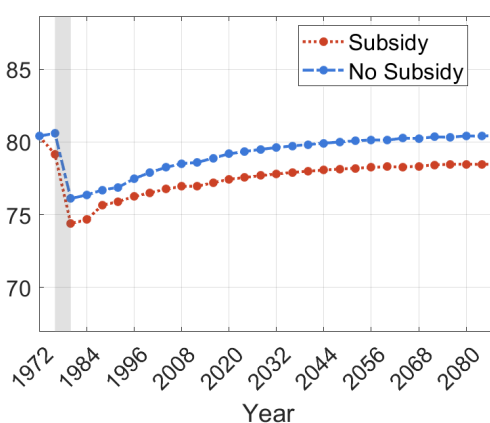
A. Heavy mfg. GDP share (%)



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 lines plot 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: $\{\lambda_{njt_0}^T\}_{n \in \mathcal{N}, j \in \mathcal{J}^T, t_0=1968}$
 - Initial population distribution: $\{L_{nt_0}^{Data}\}_{n \in \mathcal{N}, t_0=1968}$
 2. Sectoral gross output of each region: $\{GO_{njt}^{Data}\}_{n \in \mathcal{N}, j \in \mathcal{J}, t \in \{1972, 1976, 1980\}}$
 3. Regional population: $\{L_{nt}^{Data}\}_{n \in \mathcal{N}, t \in \{1972, 1976, 1980\}}$
 4. Sectoral export shares at national level: $\{EX_{jt}^{Data}/GO_{jt}^{Data}\}_{j \in \mathcal{J}, t \in \{1972, 1976, 1980\}}$ where EX_{jt}^{Data} and GO_{jt}^{Data} are sector j 's exports and gross output at national level
 5. Sectoral import shares at national level: $\{IM_{jt}^{Data}/E_{jt}^{Data}\}_{j \in \mathcal{J}, t \in \{1972, 1976, 1980\}}$ where IM_{jt}^{Data} and E_{jt}^{Data} are imports and the total expenditure on sector j goods at national level
 6. Import and export tariffs: $\{t_{jt}^{im}\}_{j \in \mathcal{J}, t \in \{1972, 1976, 1980\}}$ and $\{t_{jt}^{ex}\}_{j \in \mathcal{J}, t \in \{1972, 1976, 1980\}}$
- Micro moments
 1. Identifying moment $\hat{\beta}^{policy}$ (Equation (6.4))
 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 Θ^E and data inputs as given, we obtain the values of Θ^M , $\{\bar{s}\}_{t \in \{1976, 1980\}}$, and Ψ_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 $\{\phi_{nj}^{min}\}_{n \in \mathcal{N}, j \in \mathcal{J}}$
3. Given parameters $\{\Theta^M, \bar{s}_t\}$, we solve the model and update the fundamentals Ψ_t for each period. We fit region-sector level aggregate outcomes to the data counterparts. This step corresponds to the constraints of Equation (6.2). 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 $\{D_{jt}^{f'}\}$ using the following equation:

$$\underbrace{\frac{EX_{jt}^{Data}}{GO_{jt}^{Data}}}_{\text{Data}} = \frac{\sum_{n \in \mathcal{N}} \left(\frac{\sigma}{\sigma-1} \frac{c_{njt} t_{jt}^{ex} \tau_{nj}^x}{\phi_{njt}^{avg, x}} \right)^{1-\sigma} D_{jt}^{f'}}{\underbrace{\sum_{n \in \mathcal{N}} \left(\frac{\sigma}{\sigma-1} \frac{c_{njt}}{\phi_{njt}^{avg}} \right)^{1-\sigma} \left(\sum_{m \in \mathcal{N}} \tau_{nmj} P_{mjt}^{\sigma-1} E_{mjt} \right) + \left(\frac{\sigma}{\sigma-1} \frac{c_{njt} t_{jt}^{ex} \tau_{nj}^x}{\phi_{njt}^{avg, x}} \right)^{1-\sigma} D_{jt}^{f'}}_{\text{Model}}}$$

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

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

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

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

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:

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

where

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

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

$\tilde{\lambda}_{njt}^T = (1 - \kappa^{-\theta})\lambda_{njt}^T + \kappa^{-\theta}$, 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 geographic fundamentals, given values of parameters and subsidies, evaluate the following objective function:

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

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

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

E.3 Construction of Data Inputs

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

E.3.1 Aggregate Data

Initial Shares of Adopters in 1968. Our firm-level balance sheet data covers between 1970 and 1982, whereas technology adoption contracts cover 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.⁸¹

Regional Population Distributions in 1968, 1972, 1976, and 1980. The regional population data comes from the Population and Housing Census, the 2% random sample of the total population. The survey was conducted in 1966, 1970, 1975, and 1980. For the years not covered by this Census survey, we impute population using the geometric average using the two observed samples. For example, the population share of region n in 1973 is imputed as $\text{Pop. share}_{n,1973} = (\text{Pop. share}_{n,1970})^{\frac{3}{5}} \times (\text{Pop. share}_{n,1975})^{\frac{2}{5}}$. 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.

⁸¹Given 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.

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

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

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

Aggregate Export and Import Shares in 1972, 1976, and 1980. Both aggregate export and import shares are obtained from the national-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 tables. In the model, we treat the service sector as a non-tradable sector, so we assume that exports and imports of the service sector 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 follows:

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

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

Export and Import Tariffs Data in 1972, 1976, and 1980. Export and import tariffs data are not used for the reduced-form empirical analysis but only for the quantitative exercises. The export tariffs data are obtained from Magee (1986).⁸² 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.⁸³

Import tariff data are 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.

⁸²The US export tariff data was downloaded from <https://cid.econ.ucdavis.edu/ust.html>.

⁸³The concordance between 1972 SIC and 1987 SIC is obtained from "www.nber.com."

E.3.2 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, using 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.

E.4 Identifying Moment for Subsidy

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:

$$\ln \lambda_{n,heavy,t}^T = \alpha + \beta^{policy} \times D_t^{policy} \beta_1 \lambda_{n,heavy,t-1}^T + \epsilon_{n,heavy,t}, \quad (\text{E.1})$$

where D_t^{policy} is a dummy variable which equals one in 1980; and $\lambda_{n,heavy,t}^T$ 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 $\lambda_{n,heavy,t-1}^T$ 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. $\lambda_{n,heavy,t}^T$	(1)
D_t^{policy}	0.65** (0.25)
$\lambda_{n,heavy,t-1}^T$	5.62*** (0.80)
# of clusters (region)	42
N	84

Notes. The table reports the 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:

$$\sum_{n \in \mathcal{N}} \mathbf{X}_{nt} (\lambda_{n,heavy,t}^T - \exp(\mathbf{X}_{nt}' \boldsymbol{\beta})), \quad (\text{E.2})$$

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

Connection with Proposition 2. Note that the assumptions of Proposition 2 are not satisfied in both data and model, and we run the regression only for the heavy manufacturing sector. However, we show that $\hat{\beta}^{policy}$ is still informative on \bar{s} .

Under the unbounded Pareto distributional assumption, we can derive the following relationship

from the model for the heavy manufacturing sector without any additional assumptions:

$$\begin{aligned}
& \ln \lambda_{n,heavy,t}^T - \theta \delta \lambda_{n,heavy,t-1}^T \\
&= \underbrace{\frac{\theta}{\sigma-1} \ln \left(\left(\frac{\eta}{1-s_{n,heavy,t}} \right)^{\sigma-1} - 1 \right) - \theta \ln \left(\frac{\mu c_{n,heavy,t} (\sigma c_{n,heavy,t} F_{heavy}^T)^{\frac{1}{\sigma-1}}}{\left(\sum_{m \in \mathcal{N}} P_{heavy,t}^{\sigma-1} E_{m,heavy,t} + D_{heavy,t}^f \right)^{\frac{1}{\sigma-1}}} \right)}_{=D_t^{policy}} \\
&\quad \underbrace{+ \theta \ln \phi_{n,heavy,t}^{min}}_{=\epsilon_{n,heavy,t}}.
\end{aligned}$$

Because the heavy manufacturing sector is the only sector where technology adoption is available and targeted by the government, D_t^{policy} are not separately identified from additional time fixed effects.

D_t^{policy} captures both the subsidies in the second term of the RHS, and the general equilibrium effects in the third term of the RHS. $\mathbf{GE}_{n,heavy,t}(\Psi_t, \mathbf{s}_t)$ which depend all other regions' geographic fundamentals Ψ_t and subsidies \mathbf{s}_t . $\mathbf{GE}_{n,heavy,t}(\Psi_t, \mathbf{s}_t)$ is a function of own exogenous natural advantage in the error term and therefore is correlated with the error term, which leads to the endogeneity problem of the above regression model. In Proposition 2, by imposing the additional assumptions, these general equilibrium effects could be absorbed out by region fixed effects, which is not the case in the above regression model.

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

E.5 Gravity Equation of Migration Flows

The data on migration shares comes from the 1995 Population and Housing Census, which is the closest to the sample period of our data set among the accessible population census data. Because of the data availability, regions are aggregated up to 35 regions. μ_{nm}^{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 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_{mnt}^d that is orthogonal to distance between two regions: $d_{mn} = (dist_{mn})^{-\zeta} \epsilon_{mnt}^d$. Taking log of Equation (5.10), we can derive the following regression model:

$$\begin{aligned} \ln \mu_{mn}^{1995} = & -\nu\zeta \log dist_{mn} + \underbrace{\ln \left(V_{n,1995} \frac{(1 - \bar{\tau}_{1995}^w + \bar{\pi}_{1995}^h) w_{nt}}{P_{n,1995}} \right)}_{=\delta_n} \\ & + \underbrace{\ln \left(\sum_{n'=1}^N \left(V_{n',1995} \frac{(1 - \tau_{1995}^w + \bar{\pi}_{1995}^h) w_{n',1995}}{P_{n',1995}} d_{mn'} \right)^\nu \right)}_{\delta_m} + \epsilon_{mnt}^d, \end{aligned}$$

which gives Equation (6.1). We estimate the above equation using OLS and PPML. The results are reported in Table E2. The estimated coefficient is around -1.30. The magnitude of the estimate implies that one percent increase in distance decreases outward migration share by 1.3%.

Table E2: Gravity Equation of Migration Shares

Dep. Var.	Migration Shares between 1990 and 1995	
	OLS	PPML
	(1)	(2)
$\log Dist_{mn}$	-1.30*** (0.06)	-1.39*** (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 (6.1). 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. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.