Technology Adoption and Late Industrialization *

Jaedo Choi
Federal Reserve Board

Younghun Shim
University of Chicago

July, 2023

Abstract

We study how foreign technology adoption contributed to late industrialization in developing countries. Using novel historical data for South Korea, we provide three empirical findings on firm-level effects of technology adoption: direct effects on adopters, local spillovers, and local complementarity in the adoption. We develop a dynamic spatial general equilibrium model consistent with these findings. Due to the complementarity, the model potentially features multiple steady states. Using this model, we evaluate the policy that temporarily provided adoption subsidies to heavy manufacturing firms. Our results suggest such a big push policy could have had permanent effects by moving the economy to a more industrialized steady state with higher heavy manufacturing GDP shares and larger amounts of the adoption.

Keywords: technology adoption, industrialization, complementarity, big push, knowledge spillover
JEL Codes: O14, O33, O53, R12

*The views expressed in this paper are our own, and do not represent the views of the Board of Governors of the Federal Reserve, nor any other person associated with the Federal Reserve System. We are grateful to Andrei Levchenko, Dominick Bartelme, Jagadeesh Sivadasan, and Sebastian Sotelo for their guidance and invaluable suggestions. We also thank Costas Arkolakis, Maria Aristizabal-Ramírez, David Atkin, Natalie Bau, Barthélémy Bonadio, Yongsung Chang, Victor Couture, Javier Cravino, Dave Donaldson, Elisa Giannone (discussant), Kyle Handley, Ricardo Hausmann, Oleg Itskhoki, Yongjun Kim, Sam Kortum, Mos Laoprapassorn, Kevin Lim, Kiminori Matsuyama, Isabelle Mejean (discussant), Peter Morrow, Nathan Nunn, Nishaad Rao, Ulrich Schetter, Yongseok Shin, Michael Song, Dan Trefler, Nick Tsivanidis, Jon Vogel, Kei-Mu Yi, Xiaodong Zhu, and seminar participants at Berkeley Haas, CUHK, ETSG, EWMES, FRB, Harvard Growth Lab, HKU, HKUST, IGC-Stanford Conference on Firms, Trade, and Development, JHU-SAIS, KDI, KIEP, KIF, KIPF, LSE, Maastricht, Michigan, MWIT, NTU, PSU (New Faces in International Economics), Rochester, STEG, SNU, Toronto, UBC, UCLA, USC/UNIL/Bank-al-Maghrib Workshop, VEAMS, Yale, and Yonsei for helpful comments and discussions. We gratefully acknowledge research support from the Kwanjeong Educational Foundation. E-mail: jaedo.choi@frb.gov.

1 Introduction

In the postwar period, patterns of industrialization among developing countries diverged. The economic base of some latecomers such as South Korea, Taiwan, and Turkey transformed from agriculture to manufacturing, while many others remained stagnant. These latecomers achieved industrialization by adopting foreign technology rather than developing their own technology, known as late industrialization (Amsden, 1989).[1] Although late industrialization among these latecomers provides suggestive evidence about the importance of technology adoption for economic development (e.g. Parente and Prescott, 2002), little is known about the role of the adoption during their industrialization. The key challenge is that technology adoption is typically not observed directly due to the unavailability of detailed data.

This paper studies how the adoption of foreign technology contributes to late industrialization. Out setting is South Korea in the 1970s. There are two reasons why our setting is of interest. First, South Korea is known for the most remarkable economic transformations among the latecomers (Lucas, 1993). Second, during our sample period, the government implemented a policy that temporarily provided subsidies for the adoption. When explaining the economic development of South Korea, the abrupt industrialization that coincided with this one-time policy has led to conjectures on the role of the big push (Rosenstein-Rodan, 1943). Our setting provides a good laboratory to explore the effects of such a big push policy for technology adoption.

This paper makes three contributions related to technology adoption: measurement, empirics, and quantification. First, to directly measure firm-level adoption activities, we construct a novel historical dataset that covers the universe of technology adoption activities by South Korean firms. Second, we provide three novel empirical evidence on firm-level effects of technology adoption: direct effects on adopters, local spillovers, and complementarity in firms' adoption decisions. Third, we set up a quantitative model consistent with these three empirical findings. We analytically show that these three findings embedded in the model potentially generate multiple equilibria and use this model to evaluate the big push policy implemented by the Korean government.

Our dataset covers the universe of technology adoption contracts between South Korean and foreign firms. When making adoption contracts, the government required firms to report related contract documents to its authorities. We hand-collected and digitized these documents from the historical archive. Most of the adopted technology was related to knowledge about how to build and operate plants and capital equipment related to mass production. The data uncover a novel pattern that while the aggregate heavy manufacturing shares of the GDP of South Korea increased from 6% to 14%, there were large inflows of new technologies among heavy manufacturing firms through contracts, with yearly new contracts made being more than quadrupled.

[1] 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. "If industrialization first occurred in England on the basis of invention, and if it occurred in Germany and the US on the basis of innovation, then it occurs now among "backward" countries on the basis of learning" (Amsden, 1989, p. 4).

Using this constructed dataset, we provide three empirical evidence on the firm-level effects of technology adoption. These three findings uncover how technology adoption affected firm performance in the late-industrializing economy. Our first finding is the direct effects on adopters. An empirical challenge to identify the direct effects is selection bias due to endogenous adoption decisions. We overcome this challenge by using winners vs. losers research design (Greenstone et al., 2010). We compare firms that successfully adopted technology (winners) and firms that received the approval from the government to pursue foreign technology and made a contract with a foreign firm but *failed* to or *got delayed* to adopt technology because the foreign firm canceled the contract due to circumstances seemingly unrelated to the South Korean firm (losers). We construct matches of winners and losers by matching each loser to winners that are observationally similar. Using these matches, we adopt the stacked-by-event design (Deshpande and Li, 2019; Cengiz et al., 2019) in which the treatment effects are estimated only based on comparisons between winners and never-treated or not-yet-treated losers. We find technology adoption increased winners' sales and revenue-based total factor productivity (TFP) by around 100% and 65%, respectively.

Our second finding is local spillovers of the adoption. We regress growth in sales or revenue-based TFP on changes in local region-sector level shares of adopters, conditioning on fixed effects and other controls. To identify the local spillovers, we propose an instrumental variable (IV) strategy based on business groups' spatial networks of affiliated firms across region-sectors. We use variation in changes in the adoption status of firms outside of a region that are affiliated with business groups that own at least one firm in that local region. We estimate semi-elasticities of sales and revenue-based TFP with respect to local shares around 4% and 1.2%, respectively.

Our third finding is local complementarity in firms' adoption decisions. We regress a dummy variable of making a new adoption contract on local region-sector level shares of adopters. We use the same IV strategy for the spillover regression. We find that increases in local adopter shares cause increases in the probabilities of making new contracts.

Motivated by these three empirical findings, we develop a simple model with firms' technology adoption decisions and spillovers, which rationalizes the possibility of the big push policy. Firms can adopt a more productive modern technology after incurring a fixed adoption cost in units of final goods. The spillover operates with a one-period lag, where the current productivity increases in the shares of adopters in the previous period. This lag is a source of dynamics of the model and the share of adopters becomes a time-varying state variable. Consistent with the third empirical finding, the model features dynamic complementarity in firms' adoption decisions, that is, a higher share of adopters in the previous period leads to a higher share in the current period. The source of the complementarity comes from that fixed adoption costs are in units of final goods because the lagged spillover lowers fixed adoption costs of the current period. We analytically show that dynamic complementarity can lead to multiple steady states. There are pre-industrialized and industrialized steady states with low and high shares of adopters. An initial condition determines which steady state is realized in the long run through path-dependence. A big push policy that provides a one-time subsidy for the adoption can have permanent effects by moving an economy out of initial conditions that make the economy converge to the pre-industrialized steady state.

To conduct the counterfactual analysis of the policy, we extend this simple model with multiple elements including internal and international trade, input-output (IO) linkages, and migration. We calibrate the model to both micro and regional data. The model is tightly connected to the data. The model delivers structural equations that can be mapped to our reduced-form regression specifications. From this mapping, we pin down two key parameters that govern the strength of the direct productivity gains and the spillover.

Using the calibrated model, we evaluate how the pattern of industrialization in South Korea would have evolved differently, had the government not implemented the big push policy. Our results show that without the big push, South Korea could have converged to an alternative less-industrialized steady state. In this steady state, the heavy manufacturing sector's share of GDP would have decreased by 1.4 percentage points lower, and its export shares to total exports would have been 12.5 percentage points lower than the steady state of the baseline economy in which the policy had been implemented.

Related literature Our paper contributes to three strands of the literature. The first is the empirical literature that studies firm-level effects of industrial technology adoption in developing countries (e.g., Atkin et al., 2017; Juhász, 2018; Juhász et al., 2020; de Souza, 2021; Giorcelli and Li, 2021; Hardy and McCasland, 2021). We contribute to this literature by providing novel empirical findings on firm-level effects of technology adoption. Our data set covers the universe of technology adoption activities in South Korea known for its growth miracle, which allows us to derive aggregate implications from micro-level findings. Moreover, our spillover findings are consistent with previous studies on spillover effects of foreign direct investment and new technologies (e.g., Keller, 2002; Javorcik, 2004; Giorcelli, 2019; Bai et al., 2020; Alfaro-Ureña et al., 2022; Bianchi and Giorcelli, 2022). Unlike these studies, we provide new evidence on local complementarity in firms' technology adoption decisions.

Second, this paper is related to the literature on multiple equilibria and the big push that studies underdevelopment due to coordination failures (e.g., Rosenstein-Rodan, 1943; Hirschman, 1958; Murphy et al., 1989; Matsuyama, 1991, 1995; Redding, 1996; Rodríguez-Clare, 1996; Ciccone, 2002). Although the possibility of the big push has been theoretically examined in previous studies, its quantitative aspects are not well-known. We contribute to this literature by quantitatively exploring the possibility of the big push and evaluating the actual policy. Kline and Moretti (2014) and Buera et al. (2021) are the two most closely related papers that study Tennessee Valley Authority program in the US and the complementarity in technology adoption decisions and its interaction with distortions, respectively. Unlike Kline and Moretti (2014), our results suggest that the big push could have played a role in South Korea's late industrialization, although multiple equilibria arise in our model due to path-dependence as in their paper. Unlike Buera et al. (2021), the complementarity in our model is generated by local spillovers of the adoption. We also touch on the international trade literature that studies the evolution of comparative advantage (e.g., Hanson et al., 2015; Levchenko and Zhang, 2016; Arkolakis et al., 2019; Schetter, 2019; Atkin et al., 2021; Cai et al., 2022; Pellegrina and Sotelo, 2021) by showing that technology adoption shaped South Korea's comparative advantage.

Third, this paper relates to the literature on South Korea's growth miracle (e.g., Westphal, 1990; Young, 1995; Lee, 1996; Ventura, 1997; Connolly and Yi, 2015; Itskhoki and Moll, 2019). We contribute to this literature by studying the role of technology adoption in South Korea's industrialization. The most closely related papers are Lane (2022) and Choi and Levchenko (2023). Unlike these papers whose empirical analysis studies sector- or firm-level long-run effects of subsidies provided by South Korea's industrial policy, our empirical analysis focuses on firm-level effects of technology adoption rather than the policy itself, and subsidies from the policy are a source of endogeneity concern in our analysis. Also, our quantitative exercises focus on one particular channel of the policy through technology adoption and show that the big push can be one potential explanation for such long-run effects of the policy documented in Lane (2022) and Choi and Levchenko (2023). In the words of Rodrik (1995), this paper studies how South Korea "got its interventions right" by promoting technology adoption. Another closely related paper is Choi and Shim (2022) who study the role of technology adoption and innovation on South Korea's catching-up growth focusing on the post-1980s, whereas this paper studies the role of technology adoption on industrialization and the sectoral industrial policy in the 1970s.

Structure The rest of this paper is organized as follows. Section 2 describes the data and the historical background of South Korea's late industrialization. Section 3 presents three empirical findings on firm-level effects of technology adoption. Section 4 presents a simple model consistent with these empirical findings, and analytically characterizes the potential multiple equilibria and the possibility of the big push. Section 5 describes the full quantitative model and the calibration procedure of the model. Section 6 presents quantitative results. Section 7 concludes.

2 Data and Historical Background

2.1 Data

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

Technology adoption We hand-collected and digitized firm-level data on technology adoption from official documents related to domestic firms' technology contracts with foreign firms from the

National Archives of Korea and from the Korea Industrial Technology Association (1988).[2] These documents have information about names of domestic and foreign contractors and calendar years in which contracts were made from 1962 to 1988. The dataset includes 1,698 contracts made by 628 unique firms. Of these, 1,361 contracts and 457 firms were in heavy manufacturing sectors. About 74% of technology adoption contracts provided the know-how.[3]

Balance sheet and geographic information 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 but employment is only available after 1972. Using the addresses of plants, we map firms' adoption activities to their location of production. Firm balance sheet information is representative at the national level. On average, the dataset covers 67% of sectoral gross output from the input-output (IO) tables obtained from the Bank of Korea.

Subsidy One of the main policy instruments was directed foreign credit (Jones and Sakong, 1980; Amsden, 1989; Rodrik, 1995). We obtain firm-level credit data from Choi and Levchenko (2023). The government selectively granted access to foreign credit to targeted firms and provided guarantees to this credit once it granted the access.[4] Because of the government guarantee, targeted firms could borrow at a much lower interest rate than loans from domestic sources. A large portion of this credit was allocated to subsidize heavy manufacturing firms' purchases of expensive capital equipment related to newly adopted technologies. We use this subsidy as a control or to test identifying assumptions of our empirical analysis.

Sectoral and regional data We obtain South Korea's import tariffs from Luedde-Neurath (1986). We obtain IO tables from the Bank of Korea. The regional population data comes from the Population and Housing Census.

2.2 Late Industrialization in South Korea

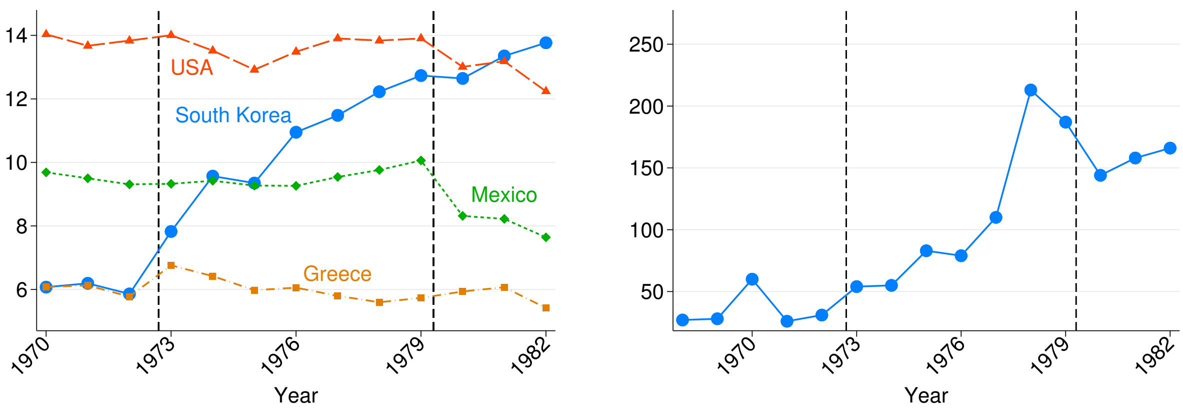
In late 1972, the Korean government launched the HCI Drive to modernize and promote heavy manufacturing sectors, including chemicals, electronics, machinery, steel, non-ferrous metal, and transport

[2] Any domestic firms' transactions with foreign firms, including technology adoption contracts, were strictly regulated under the Foreign Capital Inducement Act, first enacted in 1962. 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 in the economic policy-making process in South Korea during the sample period. Beginning in 1961 and continuing until the mid-1980s, the EPB met every month and discussed new technology contracts. The National Archives of Korea collected and preserved the documents the EPB examined in its monthly meetings.

[3] 21.2% granted licenses, and 4% permitted the use of trademarks. For example, Figure A1 is one page of the contract document between Kolon (South Korean) and Mitsui Toatsu (Japanese), both of which are chemical manufacturers. The contract shows that Mitsui had to provide blueprints, send skilled engineers to train South Korean workers, and provide training service by inviting South Korean engineers to its plants in Japan.

[4] To assert control over the balance of payment, the Korean government enacted the 1962 Foreign Capital Inducement Act and restricted domestic firms' direct financial transactions with foreign firms. See Choi and Levchenko (2023) for detail.

Figure 1. Late Industrialization and Technology Adoption in South Korea



A. Heavy mfg. share of GDP (%)

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

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 STAN Structural Analysis Database and the OECD National Accounts Statistics database.

equipment. The timing of the policy and selection of the targeted sectors were driven by a political shock rather than economic conditions due to the Nixon Doctrine (1969) and the military tension with North Korea (Lane, 2022).[5] The heavy manufacturing sectors, related to the arms industry, were targeted to modernize South Korea's military forces and achieve self-reliant defense. The HCI Drive was temporary because it ended in 1979 after President Park was assassinated.

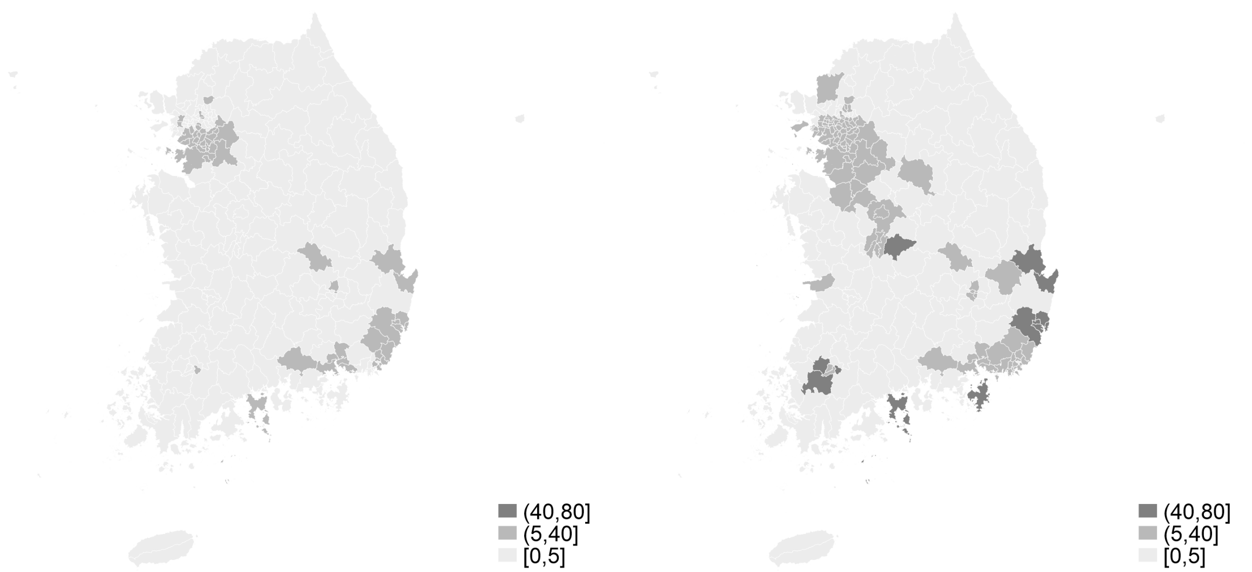
When promoting the heavy manufacturing sectors, the government heavily subsidized the adoption of foreign industrial technology. The government considered South Korea's underdeveloped technology in heavy manufacturing sectors as one of the national threats and given its large technology gap with the world frontier, the government deemed technology adoption to be the most effective way to catch up with the frontier (Ministry of Science and Technology, 1972).[6] The adoption was the main means of technology transfers from foreign developed economies to South Korea.[7]

[5]After the Vietnam War, in the Nixon Doctrine (1969), President Nixon demanded more responsibility from its East Asian allies for their self-defense instead of relying on the US military. The doctrine posed a threat to the national defense of South Korea because of rising military tension with North Korea and its heavy reliance on the US military.

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

[7]Another commonly used means of technology transfer in developing countries is the foreign direct investment

Figure 2. Geographic Concentration of Technology Adoption Activity



Local heavy mfg. shares of adopters (%)

A. 1972 B. 1979

Notes. The figure illustrates each region's heavy manufacturing adopter shares in 1972 and 1979. The cutoffs of 40% and 70% which correspond to the 99th percentile of a distribution of the 1972 shares and the maximum of the 1979 shares.

While at the beginning of the period of our analysis, South Korea's GDP share of the heavy manufacturing sector was only 6%, it achieved a remarkable takeoff during the sample period, surpassing Mexico by the mid-1970s and the US by 1982 (Panel A of Figure 1).[8] Our data shows that this industrialization toward heavy manufacturing sectors is accompanied by inflows of new foreign technologies, with the yearly number of contracts being quadrupled during the same period. (Panel B). Consistent with the narrative of the policy, this sudden increase in amounts of the adoption coincided with the government policy from 1973 to 1979. Even after the policy ended, the economy continued to specialize in the heavy manufacturing sectors.

The narrative of the one-time policy and the rapid pattern of industrialization have led to conjectures on the big push behind South Korea's economic development. Later, we show that the local complementarity and the local spillovers can rationalize this possibility of the big push. In fact, these local effects are consistent with a spatially uneven rise in the adoption activities that are concentrated in the Northwestern and Southeastern regions (Figure 2).

(FDI). In South Korea, however, FDI did not play a big role because of the government regulation on FDI (Kim, 1997, p.42-43).

[8] Consistent with the GDP shares, employment and export shares of the heavy manufacturing sectors also increased from 4 to 8% and 13.7 to 35% between 1972 and 1982.

3 Empirical Evidence on Firm-Level Effects of Technology Adoption

In this section, we provide three empirical findings on firm-level effects of technology adoption in the late industrializing economy: direct effects on adopters, local spillovers, and complementarity in firms' adoption decisions. In Appendix B.1, we provide a detailed example of how technology adoption benefited Korean firms through these three findings.[9]

3.1 Direct Effects on Adopters

Winners vs. losers research design When estimating the direct effects on adopters, one of the key econometric challenges is that firms make adoption decisions endogenously. Unobservable systematic differences between adopters and non-adopters may result in a spurious correlation between adoption status and adopters' performance, leading to the selection bias problem. To overcome this challenge, we implement winners vs. losers research design, drawing on Greenstone et al. (2010), that generates quasi-experimental variation in both adoption status and timing. Because both winners and losers attempted to adopt technology, by comparing losers and matched winners, we can control for underlying unobservables that made these firms self-select into the adoption.

We define winners (the treated) as firms that successfully adopted technology from foreign firms. We define losers (the comparison) as firms 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 foreign firms' bankruptcy, changes in management team of foreign firms, or foreign firms' sudden requests to change contractual clauses after making a deal. We exclude cancellations by domestic firms, such as domestic firms' sudden decreases in cash flow. 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 cancellations by reading these documents from the archive.

Among these losers, there are two types of losers, delayed-adopters and never-adopters. The delayed-adopters are firms that eventually adopted foreign technology but the timing of the adoption got delayed due to the cancellation. The never-adopters are firms that never adopted technology after the cancellation. Therefore, the cancellations generate exogenous variation in adoption timing for some losers or status for others.

We match each loser with at most three winners that made a contract in the same year in which losers made a contract that got eventually canceled. The matching proceeds in two steps. First, we exactly match on region-sectors to absorb shocks common within region-sectors, such as market size

[9] We provide an example of POSCO, the first integrated still mill in South Korea and now one of the top five steel producers in the world. It started operating after adopting foreign technology. Then, labor mobility of its engineers to local smaller-sized firms diffused knowledge acquired from the adoption (Enos and Park, 1988, p.210-211). Later, the knowledge diffusion to local firms facilitated POSCO to adopt more advanced technologies because the availability of cheaper domestic capital inputs produced by local firms lowered the setup costs of the new adoption (POSCO, 2018, p.138-141).

or local labor market conditions. Second, within region-sectors, we choose winners that were most similar to a loser in terms of firm size or growth measured by log assets, log fixed assets, and one-year growth in assets and fixed assets, where the similarity is measured by the Mahalanobis distance. We match losers and winners with replacements, so we can match one winner to multiple losers. If more than three winners are available for a match in a given region-sector, we keep three winners that are most similar to losers in terms of observables. With fewer than three available winners, we keep all winners. The matching procedure gives us 35 matches among 91 unique firms.

Using the matched winners and losers, we estimate the following event study specification:

$$y_{imt} = \sum_{\tau=-4}^{7} \beta_\tau (D_{mt}^\tau \times \mathbb{1}[\text{Winner}_{it}]) + \delta_{im} + \delta_{mt} + \epsilon_{imt}, \tag{3.1}$$

where i denotes firm, m match, and t year. y_{imt} is a firm outcome. D_{mt}^τ are event-study dummies defined as $D_{mt}^\tau := \mathbb{1}[t - \tau = t(m)]$, where $t(m)$ is event year of match m. $\mathbb{1}[\text{Winner}_{it}]$ is a dummy variable of winners. We normalize β_{-1} to one. δ_{im} and δ_{mt} are match-firm and match-year fixed effects. ϵ_{imt} is an error term. Matching with replacement introduces mechanical correlation across residuals, because of the possible appearance of the same firm. Thus, we cluster standard errors at the firm level.

One issue with estimating Equation (3.1) is staggered rollout design that leverages comparison between already-treated adopters and delayed-losers, which induces a bias under the presence of heterogeneous treatment effects across cohorts (e.g. Goodman-Bacon, 2021; Sun and Abraham, 2021; Borusyak et al., 2023). To deal with this issue, we adopt a stacked-by-event design (Cengiz et al., 2019; Deshpande and Li, 2019). We construct our estimation dataset based on rolling control groups as follows. Within each match, we drop matches when delayed-losers adopt technology in later periods. By doing so, we limit attention to comparisons between treated winners and not-yet-treated or never-treated losers, circumventing the forbidden comparison problem.

Identifying assumption Our identifying assumption is that losers form valid counterfactuals for winners. We require that losers and winners should be ex-ante similar in terms of both observables and unobservables prior to an event conditional on matched controls and fixed effects, and cancellations should be uncorrelated with domestic firms' unobservables. Raw plots of the data support this assumption (Online Appendix Figure B1). Average log sales of winners had similar trends with losers but increased only after cancellations, whereas the average of losers evolved similarly to their pre-trends. Also, despite the small number, the distribution of cancellations by sectors closely resembles that of total contracts, which supports that cancellations were random events.

To test this identifying assumption, we conduct three exercises. First, we assess covariate balance by comparing levels of outcomes between two groups before the cancellations and find that both groups are well-balanced.[10] Using the US patent data obtained from US Patent and Trademark Office (USPTO), we compare patenting activities of two groups of foreign firms that made contracts with winners and losers. We interpret patent activities as indicators of good foreign firms. We find that various measures of patent activities of two foreign firm groups are similar, which rules out matching between losers and bad foreign firms that were more likely to cancel contracts (Appendix Table B1). Second, we regress pre-event observables on a dummy of losers. We find that these observables do not predict cancellations regardless of whether they are controlled individually or jointly (Appendix Table B2). Third, and most importantly, to check the parallel trend assumptions, we inspect pre-trends before the cancellations.

Comparison with the full-sample TWFE estimator To assess the implications of correcting for the endogeneity issue, we compare the baseline estimates to those from the following two-way fixed effect (TWFE) event study specifications using the full-sample:

$$y_{it} = \sum_{\tau=-4}^{7} \beta_\tau (D_{it}^\tau \times \mathbb{1}[\text{Adopt}_{it}]) + \delta_i + \delta_{njt} + \epsilon_{it}. \tag{3.2}$$

$\mathbb{1}[\text{Adopt}_{it}]$ is a first-time adoption dummy. We control for time-varying region-sector fixed effects δ_{njt}, so variation comes from differences between adopters and non-adopters within region-sectors.

Baseline results We consider two standard measures for firm performance, log sales and revenue-based TFP, as outcomes. Our revenue-based TFP, TFP^{rr}, is obtained as residuals after estimating the production function using the control function approach (Olley and Pakes, 1996; Ackerberg et al., 2015), where investment is used as a proxy variable.[11] We accommodate the possibility that the adoption may affect the underlying TFP process by adapting the estimation procedure of De Loecker (2013).

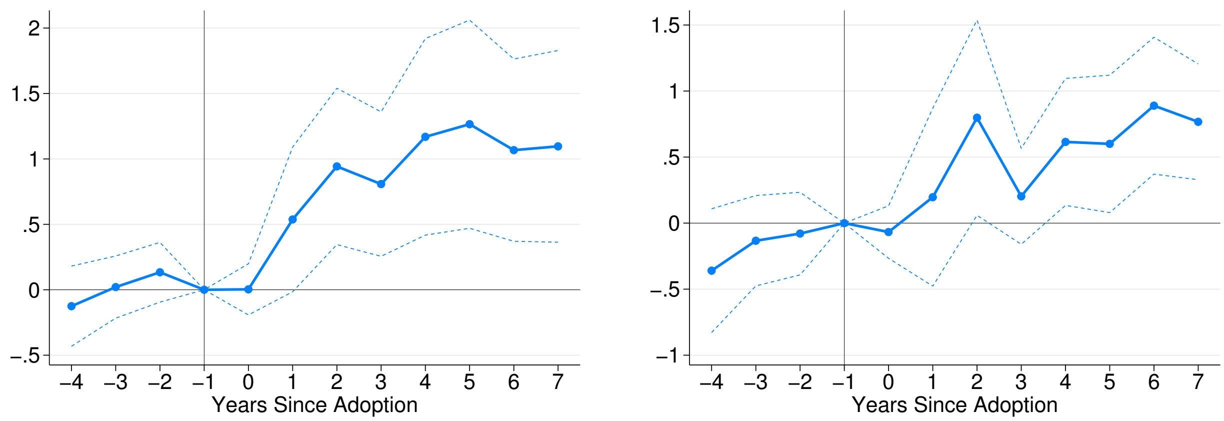
Table 1 and Figure 3 report the estimated coefficients in Equation (3.1). There were no pre-trends and winners' sales and TFP^{rr} begin to increase only after the adoption. 4 years after the adoption, winners' sales and TFP^{rr} increased by 120% and 68%, respectively, and these effects were persistent. The magnitude of our estimates is also consistent with the recent work by Giorcelli and Li (2021) who study the effects of technology transfers from the Soviet Union on Chinese steel plants during China's early industrial development.[12]

[10] Both winners and losers were larger than the average of all heavy manufacturing firms. For example, the average log sales of all heavy manufacturing firms were 15.54, but the averages of winners of losers were 17.80 and 18.46, respectively (column 2 of Table B1). Therefore, non-adopters may not represent a valid counterfactual for adopters, and naive comparisons between them may lead to biased estimates.

[11] We highlight that TFP^{rr} differs from TFPR, as emphasized by Blackwood et al. (2021). For example, under monopolistic competition without distortions, TFPR which is calculated based on cost-shares is equalized across firms whereas TFP^{rr} is proportional to productivity. We obtain investment as differences between fixed assets of two consecutive periods after imposing a depreciation rate of 0.06.

[12] They find that the technology transfers increased the TFPQ of Chinese steel plants by 25%, 6 years after the adoption. Under monopolistic competition, TFPQ $\propto \frac{\ln \text{Sale}}{\sigma - 1}$ holds, where σ is the elasticity of substitution. With

Figure 3. Direct Effects on Adopters: Winners vs. Losers Design



Notes. This figure illustrates the estimated β_τ in Equation (3.1) based on winners vs. losers research design. In Panels A and B, the dependent variables are log sales and TFP$^{\mathrm{rr}}$. All specifications control for match-year and match-firm fixed effects. The plotted coefficients are in columns 1-2 of Table 1. The dotted lines are the 90 percent confidence intervals based on standard errors clustered at the firm-level.

The TWFE estimator also shows that adopters' sales increased after the adoption, but it exhibits increasing pre-trends at -4 and its magnitude was 75% smaller than the baseline. Despite this increase in sales, however, the TWFE estimates for TFP$^{\mathrm{rr}}$ remain flat after the adoption. The discrepancies between the baseline and the TWFE estimators arise because the baseline corrects the endogeneity problem. In fact, we provide evidence that subsidies are a source of the endogeneity that leads to the discrepancies. We consider a dummy of receiving credit (subsidy) from the government as an outcome. The TWFE coefficients are positive and statistically significant under the 1%, implying that adopters were more likely to receive credit (column 6). However, the baseline estimators do not show such a pattern (column 3). These results are important for two reasons. First, they indicate that our winners vs. research design corrects the endogeneity problem due to subsidies. Second, we can interpret the increases in sales and TFP$^{\mathrm{rr}}$ from the baseline as the *pure* effects of the adoption rather than the *joint* effects of the adoption and subsidies. Suppose the government took subsidies back from losers after cancellations. In such a case, we would expect that winners receive more credit and the estimated coefficients from the subsidy dummy to become statistically significantly positive after the adoption as in the TWFE specification. However, we do not find such a pattern.

Robustness We rule out alternative hypotheses against our findings. One possibility is that if foreign firms sold technology tailored for their inputs, winners could have increased sales or TFP$^{\mathrm{rr}}$

commonly calibrated values of 3–4 for σ, our estimates for sales imply that TFPQ increased by 35–57%.

Table 1: Direct Effects on Adopters

Research Design	Winners vs. Losers			Full-sample TWFE		
Dep. Var.	Sale	TFP^{rr}	Subsidy	Sale	TFP^{rr}	Subsidy
	(1)	(2)	(3)	(4)	(5)	(6)
4 years before	–0.13	–0.36	0.08	–0.17***	0.02	0.01
	(0.18)	(0.28)	(0.08)	(0.05)	(0.07)	(0.01)
3 years before	0.02	–0.13	–0.01	–0.05	–0.01	0.00
	(0.14)	(0.21)	(0.08)	(0.05)	(0.06)	(0.01)
2 years before	0.13	–0.08	0.04	–0.03	0.02	0.02
	(0.14)	(0.19)	(0.08)	(0.05)	(0.06)	(0.01)
1 year before						
Year of event	0.00	–0.07	0.04	0.02	–0.03	0.02**
	(0.12)	(0.12)	(0.09)	(0.06)	(0.06)	(0.01)
1 year after	0.54	0.20	0.01	0.13	–0.05	0.04**
	(0.33)	(0.40)	(0.10)	(0.09)	(0.08)	(0.02)
2 years after	0.94**	0.80*	–0.04	0.23**	–0.02	0.03**
	(0.36)	(0.44)	(0.11)	(0.10)	(0.08)	(0.01)
3 years after	0.81**	0.20	0.13	0.18	–0.02	0.02
	(0.33)	(0.22)	(0.12)	(0.13)	(0.12)	(0.01)
4 years after	1.17**	0.62**	–0.04	0.26**	0.00	0.02
	(0.45)	(0.29)	(0.09)	(0.13)	(0.13)	(0.02)
5 years after	1.27***	0.60*	–0.03	0.30**	0.03	–0.02
	(0.48)	(0.31)	(0.09)	(0.14)	(0.17)	(0.02)
6 years after	1.07**	0.89***	–0.04	0.30**	0.08	0.02
	(0.42)	(0.31)	(0.10)	(0.14)	(0.11)	(0.02)
7 years after	1.10**	0.77***	–0.04	0.28*	0.07	–0.00
	(0.44)	(0.26)	(0.10)	(0.15)	(0.14)	(0.03)
Match×Firm FE	✓	✓	✓			
Match×Year FE	✓	✓	✓			
Firm FE				✓	✓	✓
Region×Sector×Year FE				✓	✓	✓
# Cl (Firm or Region)	91	80	91	77	55	77
N	644	484	644	24131	12657	24131

Notes. Standard errors in parenthesis are clustered at the firm level or the region level in columns 1-3 or 4-6, respectively. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Columns 1-3 and 4-6 report the estimated event study coefficients β_τ from winners vs. research design (Equation (3.1)) and full-sample TWFE (Equaton (3.2)), respectively. β_{-1} is normalized to zero. The dependent variables are log sales, TFP^{rr}, or a dummy of receiving subsidy. Columns 1-3 control for match-firm and match-year fixed effects, and columns 4-6 control for firm and region-sector-year fixed effects.

due to technology-driven demand shocks by selling more inputs to these foreign technology sellers rather than physical productivity gains. However, aggregate trade patterns suggest this alternative story is unlikely, because import or export shares from Japan and the US, the two largest sources of foreign technology, decreased during the sample period (Appendix Figure B2). We also can rule out demand shocks due to the government military spending as these matched winners and losers were not military firms. Finally, our estimates are driven by positive gains of adopters rather than negative effects of cancellations to losers because the raw plot of losers' sales exhibits similar trends before and after cancellations (Appendix Figure B1).[13]

Appendix Table B3 reports additional robustness exercises. We consider alternative outcomes. The adoption had positive impacts on labor productivity defined as sales divided by employment, and marginally increased the probability of exporting. Although marginally significant, the export results suggest that the adopters became productive enough to compete in global markets. We also consider alternative numbers of winners matched for each loser and two-way clustering at the levels of match and firm.

3.2 Local Spillover

We define shares of adopters in region-sector nj in year t as

$$\text{Share}_{(-i)nj,t-h} = \frac{N^{\text{T}}_{(-i)nj,t-h}}{N_{(-i)njt}}. \tag{3.3}$$

$N_{(-i)njt}$ is the total number of firms in region-sector nj in t excluding firm i. $N^{\text{T}}_{(-i)nj,t-h}$ is the number of firms operating in t that were in contract with any foreign firms in year $t-h$ excluding firm i. We construct $N^{\text{T}}_{(-i)nj,t-h}$ using information on contract years. We exclude i to rule out the mechanical correlation. Lagging by h years allows for the possibility that it took some time for local diffusion of new knowledge from adopted technologies. We set the value of h to 2 as a baseline.

We consider a following long difference specification

$$\triangle y_{it} = \beta \triangle \text{Share}_{(-i)nj,t-2} + y_{it_0} + \mathbf{X}'_{injt}\gamma + \delta_n + \delta_j + \sum_g D_g \delta_{jg} + \triangle \epsilon_{it}, \tag{3.4}$$

where $\triangle$ is a time difference operator and i denotes firm, g business group, j sector, n region, and t year. Dependent variables are changes in log sales, TFP^{rr}, or log fixed asset. Firm time-invariant factors are differenced out. δ_n and δ_j are region and sector fixed effects. D_g is a dummy of whether a firm is affiliated with a business group g that may own multiple firms across region-sectors. For firms affiliated with group g ($D_g = 1$), we control for group-sector fixed effects δ_{jg} which absorb common factors within group-sectors, such as within group-sector spillovers. In all specifications, we control

[13] For example, increased local competition can exert negative effects on losers. However, manufacturing sectors are highly tradable and a spatial unit of analysis is very granular. South Korea is about the size of Indiana in the US and we match firms within 135 sub-divided regions.

for the initial level of a dependent variable y_{injt_0} because of a well-documented fact that larger firms grow less fast. Some specifications include additional observables $\mathbf{X}_{injt}$. We two-way cluster standard errors at the levels of regions and business groups. Individual firms not affiliated with any groups are subject to their own group-level clusters.

Note that the adopter shares can affect firm performance through the spillovers but also through their influences on firms' adoption decisions. To restrict our attention to the former channel, the estimation sample only includes firms that never adopted technology. The estimates based on the never-adopter sample reflect only the spillovers because, by definition, they had not benefited from any direct effects of the adoption.

To use the data more efficiently, we use overlapping 7-year long-differences between 1972 and 1979 or 1973 and 1980, which covers the policy period. To deal with potential sorting, we estimate Equation (3.4) only for continuing firms, but firm entry and exit affect $\text{Share}_{(-i)nj,t-2}$.

IV strategy OLS estimates of Equation (3.4) may suffer from endogeneity due to correlations between the error term and region-sector level adopter shares. For example, unobserved region-sector level productivity or subsidy shocks that affect both firm growth and other local firms' adoption decisions can lead to such correlations. Moreover, the restriction to the never-adopter sample can cause the selection bias. However, the direction of the bias of OLS is a priori unclear. On the one hand, positive productivity shocks lead to an upward bias. On the other hand, if adoption subsidies were systematically provided to less productive but more politically-connected firms, subsidy shocks can lead to a downward bias. Also, as our data do not cover the universe of firms, measurement errors in local shares can be another source of the downward bias.

We use geographical structure of business groups with multiple firms across regions to construct an IV that isolates variation in local adopter shares arguably exogenous to firm-level unobserved factors, following the IV strategy developed by Moretti (2021).[14] Let $N^{\mathrm{T},\geq 25\mathrm{km}}_{g(-n)jt}$ be the total number of sector j adopters affiliated with business group g in year t excluding firms that are located in region n or within 25km radius circles around region n. Also, we define

$$Z^{\geq 25\mathrm{km}}_{inj,t-h} = \sum_{\tilde{g}\neq g(i)} D_{\tilde{g}njt_0} \times \frac{N^{\mathrm{T},\geq 25\mathrm{km}}_{\tilde{g}(-n)j,t-h}}{\tilde{N}^{\mathrm{p}}_{(-i)njt}},$$

where $D_{\tilde{g}njt_0}$ is a dummy variable of whether business group $\tilde{g}$ has at least one firm in region-sector nj in the initial year, and $\tilde{N}^{\mathrm{p}}_{(-i)njt}$ is the predicted number of firms in region-sector nj: $\tilde{N}^{\mathrm{p}}_{(-i)njt} \equiv g_{(-n)jt} \times N_{(-i)njt_0}$, where $g_{(-n)jt}$ is national-level growth of the number of sector j firms excluding firms in region n and $N_{(-i)njt_0}$ is the number of firms in region-sector nj excluding firm i

[14] Moretti (2021) uses spatial network of firms with multiple locations to construct exogenous shifters for local inventor cluster size. Specifically, he exploits variation in the number of inventors outside of a local region hired by firms that have a presence in a local region.

in the initial year. We construct the IV as

$$\mathrm{IV}^{\geq 25\mathrm{km}}_{inj,t-h} = \triangle Z^{\geq 25\mathrm{km}}_{inj,t-h}. \tag{3.5}$$

We exclude firms located within 25km because of potential spatial correlation with neighboring firms, for example, through input-output linkages or spatially correlated unobservables. Our IV varies at the level of business groups within region-sectors. Individual firms not affiliated with any groups share the same values of the IV but firms affiliated with groups have different values from these individual firms because the summation excludes their own groups.

The explicit identifying assumption is that for firm i, variation in the number of adopters outside i's region affiliated with business groups that owned a firm located in i's region is orthogonal to i's unobservable. To illustrate the intuition of the IV, consider the Samsung Group as an example, which owned six electronics sector firms. Among these six firms, four of them were located in Suwon and one in Ulsan, the Northwestern and the Southeastern parts of the country (Figure 2), respectively. The idea is that Samsung's group-level adoption decisions outside of Ulsan may increase amounts of adoption in Ulsan through its affiliate in Ulsan, but may not be correlated with productivity or subsidy shocks of other firms in Ulsan not affiliated with Samsung Group.

Baseline results Table 2 reports the estimates. In Panel A, the dependent variable is sales growth. Column 1 reports the OLS estimates, which are marginally significant. Column 2 reports the IV estimates. The IV estimate is 4, implying that one percentage point increase in adopter shares leads to 4% higher sales. The magnitude of the IV estimate is larger than the OLS estimate because the IV corrects for the measurement error and endogeneity problems. The IV is strong with a first-stage coefficient of 0.17 and a Kleibergen-Paap F-statistics (KP-F) of 41. In Panel B, the dependent variable is $\mathrm{TFP}^{\mathrm{rr}}$ growth. Although the statistical significance becomes weaker as the number of samples decreases due to missing employment, we find that one percentage point increase in adopter shares leads to 1.2% higher $\mathrm{TFP}^{\mathrm{rr}}$ growth. Our firm-level local spillover analysis is broadly consistent with the findings from the previous studies (e.g. Greenstone et al., 2010; Giorcelli and Li, 2021). Also, the limited competition effect is consistent with increased foreign demand and large labor supply due to reallocation from the agricultural sector to the manufacturing sectors during the industrialization (Vogel, 1991; Lucas, 2004)

Additional controls Our main findings are robust to additional controls. Because firms outside of nj can affect i's growth through input-output linkages, in column 3, we control for market access measure defined as a weighted sum of other firms' sales weighted by the inverse of distance between firms (Donaldson and Hornbeck, 2016):

$$\triangle \ln \mathrm{MA}_{i(-nj)t} = \triangle \ln \Big(\sum_{m,k,mk \neq nj} \sum_{i' \in \mathcal{F}_{mkt}} \mathrm{Dist}^{-\chi}_{nm} \times \gamma^{j}_{k} \mathrm{Sales}_{i't} \Big), \tag{3.6}$$

Table 2: Local Spillover

	OLS	IV						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. Dep. $\triangle \ln \text{Sale}_{it}$ 1972-1979 or 1973-1980								
$\triangle\text{Share}_{(-i)nj,t-2}$	0.77^{*}	3.99^{***}	3.90^{***}	4.16^{***}	3.81^{***}	3.88^{***}	4.11^{***}	3.88^{***}
	(0.46)	(1.00)	(1.04)	(1.09)	(0.98)	(0.97)	(0.96)	(1.02)
KP-F		67.59	68.85	82.76	66.02	61.07	89.20	104.49
# Cl. (Region)	79	79	79	79	79	79	79	79
# Cl. (Group)	1294	1294	1294	1294	1294	1294	1294	1294
N	1492	1492	1492	1492	1492	1492	1492	1492
Panel B. Dep. $\triangle \ln \text{TFP}^{\text{rr}}_{it}$ 1972-1979 or 1973-1980								
$\triangle\text{Share}_{(-i)nj,t-2}$	-0.23	1.15^{*}	1.00	1.26^{**}	1.16^{**}	1.11^{*}	1.09^{*}	0.95
	(0.29)	(0.58)	(0.60)	(0.60)	(0.58)	(0.57)	(0.60)	(0.68)
KP-F		64.68	69.11	79.10	56.32	62.20	93.41	114.49
# Cl. (Region)	67	67	67	67	67	67	67	67
# Cl. (Group)	742	742	742	742	742	742	742	742
N	824	824	824	824	824	824	824	824
Region FE	✓	✓	✓	✓	✓	✓	✓	✓
Sector FE	✓	✓	✓	✓	✓	✓	✓	✓
Sector-group FEs	✓	✓	✓	✓	✓	✓	✓	✓
Market access			✓					✓
Own region-sector GO				✓				✓
Directed credit					✓			✓
Complex controls						✓		✓
Tariff controls							✓	✓

Notes. Standard errors two-way clustered at the levels of regions and business groups are reported in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. This table reports the OLS and IV estimates of Equation (3.4). The adoption shares and IV are defined in Equations (3.3) and (3.5). In Panels A and B, dependent variables are changes in log sales or TFP^{rr} between 1972 and 1979 or 1973 and 1980. In column 3, we control for the market access defined in Equation (3.6). In column 4, we control for own region-sector gross output defined in Equation (3.7). In column 5, we control for the inverse hyperbolic sine transformation of cumulative credit received between 1973 and 1979. In column 6, we control for an industrial complex dummy. In column 7, we control for interaction terms between port dummies and import and input tariffs. In column 8, we control for all additional controls. All specifications include region, sector, and sector-group fixed effects, and initial levels of dependent variables. KP-F is the Kleibergen-Paap F-statistics.

where $\mathcal{F}_{mkt}$ is the set of firms in region-sector mk operating in year t. We proxy internal trade costs using distance between regions Dist_{nm}, and set χ to 1.1 (Costinot and Rodríguez-Clare, 2014). To mitigate the endogeneity concern, we exclude i's own region-sector.

It is possible that firms in regions with larger adopter shares grew faster because they were co-located with larger-sized firms. To isolate variation of the adopter shares from variation of being co-located with large-sized firms, in column 4, we control for the sum of sales of firms within regions-sectors defined as

$$\triangle \ln \mathrm{GO}_{(-i)njt} = \triangle \ln \Big(\sum_{i' \in \mathcal{F}_{(-i)njt}} \mathrm{Sales}_{i't} \Big), \tag{3.7}$$

where $\mathcal{F}_{(-i)njt}$ is the set of firms in nj in t excluding i.

In column 5, we control for the inverse hyperbolic sine transformation of the sum of directed credit received between 1972 and 1979 or 1973 and 1980. During the policy period, the government constructed industrial complexes in the Southeastern regions and promoted heavy manufacturing firms in these complexes (Choi and Levchenko, 2023). From the 1980 Yearbooks of Industrial Complexes published by the Korea Industrial Complex Corporation, we obtain information on which sector firms can be located in these complexes. Using this information and firms' location of production, we construct a dummy of whether firms were located in these complexes and control for this dummy in column 6.

The government has strongly promoted export-oriented development through trade policy (Connolly and Yi, 2015). Common effects of these trade policies are absorbed by the sector-fixed effects, but the policies can have differential impacts across regions depending on how internal trade costs shield them from foreign competition. Reductions in import tariffs will increase the degree of foreign competition of firms located near ports relatively more than those located inland. In column 7, we control for a port dummy interacted with the changes in import tariffs. Also, because import tariffs can affect firm performance through the costs of imported intermediates, we control for the port dummies interacted with the changes of input tariffs. We construct input tariffs as the weighted average of import tariffs, where the weights are given by the value share of inputs from the 1970 IO table. In column 8, we control for all these additional controls jointly.

Placebo To examine whether our results are driven by a spurious correlation between unobservables and the IV, we conduct the placebo exercise. We re-estimate the regression model using as dependent variables sales growth between 1970 and 1972 or 1971 and 1973. The intuition is that because the IV is an exogenous shifter for adopter shares between 1972 and 1979 or 1973 and 1980, the IV should not affect firm growth before these periods. We find that the coefficients are statistically insignificant, suggesting that the IV or future changes in the adopter shares do not predict the past sales growth (Appendix Table B4).

Functional form Our baseline specifications impose a linear relationship between log sales and the adopter shares. Also, the adopter shares are scale-free and the spillover effects do not vary across firms depending on their size. To explore the linearity, the scale-freeness, and the firm size heterogeneity, we add interaction terms between the changes in the adopter shares and a dummy whether initial levels of adopter shares, region-sectors' number of firms, and firm sales are above the 90th percentile, respectively. We instrument these additional interaction terms with interaction terms between our IV and the corresponding initial dummies. None of the interaction terms are precisely estimated (Appendix Table B5).[15]

Additional robustness checks We consider a battery of robustness checks. We consider alternative outcomes. We find the positive effects of the adopter shares on the probability of exporting and labor productivity, which supports that never-adopters' productivity improved due to local spillovers. We consider not controlling for y_{it_0} and an alternative lag of 3.

Variation of the IV comes from business groups that own multiple firms across regions within sectors. We push this leave-out logic further and re-estimate the regression model with the same IV only for a subsample of firms not affiliated with any business groups that are located only in one region by definition. We also consider subsamples excluding firms in regions in which heavy manufacturing industrial complexes are constructed, a single difference between 1973 and 1980, and the full sample instead of the never-adopter sample.

When constructing the IV, we consider alternative radius circles with distances ranging from 0km to 150km.

3.3 Complementarity in Adoption Decisions

To examine the complementarity, we consider a similar overlapping long-difference regression model between 1972 and 1979 or 1973 and 1980:

$$\triangle \mathbb{1}[\text{New Contract}_{i,t+1}] = \beta \triangle \text{Share}_{(-i)nj,t-2} + \mathbf{X}'_{injt}\boldsymbol{\gamma} + \delta_n + \delta_j + \sum_g D_g \delta_{jg} + \triangle \epsilon_{it}, \tag{3.8}$$

where a dependent variable is a dummy variable of making new adoption contracts in a given year, $\mathbb{1}[\text{New Contract}_{i,t+1}]$. $\mathbb{1}[\text{New Contract}_{it}]$ differs from $\mathbb{1}[\text{Adopt}_{it}]$ that is used to construct the adopter shares. For example, if a firm has not adopted any foreign technologies previously and makes a new contract in time t, both $\mathbb{1}[\text{New Contract}_{it}]$ and $\mathbb{1}[\text{Adopt}_{it}]$ become 1 in t. If a firm made a contract in $t-3$ but did not make a new contract in t, then only $\mathbb{1}[\text{Adopt}_{it}]$ takes a value of 1.

Using the full sample including both never- and ever-adopters, we estimate Equation (3.8) with the same IV and the set of fixed effects to those of the spillover regression. The identifying assumption is the same as that of the spillover IV regression. The positive β implies that firms were more likely to adopt new foreign technology if more local firms adopted technology. Standard errors are two-way

[15] Kline and Moretti (2014) also test nonlinearities in the agglomeration effects of manufacturing density using interaction terms between changes in the density and the initial density. They find the log-linear relationship between log employment and the density, implying that the agglomeration effects are log-linear or concave in the density. Unlike their paper, our data support the linear relationship between log sales and adopter shares, possibly due to the fact that the spillover effects of technology adoption may have a different functional form from the density.

Table 3: Complementarity in Technology Adoption Decisions

Dep.	$\triangle\mathbb{1}[\text{New Contract}_{i,t+1}]$							
	OLS	IV						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\triangle\text{Share}_{(-i)nj,t-2}$	-0.06	0.70***	0.69***	0.73**	0.73***	0.69***	0.69***	0.71**
	(0.11)	(0.26)	(0.26)	(0.28)	(0.27)	(0.25)	(0.26)	(0.28)
KP-F		35.09	33.69	39.79	34.97	31.71	45.79	47.42
# Cl. (Region)	86	86	86	86	86	86	86	86
# Cl. (Group)	1548	1548	1548	1548	1548	1548	1548	1548
N	1977	1977	1977	1977	1977	1977	1977	1977
Region FE	✓	✓	✓	✓	✓	✓	✓	✓
Sector FE	✓	✓	✓	✓	✓	✓	✓	✓
Sector-group FEs	✓	✓	✓	✓	✓	✓	✓	✓
Market access			✓					✓
Own region-sector GO				✓				✓
Directed credit					✓			✓
Complex controls						✓		✓
Tariff controls							✓	✓

Notes. Standard errors are two-way clustered at the levels of regions and business groups are reported in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. This table reports the OLS and IV estimates of Equation (3.8). Adopter shares and IV are defined in Equations (3.3) and (3.5), respectively. The dependent variable is a dummy of making a new adoption contract in $t+1$ between 1972 and 1979 or 1973 and 1980. In column 3, we control for the market access defined in Equation (3.6). In column 4, we control for own region-sector gross output defined in Equation (3.7). In column 5, we control for the inverse hyperbolic sine transformation of cumulative credit received between 1973 and 1979. In column 6, we control for an industrial complex dummy. In column 7, we control for interaction terms between port dummies and import and input tariffs. In column 8, we control for all additional controls. All specifications include region, sector, and sector-group fixed effects, and initial levels of dependent variables. KP-F is the Kleibergen-Paap F-statistics.

clustered at the levels of regions and business groups.

Table 3 reports the results. Columns 1 and 2 report the OLS and IV estimates. Once the endogeneity is corrected, the estimate becomes positive and statistically significant. The IV estimate implies that a one percentage point increase in the adopter shares leads to a 0.7 percentage point increase in the probability of making a new contract. The 0.7 percentage point is about 12 percent of the average probability of new contracts in 1979 and 1980 (6 percentage points). In columns 3-8, we sequentially include the same set of additional controls in the spillover regression. Across specifications, the estimates are positive and statistically significant, and their magnitude remains stable.

IV validity and robustness checks We conduct the placebo test with changes in the new contract dummy before 1973. We do not find a statistically significant relationship (Columns 4-6 of Appendix Table B4). We also conduct a set of robustness checks similar to those of the spillover regression (Appendix Table B7).

3.4 Summary and Discussion

To summarize, the adoption of foreign technologies increased sales and TFP^{rr} of adopters but also those of non-adopters through the local spillovers. Moreover, the local complementarity suggests a vicious circle of technology adoption. If a one-time big push triggered this vicious circle, firms would have kept adopting foreign technologies through the complementarity even after no more subsidies were provided. This big push story and the local effects of the adoption are consistent with the rapid industrialization process and the geographical concentration of adoption activities illustrated in Figures 1 and 2. We formally show this in the next section.

4 A Simple Model of Technology Adoption and Multiple Equilibria

We present a simple dynamic model with firms' technology adoption decisions. The model generates features that are consistent with the three findings. We analytically show that these three findings embodied in the model can induce an economy to have multiple steady states. When multiple steady states exist, a big push that temporarily provides subsidies for the adoption can have large impact by shifting the economy from one steady state to the other. Later, we extend this simple model and quantitatively explore the effects of the big push.

Environment We consider a closed economy with one sector and one region. Time is discrete and indexed by $t \in \{1, 2, \dots\}$. There is a fixed mass of monopolistically competitive firms indexed by i, whose mass M is normalized to one. Each firm produces a unique variety ω. A final goods producer aggregates these varieties using the CES aggregator and produces final consumption goods. Labor is the only factor of production. Households inelastically supply labor.

Firm Each firm faces demand curves $q_{it} = p_{it}^{-\sigma} P_t^{\sigma} Q_t$ where q_{it} is firms' quantity demanded, p_{it} is price charged by firms, $Q_t = (\int (q_{it}(\omega))^{\frac{\sigma-1}{\sigma}} \mathrm{d}\omega)^{\frac{\sigma}{\sigma-1}}$ is the aggregate quantity, and $P_t = (\int_{\omega} (p_{it}(\omega))^{1-\sigma} \mathrm{d}\omega)^{\frac{1}{1-\sigma}}$ is the ideal price index. $\sigma > 1$ is the elasticity of substitution across varieties. Firms charge constant markups $\mu = \sigma/(\sigma - 1)$ over its unit costs $p_{it} = w_t/z_{it}$ where z_{it} is firm productivity.

Firms are heterogeneous in productivity. Firms' decisions to adopt modern technology and spillovers endogenously determine their productivity in the equilibrium. Firm productivity is composed of three terms:

$$z_{it} \equiv z_{it}(T_{it}, \lambda_{t-1}^{\mathrm{T}}) = \eta^{T_{it}} \times f(\lambda_{t-1}^{\mathrm{T}}) \times \phi_{it},$$

where T_{it} is a binary variable that equals one if a firm adopts modern technology. The first component $\eta > 1$ governs direct productivity gains from the adoption. The second component $f(\lambda_{t-1}^{\mathrm{T}})$ common across firms is adoption spillovers that increase in the share of adopters in the previous period

$\lambda_{t-1}^{\mathrm{T}}$. In Appendix Section C.4, we provide two sets of microfoundations that generate the spillover through labor mobility and knowledge transfers.[16] The third component ϕ_{it} is exogenous productivity, independently and identically distributed across firms and periods. The first and the second terms are motivated by the first and the second empirical evidence.

We consider the following functional form for the spillover:

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

where δ is a parameter that governs the strength of the spillover. The functional form is supported by the empirical evidence from Section 3.2. Following Kline and Moretti (2014) and Allen and Donaldson (2020), we allow the spillover to operate with a one-period lag rather than contemporaneously, consistent with the specification of the spillover regression model (Equation (3.3)). We view that allowing for a lag is more realistic given that it takes some time for knowledge to be locally diffused. Also, as discussed by Adserà and Ray (1998), allowing for a lag yields an economy to have a deterministic outcome each period, which rules out implausible situations in which an economy wildly swings between different outcomes each period.

The adoption incurs fixed costs F^{T} in units of final goods (Buera et al., 2021). Firms adopt technology when additional operating profits from the adoption are larger than the fixed costs:

$$\pi_{it} = \max_{T_{it}\in\{0,1\}} \left\{ \frac{1}{\sigma} \frac{\mu w_t}{z_{it}(T_{it}, \lambda_{t-1}^{\mathrm{T}})} P_t^{\sigma} Q_t - T_{it} P_t F^{\mathrm{T}} \right\},$$

where π_{it} is i's final profits. When making adoption decisions, firms internalize the direct productivity gains η but not the spillovers $f(\lambda_{t-1}^{\mathrm{T}})$ and take $\lambda_{t-1}^{\mathrm{T}}$ as given in t. Due to these externalities, social returns to the adoption are larger than private returns and adoption rates become lower than the socially optimum level. With heterogeneous productivity, firms' adoption decisions are characterized by the following cutoff productivity:

$$(\bar{\phi}_t^{\mathrm{T}})^{\sigma-1} = \frac{\sigma P_t F^{\mathrm{T}}}{(\eta^{\sigma-1}-1)(\mu w_t)^{1-\sigma} f(\lambda_{t-1}^{\mathrm{T}})^{\sigma-1} P_t^{\sigma} Q_t}. \tag{4.1}$$

Only firms with productivity higher than the cutoff adopt technology. The probability of adoption is $\lambda_t^{\mathrm{T}} = \mathbb{P}[\phi_{it} \geq \bar{\phi}_t]$ equivalent to the adopter shares with the normalized firm mass.

Equilibrium In each period, given $\lambda_{t-1}^{\mathrm{T}}$, firms adopt technology to maximize their profits and goods and factor markets clear (static equilibrium). λ_t^{T} is a state variable that endogenously evolves based on firms' adoption decisions (dynamic equilibrium). Given $\lambda_{t-1}^{\mathrm{T}}$, the equilibrium share $\lambda_t^{\mathrm{T}*}$ is

[16] In the first setup, we consider a setup in which engineers and firms are randomly matched (Acemoglu, 1996) and engineers carry new knowledge learned from adopted technologies when matched with a new firm in the next period. In the second setup, we present a model of Desmet and Rossi-Hansberg (2014) in which own innovation costs become lower with higher adopter shares due to knowledge transfers.

determined in t; and then given $\lambda_t^{\mathrm{T}*}$, $\lambda_{t+1}^{\mathrm{T}*}$ is determined in $t+1$; and so on.

We impose the following assumptions:

Assumption 1. *(i) $\sigma > 2$; and (ii) ϕ_{it} follows the Pareto distribution with the location parameter normalized to be one and the shape parameter θ.*

Under the Pareto productivity distribution, the cutoff can be expressed as

$$\bar{\phi}_t^{\mathrm{T}} = (\lambda_t^{\mathrm{T}})^{-\frac{1}{\theta}}. \tag{4.2}$$

By combining Equations (4.1) and (4.2), we can derive the analytical expression of the equilibrium adopter shares $\lambda_t^{\mathrm{T}*} = \lambda_t^{\mathrm{T}*}(\lambda_{t-1}^{\mathrm{T}}; \eta, \delta)$ in each period conditional on the shares in the previous period $\lambda_{t-1}^{\mathrm{T}}$. The equilibrium shares are determined at

$$\lambda_t^{\mathrm{T}*}(\lambda_{t-1}^{\mathrm{T}}; \eta, \delta) = \min\{\hat{\lambda}_t^{\mathrm{T}}(\lambda_{t-1}^{\mathrm{T}}; \eta, \delta), 1\}, \tag{4.3}$$

where $\hat{\lambda}_t^{\mathrm{T}}(\lambda_{t-1}^{\mathrm{T}}; \eta, \delta)$ is implicitly defined by

$$\hat{\lambda}_t^{\mathrm{T}}(\lambda_{t-1}^{\mathrm{T}}; \eta, \delta) = \left[A(\hat{\lambda}_t^{\mathrm{T}}(\lambda_{t-1}^{\mathrm{T}}; \eta, \delta))^{2-\sigma} \frac{(\eta^{\sigma-1} - 1)}{\sigma F^{\mathrm{T}}} f(\lambda_{t-1}^{\mathrm{T}})\right]^{\frac{\theta}{\sigma-1}},$$

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

The time-invariant steady state adopter shares ($\lambda^{\mathrm{T}*} = \lambda_t^{\mathrm{T}*} = \lambda_{t-1}^{\mathrm{T}*}$) satisfy $\lambda^{\mathrm{T}*} = \lambda^{\mathrm{T}*}(\lambda^{\mathrm{T}*}; \eta, \delta)$.

Equilibrium properties and multiple equilibria Assumption (i) is a sufficient condition to guarantee a unique static equilibrium each period.[17] Given any initial adopter shares λ_{t_0}, because static equilibrium is unique each period, there exists a unique sequence of static equilibrium that forms a unique deterministic dynamic path from λ_{t_0} to a steady state.

There is dynamic complementarity in firms' adoption decisions, that is, $\lambda_t^{\mathrm{T}*}$ increases in $\lambda_{t-1}^{\mathrm{T}}$. Firms are more likely to adopt technology when more other firms adopted technology in the previous period, generating the feature consistent with the third empirical finding. The fact that the fixed adoption costs are in units of final goods is the source of this complementarity. The spillover from the adopter shares in the previous period lowers the fixed adoption costs in the current period, which

[17] When σ is sufficiently low, even when $\delta = 0$, because firms do not internalize P_t, two static equilibria, one with higher adopter shares and the other with lower shares, can arise. Higher shares increase competition but also decrease fixed adoption costs by lowering P_t, which incentivizes more and less adoption, respectively. $A(\lambda_t^{\mathrm{T}})^{2-\sigma} = A(\lambda_t^{\mathrm{T}})^{1-\sigma} \times A(\lambda_t^{\mathrm{T}})$ is related to these two general equilibrium effects that operate in the opposite directions. $A(\lambda_t^{\mathrm{T}})^{1-\sigma}$ captures the former and $A(\lambda_t^{\mathrm{T}})$ the latter. Lower σ makes competition effects weaker and more firms adopt technology with higher shares, which generates static complementarity and, in turn, potential multiple static equilibria. By imposing $\sigma > 2$, we make the competition effects sufficiently strong to rule out such possibility studied by Matsuyama (1995) and Buera et al. (2021). Unlike these papers, in our model, because we impose $\sigma > 2$, multiple long-run steady states arise only because of the spillovers. Also, commonly calibrated parameter values for σ are larger than 2.

further incentivizes more firms to adopt technology in the current period. Also, $\lambda_t^{\mathrm{T}*}$ increases in η and δ, because higher η and δ magnify the direct gains and the strength of the dynamic complementarity, respectively.

Most importantly, we show that multiple steady states can arise due to dynamic complementarity. When these steady states exist, they can be Pareto-ranked based on the steady state share of adopters. The initial share of adopters determines which steady state to be realized in the long run, that is, there is path dependence. Proposition 1 summarizes these results.

Proposition 1. *Under Assumption 1,*

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

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

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

(iv) (Multiple equilibria) There exists an interval $[\underline{\delta},\bar{\delta}]$ ($[\underline{\eta},\bar{\eta}]$) such that holding other parameters constant, multiple steady states arise only for $\delta \in [\underline{\delta},\bar{\delta}]$ ($\eta \in [\underline{\eta},\bar{\eta}]$);

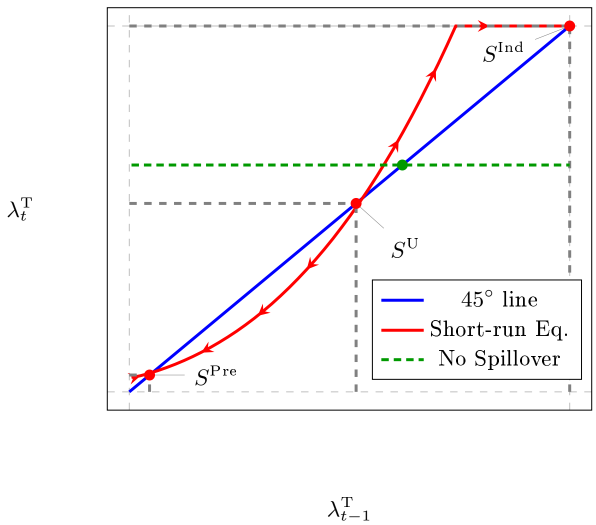
and (v) (Pareto-ranked) If multiple steady states exist, these steady states can be Pareto-ranked based on the equilibrium share of adopters.

The case of multiple steady states is illustrated in Panel A of Figure 4 in which there are three different steady states with two basins of attraction.[18] The red locus is defined by Equation (4.3). Each point on the locus is a short-run equilibrium given $\lambda_{t-1}^{\mathrm{T}}$ and the equilibrium moves along the red locus as time passes. The steady state is determined at the point where $\lambda_{t-1}^{\mathrm{T}*} = \lambda_t^{\mathrm{T}*}, \forall t$ holds, that is, where the red locus intersects with the 45-degree blue line. There are three intersection points, S^{Pre}, S^{U}, and S^{Ind}, which we label as the pre-industrialized, unstable, and industrialized steady states, respectively. S^{U} is unstable in that the economy converges to S^{U} only when the initial condition is given by the value of S^{U}, so S^{U} is out of our focus.

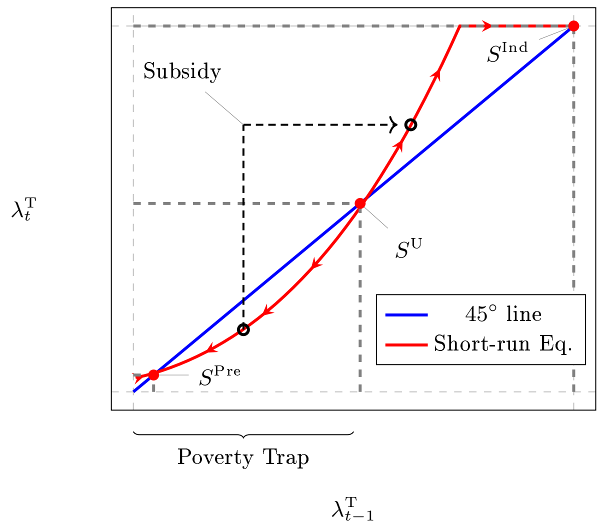
These steady states can be Pareto-ranked depending on the steady state shares. At S^{Ind}, all firms adopt technology, and S^{Pre} has a smaller adopter share than the other two, so S^{Ind} Pareto-dominates S^{Pre}. The nonlinearity of the red locus is the key to generating multiple steady states because the nonlinearity makes the locus intersect with the 45-degree line multiple times.[19] The spillover effect ($\delta > 0$) generates such nonlinearity. If there is no spillover ($\delta = 0$), the equilibrium share of adopters is determined each period regardless of the previous share and there is always a unique steady state illustrated by the intersection of the green dashed horizontal line and the 45-degree line.

[18] In this economy, there are at most three multiple steady states because of the strict convexity imposed by the assumed spillover functional form. The functional form makes λ_t^{T} be strictly convex in $\lambda_{t-1}^{\mathrm{T}}$ so that the red locus in Figure 4 intersects with the 45-degree line two times at the most. With alternative functional forms that generate a higher degree of nonlinearity, it is possible to have more multiple steady states than three.

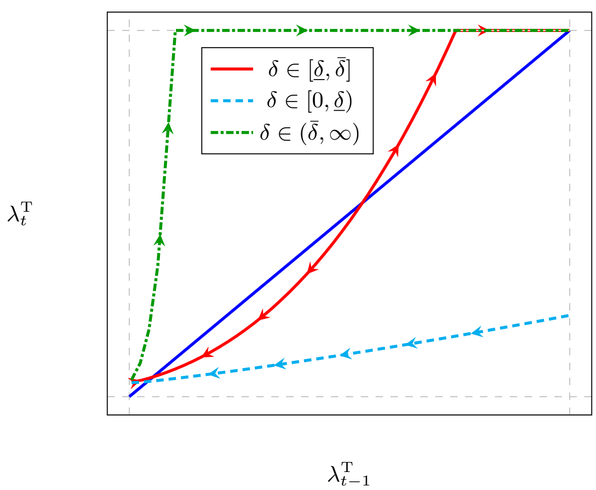
[19] Kline and Moretti (2014) did not detect nonlinearities in the agglomeration function, so they concluded that the program did not have permanent effects because the agglomeration function was not nonlinear enough to generate multiple steady states.



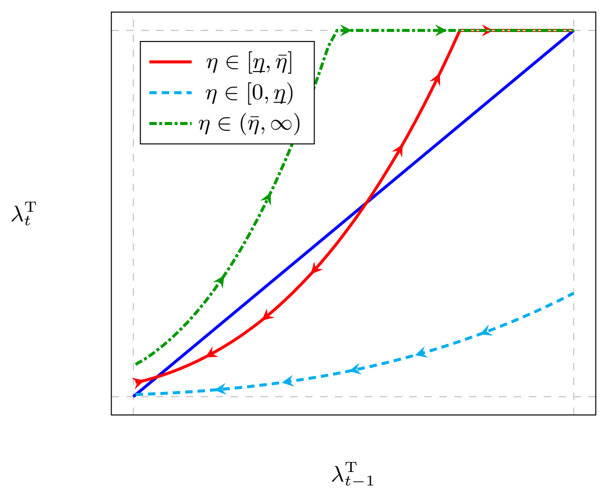
Panel A. Multiple steady states and Nonlinearity



Panel B. Poverty Trap and the Big Push



Panel C. Comparative Statistics of δ



Panel D. Comparative Statistics of η

Figure 4. Multiple steady states and the Big Push

Notes. Panel A illustrates that multiple steady states arise when the short-run equilibrium curve is sufficiently non-linear. Panel B illustrates that the big push can move an economy out of the poverty trap. Panels C and D illustrate that multiple steady states arise only for the medium range of values of η and δ, respectively.

Poverty trap and big push When multiple steady states exist, initial conditions determine which steady state to be realized. When initial conditions are given by $\lambda^{\mathrm{T}}_{njt_0} \in [0, S^{\mathrm{U}})$ and $\lambda^{\mathrm{T}}_{njt_0} \in (S^{\mathrm{U}}, 1]$, the economy converges to S^{Pre} and S^{Ind}, respectively. The range of $[0, S^{\mathrm{U}})$ is known as a poverty trap in the literature (Azariadis and Stachurski, 2005). Suppose multiple steady states exist and an initial condition is stuck in the poverty trap as in Panel B. In such a case, a big push policy that provides one-time subsidies for adopters' costs of production or fixed adoption costs can have permanent effects if it moves an economy out of the poverty trap, summarized in Proposition 2. The possibility of the big push is consistent with South Korea's big push narrative. Note that in this model, only multiple steady states can rationalize the permanent effects of the one-time policy. With a unique steady state, one-time subsidies shift the short-run equilibrium curve only temporarily and the economy converges to its original steady state after these subsidies end.

Proposition 2. *(Big push) Suppose the multiple steady states exist and the economy is initially in the poverty trap, $\lambda^{T}_{t_0} \in [0, S^{U})$. There exists $\underline{s}$ such that a one-time subsidy for adopters' input costs or fixed adoption costs that satisfy $s_t > \underline{s}$ can move an economy out of the poverty trap.*

Comparative statistics What determines this multiplicity? The existence of multiple steady states depends on values of the two key parameters δ and η (Proposition 1(iv)). 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 the spillovers or the direct productivity gains are neither too strong nor too weak (Panels C and D). Either too high or low δ leads to too strong or weak dynamic complementarity, which makes the short-run locus become not nonlinear enough and intersect with the 45-degree line only once. If η is too high, firms have large private returns from the adoption, making more firms adopt technology regardless of the previous shares and vice versa, making the locus intersect only once. These comparative statistics of δ and η offer one potential explanation for why the South Korean economy experienced a remarkable transformation toward heavy manufacturing sectors after the big push while other developing countries did not. The values of δ and η may depend on country-specific features and South Korea could have been a special case in which their values were in a range that generated multiple steady states.[20]

5 Taking the Model to the Data

5.1 Quantitative Model

We extend the simple model and develop a quantitative framework to quantify the effects of the big push policy. Appendix Section D provides further details.

[20] η can be related to the absorptive capacity of new technology and δ to a degree of barriers for knowledge diffusion. For example, countries with lower amounts of skilled labor endowments or higher language barriers may have lower η or δ, respectively. Compared to other developing countries, South Korea had higher amounts of skilled labor and used the same language (Rodrik, 1995), which could have made South Korea have higher η and δ.

Geography, sectors, and trade We divide the world into Home and Foreign (H and F). Home is a small open economy that cannot affect Foreign aggregates. Home has multiple regions indexed by $n, m \in \{1, \ldots, N\} \equiv \mathcal{N}$ and multiple sectors indexed by $j, k \in \{1, \ldots, J\} \equiv \mathcal{J}$. Each sector j variety is tradable across regions and countries, subject to iceberg costs $\tau_{nmj} \geq 1$ and $\tau^x_{nj} \geq 1$, respectively.

In each region-sector, there is a fixed mass of monopolistically competitive firms M_{nj} and perfectly competitive final goods producers produce nontradable local sectoral aggregate goods Q_{njt} used for final consumption and for intermediate inputs. They aggregate all available varieties from all regions and countries using a CES aggregator with the price index given by

$$P_{njt} = \left[\sum_m \int_{\omega \in \Omega_{mj}} (p_{njt}(\omega))^{1-\sigma} \mathrm{d}\omega + (\tau^x_{nj}(1 + t^x_{jt}) P^f_{jt})^{1-\sigma} \right]^{\frac{1}{1-\sigma}}. \tag{5.1}$$

$p_{njt}(\omega)$ is a price charged by firms. Ω_{mj} is the set of available sector j varieties in region m. Because there are no fixed export costs for internal trade, each region has the same set of available varieties. P^f_{jt} is a exogenous price of foreign variety and t^x_{jt} is a import tariff.

Home firms take foreign demands D^x_{jt} as exogenously given and face the demand schedule of $p_{it}^{-\sigma} D^x_{jt}$. When exporting to Foreign, firms incur fixed export costs F^x_j in units of labor (Melitz, 2003).[21]

Production Firms have constant return to scale (CRS) Cobb-Douglas production which requires intermediate inputs for their production: for firm i in sector j,

$$y_{it} = z_{it} L_{it}^{\gamma^L_j} \prod_k (M^k_{it})^{\gamma^k_j}, \qquad \gamma^L_j + \sum_k \gamma^k_j = 1,$$

where L_{it} are labor inputs, and M^k_i are sector k intermediate inputs.

z_{it} consists of three components as in the simple model, but the spillover $f(\lambda_{nj,t-1})$ increases in the previous region-sector adopter shares and ϕ_{it} follows the bounded Pareto distribution:

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

parametrized by $\phi^{\max}_{njt}$, $\phi^{\min}_{njt}$, and θ. We assume that the gap between the lower and upper bounds of the distribution is constant across regions, sectors, and periods: $\phi^{\max}_{njt} = \kappa \phi^{\min}_{njt}$, parametrized by κ. The lower bounds vary across regions, sectors, and periods, but the upper bounds are always proportional to the lower bounds by κ.[22] $\phi^{\min}_{njt}$ is related to natural advantage. Any region-sector level productivity shifters that cannot be explained by technology adoption, such as the construction

[21] Note that unlike the fixed adoption costs, fixed export costs are not subject to the dynamic complementarity because they are not in units of final goods.

[22] If $\kappa \to \infty$, the bounded Pareto collapses to the unbounded Pareto.

of industrial complexes, are rationalized by $\phi_{njt}^{\min}$

Adoption cost and subsidy We model the adoption subsidies as input subsidies $0 \leq s_{njt} \leq 1$ potentially varying across regions, sectors, and periods.[23] With subsidies, adopters' input costs are

$$(1 - s_{njt})[w_{nt}L_{it} + \sum_j P_{njt}M_{it}].$$

The government imposes a common labor tax τ_t^w to finance these subsidies and the after-tax wage in region n is $(1 - \tau_t^w)w_{nt}$.[24] The government budget is balanced every period.

We assume that adoption goods are produced using the same Cobb-Douglas production technology. The fixed adoption costs are in units of input bundles: $c_{njt}F_{nj}^{\mathrm{T}}$ where the amounts of input bundles required for the adoption F_{nj}^{T} potentially vary across regions.

Household preference and migration In each region, there is a competitive labor market and wages are equalized across sectors. Households inelastically supply labor. We normalize the total population of the Home regions to 1. Households have Cobb-Douglas preferences over final consumption baskets with the Cobb-Douglas shares $\sum_j \alpha_j = 1$. They are subject to their budget constraints: $P_{nt}C_{nt} = (1 - \tau_t^w + \bar{\pi}_t)w_{nt}$, where C_{nt} is Cobb-Douglas consumption baskets and P_{nt} is the price index for these baskets. $(1 - \tau_t^w + \bar{\pi}_t)w_{nt}$ is the total income of households, which is the sum of after-tax wages $(1 - \tau_t^w)w_{nt}$ and income from dividends $\bar{\pi}_t^h w_{nt}$.

At the beginning of each period, households make *myopic* migration decisions and then, they supply labor and earn wages in new regions in which they decided to live. They maximize their static utility by choosing a location that maximizes their static utility each period: $\max_n\{U_{mnt}^h(\epsilon_{mnt}^h)\}$, where $\mathcal{U}_{mnt}^h(\epsilon_{mnt}^h)$ is the utility of a household h that lived in n and move to m in t:

$$\mathcal{U}_{nmt}^h(\epsilon_{nmt}^h) = V_{mt}\frac{(1 - \tau_t^x + \bar{\pi}_t^h)w_{mt}}{P_{mt}}d_{nm}\epsilon_{nmt}^h.$$

V_{mt} is an exogenous amenity in m that captures characteristics that make regions more or less attractive to live in. d_{nm} are the utility costs of moving from n to m. ϵ_{nmt}^h is a preference shock drawn i.i.d. from Fréchet distribution with the shape parameter ν: $F(\epsilon) = exp(\epsilon^{-\nu})$ (Eaton and Kortum, 2002). A share of households moving from n to m in t is

$$\mu_{nmt} = \frac{\left(V_{mt}\frac{(1-\tau_t^x+\bar{\pi}_t^h)w_{mt}}{P_{mt}}d_{nm}\right)^\nu}{\sum_{m'}\left(V_{m't}\frac{(1-\tau_t^x+\bar{\pi}_t^h)w_{m't}}{P_{m't}}d_{nm'}\right)^\nu}.$$

[23]This is based on the fact that the government provided subsidies to large adopters so they could purchase capital equipment related to adopted technologies, and we interpret new capital equipment as intermediate inputs in our model.

[24]The assumption that the government finances its adoption subsidies through a labor tax is based on 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; Itskhoki and Moll, 2019).

ν is migration elasticity that governs the responsiveness of migration flows to real income changes of destination. The population of each region evolves according to $L_{mt} = \sum_n \mu_{nmt} L_{n,t-1}$.

We define the regional welfare of households living in region n in time t as the expected static utility before the realization of the preference shocks:

$$U_{nt} = \left[\sum_m \left(V_{mt} \frac{(1 - \tau_t^x + \bar{\pi}_t^h) w_{mt}}{P_{mt}} d_{nm} \right)^\nu \right]^{\frac{1}{\nu}}. \tag{5.2}$$

The aggregate welfare is defined as the population-weighted average of U_{nt}: $U_t^{\text{agg}} \equiv \sum_n \frac{L_{nt}}{L_t} U_{nt}$.

Equilibrium In the equilibrium, given initial conditions $\{\lambda^{\text{T}}_{njt_0}, L_{nt_0}\}$ and a path of the fundamentals $\{\phi^{\min}_{njt}, V_{nt}, P^f_{jt}, D^x_{jt}\}$, subsidies $\{s_{njt}\}$, and tariffs $\{t^x_{jt}\}$, firms maximize profits; households maximize utility; labor and goods markets clear; trade is balanced, and the government budget is balanced; and firms' adoption and households' migration decisions endogenously determine the path of state variables λ_{njt} and L_{nt}.

5.2 Calibration

Each period corresponds to 4 years in the data. We aggregate sectors into four categories: commodity, light and heavy manufacturing, and service sectors. Commodity and manufacturing sectors are tradable both internally and internationally, whereas the service sector is tradable only internally. Because most of the adoption occurred in the heavy manufacturing sectors, we assume that technology adoption is available only for the heavy manufacturing sector.

We calibrate our model to the period between 1972 and 1980. We take initial adopter shares $\lambda^{\text{T}}_{nj,68}$ and population $L_{n,68}$ directly from the data.[25] Given these initial values, we solve the model for $t = 1$ corresponding to 1972 in the data. After solving for $t = 1$, given values of λ_{nj1} and L_{n1}, we obtain the equilibrium λ_{nj1} and L_{n1}, and solve for $t = 2$, and so on. After $t = 3$, fundamentals and tariffs are held constant at the level of 1980. We sequentially solve the model period by period for large enough T until the model converges to a steady state.

We calibrate subsidy s_{njt}, fundamentals $\mathbf{\Psi}_t$, tariffs t^x_{jt}, and the following set of structural parameters

$$\mathbf{\Theta} = \{ \underbrace{M_{nj}}_{\substack{\text{Fixed} \\ \text{firm mass}}}, \underbrace{\theta, \kappa}_{\substack{\text{Pareto} \\ \text{distribution}}}, \underbrace{\eta, \delta, F^{\text{T}}_{nj}}_{\substack{\text{Technology} \\ \text{adoption}}}, \underbrace{\sigma, \gamma^k_j, \gamma^L_j}_{\text{Production}}, \underbrace{\tau_{nmj}, \tau^x_{nj}, F^x_j}_{\text{Trade costs}}, \underbrace{\nu, d_{nm}}_{\text{Migration}}, \underbrace{\alpha_j}_{\text{Preference}} \}.$$

We divide $\mathbf{\Theta}$ into two subgroups, $\mathbf{\Theta}^{\text{E}} = \{\eta, \delta, M_{nj}, \theta, \sigma, \gamma^L_j, \gamma^k_j, \nu, d_{nm}, \tau_{nmj}, \tau^x_{nj}, \alpha_j\}$ and $\mathbf{\Theta}^{\text{M}} = \{\kappa, F^x_j, F^{\text{T}}_{nj}\}$ depending on whether they are externally or internally calibrated, respectively. We externally calibrate $\mathbf{\Theta}^{\text{E}}$ and $\boldsymbol{t}^x_t$ and internally calibrate $\bar{s}$, $\mathbf{\Psi}_t$, and $\mathbf{\Theta}^{\text{M}}$ by indirect inference. Table 4 summarizes our calibration strategy. Appendix Section E.1 explains the procedure in detail.

[25] While our firm balance sheet data covers from 1970 to 1982, technology adoption contracts cover from 1962 to

Table 4: Calibration Strategy

	Description	Value	Identification / Moments
		External calibration	
η	Direct productivity gains	1.35	Winners vs. losers, Table 1
δ	Spillover semi-elasticity	1.33	Spillover estimate, Table 2
σ	Elasticity of substitution	4	Broda and Weinstein (2006)
θ	Pareto shape parameter	3.18	Axtell (2001)
ν	Migration elasticity	2	Peters (2021)
ζ	Distance migration cost elasticity	0.78	Gravity estimates
ξ	Distance trade cost elasticity	0.43	Monte et al. (2018)
α_j	Preferences		IO table
γ_j^k	Production		IO table
M_{nj}	Exogenous firm mass		Value added (Chaney, 2008)
		Internal calibration	
$\varphi_{j0}^{\mathrm{T}}$	Fixed adoption cost	1e-3	Adopter share, heavy mfg.
$\varphi_{j1}^{\mathrm{T}}$	Fixed adoption cost, dist. to port	8e-4	PPML, adopter share & dist. to port
F_j^x	Fixed export cost, comm., light mfg.	0.29	Exporter share, light mfg.
F_j^x	Fixed export cost, heavy mfg.	0.03	Exporter share, heavy mfg.
κ	Pareto upper bound	1.41	Share of regions with adoption
$\bar{s}$	Subsidy rate	0.08	
$\phi_{njt}^{\min}$	Natural advantage		Region-sector GO
D_{jt}^x	Foreign market size		Export intensity
P_{jt}^f	Foreign import cost		Import share
V_{nt}	Amenity		Pop. dist.

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

5.2.1 External Calibration

Elasticity of substitution We set the elasticity of substitution σ to 4 following Broda and Weinstein (2006).

Technology adoption Taking a log of adopters' sales, we derive the following relationship that can be mapped to the winners vs. losers specification (Equation (3.1)):

$$\ln Sales_{it} = (\sigma - 1)\ln(\eta)T_{it} + \delta_{mt} + (\sigma - 1)\ln\phi_{it}.$$

1985. Using the information on the start year of firms, we construct the adopter shares in 1968.

Match-year fixed effects δ_{mt} capture variables common at the match levels, including local spillovers, unit costs of production, and market size common across firms within region-sectors.[26] Also, based on the lack of evidence that winners received more subsidies relative to losers, we map the estimates to pure effects of technology adoption rather than joint effects including subsidies, and let subsidies be absorbed out by δ_{mt}.[27] From this mapping, we set $\eta = \exp(\frac{0.9}{\sigma-1}) = 1.6$, where 0.9 corresponds to the average of the first four year coefficients after the adoption (column 1 of Table 1).

Taking log on non-adopters' sales, we obtain the following relationship that can be mapped to the spillover regression (Equation (3.4)):

$$\ln Sales_{it} = (\sigma-1)\delta\lambda^{\mathrm{T}}_{njt} + \mathbf{X}'_{njt}\gamma + (\sigma-1)\ln\phi_{it},$$

where $\mathbf{X}_{njt}$ are region-sector controls including unit cost and market access terms. From this relationship, we pin down δ to be $4/(\sigma-1) = 2$ (column 2. of Table 2). An alternative mapping based on $\mathrm{TFP}^{\mathrm{rr}}$ gives similar values for η and δ. [28]

Migration We parametrize migration costs as $d_{nm} = (\mathrm{Dist}_{nm})^{\zeta}$, where Dist_{nm} is the distance between regions n and m. We set ν to 2 (Peters, 2021).[29] We derive a gravity equation for migration flows and estimate the equation using migration flows of people aged 20 to 55 from 1990 to 1995 obtained from the 1995 Population and Housing Census, which was the closest to the sample period among the accessible population census data: $\mu_{nm} = exp(\nu\zeta \times \mathrm{Dist}_{nm} + \delta_n + \delta_m) \times \epsilon_{nmt}$, where standard errors are two-way clustered at the origin and destination levels. 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). The gravity estimate is $\nu\zeta = 1.39$ and statistically significant under the 1%.

Iceberg costs and tariffs We parametrize internal iceberg costs as $\tau_{nmj} = (\mathrm{Dist}_{nm})^{\xi_j}$ where Dist_{nm} is distance between n and m and ξ_j is sector-specific distance elasticity. We set $\xi_j = 1.29/(\sigma - 1)$ for commodity and manufacturing sectors (Monte et al., 2018). For international iceberg costs, we assume that firms have to ship their products to the nearest port and then pay both iceberg and fixed trade costs at the port when they export or import and parametrize international iceberg costs as $\tau^x_{nj} = (\mathrm{Dist}^{\mathrm{port}}_n)^{\xi_j}$, where $\mathrm{Dist}^{\mathrm{port}}_n$ is the distance between region n and the nearest port among the seven largest ports in South Korea. Any common components of the iceberg costs are not separately identifiable from D^x_{jt}, so we normalize $\tau^x_{nj} = 1$ for regions with ports. We take tariffs directly from the data, obtained as the average within aggregated sectors.

[26] Specifically, δ_{mt} absorb out $(1-\sigma)\ln c_{njt} + (\sigma-1)\delta\lambda^{\mathrm{T}}_{njt} + \ln(\sum_m \tau_{nmj} P^{\sigma-1}_{mjt} E_{mjt} + \tau^x_{nj} D^x_{jt})$.

[27] If we had found that winners were more likely to receive subsidies, the sales estimates should have been mapped to the joint effects $(\sigma-1)\ln(\frac{\eta}{1-s_{it}})$.

[28] Using that $\mathrm{TFP}^{\mathrm{rr}} \propto \frac{\sigma-1}{\sigma}\ln z_{it}$, we obtain $\eta = 2.2$ and $\delta = 1.8$.

[29] The value of 2 is also in line with the migration elasticity of 0.7 at the annual frequency by Choi (2022) based on the South Korean migration flows.

The remaining parameters We set the Pareto shape parameter θ to $1.06 \times (\sigma - 1)$ (Axtell, 2001; di Giovanni et al., 2011). We set M_{nj} to be proportional to the GDP share of each region and sector and set $\sum_{n,j} M_{nj} = 1$ (Chaney, 2008). The Cobb-Douglas shares of preference and production function, α_j, γ_j^k and γ_j^L, are taken from the IO tables.

5.2.2 Internal Calibration

Adoption subsidy The adoption subsidies are provided in $t = 2, 3$, corresponding to 1976 and 1980, and to firms in regions with at least one firm that ever received directed credit $\mathcal{N}^{\mathrm{T}}$. We assume the same subsidy level $\bar{s}$ across these regions and periods:

$$s_{njt} = \begin{cases} \bar{s} & \text{if} \quad t \in \{2, 3\}, \quad \forall n \in \mathcal{N}^{\mathrm{T}}, \quad j = \{\text{heavy mfg.}\} \\ 0 & \text{otherwise.} \end{cases} \tag{5.3}$$

Adoption cost We parametrize adoption costs as a function of distance to the nearest port: $F_{njt}^{\mathrm{T}} = \varphi_{j0}^{\mathrm{T}} + \varphi_{j1}^{\mathrm{T}} \ln \mathrm{Dist}_n^{\mathrm{port}}$. $\varphi_{j0}^{\mathrm{T}}$ governs the common costs across regions and $\varphi_{j1}^{\mathrm{T}} > 0$ captures the notion that if firms were located farther away from the ports, knowledge transfer becomes more costly.[30]

Constrained minimum distance We calibrate $\mathbf{\Theta}^{\mathrm{M}}$, $\bar{s}$, and $\mathbf{\Psi}_t$ by minimizing the distance between the data moments and the model counterparts. Our calibration procedure requires moments from micro data and a set of cross-sectional aggregate variables in 1972, 1976, and 1980. Let $g(\mathbf{\Theta}^{\mathrm{M}}, \bar{s}, \mathbf{\Psi}_t) \equiv \bar{\mathbf{m}} - \mathbf{m}(\mathbf{\Theta}^{\mathrm{M}}, \bar{s}, \mathbf{\Psi}_t)$ be distance between a vector of the model moments $\bar{\mathbf{m}}$ and the data counterparts $\mathbf{m}(\mathbf{\Theta}^{\mathrm{M}}, \bar{s}, \mathbf{\Psi}_t)$ and let $c(\mathbf{\Theta}^{\mathrm{M}}, \bar{s}, \mathbf{\Psi}_t) \equiv \mathbf{C}(\mathbf{\Theta}^{\mathrm{M}}, \bar{s}, \mathbf{\Psi}_t) - \mathbf{C}_t$ be the imposed constraints, where $\mathbf{C}(\mathbf{\Theta}^{\mathrm{M}}, \bar{s}, \mathbf{\Psi}_t)$ and $\mathbf{C}_t$ are a vector of the model moments and data counterparts. We calibrate $\mathbf{\Theta}^{\mathrm{M}}$, $\mathbf{\Psi}_t$, and $\bar{s}$ by solving the following constrained minimization problem:

$$\{\hat{\mathbf{\Theta}}^{\mathrm{M}}, \hat{\bar{s}}\} \equiv \underset{\{\mathbf{\Theta}^{\mathrm{M}}, \bar{s}\}}{\arg\min} \left\{ g(\mathbf{\Theta}^{\mathrm{M}}, \bar{s}, \mathbf{\Psi}_t)' g(\mathbf{\Theta}^{\mathrm{M}}, \bar{s}, \mathbf{\Psi}_t) \right\} \qquad \text{s.t.} \qquad c(\mathbf{\Theta}^{\mathrm{M}}, \bar{s}, \mathbf{\Psi}_t) = 0. \tag{5.4}$$

We identify $\varphi_{j0}^{\mathrm{T}}$ and $\varphi_{j1}^{\mathrm{T}}$ using the median adopter shares conditional on $\lambda_{njt}^{\mathrm{T}} > 0$ in 1972 and the estimates obtained from the PPML regression where we regress the 1972 adopter shares on log of the nearest distance to port: $\lambda_{njt}^{\mathrm{T}} = \exp(\beta_0 + \beta_1 \ln \mathrm{Dist}_n^{\mathrm{port}}) \times \epsilon_{njt}$. The estimated value of β_1^{T} is $\hat{\beta}_1^{\mathrm{T}} = -0.35$, statistically significant under the 5%, implying that regions farther away from ports had lower adopter shares. $\hat{\beta}_1^{\mathrm{T}} = -0.35$ is consistent with Comin et al. (2012) who find technology diffuses slower to locations that are farther away from origins of new technologies. We run the same regression using the model-generated data and calibrate $\varphi_{j1}^{\mathrm{T}}$ to match $\hat{\beta}_1^{\mathrm{T}}$.

We calibrate $\bar{s}$ by targeting the median of adopter shares conditional on $\lambda_{njt}^{\mathrm{T}} > 0$ in 1976 and 1980. Conditional on the benefits from the adoption (direct and spillover effects) and F_{nj}^{T}, because $\bar{s}$

[30] For example, training service provided by foreign engineers could have been more costly if firms were located farther away from ports.

Table 5: Model Fit

Moment	Model	Data
Mean $\lambda^{x}_{nj,72}$, light mfg.	0.24	0.22
Mean $\lambda^{x}_{nj,72}$, heavy mfg.	0.10	0.09
Med. $\lambda^{\mathrm{T}}_{nj,72} \vert \lambda^{\mathrm{T}}_{nj,72} > 0$	0.07	0.10
Med. $\lambda^{\mathrm{T}}_{nj,76} \vert \lambda^{\mathrm{T}}_{nj,76} > 0$	0.10	0.13
Med. $\lambda^{\mathrm{T}}_{nj,80} \vert \lambda^{\mathrm{T}}_{nj,80} > 0$	0.12	0.25
Shares of regions with adoption, 1972	0.07	0.24
Shares of regions with adoption, 1976	0.36	0.30
Shares of regions with adoption, 1980	0.58	0.36
PPML estimate, $\lambda^{\mathrm{T}}_{nj,72}$ & dist. to port	-0.35	-0.36

Notes. This table presents the fit of the model.

only enters in 1976 and 1980, the increases in the adopter shares in 1976 and 1980 relative to those in 1972 are informative about the subsidies.

With lower κ, the cutoff adoption productivity becomes more likely to be above the Pareto upper bound, leading to zero adoption, so we identify κ using the share of regions with positive adoption in 1972, 1976, and 1980. We calibrate F^{x}_{j} of the light and heavy manufacturing sectors to match the mean shares of exporters across regions in 1972. Due to the lack of data on commodity sector firms, we set F^{x}_{j} of the commodity sector to be the same as that of the light manufacturing sector.

Conditional on $\mathbf{\Theta}^{\mathrm{M}}$ and $\bar{s}$, the constraints in Equation (5.4) identify $\mathbf{\Psi}_t$ based on the model-inversion logic (Allen and Arkolakis, 2014). We impose the constraints such that sectoral export and import shares, regional distribution of sectoral gross output, and regional population distribution of the model are exactly fitted to the data counterpart of 1972, 1976, and 1980. The number of the constraints is the same as the dimension of the fundamentals, so for any given $\mathbf{\Theta}^{\mathrm{M}}$ and $\bar{s}$, the fundamentals are exactly identified by these constraints and there exists a set of the fundamentals (up to normalization) that rationalizes the data.[31] D^{x}_{jt} and P^{f}_{jt} are identified by sectoral export and import shares; $\phi^{\min}_{njt}$ by the regional sectoral gross output distribution; and V_{nt} by population distribution.

Estimation results Table 4 and 5 present the estimated values and the model fit. The data moments are well-approximated in the model. The estimated subsidy rate is 0.08, which indicates

[31] We only identify relative productivity differences across regions within sectors and periods and relative amenity differences. We normalize $\phi^{\min}_{njt}$ of the reference region to 1 for each sector and period, and V_{nt} of the reference region to 1 for each period. This normalization is unlikely to be a big concern because our interest is the comparison between the baseline economy and the counterfactual economy, which differences out the common aggregate components.

that adopters are subsidized with 8% of input expenditures. In 1976 and 1980, the average ratio between total subsidies provided to adopters and GDP is 1.2%. The adoption was more costly than exporting. On average, the calibrated adoption cost is about 120 times larger than the fixed export cost in 1972, calculated as the mean of $w_{n,72}F_j^x/c_{nj,72}F_{njt}^x$ across regions.

6 Quantification of the Effects of the Big Push Policy

Using the calibrated model, we ask how the economy would have evolved differently, had the policy not been implemented. We compare the outcomes of the baseline economy in which the big push policy is implemented to those of the counterfactual economy in which the policy is not implemented.

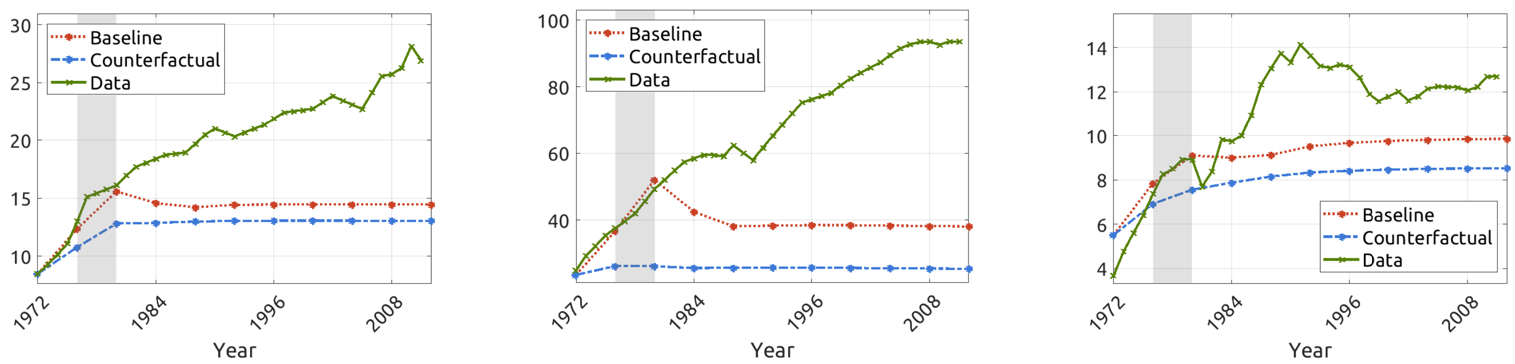
Figure 5 reports this comparison. Had the policy not been implemented, the pattern of industrialization and its comparative advantage would have evolved differently because the counterfactual economy converges to an alternative less-industrialized steady state.[32] We compare the heavy manufacturing's shares of value added to GDP, shares of employment to total employment, and shares of exports to total exports. In this alternative steady state, the GDP share would have decreased by 1.4 percentage points, the employment share by 1.3 percentage points, and the export shares by 12.5 percentage points. Although we do not directly target the employment shares, the calibrated model approximates the evolution of the employment shares between 1972 and 1980 quite well, which is the non-targeted moment of the model. However, because we do not directly target data moments after 1980, our model does not explain the evolution of these outcomes after 1980 well.

Figure 6 illustrates productivity and welfare results. In Panel A, the x and y axes are each region's steady state productivity , $M_{nj}[\int z_{it}(\phi)^{\sigma-1}dG_{njt}(\phi)]^{1/(\sigma-1)}$, in the baseline and counterfactual economies. Each dot represents each region and dots located below the 45-degree line, denoted as red stars, represent regions that had higher productivity in the baseline than the counterfactual. Only three regions had higher productivity, and the others experienced productivity loss due to intensified competition with firms in these three regions. This fact implies that the aggregate industrialization patterns documented in Figure 5 were driven by local productivity improvement of these three regions rather than the uniform improvement across the whole country. These uneven local changes are also consistent with Figure 2.

In Panel B, the x and y axes refer to the steady state regional welfare in the baseline and counterfactual economies. In the steady states, all regions had higher welfare levels in the baseline, because large productivity gains of the three regions were shared with households in other regions through trade and migration linkages. However, the three regions with productivity improvement had larger welfare gains. The aggregate welfare gains of the baseline are 4% permanently higher than the counterfactual once the economies reach steady states (Panel C). With the discount factor of

[32]Unlike the simple model that has a maximum of three steady states, the quantitative model potentially admits a larger number of steady states due to complex interaction across regions through costly trade and migration (Allen and Donaldson, 2020).

Figure 5. Industrialization

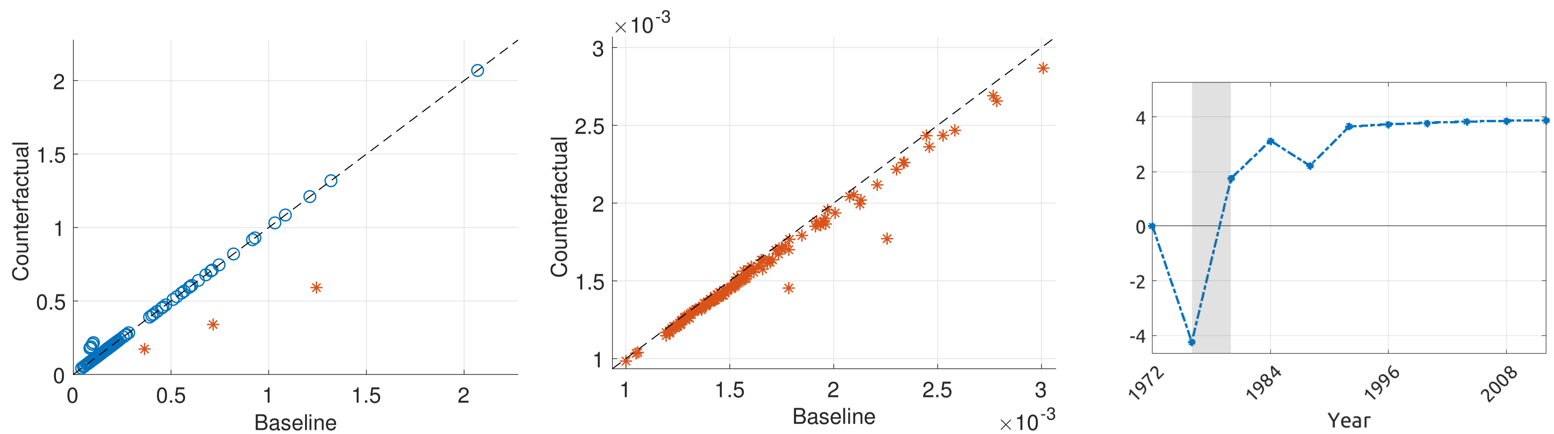


Heavy mfg. shares (%)

A. GDP B. Export C. Employment

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

Figure 6. Productivity and Welfare



A. steady state reg. prod. B. Reg. welfare C. Agg. welfare gain (%)

Notes. Panels A and B illustrate each region's productivity and welfare under the baseline and counterfactual economies (x and y axes). Each dot represents each region and is colored red (blue) if a corresponding region had higher productivity and welfare in the baseline (counterfactual) economy. Panel C illustrates the aggregate welfare gain in the baseline relative to the counterfactual economy.

0.81, the discounted utility, $\sum_{t=1}^{\infty} \beta^{t-1} U_t^{agg}$, was 7.7% higher in the baseline. At the beginning of the implementation of the policy, however, the aggregate welfare of the baseline relatively decreased. The calibrated subsidies are not optimally designed, so there is room for welfare improvement.[33]

International trade In this model, due to the Cobb-Douglas production and utility, consumers and firms spend a constant fraction of their total expenditures. Therefore, in the closed economy which is the limiting case of the open economy model that can be achieved by letting $P_{jt}^f \to \infty$ and $D_{jt}^x \to 0$, although the big push induces the economy to reach an alternative steady state, the gross output shares would be constant across steady states despite different amounts of the adoption. Changes in the gross output shares come from changes in South Korea's export patterns in the open economy. In the industrialized steady state, higher productivity in the heavy manufacturing sector increases its exports, which leads to higher gross output shares when compared to the less-industrialized steady state.

Scale complementarity The newly added features of the quantitative model introduce additional complementarities between firm scale and the adoption because firms have larger profit gains from the adoption with a larger scale. The scale complementarities interact with the dynamic complementarity and potentially amplify the latter.[34] First, international trade makes firms scale larger through market size effects (Yeaple, 2005; Verhoogen, 2008; Lileeva and Trefler, 2010; Bustos, 2011). Second, forward and backward linkages due to roundabout production is another source of the scale complementarity (Krugman and Venables, 1995). Third, migration amplifies the scale complementarity in regions with higher adopter shares because these regions attract higher migration inflows, which lowers the labor costs of production.

The assumptions of static technology adoption decisions and myopic migration The assumptions of myopic migration by households and static technology adoption decisions by firms make state variables $\{L_{nt}, M_{njt}^{\mathrm{T}}\}$ backward-looking. This simplification allows us to preserve the rich spatial heterogeneity and connect the model to the empirical findings while facilitating computational implementation.[35] If adoption costs are sunk rather than fixed, adoption decisions become forward-looking and depend on the entire path of future wages and prices. Even with forward-looking decisions, the dynamic complementarity potentially generates multiple steady states, the setup studied by Alvarez et al. (2023). Because static equilibrium, such as employment, gross output, and export shares, is not affected by these simplifying assumptions, if we target the same path of state variables using such a forward-looking model with multiple steady states, our results do not change qualitatively.

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

[34] Note that the scale complementarity differs from the dynamic complementarity. When fixed adoption costs are in units of labor, regardless of market size, the simple model does not feature dynamic complementarity and therefore multiple steady states. See Section C.3.

[35] Desmet and Rossi-Hansberg (2014), Desmet et al. (2018), Nagy (2020), and Peters (2021) simplify forward-looking decisions of agents to make models more tractable while preserving rich spatial heterogeneity.

However, the values of the calibrated models would be different, and the counterfactual results from the forward-looking model would be quantitatively different.

7 Conclusion

We empirically and quantitatively examine the impact of technology adoption on late industrialization in South Korea. We find that technology adoption not only directly benefited adopters but also had local spillover effects, and firms were more likely to adopt new technologies when more local firms adopt them. Based on these findings, we build a dynamic spatial model to conduct a counterfactual analysis of the big push policy for the adoption implemented by the Korean government. Using a quantitative model calibrated to firm-level data and econometric estimates, we show that the big push could have a permanently large impact on the economy by moving it to a more-industrialized steady state. Our empirical findings and quantitative exercises highlight that coordination failures can explain why technologies diffuse slowly to developing countries, and knowledge flows from developed to developing countries can be an important source of economic development.

References

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

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

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

Alfaro-Ureña, Alonso, Isabela Manelici, and Jose P. Vasquez, "The Effects of Joining Multinational Supply Chains: New Evidence from Firm-to-Firm Linkages," *Quarterly Journal of Economics*, 2022, *137* (3), 1495–1552.

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.

Alvarez, Fernando E., David Argente, Francesco Lippi, Esteban Méndez, and Diana Van Patten, "Strategic Complementarities in a Dynamic Model of Technology Adoption: P2P Digital Payments," *NBER Working Paper 31280*, 2023.

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

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

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.

Bai, Jie, Panle Jia Barwick, Shengmao Cao, and Shanjun Li, "Quid Pro Quo, Knowledge Spillover, and Industrial Quality Upgrading: Evidence from the Chinese Auto Industry," *NBER Working Paper 27644*, 2020.

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

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

Bianchi, Nicola and Michela Giorcelli, "The Dynamics and Spillovers of Management Interventions: Evidence from the Training Within Industry Program," *Journal of Political Economy*, 2022, *130* (6), 1630–1675.

Blackwood, G. Jacob, Lucia S. Foster, Cheryl A. Grim, John Haltiwanger, and Zoltan Wolf, "Macro and Micro Dynamics of Productivity: From Devilish Details to Insights," *American Economic Journal: Macroeconomics*, 2021, *13* (3), 142–172.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess, "Revisiting Event Study Designs: Robust and Efficient Estimation," *arXiv preprint arXiv:2108.12419*, 2023.

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

Buera, Francisco J., 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 Santacreu, "Knowledge Diffusion, Trade, and Innovation across Countries and Sectors," *American Economic Journal: Macroeconomics*, 2022, *14* (1), 104–145.

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer, "The Effect of Minimum Wages on Low-Wage Jobs," *The Quarterly Journal of Economics*, 2019, *134* (3), 1405–1454.

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

Choi, Jaedo, "Internal Migration, Sectoral Reallocation, and Large Devaluation," *Working Paper*, 2022.

— **and Andrei A. Levchenko**, "The Long-Term Effects of Industrial Policy," *NBER Working Paper 29263*, 2023.

— **and Younghun Shim**, "From Adoption to Innovation: State-Dependent Technology Policy in Developing Countries," *Working Paper*, 2022.

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

Comin, Diego, Mikhail Dmitriev, and Esteban Rossi-Hansberg, "The Spatial Diffusion of Technology," *NBER Working Paper 18534*, 2012.

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).

Costinot, Arnaud and Andrés Rodríguez-Clare, "Trade Theory with Numbers: Quantifying the Consequences of Globalization," in "Handbook of International Economics," Vol. 4 2014, pp. 197–261.

De Loecker, Jan, "Detecting Learning by Exporting," *American Economic Journal: Microeconomics*, 2013, *5* (3).

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

Deshpande, Manasi and Yue Li, "Who is Screened Out? Application Costs and the Targeting of Disability Programs," *American Economic Journal: Economic Policy*, 2019, *11* (4), 213–248.

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

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

di Giovanni, Julian, Andrei A. Levchenko, 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.

Giorcelli, Michela, "The Long-Term Effects of Management and Technology Transfers," *American Economic Review*, 2019, *109* (1), 121–152.

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

Goodman-Bacon, Andrew, "Difference-in-Differences with Variation in Treatment Timing," *Journal of Econometrics*, 2021, *225* (2), 254–277.

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.

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

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

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

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

Javorcik, Beata Smarzynska, "Does Foreign Direct Investment Increase the Productivity of Domestic Firms? in Search of Spillovers through Backward Linkages," 2004.

Jones, Leroy P. and IL Sakong, *Government, Business, and Entrepreneurship in Economic Development*, Cambridge, MA: Harvard University Press, 1980.

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.

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

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.

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 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*, 2022.

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

Lee, Jong Wha, "Government Interventions and Productivity Growth," *Journal of Economic Growth*, 1996, *1* (3), 391–414.

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

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

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

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

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

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

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

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

Olley, G. Steven and Ariel Pakes, "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 1996, *64* (6), 1263–1297.

Parente, Stephen L. and Edward C. Prescott, *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.

Peters, Michael, "Market Size and Spatial Growth - Evidence from Germany's Post-War Population Expulsions," *NBER Working Paper w29329*, 2021.

POSCO, *History of POSCO, 1968-2018 (in Korean)*, Pohang: POSCO, 2018.

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

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

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

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

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

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

Sun, Liyang and Sarah Abraham, "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects," *Journal of Econometrics*, 2021, *225* (2), 175–199.

Ventura, Jaume, "Growth and Interdependence," *Quarterly Journal of Economics*, 1997, *112* (1), 57–84.

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

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

Westphal, Larry E, "Industrial Policy in an Export-Propelled Economy: Lessons from South Korea's Experience," *Journal of Economic Perspectives*, 1990, *4* (3), 41–59.

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

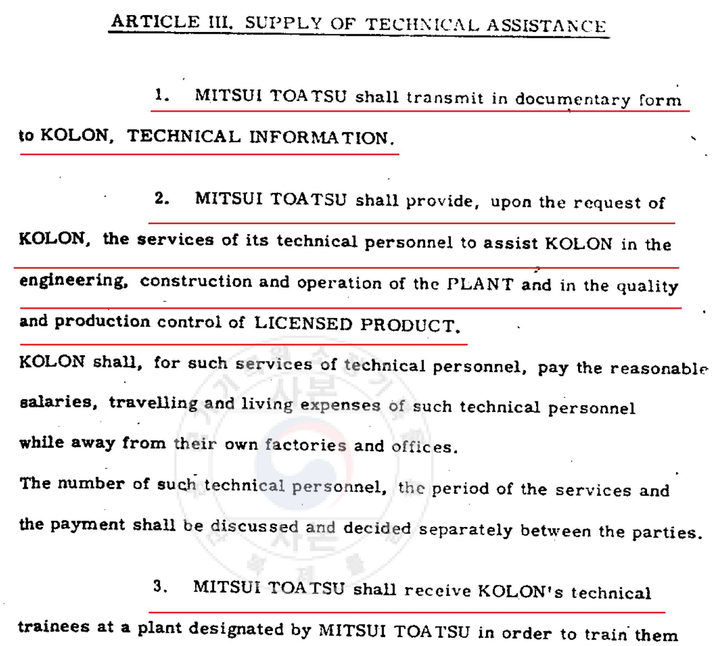
Young, Alwyn, "The Tyranny of Numbers: Confronting the Statistical Realities of the East Asian Growth Experience," *Quarterly Journal of Economics*, 1995, *110* (3), 641–680.

APPENDIX

Appendix A Data

Firm-level data From contract documents, we obtain three main pieces of information: names of domestic firms, names of foreign firms, and years contracts were made. We convert all monetary values into 2015 US dollars. The dataset covers firms with more than 50 employees. When a firm merged with another firm, we counted that as an exit. For firms with missing sales, we impute sales using assets. We convert the addresses of the location of production to the 2010 administrative divisions of South Korea. We classify firms into 10 manufacturing sectors, four of which are classified as heavy manufacturing, reported in Table A1. The numbers inside the parenthesis are ISIC Revision 3.1 codes.

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


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

Foreign firms' patent We match the USPTO with foreign firms in our dataset based on foreign firms' names. Then, we merge assignee IDs with the IDs from the Global Compustat (gvkey) based on the matching constructed by Bena et al. (2017). For foreign firms that have different assignee IDs but with the same Compustat ID, we give them a unique assignee ID and sum the numbers of

Table A1: Sector Classification

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

Notes. * denotes for heavy manufacturing sectors. The numbers inside parenthesis denote ISIC Rev 3.1 codes.

patents and citations up to the Compustat ID level.

Other regional and sectoral data The regional population data comes from the Population and Housing Census, the 2% random sample of the total population. We digitize import tariff data from Luedde-Neurath (1986) for 1968, 1974, 1976, 1978, 1980, and 1982. The tariffs are in the Customs Cooperation Council Nomenclature (CCCN). We convert CCCN to ISIC and then average the results across four-digit ISIC codes. For missing years, we impute values using the geometric average. We obtain IO tables from the Bank of Korea. We convert the codes of the IO tables into the ISIC codes.

Appendix B Empirics

B.1 An Example of POSCO

Before presenting formal econometric evidence, we give an example of POSCO to illustrate how technology adoption benefited firms through these three channels. POSCO, now one of the top five steel producers in the world, was the first integrated steel mill in South Korea. Integrated steel mills are known for their integral role during industrialization because they produce high-quality steel used as inputs to other manufacturing sectors.

In 1968, POSCO made technology adoption contracts for the first time with a Japanese company, Nippon Steel Corporation (NSC).[36] As part of the contract, NSC transferred blueprints of the factory and capital equipment and sent its engineers to POSCO to train Korean engineers and to guide the construction of the factory and operation of capital equipment. The Korean government subsidized the costs of capital equipment related to the newly adopted technology through directed credit. Due to this contract, in 1973, POSCO could initiate its production. The initiation of production due to the new adoption is related to our first finding on the direct effects on adopters.

Because of the local labor mobility of engineers across firms, POSCO's newly acquired technology was not confined to itself. After the first adoption, its engineers acquired new knowledge from training from the Japanese engineers, learning by doing and reverse engineering. Soon after, they moved to local smaller-sized mills or capital goods producers and diffused this knowledge, which improved the performance of these local firms (Enos and Park, 1988).[37] This knowledge diffusion through labor mobility is related to our second finding on local spillovers.

Later, the knowledge diffusion to local smaller-sized firms facilitated POSCO to adopt more advanced technologies. In 1980, POSCO planned to adopt new technology related to the computerization of the production process, which entailed large setup costs for installing new capital equipment and expanding its existing plants. Even though no more credit was provided by the government, POSCO decided to adopt the new technology because the availability of cheaper domestic capital inputs produced by local firms lowered the setup costs of the new adoption (POSCO, 2018, p.138-141).[38]

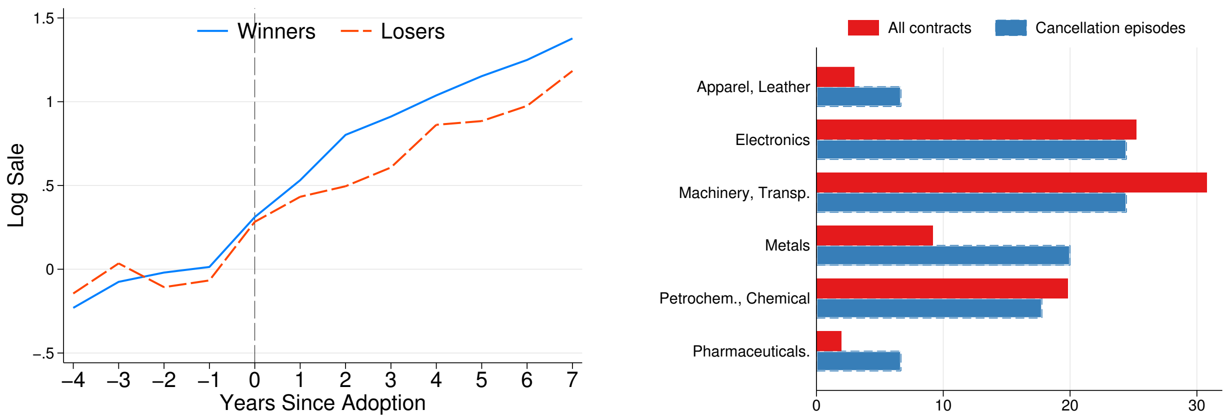
B.2 Additional Figures and Tables

[36]NSC made the contract because of the profits. The fixed fee that POSCO had to pay for the contract accounted for 20% of the total annual exports of plant engineering of NSC.

[37]Local smaller-sized mills benefited from new skills brought by POSCO engineers. Also, local capital goods producers started producing new types of more complicated equipment installed in the POSCO plants, such as equipment for water treatment and dust collection, and large magnetic cranes, whereas, in the early 1970s, these types of capital equipment were all imported (Enos and Park, 1988, p.210-211).

[38]For the new expansion of production facilities in 1980, shares of expenditures on locally-produced capital equipment were 35%, which was 12% when they first adopted technology in 1968.

Figure B1. Raw Plots of the Data that Support the Identifying Assumption

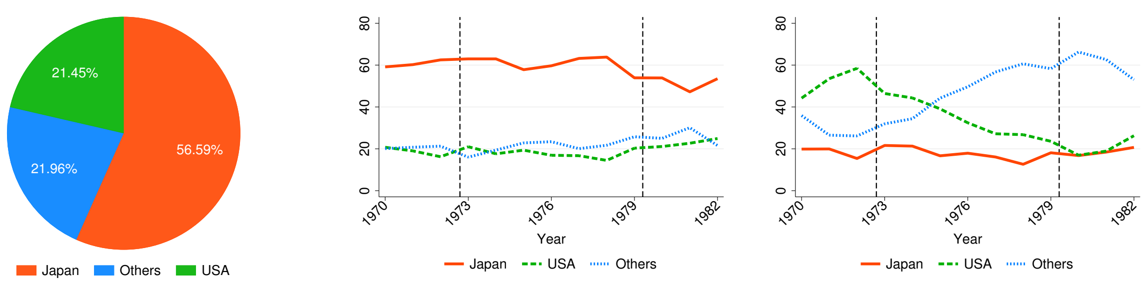


A. Trends around the event

B. Shares of contracts by sector (%)

Notes. Panel A plots the mean of log sales of winners and losers, normalized by the average before the event, respectively. Panel B illustrates the sectoral distribution of all contracts and cancellation episodes.

Figure B2. Technology Adoption, Export, and Import Shares by Country



Shares by country (%)

A. Technology adoption

B. Export

C. Import

Notes. This figure plots heavy manufacturing shares of technology adoption, export, and import by countries. Technology adoption shares are defined as the number of contracts from each country divided by the total number of contracts. The export and import shares are similarly defined in terms of values.

Table B1: Descriptive Statistics: Winners vs. Losers Design Samples from the Year of the Cancellation to 4 Years before the Cancellation

	Winner				Loser				t-Stat.	
	Mean	Med.	SD	Obs.	Mean	Med.	SD	Obs.	(Col. 1 - Col. 5)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	
Panel A. Domestic firm balance										
Log sales	17.56	17.40	2.0	319	17.97	18.06	1.82	194	0.99	[0.32]
Log emp.	6.98	7.09	1.23	237	7.05	7.21	1.49	153	0.04	[0.84]
Log fixed assets	16.82	16.74	2.18	319	16.98	16.96	2.23	194	0.10	[0.75]
Log assets	17.78	17.57	2.0	319	17.98	18.12	1.96	194	0.18	[0.67]
Panel B. Foreign firm patent activities										
Ihs # cum. patents	1.88	0	3.19	72	1.06	0	2.44	35	1.95	[0.17]
Ihs # cum. citations	2.0	0	3.4	72	1.14	0	2.63	35	1.86	[0.18]
$\mathbb{1}$[# cum. patents $\geq$ 0]	0.31	0	0.46	72	0.2	0	0.41	35	1.31	[0.26]
$\mathbb{1}$[# cum. citations $\geq$ 0]	0.31	0	0.46	72	0.2	0	0.41	35	1.31	[0.26]

Notes. Panel A reports the descriptive statistics of the winners vs. losers design samples from 4 years before the cancellations to the year of the cancellation. Panel B reports the descriptive statistics of patent activities by foreign firms matched with winners and losers. We report inverse hyperbolic sine transformation and a dummy of cumulative numbers of patents and citations. Column 9 reports the t-statistics of the mean difference between winners and losers with its p-value in brackets.

Table B2: Robustness. Covariate Balance Test

Var.	Log sale (N = 513)	Log emp. (N=390)	Log fixed assets (N=513)	Log assets (N=513)	Joint F-stat.
	(1)	(2)	(3)	(4)	(5)
Individually	-0.03 (0.03)	-0.01 (0.04)	-0.01 (0.03)	-0.01 (0.03)	NA
Jointly	-0.10 (0.07)	-0.02 (0.05)	-0.04 (0.09)	0.14 (0.14)	0.62 [0.65]

Notes. Standard errors in parenthesis are clustered at the firm level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. This table reports the covariate balance test of the winners vs. losers design samples from 4 years before the cancellation to the year of the cancellation. In the first and second rows, we regress a dummy of winners on observable individually and jointly, respectively. For the joint specification, we report the F-statistics that test whether the observables are jointly zero.

Table B3: Robustness. Direct Effects on Adopters

	Alternative TFP		Matching # = 2			Matching # = 4			Two-way clustering		
Dep. Var.	Labor prod.	Export dummy	Sale	TFP^{rr}	Subsidy	Sale	TFP^{rr}	Subsidy	Sale	TFP^{rr}	Subsidy
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
4 years before	–0.35	0.07	–0.15	–0.39	0.08	–0.08	–0.33	0.09	–0.08	–0.33	0.09
	(0.45)	(0.15)	(0.21)	(0.31)	(0.08)	(0.18)	(0.27)	(0.08)	(0.18)	(0.29)	(0.09)
3 years before	0.05	–0.00	0.05	–0.15	–0.00	0.03	–0.17	–0.01	0.03	–0.17	–0.01
	(0.32)	(0.10)	(0.17)	(0.23)	(0.09)	(0.14)	(0.21)	(0.07)	(0.13)	(0.22)	(0.09)
2 years before	0.05	–0.08	0.07	–0.08	0.03	0.15	–0.09	0.03	0.15	–0.09	0.03
	(0.29)	(0.10)	(0.14)	(0.20)	(0.08)	(0.13)	(0.18)	(0.08)	(0.13)	(0.18)	(0.08)
1 year before											
Year of event	0.04	–0.05	0.03	–0.04	0.03	0.01	–0.07	0.04	0.01	–0.07	0.04
	(0.15)	(0.08)	(0.12)	(0.12)	(0.09)	(0.11)	(0.12)	(0.09)	(0.11)	(0.13)	(0.09)
1 year after	0.58	0.16	0.67*	0.53	0.00	0.53	0.14	0.01	0.53*	0.14	0.01
	(0.46)	(0.19)	(0.36)	(0.36)	(0.10)	(0.32)	(0.39)	(0.09)	(0.29)	(0.19)	(0.10)
2 years after	1.15*	0.17	0.84***	0.72*	–0.05	0.96***	0.75*	–0.04	0.96***	0.75*	–0.04
	(0.63)	(0.19)	(0.28)	(0.37)	(0.12)	(0.35)	(0.45)	(0.11)	(0.33)	(0.43)	(0.11)
3 years after	0.49*	0.34*	0.71***	0.25	0.13	0.82**	0.15	0.09	0.82***	0.15	0.09
	(0.28)	(0.19)	(0.27)	(0.26)	(0.13)	(0.32)	(0.23)	(0.11)	(0.29)	(0.24)	(0.11)
4 years after	0.74**	0.09	1.00***	0.64**	–0.04	1.18***	0.59**	–0.04	1.18***	0.59*	–0.04
	(0.29)	(0.28)	(0.35)	(0.32)	(0.10)	(0.44)	(0.28)	(0.09)	(0.42)	(0.30)	(0.10)
5 years after	0.58*	–0.13	1.09**	0.59*	–0.04	1.28***	0.58*	–0.04	1.28**	0.58*	–0.04
	(0.34)	(0.19)	(0.41)	(0.35)	(0.09)	(0.47)	(0.31)	(0.09)	(0.47)	(0.34)	(0.09)
6 years after	1.30***	–0.18	1.02**	0.91***	–0.05	1.08**	0.86***	–0.05	1.08**	0.86**	–0.05
	(0.43)	(0.18)	(0.43)	(0.34)	(0.10)	(0.41)	(0.30)	(0.09)	(0.43)	(0.33)	(0.10)
7 years after	0.98**	–0.06	1.05**	0.78***	–0.05	1.11**	0.74***	–0.05	1.11**	0.74**	–0.05
	(0.41)	(0.28)	(0.45)	(0.29)	(0.10)	(0.43)	(0.26)	(0.09)	(0.42)	(0.28)	(0.10)
Match-firm FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Match-year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
# Cl. (Firm)	80	91	82	72	82	95	84	95	95	84	95
# Cl. (Match)									35	33	35
N	484	644	565	425	565	690	515	690	690	515	690

Notes. Standard errors in parenthesis are clustered at the firm level or two-way clustered at the firm and match levels in columns 1-8 or 9-11, respectively. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. This table reports the estimated event study coefficients β_τ. Columns 1-3 and 4-6 report the estimates from winners vs. research design (Equation (3.1)). β_{-1} is normalized to zero. The dependent variables are log labor productivity, export dummy, log sales, TFP^{rr}, and a dummy of receiving a subsidy (credit). All specifications control for match-firm and match-year fixed effects. In columns 3-5 and 6-8, we consider alternative numbers of matched winners of 2 and 4, respectively.

Table B4: Robustness. Local Spillover. Placebo Test

Dep.	1970-1972 or 1971-1973					
	$\triangle \ln \text{Sales}_{it}$			$\triangle \mathbb{1}[\text{New Contract}_{i,t+1}]$		
	OLS	RF	IV	OLS	RF	IV
	(1)	(2)	(3)	(4)	(5)	(6)
$\triangle \text{Share}_{nj,t-2}$	-0.18		1.27	0.17		0.91
	(0.33)		(1.48)	(0.16)		(1.14)
$\text{IV}^{25\text{km}\geq}_{nj,t-2}$		0.37			0.25	
		(0.39)			(0.26)	
KP-F			8.26			7.86
Region FE	✓	✓	✓	✓	✓	✓
Sector FE	✓	✓	✓	✓	✓	✓
Sector-group FE	✓	✓	✓	✓	✓	✓
# Cl. (Region)	73	73	73	73	73	73
# Cl. (Group)	830	830	830	830	830	830
N	1004	1004	1004	1004	1004	1004

Notes. Standard errors two-way clustered at the levels of regions and business groups are reported in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. This table reports the OLS, reduced-form, and IV estimates of Equation (3.4). The adoption shares and IV are defined in Equations (3.3) and (3.5). In columns 1-3 and 4-6, dependent variables are changes in log sales or a dummy of making a new adoption contract between 1970 and 1972 or 1971 and 1973. All specifications include region, sector, and sector-group fixed effects, and initial levels of dependent variables. KP-F is the Kleibergen-Paap F-statistics. In columns 1-3, we control for initial log sales in 1970 or 1971.

Table B5: Robustness. Functional Form. Local Spillover

Dep. $\triangle \ln \text{Sales}_{it}$ 1970-1972 or 1973-1980			
	(1)	(2)	(3)
$\triangle\text{Share}_{nj,t-2}$	3.29***	2.68**	2.85**
	(1.14)	(1.23)	(1.19)
$\triangle\text{Share}_{(-i)nj,t-2} \times \mathbb{1}[\text{Share}_{(-i)njt_0} \geq \text{p90}]$	−0.57		
	(2.89)		
$\triangle\text{Share}_{(-i)nj,t-2} \times \mathbb{1}[\#\ \text{firms}_{njt_0} \geq \text{p90}]$		−0.70	
		(2.04)	
$\triangle\text{Share}_{(-i)nj,t-2} \times \mathbb{1}[\text{Sale}_{it_0} \geq \text{p90}]$			4.25
			(11.10)
SW-F, Share	8.93	37.45	40.61
SW-F, Interaction	42.96	15.61	60.31
Region FE	✓	✓	✓
Sector FE	✓	✓	✓
Sector-group FE	✓	✓	✓
# Cl. (Region)	78	78	78
# Cl. (Group)	1412	1412	1412
N	1644	1644	1644

Notes. Standard errors two-way clustered at the levels of regions and business groups are reported in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. This table reports the IV estimates of Equation (3.4). In columns 1, 2, and 3, we include interaction terms between the adopter shares and dummies of whether the initial adopter shares, the initial number of firms, and the initial sales are above the 90th percentile, respectively. We instrument these terms with interaction terms between the IV in Equation (3.5) and the corresponding initial dummies. All specifications include region, sector, and sector-group fixed effects, and initial levels of dependent variables. SW-F is the Sanderson-Windmeijer F-statistics.

Table B6: Robustness. Local Spillover

Robustness.	Alternative outcomes/controls				Alternative samples				Alternative IV distances			
Dep.	△ Export dummy	△ Log labor prod.			△ ln Sales							
Sample		Baseline			Excl. firms affil. with business grp.	Excl. regions with heavy mfg. ind. complex	Single diff. 1973–1980	Full-sample	Baseline			
IV		$IV_{inj,t-2}^{\geq 25km}$		$IV_{inj,t-3}^{\geq 25km}$		$IV_{inj,t-2}^{\geq 25km}$			$IV_{inj,t-2}^{\geq 0km}$	$IV_{inj,t-2}^{\geq 10km}$	$IV_{inj,t-2}^{\geq 50km}$	$IV_{inj,t-2}^{\geq 150km}$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$\triangle \text{Share}_{(-i)nj,t-2}$	1.03^{*}	1.70^{**}	2.88^{***}		3.81^{***}	3.69^{***}	4.00^{***}	3.93^{**}	3.73^{***}	4.06^{***}	3.63^{***}	3.75^{***}
	(0.53)	(0.74)	(1.07)		(1.03)	(0.93)	(1.09)	(1.53)	(1.10)	(1.16)	(1.01)	(1.05)
$\triangle \text{Share}_{(-i)nj,t-3}$				4.88^{***}								
				(1.73)								
KP-F	70.62	65.08	68.88	19.66	69.71	102.18	33.24	34.34	76.92	67.69	63.59	77.13
Region FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Sector FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Sector-group FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Initial y_{it_0}	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓
# Cl. (Region)	79	76	79	79	79	73	62	86	79	79	79	79
# Cl. (Group)	1294	744	1294	1294	1221	1241	724	1548	1294	1294	1294	1294
N	1492	826	1492	1492	1360	1422	734	1977	1492	1492	1492	1492

Table B7: Robustness. Local Complementarity

Dep.	$\triangle\mathbb{1}[\text{New Contract}_{i,t+1}]$							
Sample	Baseline	Excl. firms affil. with business grp.	Excl. regions with heavy mfg. ind. complex	Single diff. 1973–1980	Baseline			
IV	$\text{IV}^{\geq 25km}_{inj,t-3}$	$\text{IV}^{\geq 25km}_{inj,t-2}$			$\text{IV}^{\geq 0km}_{inj,t-2}$	$\text{IV}^{\geq 10km}_{inj,t-2}$	$\text{IV}^{\geq 50km}_{inj,t-2}$	$\text{IV}^{\geq 150km}_{inj,t-2}$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\triangle\text{Share}_{(-i)nj,t-2}$		0.67^{**}	0.71^{***}	0.97^{***}	0.78^{**}	0.80^{**}	0.68^{**}	0.62^{**}
		(0.27)	(0.26)	(0.31)	(0.37)	(0.38)	(0.27)	(0.27)
$\triangle\text{Share}_{(-i)nj,t-3}$	0.61^{*}							
	(0.33)							
KP-F	7.70	34.86	52.50	19.08	36.51	31.80	34.51	42.95
Region FE	✓	✓	✓	✓	✓	✓	✓	✓
Sector FE	✓	✓	✓	✓	✓	✓	✓	✓
Sector-group FE	✓	✓	✓	✓	✓	✓	✓	✓
# Cl. (Region)	86	83	79	68	86	86	86	86
# Cl. (Group)	1548	1454	1468	923	1548	1548	1548	1548
N	1977	1701	1820	974	1977	1977	1977	1977

Appendix C Model

C.1 Derivation of Equation (4.3)

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

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

and the probability of adoption is $\lambda_t^{\mathrm{T}} = (\bar{\phi}_t^{\mathrm{T}})^{-\theta}$, which can be re-written as $(\lambda_t^{\mathrm{T}})^{-\frac{1}{\theta}} = \bar{\phi}_t^{\mathrm{T}}$

We first show that $Q_t = A(\lambda_t^{\mathrm{T}})f(\lambda_{t-1}^{\mathrm{T}})$ and $\frac{w_t}{P_t} = \frac{1}{\mu}A(\lambda_t^{\mathrm{T}})f(\lambda_{t-1}^{\mathrm{T}})$, where

$$A(\lambda_t^{\mathrm{T}}) = \left[\frac{\theta}{\tilde{\theta}}\Big((\eta^{\sigma-1}-1)(\lambda_t^{\mathrm{T}})^{\frac{\tilde{\theta}}{\theta}}+1\Big)\right]^{\frac{1}{\sigma-1}}$$

and $\tilde{\theta} = \theta - (\sigma-1)$. Note that $\frac{L_t}{Q_t} = \frac{\int l(\omega)\mathrm{d}\omega}{Q_t} = \int \frac{y(\omega)}{Q}\frac{1}{z(\omega)}\mathrm{d}\omega = \int \frac{1}{z(\omega)}\big(\frac{p(\omega)}{P_t}\big)^{-\sigma}\mathrm{d}\omega$, where $z(\omega) = \eta(\omega)f(\lambda_{t-1}^{\mathrm{T}})\phi(\omega)$ for adopters and $z(\omega) = f(\lambda_{t-1}^{\mathrm{T}})\phi(\omega)$ for non-adopters. Using that $p(\omega) = \mu\frac{w_t}{z(\omega)}$ and $P_t = \mu w_t[\int z(\omega)^{\sigma-1}\mathrm{d}\omega]^{\frac{1}{1-\sigma}}$, we obtain $Q_t = [\int z(\omega)^{\sigma-1}\mathrm{d}\omega]^{\frac{1}{\sigma-1}}L_t$. Using the assumption of Pareto distribution and the cutoff property, we can further derive that

$$Q_t = \left[\frac{\theta}{\tilde{\theta}}\Big((\eta^{\sigma-1}-1)(\bar{\phi}_t^{\mathrm{T}})^{\tilde{\theta}}+1\Big)\right]^{\frac{1}{\sigma-1}} f(\lambda_{t-1}^{\mathrm{T}})L_t = \underbrace{\left[\frac{\theta}{\tilde{\theta}}\Big((\eta^{\sigma-1}-1)(\lambda_t^{\mathrm{T}})^{\frac{\tilde{\theta}}{\theta}}+1\Big)\right]^{\frac{1}{\sigma-1}}}_{=A(\lambda_t^{\mathrm{T}})} f(\lambda_{t-1}^{\mathrm{T}})L_t, \tag{C.2}$$

where the second equality is derived from $(\lambda_t^{\mathrm{T}})^{-\frac{1}{\theta}} = \bar{\phi}_t^{\mathrm{T}}$. Using that $P_t = \mu w_t[\int z(\omega)^{\sigma-1}\mathrm{d}\omega]^{\frac{1}{1-\sigma}}$, we can obtain that

$$\frac{w_t}{P_t} = \frac{w_t}{[\int(\mu w_t/z_{it}(\omega))^{1-\sigma}]^{\frac{1}{1-\sigma}}} = \frac{1}{\mu}A(\lambda_t^{\mathrm{T}})f(\lambda_{t-1}^{\mathrm{T}}), \tag{C.3}$$

where the second equality is derived from $(\lambda_t^{\mathrm{T}})^{-\frac{1}{\theta}} = \bar{\phi}_t^{\mathrm{T}}$.

Substituting Equations (C.2) and (C.3) into Equation (C.1), we obtain that

$$\lambda_t^{\mathrm{T}} = \left(\frac{(\eta^{\sigma-1}-1)}{\sigma F^{\mathrm{T}}}A(\lambda_t^{\mathrm{T}})^{2-\sigma}f(\lambda_{t-1}^{\mathrm{T}})L_t\right)^{\frac{\theta}{\sigma-1}}. \tag{C.4}$$

Let $\hat{\lambda}_t^{\mathrm{T}}$ be the solution of Equation (C.4). Because the equilibrium share is bounded by 1, the equilibrium share is defined as follows:

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

C.2 Proofs of Propositions

Proposition 1(i) Because the left hand side of Equation (C.4) strictly increases in λ_t^{T} but the right hand side strictly decreases in λ_t^{T} due to Assumption 1(v), there exists a unique value of λ_t^{T} that satisfies this equation. If the obtained λ_t^{T} from this equation is greater than 1, $\lambda_t^{\mathrm{T}} = 1$.

Proposition 1(ii) We use the implicit function theorem. Let

$$G(\lambda_t^{\mathrm{T}};\eta,\delta,\lambda_{t-1}^{\mathrm{T}}) = A(\lambda_t^{\mathrm{T}})^{2-\sigma} f(\lambda_{t-1}^{\mathrm{T}})\frac{(\eta^{\sigma-1}-1)}{\sigma F^{\mathrm{T}}} - (\lambda_t^{\mathrm{T}})^{\frac{\sigma-1}{\theta}}. \tag{C.5}$$

Taking the derivative of Equation (C.5) with respect to λ_t^{T}, we obtain

$$\frac{\partial G}{\partial \lambda_t^{\mathrm{T}}} = \underbrace{\frac{2-\sigma}{\sigma-1}}_{<0} \times \underbrace{A(\lambda_t^{\mathrm{T}})^{3-2\sigma}(\eta^{\sigma-1}-1)(\lambda_t^{\mathrm{T}})^{-\frac{\sigma-1}{\theta}} f(\lambda_{t-1}^{\mathrm{T}})\frac{(\eta^{\sigma-1}-1)}{\sigma F^{\mathrm{T}}}}_{>0} - \underbrace{\frac{\sigma-1}{\theta}(\lambda_t^{\mathrm{T}})^{-\frac{\tilde{\theta}}{\theta}}}_{<0} < 0, \tag{C.6}$$

where the last inequality comes from the fact that $\frac{2-\sigma}{\sigma-1} < 0$ due to $\sigma > 3$ (Assumption 1). Taking the derivative with respect to $\lambda_{t-1}^{\mathrm{T}}$,

$$\frac{\partial G}{\partial \lambda_{t-1}^{\mathrm{T}}} = A(\lambda_t^{\mathrm{T}})^{2-\sigma}\frac{\eta^{\sigma-1}-1}{\sigma F^{\mathrm{T}}} f(\lambda_{t-1}^{\mathrm{T}})\delta > 0. \tag{C.7}$$

Applying the implicit function theorem and using the signs of Equations (C.6) and (C.7), we obtain $\frac{\partial \lambda_t^{\mathrm{T}}}{\partial \lambda_{t-1}^{\mathrm{T}}} = -\frac{\partial G/\partial \lambda_{t-1}^{\mathrm{T}}}{\partial G/\partial \lambda_t^{\mathrm{T}}} > 0$, which proves that λ_t^{T} strictly increases in $\lambda_{t-1}^{\mathrm{T}}$.

Proposition 1(iii) Taking the derivative of Equation (C.5) with respect to η and δ, we obtain

$$\frac{\partial G}{\partial \eta} = A(\lambda_t^{\mathrm{T}})^{3-2\sigma} f(\lambda_{t-1}^{\mathrm{T}})\frac{(\sigma-1)\eta^{\sigma-2}}{\sigma F^{\mathrm{T}}}\frac{\theta}{\tilde{\theta}}\left[\frac{1}{\sigma-1}(\eta^{\sigma-1}-1)(\lambda_t^{\mathrm{T}})^{\frac{\theta}{\tilde{\theta}}} + 1\right] > 0, \tag{C.8}$$

and

$$\frac{\partial G}{\partial \delta} = A(\lambda_t^{\mathrm{T}})^{2-\sigma}\frac{\eta^{\sigma-1}-1}{\sigma F^{\mathrm{T}}} f(\lambda_{t-1}^{\mathrm{T}})\lambda_{t-1}^{\mathrm{T}} > 0, \tag{C.9}$$

respectively. Applying the implicit function theorem and using the signs of Equations (C.6), (C.9), and (C.8), we obtain $\frac{\partial \lambda_t^{\mathrm{T}}}{\partial \eta} = -\frac{\partial G/\partial \eta}{\partial G/\partial \lambda_t^{\mathrm{T}}} > 0$ and $\frac{\partial \lambda_t^{\mathrm{T}}}{\partial \delta} = -\frac{\partial G/\partial \delta}{\partial G/\partial \lambda_t^{\mathrm{T}}} > 0$.

Proposition 1(iv) First, we show that λ_t^{T} is strictly convex in $\lambda_{t-1}^{\mathrm{T}}$, $\frac{\partial^2 \lambda_t^{\mathrm{T}}}{\partial (\lambda_{t-1}^{\mathrm{T}})^2} > 0$. Applying the implicit function theorem,

$$\frac{\partial^2 \lambda_t^{\mathrm{T}}}{\partial (\lambda_{t-1}^{\mathrm{T}})^2} = -\frac{1}{(\partial G/\partial \lambda_t^{\mathrm{T}})^3} \times \left[\frac{\partial G}{\partial \lambda_{t-1}^{\mathrm{T}}}\left(\frac{\partial G}{\partial \lambda_t^{\mathrm{T}}}\right)^2 - 2\frac{\partial^2 G}{\partial \lambda_t^{\mathrm{T}}\partial \lambda_{t-1}^{\mathrm{T}}}\frac{\partial G}{\partial \lambda_{t-1}^{\mathrm{T}}}\frac{\partial G}{\partial \lambda_t^{\mathrm{T}}} + \frac{\partial^2 G}{\partial (\lambda_t^{\mathrm{T}})^2}\left(\frac{\partial G}{\partial \lambda_{t-1}^{\mathrm{T}}}\right)^2\right]. \tag{C.10}$$

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

$$\frac{\partial^2 G}{\partial(\lambda_{t-1}^{\mathrm{T}})^2} = A(\lambda_t^{\mathrm{T}})^{2-\sigma}\frac{(\eta^{\sigma-1}-1)}{\sigma F^{\mathrm{T}}} f(\lambda_{t-1}^{\mathrm{T}})\delta^2 > 0. \tag{C.11}$$

$$\frac{\partial^2 G}{\partial\lambda_t^{\mathrm{T}}\partial\lambda_{t-1}^{\mathrm{T}}} = \frac{\partial^2 G}{\partial\lambda_{t-1}^{\mathrm{T}}\partial\lambda_t^{\mathrm{T}}} = \frac{2-\sigma}{\sigma-1}A(\lambda_t^{\mathrm{T}})^{3-2\sigma}\left[(\eta^{\sigma-1}-1)(\lambda_t^{\mathrm{T}})^{-\frac{\sigma-1}{\theta}}\right] f(\lambda_{t-1}^{\mathrm{T}})\delta < 0. \tag{C.12}$$

$$\begin{aligned}\frac{\partial^2 G}{\partial(\lambda_t^{\mathrm{T}})^2} = {} & \frac{(2-\sigma)(3-\sigma)}{(\sigma-1)^2}A(\lambda_t^{\mathrm{T}})^{4-3\sigma}\left[\frac{\tilde{\theta}}{\theta}(\lambda_t^{\mathrm{T}})^{-\frac{\sigma-1}{\theta}}(\eta^{\sigma-1}-1)\right]^2 f(\lambda_{t-1}^{\mathrm{T}})\frac{(\eta^{\sigma-1}-1)}{\sigma F^{\mathrm{T}}} \\ & + \frac{\sigma-2}{\theta}A(\lambda_t^{\mathrm{T}})^{3-2\sigma}(\lambda_t^{\mathrm{T}})^{-\frac{\sigma-1}{\theta}-1} f(\lambda_{t-1}^{\mathrm{T}})\frac{(\eta^{\sigma-1}-1)^2}{\sigma F^{\mathrm{T}}} + \frac{\sigma-1}{\theta}\frac{\tilde{\theta}}{\theta}(\lambda_t^{\mathrm{T}})^{-\frac{\tilde{\theta}}{\theta}-1} > 0,\end{aligned} \tag{C.13}$$

where each term of the right hand side is positive due to the assumption $\sigma > 3$. Substituting Equations (C.6), (C.7), (C.11), (C.12), and (C.13) in Equation (C.10), we obtain $\frac{\partial^2\lambda_t^{\mathrm{T}}}{\partial(\lambda_{t-1}^{\mathrm{T}})^2} > 0$, 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 $\lambda_{t-1}^{\mathrm{T}}$, the locus defined by $(\lambda_{t-1}^{\mathrm{T}}, \lambda_t^{\mathrm{T}})$ that satisfies Equation (4.3) can intersect with the 45-degree line two times at most. Note that the intercept is always positive because of the assumption of unbounded Pareto distribution which always guarantees a positive share of adopters.

Because λ_t^{T} strictly increases in δ, there exists $\underline{\delta}$ such that the 45-degree line and the short-run locus meet at $\lambda_{t-1}^{\mathrm{T}} = 1$. In other words, holding other parameters constant including η, $\underline{\delta}$ satisfies $A(1;\eta)^{2-\sigma}f(1;\underline{\delta})\frac{(\eta^{\sigma-1}-1)}{\sigma F^{\mathrm{T}}} - 1 = 0$. Similarly, there exists $\underline{\eta}$ that satisfies $A(1;\underline{\eta})^{2-\sigma}f(1;\delta)\frac{(\underline{\eta}^{\sigma-1}-1)}{\sigma F^{\mathrm{T}}} - 1 = 0$. Also, because λ_t^{T} is strictly convex in $\lambda_{t-1}^{\mathrm{T}}$, there exists $\bar{\delta}$ and $\bar{\eta}$ such that the 45-degree line is tangent to the short-run locus implicitly defined by Equation (C.5). In other words, $\bar{\delta}$ and $\bar{\eta}$ satisfy $A(\lambda^{\mathrm{T}};\eta)^{2-\sigma}f(\lambda^{\mathrm{T}};\bar{\delta})\frac{(\eta^{\sigma-1}-1)}{\sigma F^{\mathrm{T}}} - \lambda^{\mathrm{T}} = 0$ and $A(\lambda^{\mathrm{T}};\bar{\eta})^{2-\sigma}f(\lambda^{\mathrm{T}};\delta)\frac{(\bar{\eta}^{\sigma-1}-1)}{\sigma F^{\mathrm{T}}} - \lambda^{\mathrm{T}} = 0$, respectively, where $\lambda^{\mathrm{T}} = \lambda_{t-1}^{\mathrm{T}} = \lambda_t^{\mathrm{T}}$.

For $\delta \in [0, \underline{\delta})$ or $\eta \in [0, \underline{\eta})$, the equilibrium share is always below one and the short-run locus implicitly defined by Equation(4.3) intersects with the 45-degree line only once. For $\delta \in (\bar{\delta}), 1]$ or $\eta \in (\bar{\eta}), 1]$, the short-run locus intersect with the 45-degree line at $\lambda^{\mathrm{T}*} = \lambda_t^{\mathrm{T}*} = \lambda_{t-1}^{\mathrm{T}*} = 1$. For $\delta \in (\underline{\delta}, \bar{\delta})$ or $\eta \in (\underline{\eta}, \bar{\eta})$, the short-run locus and the 45-degree line intersect three times, leading to three multiple steady states. At the boundary values, the short-run locus intersects with the 45-degree line twice, leading to two multiple steady states.

Proposition 1(v) The welfare of household is $\frac{w_t+\Pi_t}{P_t}$ where Π_t are the aggregate profits summed across all firms in the economy. Using Equations (C.2) and (C.3) and the following expression

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

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

Proposition 2 Suppose an economy features multiple steady states S^{Pre}, S^{U}, and S^{Ind} and is initially stuck in the poverty trap.

First, consider input subsidies. Firms' costs of production are $(1-s_{it})w_t l_{it}$ where $s_{it} = \bar{s}_t$ for $T_{it} = 1$ and 0 otherwise. $0 < \bar{s}_t < 1$ is the subsidy rate for adopters. Firm charges price $p(\omega) = \frac{\mu(1-s(\omega))w}{z(\omega)}$. The cutoff is

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

$Q_t = A(\lambda_t^{\mathrm{T}})f(\lambda_{t-1}^{\mathrm{T}})$ still holds with subsidies, but the expression for $\frac{w_t}{P_t}$ gets slightly modified

$$\frac{w_t}{P_t} = \frac{1}{\mu}\tilde{A}(\lambda_t^{\mathrm{T}}, \bar{s}_t) f(\lambda_{t-1}^{\mathrm{T}}), \quad \text{where} \quad \tilde{A}(\lambda_t^{\mathrm{T}}, \bar{s}_t) = \left[\frac{\theta}{\tilde{\theta}}\left(\left(\left(\frac{\eta}{1-\bar{s}_t}\right)^{\sigma-1} - 1\right)(\lambda_t^{\mathrm{T}})^{\frac{\tilde{\theta}}{\theta}} + 1\right)\right]^{\frac{1}{\sigma-1}}.$$

The equilibrium share of adopters can be expressed as

$$\lambda_t^{\mathrm{T}} = \left[\frac{\left(\frac{\eta}{1-\bar{s}_t}\right)^{\sigma-1} - 1}{\sigma F^{\mathrm{T}}} A(\lambda_t^{\mathrm{T}})\tilde{A}(\lambda_t^{\mathrm{T}}, \bar{s}_t)^{1-\sigma} f(\lambda_{t-1}^{\mathrm{T}})\right]^{\frac{\theta}{\sigma-1}}. \tag{C.14}$$

Similarly with the subsidies to the fixed adoption costs $(1-\bar{s}_t)P_t F^{\mathrm{T}}$, the cutoff becomes

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

The equilibrium shares of adopters are

$$\lambda_t^{\mathrm{T}} = \left[\frac{\eta^{\sigma-1}-1}{\sigma(1-\bar{s}_t)F^{\mathrm{T}}} A(\lambda_t^{\mathrm{T}})^{2-\sigma} f(\lambda_{t-1}^{\mathrm{T}})\right]^{\frac{\theta}{\sigma-1}}. \tag{C.15}$$

In the cases of both subsidies, the right hand side of both Equations (C.14) and (C.15) strictly increases in $\bar{s}_t$, and $\lim_{\bar{s}_t \to 0} \lambda_t^{\mathrm{T}} \to 1$. Therefore, there exists $\underline{s}$ such that $\lambda_t^{\mathrm{T}} = S^{\mathrm{U}}$. For $\bar{s}_t > \underline{s}$, $\lambda_t^{\mathrm{T}} > S^{\mathrm{U}}$ and the economy starts to converge to S^{Ind}.

C.3 Source of Dynamic Complementarity

Let L_t denote total labor endowment that can be interpreted as market size. We show that when fixed adoption costs are in units of labor, regardless of L_t, the model does not feature dynamic complementarity. When fixed adoption costs are in units of labor, the cutoff and the equilibrium shares are

$$(\bar{\phi}_t^{\mathrm{T}})^{\sigma-1} = \frac{\sigma F^{\mathrm{T}}}{(\eta^{\sigma-1}-1)\mu^{1-\sigma} f(\lambda_{t-1}^{\mathrm{T}})^{\sigma-1} P_t^{\sigma} Q_t}, \qquad \lambda_t^{\mathrm{T}} = \Big(\frac{\mu(\eta^{\sigma-1}-1)L_t}{\sigma F^{\mathrm{T}}} A(\lambda_t^{\mathrm{T}})^{1-\sigma}\Big)^{\frac{\theta}{\sigma-1}}.$$

Although larger market size L_t increases the equilibrium share due to the scale complementarity, the share is uniquely determined regardless of values of $\lambda_{t-1}^{\mathrm{T}}$. A higher adopter share in $t-1$ increases overall productivity in t through the spillover, which in turn leads to higher demand for labor. Such higher demand increases the equilibrium wage, which makes adoption costs $w_t F^{\mathrm{T}}$ higher. These increased costs exactly cancel out larger incentives for the adoption induced by the spillover.

C.4 Possible Microfoundations for Adoption Spillovers

We provide two possible microfoundations for the spillovers. For both cases, we consider a simple closed economy setup with one sector and N regions. Goods are freely tradable and labor is freely mobile across regions, so wage and price indices are equalized across regions.

Local diffusion of knowledge A firm receives exogenous productivity $\tilde{\phi}_{it}$ and makes two static decisions each period, whether to adopt advanced foreign technology T_{it} and a level of innovation a_{it} as in Desmet and Rossi-Hansberg (2014). Their profit maximization problem is

$$\pi_{it} = \max_{T_{it}\in\{0,1\}, a_{it}\in[0,\infty)} \left\{ \frac{1}{\sigma}\left(\frac{\mu w_t}{\tilde{\eta}^{T_{it}} a_{it}^{\gamma_1} \tilde{\phi}_{it}}\right)^{1-\sigma} P_t^{\sigma} Q_t - T_{it} P_t F^{\mathrm{T}} - w_t a_{it}^{\alpha_1} g(\lambda_{n,t-1}^{\mathrm{T}}) P_t^{\sigma} Q_t \right\}, \tag{C.16}$$

where $\tilde{\eta}$ governs direct productivity gains from adoption, and $a_{it}^{\alpha_1} g(\lambda_{n,t-1}^{\mathrm{T}}) P_t^{\sigma} Q_t$ 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 the cost of innovation is proportional to market size $P_t^{\sigma} Q_t$ and normalize $w_t = 1$ without loss of generality.

The positive externalities of adoption come from that the innovation costs are decreasing in the previous adopter shares $\partial g(\lambda_{n,t-1}^{\mathrm{T}})/\partial \lambda_{n,t-1}^{\mathrm{T}} < 0$, which captures that with higher adopter shares, other local firms are more likely to learn new ideas from these adopters and use this knowledge for their own innovation in a reduced-form. We impose that $\tilde{\alpha} = \alpha_1 - \gamma_1(\sigma-1) > 0$ holds, which guarantees the second-order condition of a firm's maximization problem.

Using the first-order condition, a firm's optimal level of a_{it} is characterized as

$$a_{it}^{*} = (\frac{\gamma_1}{\alpha_1}\mu^{-\sigma})^{\frac{1}{\tilde{\alpha}}} g(\lambda_{n,t-1}^{\mathrm{T}})^{-\frac{1}{\tilde{\alpha}}} (\tilde{\eta}^{T_{it}} \tilde{\phi}_{it})^{\frac{\sigma-1}{\tilde{\alpha}}}.$$

Because $-1/\tilde{\alpha} > 0$ and $(\sigma-1)/\tilde{\alpha} > 0$, a_{it}^{*} increases in $\lambda_{n,t-1}^{\mathrm{T}}$, T_{it}, and $\tilde{\phi}_{it}$. Substituting a_{it}^{*} into

Equation (C.16), a firm's maximization problem can be re-written as:

$$\pi_{it} = \max_{T_{it}\in\{0,1\}} \left\{ \bar{C} \left(\frac{1}{g(\lambda^{\mathrm{T}}_{n,t-1})^{-\frac{\gamma_1}{\tilde{\alpha}}} (\tilde{\eta}^{\frac{\alpha_1}{\tilde{\alpha}}})^{T_{it}} (\tilde{\phi}_{it})^{\frac{\alpha_1}{\tilde{\alpha}}}} \right)^{1-\sigma} P^{\sigma}_t Q_t - T_{it} P_t F^{\mathrm{T}} \right\},$$

where $\bar{C}$ is a collection of model parameters. $g(\lambda^{\mathrm{T}}_{n,t-1})^{-\frac{1}{\tilde{\alpha}}}$ can be mapped to $f(\lambda^{\mathrm{T}}_{n,t-1})$, $(\tilde{\phi}_{it})^{\frac{\alpha_1}{\tilde{\alpha}}}$ to ϕ_{it}, and $\tilde{\eta}^{\frac{\gamma_1}{\alpha_1-1-\gamma_1(\sigma-1)}}$ to η.

Learning externalities and labor mobility In each region, there is a unit measure of engineers and owners of firms, both of which are immobile across regions. Engineers live in two periods, child and adult. Once they become adults in the second period, they give birth to a child. They only consume and work in their adulthood. Engineers who work in firms that adopted technologies pass their knowledge to their children. This learning from parents increases the engineering skills of children when they grow up, which increases their skills by $\gamma_1 > 1$. If parents do not work in firms with foreign technology, their children's engineering skills are 1.

Engineers and owners are randomly matched one to one (Acemoglu, 1996). After a match, production happens and they jointly maximize profits. The profits this match generates are divided among engineers and owners based on Nash bargaining. Managers take a proportion of $\tilde{\beta}$. Because owners make adoption decisions before a match happens, they must make these decisions based on anticipated profits. Because of the random matching process, owners are matched with high- and low-skilled engineers with a probability of $\lambda^{\mathrm{T}}_{n,t-1}$ and $1-\lambda^{\mathrm{T}}_{n,t-1}$, respectively.

A firm's maximization problem is

$$\pi_{it} = \max_{T_{it}\in\{0,1\}} (1-\tilde{\beta}) \left\{ \lambda^{\mathrm{T}}_{n,t-1} \left[\frac{1}{\sigma} \left(\frac{\mu w_t}{\eta^{T_{it}} \gamma_1 \phi_{it}} \right)^{1-\sigma} P^{\sigma}_t Q_t \right] + (1-\lambda^{\mathrm{T}}_{n,t-1}) \left[\frac{1}{\sigma} \left(\frac{\mu w_t}{\eta^{T_{it}} \phi_{it}} \right)^{1-\sigma} P^{\sigma}_t Q_t \right] - P_t F^{\mathrm{T}} T_{it} \right\}.$$

This can be re-written as

$$\pi_{it} = \max_{T_{it}\in\{0,1\}} (1-\tilde{\beta}) \left\{ \frac{1}{\sigma} \left(\frac{\mu w_t}{\tilde{f}(\lambda^{\mathrm{T}}_{n,t-1}) \eta^{T_{it}} \phi_{it}} \right)^{1-\sigma} P^{\sigma}_t Q_t - P_t F^{\mathrm{T}} T_{it} \right\},$$

where $\tilde{f}(\lambda^{\mathrm{T}}_{n,t-1}) = [\lambda^{\mathrm{T}}_{n,t-1}(\gamma_1^{\sigma-1} - 1) + 1]^{\frac{1}{\sigma-1}}$.

Appendix D Quantitative Model

Sector A final goods producer aggregate varieties using a CES aggregator:

$$Q_{njt} = \left[\sum_m \int_{\omega\in\Omega_{mj}} (q_{njt}(\omega))^{\frac{\sigma-1}{\sigma}} \mathrm{d}\omega + (q^f_{njt})^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}},$$

where $q_{njt}(\omega)$ and q^f_{njt} are the quantities demanded of a variety produced by a domestic and a foreign firms, respectively. The exact price index is given by Equation (5.1)

Firm With CRS Cobb-Douglas production, unit costs of input bundles are

$$c_{njt} = (w_{nt}/\gamma_j^L)^{\gamma_j^L} \prod_k (P_{nkt}/\gamma_j^k)^{\gamma_j^k}.$$

Firm i's quantities demanded from region m are Foreign are $q_{inmjt} = (p_{inmjt})^{-\sigma} P_{mjt}^{\sigma-1} E_{mjt}$ and $q^x_{injt} = (p^x_{injt})^{-\sigma} D^x_{jt}$. A firm optimally charges a constant markup over its marginal cost. Thus, the prices charged by firm i in region n of sector j charged to buyers in region m are $p_{inmjt} = \mu\tau_{nmj}c_{njt}/z_{it}$ and export prices are $p^x_{injt} = \mu\tau^x_{nj}c_{njt}/z_{it}$.

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

$$\begin{aligned} \pi_{it} = \pi(\phi_{it}) &= \max_{x_{it}, T_{it}\in\{0,1\}} \{\pi(T_{it}, x_{it}; \phi_{it})\} \\ &= \max_{x_{it}, T_{it}\in\{0,1\}} \Bigg\{ \underbrace{\sum_m \left[\frac{1}{\sigma} \left(\mu \frac{\tau_{nmj}(1-s_{njt})^{T_{it}} c_{njt}}{\phi_{it}\eta^{T_{it}} f(\lambda^{\mathrm{T}}_{nj,t-1})} \right)^{1-\sigma} P_{mjt}^{\sigma-1} E_{mjt} \right]}_{:=\pi^{\mathrm{d}}(T_{it};\phi_{it}) = \sum_m \pi^m(T_{it};\phi_{it})} \\ &\qquad + x_{it} \Bigg[\underbrace{\frac{1}{\sigma} \left(\mu \frac{\tau^x_{nj}(1-s_{njt})^{T_{it}} c_{njt}}{\phi_{it}\eta^{T_{it}} f(\lambda^{\mathrm{T}}_{nj,t-1})} \right)^{1-\sigma} D^x_{jt}}_{:=\pi^x(T_{it};\phi_{it})} - w_{nt}F^x_j \Bigg] - T_{it}c_{njt}F^{\mathrm{T}}_j \Bigg\}, \end{aligned} \tag{D.1}$$

where x_{it} and T_{it} are binary export and adoption decisions, E_{mjt} are region m's total expenditures on sector j goods, D^x_{jt} are exogenous foreign demands, and F^x_j are fixed export costs in units of labor F^x_j. $\pi^m(T_{it};\phi_{it})$ are operating profits conditional on adoption status obtained from region m, and $\pi^{\mathrm{d}}(T_{it};\phi_{it}) = \sum_m \pi^m(T_{it};\phi_{it})$ 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.

Firms' adoption and export decisions are characterized by the cutoff productivities. Only firms with productivity above these cutoffs participate in adoption and exporting. To avoid a taxonomic presentation, we only consider a case in which fixed adoption costs are high enough so that the adoption cutoff is higher than the export cutoff in all regions. In the quantitative analysis, we allow for other possibilities.

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

$$\bar{\phi}^x_{njt} = \frac{\mu c_{njt}(\sigma w_{nt}F^x_j)^{\frac{1}{\sigma-1}}}{f(\lambda^{\mathrm{T}}_{nj,t-1})\big((\tau^x_{nj})^{1-\sigma}D^x_{jt}\big)^{\frac{1}{\sigma-1}}}. \tag{D.2}$$

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

$$\bar{\phi}_{njt}^{\mathrm{T}} = \frac{\mu c_{njt}(\sigma c_{njt} F_{nj}^{\mathrm{T}})^{\frac{1}{\sigma-1}}}{\left(\left(\frac{\eta}{1-s_{njt}}\right)^{\sigma-1} - 1\right)^{\frac{1}{\sigma-1}} f(\lambda_{nj,t-1}^{\mathrm{T}})\left(\sum_m \tau_{nmj}^{1-\sigma} P_{mjt}^{\sigma-1} E_{mjt} + (\tau_{nj}^x)^{1-\sigma} D_{jt}^x\right)^{\frac{1}{\sigma-1}}}. \tag{D.3}$$

A share of adopters is expressed as

$$\lambda_{njt}^{\mathrm{T}} = 1 - G_{njt}(\bar{\phi}_{njt}^{\mathrm{T}}) = \begin{cases} 1 & \text{if} \quad \bar{\phi}_{njt}^{\mathrm{T}} \leq \phi_{njt}^{\min} \\ \frac{(\bar{\phi}_{njt}^{\mathrm{T}}/\phi_{njt}^{\min})^{-\theta} - \kappa^{-\theta}}{1-\kappa^{-\theta}} & \text{if} \quad \phi_{njt}^{\min} < \bar{\phi}_{njt}^{\mathrm{T}} \leq \kappa\phi_{njt}^{\min} \\ 0 & \text{if} \quad \kappa\phi_{njt}^{\min} \leq \bar{\phi}_{njt}^{\mathrm{T}}, \end{cases} \tag{D.4}$$

where $G_{njt}(\phi)$ is productivity distribution of nj in t. A mass of adopters is $M_{njt}^{\mathrm{T}} = M_{nj}\lambda_{njt}^{\mathrm{T}}$. Similarly, a share of exporters is $\lambda_{njt}^x = 1 - G_{njt}(\bar{\phi}_{njt}^x)$ and a mass of exporters is $M_{njt}^x = M_{nj}\lambda_{njt}^x$.

Region-sector variables We define the region-sector level average firm productivity inclusive of subsidies as

$$\begin{aligned}
\bar{\phi}_{njt}^{\mathrm{avg}} &= f(\lambda_{nj,t-1}^{\mathrm{T}})\left[\int_{\phi_{njt}^{\min}}^{\bar{\phi}_{njt}^{\mathrm{T}}} \phi_{it}^{\sigma-1} dG_{njt}(\phi_{it}) + \int_{\bar{\phi}_{njt}^{\mathrm{T}}}^{\kappa\phi_{njt}^{\min}} \left(\frac{\eta}{1-s_{njt}}\phi_{it}\right)^{\sigma-1} dG_{njt}(\phi_{it})\right]^{\frac{1}{\sigma-1}} \\
&= \frac{\theta f(\lambda_{njt-1}^{T})(\phi_{njt}^{\min})^{\sigma-1}}{\tilde{\theta}(1-\kappa^{-\theta})}\left\{\left(\left(\frac{\eta}{1-s_{njt}}\right)^{\sigma-1} - 1\right)((1-\kappa^{-\theta})\lambda_{njt}^{\mathrm{T}} + \kappa^{-\theta})^{\frac{\tilde{\theta}}{\theta}} + \left(1 - \left(\frac{\eta}{1-s_{njt}}\right)^{\sigma-1}\kappa^{-\tilde{\theta}}\right)\right\}.
\end{aligned}$$

$\bar{\phi}_{njt}^{\mathrm{avg}}$ captures the average cost advantage of sector j firms in region n. $\bar{\phi}_{njt}^{\mathrm{avg}}$ decreases in $\bar{\phi}_{njt}^{\mathrm{T}}$ but increase in s_{njt} and $\lambda_{nj,t-1}^{\mathrm{T}}$. The average productivity for exporters can be expressed similarly

$$\bar{\phi}_{njt}^{\mathrm{avg,x}} = \frac{\theta f(\lambda_{njt-1}^{T})(\phi_{njt}^{x})^{\sigma-1}}{\tilde{\theta}(1-\kappa^{-\theta})}\left\{\left(\left(\frac{\eta}{1-s_{njt}}\right)^{\sigma-1} - 1\right)((1-\kappa^{-\theta})\lambda_{njt}^{\mathrm{T}} + \kappa^{-\theta})^{\frac{\tilde{\theta}}{\theta}} + \left(1 - \left(\frac{\eta}{1-s_{njt}}\right)^{\sigma-1}\kappa^{-\tilde{\theta}}\right)\right\}.$$

Aggregate variables can be expressed as a function of $\bar{\phi}_{njt}^{\mathrm{avg}}$ and $\bar{\phi}_{njt}^{\mathrm{avg,x}}$. The price index is

$$P_{njt}^{1-\sigma} = \sum_m \left[M_{mj}\left(\frac{\mu\tau_{mnj}c_{mjt}}{\bar{\phi}_{mjt}^{\mathrm{avg}}}\right)^{1-\sigma}\right] + (\tau_{nj}^x P_{jt}^f)^{1-\sigma}.$$

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

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

Regional gross output for domestic expenditures R^{d}_{njt} and the total value of exports R^{x}_{njt} are

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

The total regional gross output is $R_{njt} = R^{\mathrm{d}}_{njt} + R^{x}_{njt}$.

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

$$w_{nt}L_{nt} = \left[\sum_{j}\gamma^{L}_{j}\left(\frac{1}{\mu}R_{njt} + M^{\mathrm{T}}_{njt}c_{njt}F^{\mathrm{T}}_{nj}\right) + M^{x}_{njt}w_{nt}F^{x}_{j}\right], \tag{D.5}$$

where the right-hand side is the sum of labor used for production, fixed adoption costs, and fixed export costs. Goods market clearing implies

$$R^{\mathrm{d}}_{njt} = \sum_{m}\pi_{nmjt}(\alpha_j w_{nt}L_{nt} + \gamma^{j}_{k}\frac{1}{\mu}R_{nkt} + M^{\mathrm{T}}_{njt}c_{njt}F^{\mathrm{T}}). \tag{D.6}$$

The government budget is balanced each period:

$$\sum_{j}\frac{t^{x}_{jt}}{1+t^{x}_{jt}}\pi^{f}_{jt}E_{jt} + \tau^{w}_{t}\sum_{n}w_{nt}L_{nt} = \sum_{n}\sum_{j}\left[\frac{s_{njt}}{1-s_{njt}}M_{nj}\int_{\bar{\phi}^{\mathrm{T}}_{njt}}^{\kappa\phi^{\mathrm{min}}_{njt}}\frac{1}{\mu}r(\phi_{it})dG_{njt}(\phi)\right], \tag{D.7}$$

where the left-hand side is sum of revenues from tariffs and labor tax.

Definition 1. *Given initial conditions $\{\lambda^{T}_{njt_0}, L_{nt_0}\}$ and a path of the fundamentals $\{\phi^{min}_{njt}, V_{nt}, P^{f}_{jt}, D^{x}_{jt}\}$, subsidies $\{s_{njt}\}$, and tariffs $\{t^{x}_{jt}\}$, an equilibrium is a path of wages $\{w_{nt}\}$, price indices $\{P_{njt}\}$, a set of functions $\{p_{inmjt}(\omega), q_{inmjt}(\omega), p^{x}_{injt}(\omega), q^{x}_{injt}(\omega), T_{it}(\omega), x_{it}(\omega)\}$, labor tax $\{\tau^{w}_{t}\}$, population $\{L_{nt}\}$, and shares of adopters $\{\lambda^{T}_{njt}\}$ such that for each period t, (i) firms maximize profits; (ii) households maximize utility; (iii) labor markets clear; (iv) goods markets clear; (v) trade is balanced, and (vi) the government budget is balanced; and (vii) firms' adoption and households' migration decisions endogenously determine the path of state variables λ_{njt} and L_{nt}, respectively.*

Appendix E Quantification

E.1 Calibration Procedure

Data inputs The quantitative exercises require the following data inputs:

1. Initial adopter shares $\{\lambda^{\mathrm{T}}_{nj,68}\}_{n\in\mathcal{N},j\in\mathcal{J}^{\mathrm{T}}}$ and population $\{L^{\mathrm{Data}}_{n,68}\}_{n\in\mathcal{N}}$ in 1968
2. Region-sector gross output $\{R^{\mathrm{Data}}_{njt}\}_{n\in\mathcal{N},j\in\mathcal{J},t\in\{72,76,80\}}$
3. Population: $\{L^{\mathrm{Data}}_{nt}\}_{n\in\mathcal{N},t\in\{72,76,80\}}$
4. Sectoral exports and import shares $\{EX^{\mathrm{Data}}_{jt}, \pi^{f,\mathrm{Data}}_{jt}\}_{j\in\mathcal{J},t\in\{72,76,80\}}$
5. Import tariffs $\{t^{x}_{jt}\}_{j\in\mathcal{J},t\in\{72,76,80\}}$

6. Moments: Median of light and heavy mfg. shares of exports in 1972 across regions, Median of heavy mfg. shares of adopters in 1972 and 1982 across regions, Percent of zero adoption regions in 1972 and 1982

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

1. Guess parameters.
2. Guess fundamentals $\{c_{fj}, D_{fj}\}_{j\in\mathcal{J}}$, $\{V_{nt}\}_{n\in\mathcal{N}}$, and $\{\phi_{nj}^{\min}\}_{n\in\mathcal{N},j\in\mathcal{J}}$
3. Given parameters $\{\mathbf{\Theta}^{\mathrm{M}}, \bar{s}_t\}$, we solve the model and update the fundamentals $\mathbf{\Psi}_t$ for each period. Then, we fit region- and sector-level aggregate outcomes to the data counterparts. This step corresponds to the constraints of Equation (5.4).
   (a) Update new $\{D_{jt}^{f'}\}$ by fitting export intensities of the model to those in the data.
   (b) Update new $\{P_{jt}^{F'}\}$ by fitting import shares of the model to those in the data.
   (c) Update new $\{V'_{nt}\}$ until the population outcome of the model fits the actual distribution of the population. Only relative levels of $\{V'_{nt}\}$ are identified from the above equation, so we normalize the value of the amenity of the reference region n_0 to be 1 for each period.
   (d) Update new $\{\phi_{nj}^{min'}\}$ until shares of regional gross output are exactly fitted to the data counterparts $GO_{njt}^{\mathrm{Data}}/\sum_{m,k}GO_{mkt}^{\mathrm{Data}}$. Within each sector, $GO_{njt}^{\mathrm{Data}}/\sum_{m,k}GO_{mkt}^{\mathrm{Data}}$ only identifies the relative levels, so we normalize the Pareto lower bound parameter of the reference region to 1 for each sector and period.
4. After updating the geographic fundamentals, given values of parameters and subsidies, we evaluate the following objective function.
5. We iterate steps 1-4 until we find values of $\{\hat{\mathbf{\Theta}}^{\mathrm{M}}, \hat{\mathbf{s}}_t\}$ that minimize the objective function.