Innovation Spurred: Evidence from South Korea's Big R&D Push*

Luis F. Jaramillo†
Job Market Paper
University of Maryland

Chan Kim‡
University of Maryland

Click here for the Latest Version of the Paper

October 29, 2024

Abstract

We study how South Korea's first "mission-oriented" R&D program, implemented between 1992 and 2001, shaped innovation and economic outcomes. Using new textual data and a language model to identify targeted and control technological classes, we exploit the fact that some of the planned research projects were not implemented because of budget shocks. We use a local projections event study to compare the outcomes of targeted technological classes to those of control classes. Despite the absence of differential trends before the program, by ten years after the extension of program support, future-citation-weighed patenting output in the targeted classes doubled and real exports tripled relative to the control technology classes. These results stand when we study cross-country evidence. Technological classes with less concentrated patenting output before the program drive our results. Using market-based patent valuations, we find that the program's benefits exceeded its costs by over a factor of three. Our findings suggest that technology policy was central to South Korea's transition to a knowledge-intensive economy.

JEL: F14, O25, O32, O33.

Keywords: Technology Policy, Innovation, South Korea, Mission-Oriented Programs

*We are grateful to Sebastián F. Galiani, Nuno Limão, and John Wallis for their guidance, encouragement, and thoughtful suggestions. Jaedo Choi, Xiang Ding, Sharat Ganapati, Penny Goldberg, Ginger Zhe Jin, Rodney Ludema, Pablo Querubín, Heitor Pellegrina, Charly Porcher, Steve Redding, Lumi Stevens, Mateo Uribe-Castro, Sergio Urzúa, Qiyao Zhou, and several participants at the 2024 Mid-Atlantic Trade Workshop, the 2024 Washington Area International Trade Symposium, the 2024 Summer Meetings of the European Economic Association, Georgetown University and NABETech 2024 provided useful comments that significantly improved the paper. We thank Nayeon Kwon for her outstanding help with data collection. The National Research Foundation of Korea also provided timely support in the data collection process.

†Corresponding author. lfj@umd.edu
‡chankim@umd.edu

1 Introduction

Industrial policy is back: Programs such as the US CHIPS Act and the European Green Deal have committed trillions to transforming their respective economies' structure and direction of innovation. Despite the theoretical arguments that rationalize interventions to address market failures, there is a lack of clarity on what works in practice. As Juhász et al. (2023) note, the question is increasingly not whether but how governments should conduct industrial policy.

We contribute to this debate by studying how South Korea's first "mission-oriented" R&D initiative, the G7 Program (G7P), shaped innovation and real outcomes. The G7P was active between 1992 and 2001 and invested over $7 billion (2023 dollars) to fund government-selected R&D projects. It was the first explicit and coordinated effort to create frontier technological development capacities in South Korea and close the gap with G7 nations in selected technologies by the 2000s. The program responded to South Korea's waning catch-up strategy and rising labor costs following its return to democracy in 1987 (Ministry of Science and Technology, Republic of Korea, 1991).

The G7P aimed to solve market failures for different types of technologies. For "product" technologies, where South Korea had strengths (electronics, machines, materials), investments were too large and too risky for individual companies to pursue on their own. A targeted research subsidy and pooling mechanism might induce private firms to engage with R&D. For "base" technologies, which had environmental and national security externalities and where South Korea did not have strengths or expertise (energy, biotechnology), private firms were unlikely to provide optimal levels of R&D because of the difference between the private and social returns of these projects. A research subsidy would bridge that gap.

We exploit the government's selection of 23 megaprojects, of which only 18 were funded to address selection concerns. Despite being deemed high potential by program experts, the remaining five were not funded because of budget and program-fit considerations (KISTEP, 2002). Our treated (control) group includes technological classes related to the 18 (5) implemented (not funded) megaprojects. We use newly digitized files with detailed information on the universe of G7P-supported research projects and use a language model to link them to technological classes, the unit at which we observe the outcomes we study.

We find that the G7P shifted the direction of South Korea's innovation. The G7P substantially increased future-citation-weighed patenting output in the targeted technological classes relative to that in control classes. By the 5th year after receipt of program support, the number of patents granted in the targeted classes had increased by 64%. This effect had risen to 123% by the 10th year and to 232% by the 15th year. For both the targeted and control groups, this outcome had followed similar trends in the five years before the G7P support was extended. Moreover, the G7P targeting was unrelated to underlying economic characteristics that might have influenced sectoral choices, such as value added, output per worker, and capital intensity.

The changes in innovation had significant consequences for the real economy, although the emergence of the effects was less immediate than in the case of patenting. We focus on exports, widely acknowledged as a factor in South Korea's economic success. We find null effects on exports for the first three years after targeting. By the 5th year, however, exports in G7P-targeted technological classes had grown by 62% in comparison to those in control classes, with this figure increasing to 245% by the 10th year and 204% by the 15th year. There were no differential trends in exports before targeting. Our findings highlight that R&D programs may take time to yield tangible benefits in the real economy.

Technological classes with less concentrated scientific output drive our results. We show that classes with less concentrated citation shares in the decade leading up to the G7P experienced stronger program effects. A move from the 25th to the 75th percentile on the Hirschman Herfindahl index (HHI) for citation shares corresponds to a reduction in the program's impact of approximately three-quarters of the baseline effect. This differential impact by citation concentration suggests that preexisting structures in technological classes, which influence spillovers through knowledge networks, had a role in determining the program's effectiveness. It also underscores the need for caution with respect to policies focused solely on creating "national champions" as they may limit knowledge spillovers.

Policymakers often frame industrial policy in the context of strategic competition between countries, with some policies responding to rival programs in other places. We explore this dimension and validate our within-country findings by comparing South Korea's performance in targeted technological classes with that of other countries, which we use as placebos. Despite the trends across countries being similar before the targeting, South Korea's future-citation-weighted patents and exports in targeted classes grew significantly faster than those of other countries after the G7P. As in our within-country exercise, exports show a delayed response. Overall, however, South Korea outperformed other nations in G7P-supported technological classes, indicating that the G7P accomplished its primary goal.

Government subsidies are not a free lunch. Though substantial, our previous findings are not informative about the subsidies' cost-effectiveness. If the identified effects came at a high cost, policymakers might question the program's desirability. We use the program's R&D expenditures, our reduced-form estimates on patenting, and the method of Kogan et al. (2017) for valuing patents from stock market reactions to patent-granting, and we find that the program yielded a 21% internal rate of return and benefits 3.3 times its costs. The G7P program was a cost-effective intervention.

A substantial data effort makes our analysis possible. We obtained and digitized data on approximately 4,800 G7P research projects from the National Research Foundation of Korea through a Transparency Law request. These files lack explicit data on the technological classes targeted by each research project, which are essential for us to study outcomes such as patents and exports since the data for those outcomes are available only at the technological class level. We address this gap using a text-based approach. We observe each research project's name, description, and objectives. This information allows us to use a language model to classify research projects into technological classes. The World International Patent Organization (WIPO) developed the model that we use, the IPCCAT tool, to classify inventions into technological classes using patent descriptions and abstracts as inputs.[1]

We download the universe of USPTO patents granted between 1980 and 2015 from USPTO's PatentView. Our sample period starts twelve years before the implementation of the first G7P project and fifteen years after the last one. Though data for later years are available, we set the starting year to 2015 to avoid right-truncation issues that arise from long patent application cycles. We focus on USPTO-granted patents to keep contextual elements such as relative market attractiveness and strength of property rights protection as fixed as possible. Moreover, inventors worldwide typically file important discoveries with the USPTO (Bloom et al., 2021).

We use export data covering 1980 to 2015 from the UN-COMTRADE database. While these data are not available at the IPC-code level, we use Lybbert and Zolas (2014)'s correspondence table between the Standard International Trade Classification (SITC) Rev. 2 and IPC codes for our analysis. Additionally, given the limited timeframe of our sample period, we use South Korea's Mining and Manufacturing Survey (MMS), which provides plant-level data from 1980 to 2003, to validate our identification strategy. To inform our cost-benefit analysis, we obtain South Korean firms' balance sheets and stock

[1]IPCCAT stands for International Patent Classification (IPC) computer-assisted categorization.

price movements for 1980–2015 from DataGuide, a database akin to COMPUSTAT.

2 Contributions to the Literature

Our work relates to several branches of the literature. First, a large body of research examines the causes and consequences of R&D investment (Romer, 1990; Aghion and Howitt, 1992; Howell, 2017; Acemoglu et al., 2018; Akcigit et al., 2020; Chen et al., 2021; Dechezleprêtre et al., 2023). An important part of this literature explores how policy regimes affect innovation (Bloom et al., 2019). Most studies focus on marginal funding changes or grants that incentivize R&D projects that firms choose in a decentralized manner. In contrast, we examine a large R&D program where a public organization centrally conceived and selected projects, making the G7P a representative "mission-oriented" program (Mazzucato, 2013; Kim, 2020; Gruber and Johnson, 2023).

Our paper is close to the works of Gross and Sampat (2023) and Kantor and Whalley (2023), who study other "mission-oriented" programs in crisis moments in the US (WWII and the race to the moon, respectively). However, our research differs in several ways. First, these programs focused on base technologies with unknown commercial applications. The G7P supported research in base and product technologies, aiming to address distinct market failures affecting each type of technology. Second, we study a program in a developing economy with a strong industrial base but limited innovation activities. Third, the G7P did not occur during a crisis, when stakes and incentives might differ. Our setting might be more informative for policymaking in more mundane times in developing countries.

Second, we contribute to a growing literature on industrial policy (Kalouptsidi, 2017; Juhász, 2018; Hanlon, 2018; Criscuolo et al., 2019; Giorcelli, 2019; Mitrunen, 2021; Choi and Shim, 2023a; Lane, 2023; Barwick et al., 2023). While most work focuses on technology adoption, our paper addresses technology development in a developing country. Though technological change in developing-country contexts is often seen as exogenous Gollin et al. (2002), Moscona and Sastry (2023) suggest that developing countries might invest in R&D because of the high productivity costs of inappropriate technology. Moreover, Choi and Shim (2023b) show that advanced nations may hesitate to transfer technology as receiving countries become their competitors. We also examine how concentration in a country's innovation structure affects industrial policy effectiveness. Together, our findings speak to the demand for knowledge of technology development in developing countries, a domain that remains fundamentally underinvestigated.

Third, we contribute to an extensive literature on the role of industrial policy in East Asia's economic miracles (Johnson, 1982; Wade, 1990; Amsden, 1992; Chang, 1993; Krueger, 1995; Rodrik et al., 1995; Noland and Pack, 2003; Choi and Shim, 2023a; Lane, 2023). This work focuses on interventions such as South Korea's Heavy and Chemical Industry (HCI) drive in the 1970s. These papers explain the country's industrial rise but not its transformation into a global innovation leader, a leap that many middle-income countries fail to make. Choi and Shim (2023b) study why countries transition from adoption subsidies to R&D subsidies. Our work complements theirs by documenting South Korea's shift in industrial strategy and providing microeconometric evidence on the effectiveness of the country's first "mission-oriented" R&D program. This policy shift was significant as prior attempts with traditional R&D tax credits had proved less successful (Kwon, 2021).

The rest of the paper is organized as follows: Section 3 provides historical context and institutional detail. Section 4 discusses our data collection process. Section 5 outlines our empirical strategy. Section 6 presents and discusses our results. Section 7 shows the Cost-Benefit analysis. Section 8 concludes.

3 Historical Context and the G7 Program

Though highly successful, South Korea's insertion into the global economy was not linear. The South Korean external sector has undergone several boom-and-bust cycles over the last five decades. One materialized in the late 1980s, with the abrupt end of the period of the so-called three lows: low oil price, low interest rates, and low (weak, relative to the Japanese yen) dollar. These circumstances enabled a rapid, debt-driven expansion during the second half of the 1980s. As these external conditions changed, the external sector took a hit: exports stagnated, with their share of GDP falling from 34.8% in 1987 to 23.8% in 1991 (World Bank, 2023).

Most economists agree that the crisis revealed structural weaknesses in South Korea's industrial development strategy, which favored (debt-driven) input-oriented expansion based on comparatively low labor costs (Kwon, 2021). The end of favorable external conditions and the increase in labor costs that followed the return to democracy exhausted the strategy's sources of competitive edge. Indeed, real labor pay rose by 53% between 1987 and 1989, far surpassing the growth rate of labor productivity. These conditions made competition in relatively low-value-added markets with other Asian countries much tougher. Without the advantage of low labor costs, South Korea's prospect of competing with advanced countries was grim given its relatively poor technology capacities (Ministry of Science and Technology, Republic of Korea, 1991).

South Korean policymakers identified the need to shift the nature of the markets in which South Korea competed abroad toward ones with higher value added. This necessity, paired with the increased reluctance of developed countries to share technology with South Korean firms (Choi and Shim, 2023b), justified the development of "indigenous" innovation and R&D capacities. Following the relative lack of success of earlier promotion policies (Kwon, 2021), which included an R&D tax credit, policymakers identified the need for a more coordinated and concentrated effort (Ministry of Science and Technology, Republic of Korea, 1991). The G7 Program, announced by President Roh Tae-wooh in November 1991, responded to this challenge.

3.1 The G7 Program[2]

Also known as the Highly Advanced National Program (HAN, like the river crossing Seoul), the G7P was South Korea's first national R&D program. It invested over $7 billion (2023 dollars) and mobilized over 100,000 research staff from 1992 to 2000 (Kwon, 2021). The program aimed to bring South Korean R&D capacity in select sectors to the level in G7 countries by the 2000s.

The G7P supported research projects looking to address problems in applied technology, not basic science. Over its course, the program supported 18 megaprojects of two types—projects in "product technologies" and in "base technologies"—nine in each category. Our empirical strategy exploits the fact that budget shocks and program-fit concerns reduced the number of supported megaprojects from 23 to 18. Each megaproject comprised smaller individual projects, for which we collected data, that we map to IPC technology classes for our regression analysis.

For product-technology megaprojects, policymakers' concern was that the private sector would not undertake such projects because they were too large and risky. Indeed, and with few exceptions, the South Korean private sector had been unwilling to engage in R&D and favored instead the continuation of previous input-driven strategies (Kwon, 2021). A government subsidy and a pooling mechanism that enabled the participation of several firms in a single research project would enhance the risk–reward profile of these investments. The distinguishing features of these projects were that they had immediate commercial applications and that South Korea already had capacities in the targeted sectors. Table 1 shows the nine megaprojects that fell within this category. Projects developing some familiar products such as HDTV (a next-generation flat panel display),

[2]This section relies heavily on KISTEP (2002) and KISTEP (2003).

a high-capacity semiconductor, and an electric vehicle appear here.

Regarding base-technology megaprojects, policymakers' main concern was that the private sector would not find such projects profitable because they lacked immediate commercial applications and there was little existing underlying capacity in the country. This phenomenon might lead to a typical underprovision of public goods because private agents do not incorporate the society-wide returns that a given invention brings. Policymakers identified the need to build capacity in these sectors, as they considered self-sufficiency vital for any advanced nation. Table 1 shows the nine megaprojects supported in this category. Projects advancing technologies with significant environmental and national security externalities, such as a next-generation nuclear reactor, are among them.[3]

Table 1: G7P Megaprojects

Type	**Name**	**Implementation Period**
Product	HDTV	1992 – 1994
Product	High-capacity semiconductor	1995 – 1999
Product	Next-generation car (electric vehicle)	1992 – 2001
Product	Next-generation flat panel display	1995 – 2000
Product	B-ISDN – Broadband Comprehensive Information Network Devices for 10GB environments	1992 – 2001
Product	New medicines and agrochemicals	1992 – 1997
Product	Medical Engineering	1995 – 2001
Product	Ultra-compact precision machinery	1995 – 2001
Product	High-speed train	1996 – 2001
Base	Advanced energy and informatic materials	1992 – 2001
Base	New functional biomaterials	1992 – 2001
Base	Advanced production system	1992 – 2001
Base	Next-generation semiconductor	1993 – 1996
Base	Environmental engineering	1992 – 2001
Base	New energy (Fuel-cell)	1992 – 2001
Base	Next-generation nuclear reactor	1992 – 2001
Base	Sensorial Engineering	1995 – 2001
Base	Next-generation superconducting nuclear fusion device	1995 – 2001

The G7P megaproject Selection started with a broad search for candidate projects led by the Research Coordination Department of the Ministry of Science and Technology in conjunction with other ministries. The G7 Expert Planning Team, the government unit created to run the G7P, received this information. The G7P unit then came up with lists of candidate megaprojects and drafted preliminary plans for each. These numbered seventy four.

[3] Some of these, such as the projects related to nuclear energy, are increasingly relevant to South Korea's export basket today. For example, Korea Electric Corporation (KEPCO), a major G7P nuclear energy R&D beneficiary, provided the United Arab Emirates's civil nuclear program and was recently shortlisted or selected as a preferred supplier for Saudi Arabia's and the Czech Republic's programs (Financial Times, 2024).

The G7P unit sent a questionnaire to hundreds of sectorial experts, mainly in ministries and universities. The experts were asked to choose the most promising projects on the basis of nine dimensions related to possible externalities, potential to succeed and close the technological gap with frontier countries, market potential, and fit with the program's philosophy.[4] The experts then rated all the projects on the nine dimensions.[5]

Using the survey results as input and in consultation with other ministries, the G7P unit selected twenty three projects for funding from South Korea's General Science and Technology Council, the country's highest technology policymaking body.

Though all the candidate projects were deemed worthy of support and highly ranked by the experts in their questionnaire answers, the council did not go fund all the projects. Those not funded were were a high-speed maritime ship, an aircraft core technology, a Korean natural language processing system, an automated traffic control system, and an offshore manufacturing plant. These projects were not funded because of concerns about the ability to complete them after some a budget shock. There were also some more idiosyncratic concerns about their fit with the G7P philosophy. The council decided to form commissions to assess the possibility of independently supporting the deferred projects outside the G7P, although this support ultimately did not materialize.

We exploit the fact that these projects were selected but never funded to inform our empirical analysis. A concern is that our estimates might reflect successful selection of profitable technologies. Indeed, as our previous discussion suggests, this was precisely what experts and policymakers sought. We address this concern by using as control technologies only those that were selected but were not ultimately funded in the end because of the budget shock.

We map the megaprojects to IPC technological classes using a language model. In support of our identification assumption, we find parallel trends in exports, a variable explicitly targeted in the G7P selection process, for the targeted and control technologies over the years prior to program implementation. Moreover, we find no systematic differences between producers of technologies in the targeted and control technological classes on a variety of observables (such as value added, output per worker, and capital intensity) that might have influenced the policymakers' decision-making.

[4]The dimensions were the following: technical externalities, comparison of technology level with that in advanced countries in the early 2000s if supported, comparison of international competitiveness in the early 2000s if supported, size of the domestic market upon commercialization, size of the global market upon commercialization, contribution to general welfare, estimated R&D cost, required R&D investment, and fit with the G7P's philosophy.

[5]The final megaproject choices seem to follow the experts' choices.

Once the megaprojects were funded, the G7P unit designated a public research institute to run them. These institutes expanded and implemented the research plans developed by the G7P unit. Once these plans were completed and the specific research projects defined, the institutes issued public requests for proposals, in response to which firms, both state-owned and private, submitted budgets and research plans. The research activities ensued after approval was granted by the managing public research institutes.

4 Data

Our empirical analysis relies on newly digitized data from the G7P, a language model to classify research projects into technological classes, and patenting, citation, export, manufacturing, and balance-sheet data. The rest of this section discusses the samples that we use, the G7P files, and how we use a language model to identify the technological classes related to the targeted and control/almost-targeted projects, as well as the data sources for the different outcome variables that we look at.

We use two types of samples for the main outcomes that we study: a South Korean sample, which includes observations at the technological class level for South Korea, and a cross-country sample, which includes information for all countries for which data are available. We use the South Korean sample to perform within-country comparisons and the cross-country samples for comparisons across countries. Tables A1 and A2 show the countries we include in our analysis.

We mainly look at two outcomes. The first is future-citation-weighted patenting for the targeted and control technological classes at the 4-digit IPC level.[6] We choose this level because it is widely used in the innovation literature and because considering more disaggregated data would impose a significant cost in terms of the precision of classification by our language model.[7] Once we restrict the sample to targeted and control technological classes, we keep 520 out of a universe of 646 classes. These 520 technological classes that we consider in our empirical analysis account for 90.7% of the USPTO patents granted to South Korean assignees in 1990.[8] The data that we use start in 1980

[6] An illustrative example is the following: The IPC has 5 levels of disaggregation: 1-digit ("domain"), 3-digit ("class"), 4-digit ("subclass"), 5-digit ("main group"), and 7-digit ("subgroup"). For a hydraulic steering gear, the 1-digit IPC code would be "B – Performing operations, transporting," the 3-digit code would be "B62 – Land vehicles for travelling otherwise than on rails," the 4-digit code would be "B62D – Motor vehicles; trailers," the 5-digit code would be "B62D3 – Steering gears," and the 7-digit code would be "B62D314 – Hydraulic."

[7] The precision of the IPCCAT model is 96.2% at the 3-digit level, 94% at the 4-digit level, 89.4% at the 5-digit level, and 82% at the 7-digit level.

[8] In practice, our samples cover all treated technological classes in all domains and untreated classes

and end in 2015.

The second main outcome that we study is exports. We look at targeted and almost-targeted classes at the 3-digit IPC level. We choose this level both to follow the literature (Liu and Ma, 2023) and because of the noisy correspondence between IPC codes (which characterize our targeting variable) and SITC classes (by which our export data are categorized; (Lybbert and Zolas, 2014)) at finer levels. After restricting the sample to targeted and almost-targeted technological classes, we have 101 technological classes out of a universe of 131. These accounted for 80.3% of South Korea's exports in 1990. As with our patenting data, this sample starts in 1980 and ends in 2015.

We also use South Korea's Mining and Manufacturing Survey (MMS), currently available to us for years between 1980 and 2003. This source, which is available only for South Korea, contains yearly plant-level information on sales, inputs, and outputs for South Korean establishments involved in mining or manufacturing and employing ten or more employees. Given the limited timeframe of the data to which we have access, we use this source to assess the extent to which our identification strategy addresses selection concerns. Here we focus on observables that policymakers might have targeted while selecting the G7P projects, such as output per worker and capital intensity.

4.1 G7 Program Files and the Language Model

Our primary source for G7P information is the G7P Yearly Project List.[9] We obtained a copy for every year that the G7P was active (1992–2001) through a Transparency Law request to the National Research Foundation of Korea. We digitized and cleaned these records for information on all 4,787 G7P projects. We observe each project's G7P megaproject affiliation, name, description, objectives, managing research institute, participating firms (if any), start date, end date, and funds provided (public and private). Figure 1 shows the typical record we observe.

We would also like to observe the specific technological class targeted by each project. We do not have such data. This lack of information presents a challenge since any econometric evaluation requires a notion of the sectors the G7P targeted. How do we use the G7P information we gathered to study the important questions that motivate this paper?

We overcome this challenge with a text-based approach. We use the rich textual data

in all domains except C (Chemistry; Metallurgy) and A (Textiles; Paper). An IPCCAT search of the classes corresponding to the control megaprojects yields all classes but those two.

[9] The publication name in Korean is 선도기술개발사업 과제목록 (G7 프로젝트).

Figure 1: Sample Page of the 1995 G7P Yearly Project List

사업구분 선도기술개발사업

과제번호	과 제 명	연구기관 (책임자)	참여기업	연구기간	'95 연구개발비(단위:천원)			최종목표	연구내용
					정부	기업	계		
95-G-02-01-A	교환기술분야개발	전자통신연구소 (임주환)	한화정보통신 동아전기 삼성전자(주) 대우통신(주) 우진전자통신(주) LG정보통신(주)	'92 ~ '97 ('95/01/01 ~'95/12/31)	32,868,000	44,237,000	77,105,000		
95-G-02-01-A-01	ATM 교환기 시스템 개발	전자통신연구소 (한치문)	한화정보통신 동아전기 삼성전자(주) 대우통신(주) 우진전자통신(주) LG정보통신(주)	'92 ~ '97 ('95/01/01 ~'95/12/31)	27,520,000	38,669,000	66,189,000	정보화 사회의 구축에 핵심적인 광대역 ATM기술, 광교환기술 등 차세대 교환기술개발	ㅇ소형 ATM 교환기 개발 완료 ㅇ중형 ATM 교환기 구조 설계
95-G-02-01-A-01-A	ATM 교환기에서의 과부하제어에 관한 연구	한남대학교 (최진규)					15,000		
95-G-02-01-A-01-AA	운용메시지의 음성화에 관한 연구	과학기술원 (오영환)					15,000		
95-G-02-01-A-01-AB	ATM 교환기의 내진동 설계 및 해석에 관한연구	과학기술원 (엄윤용)					25,000		

–59–

that we digitize to classify projects into technological classes on the basis of their textual description. We feed the project's name, description, and objectives into the WIPO's IPCCAT language model to retrieve the technology classes associated with each project.

The WIPO developed the first version of IPCCAT in 2002 to assist resource-constrained patent offices in classifying inventions by IPC technological class, precisely our task. Improved and refined ever since, IPCCAT today uses data on over 49 million patent documents (abstracts and description) and their human-originated classification to support patent classification tasks.

For each G7P project, we do the following: (i) input the project's name, description, and objectives into IPCCAT, (ii) choose the level at which we want to generate our classification (3-digit, 4-digit, etc.), and (iii) choose the language in which we are inputting the text.[10] Once we set these options, IPCCAT prints the predicted IPC classifications for the input text with a degree of confidence that ranges from 0 to 5. We detail our choices on these items below.

As discussed above, we choose to generate our classifications at the IPC 4-digit level

[10] Another decision the researcher needs to make is the IPC version in which the IPCCAT prints the predictions. This decision is not relevant for us, however, since the IPC codes do not change at the levels of disaggregation that we use.

for the patenting sample because predictions at finer levels might be subject to a lower degree of confidence because of the increased number of choices that the algorithm faces. The number of categories (precision of the algorithm) rises (falls) from 646 (94%) at 4 digits to 7,437 (89.2%) at 5 digits and 65,158 (81.3%) at 7 digits (World International Patent Organization, 2024). Asking IPCCAT to generate predictions at finer levels makes the classification problem more complex with no perceptible gain from the increased granularity. We opt to classify our export and manufacturing survey sample data by IPC 3-digit code. This choice follows from the fact that the correspondence between IPC codes and real production variables (exports, in this case) is imperfect at relatively fine levels of detail, which precludes us from classifying the data at the IPC 4-digit code as we do for our patenting sample.

Finally, we would like to use high-quality IPC-code predictions only. For both of our samples, we decide to use predictions subject to a degree of confidence of 3 or higher. This level allows us to accept the classifications for projects accounting for over 97% of the total G7P funds. For those projects for which we discard the predicted classification, we impute their respective IPC codes from all the other projects in the same G7P megaproject each year. We show in the appendix that using predicted classifications subject to alternative confidence levels does not substantially change our findings.

After we perform this exercise, we have a database of 4,787 research projects with information on G7P megaproject affiliation, name, description, objectives, managing research institute, participating firms (if any), start date, end date, funds provided (public and private), and targeted technology classes at the 3-digit and 4-digit IPC code levels. We use this information to determine the targeted classes and time of targeting. We perform a similar exercise for the almost-targeted (planned but not implemented) G7P projects, which yields the technological classes that we use as controls for our targeted classes. We assume that once a class was targeted, it remained so until the end of our study period.

4.2 Patenting Data

We download the universe of patents granted by the USPTO from 1980 to 2015. For each of the over 7 million patents granted, we observe the patent's application and grant years and IPC code(s), the geographic location of the assignee (the legal entity holding ownership interest in the legal rights at the time of application) and each of the inventors, the patents it cites, and the citations from subsequent patents. Our primary outcome of interest for this dataset is the future-citation-weighted count of patents granted by the USPTO to assignees in our sample countries at the 4-character IPC technological class level.

We define a patent as coming from a given country when the assignee is in that country. We show in the appendix that our results do not change when we use more demanding definitions of patent nationality. Second, we consider only future citations coming from patents (i) classified in 3-digit IPC codes different from that of the underlying patent and (ii) from countries other than the country where the assignee is located. We do so to avoid the so-called home bias in patent citations (Kwon et al., 2017) and differential citation patterns within fields. Focusing on citations made by inventors in other countries controls for potential strategic behavior by G7P beneficiaries, which might have started citing themselves more often. Third, we divide the patent's future citations equally among all its IPC codes (that is, we use fractional citations). We add these citations at the 4-digit IPC code level for each year, referencing the application year. We then merge these data with our database on G7P-targeted technological classes described above. Table A3 shows summary statistics on the levels of South Korean future-citation-weighted patenting throughout the study period and before and after the G7P started.

4.3 Export Data

We use UN-COMTRADE export data for South Korea and the rest of the world for the period between 1980 and 2015. We gather this information at the SITC Rev. 2 4-digit level. We use Lybbert and Zolas (2014)'s SITC–IPC 3-digit correspondence table. We add exports at the IPC 3-digit level using the probability that each SITC code belongs to an IPC 3-digit code as weights. We end with a panel of exports at the IPC 3-digit level from 1980 to 2015. We merge these data with the information that we retrieved on G7P-targeted technological classes. As for the patenting sample, Table A3 shows summary statistics for exports over relevant timeframes.

4.4 Plant and Firm-Level Data

We have access to South Korea's MMS for years between 1980 and 2003. This source gives us access to plant-level information on output and input usage for all mining or manufacturing plants employing ten or more people. The MMS includes information on the Standard Industrial Classification (SIC) sector in which each plant operates. We use this information and Lybbert and Zolas (2014)'s correspondence table to determine the IPC 3-digit codes relevant to each plant. We abstract from the effects of entry by limiting our sample to plants existing before the G7P.

We turn to firm-level data to inform our cost–benefit analysis. We use DataGuide, which provides daily stock price movements for all the 3,345 publicly traded South Korean companies for 1990–2015. This input is essential for us to infer the valuations of the USPTO patents granted to South Korean firms: Kogan et al. (2017)'s method, which we use, exploits changes in the patent grantee's stock market capitalization in the days after a patent is granted. We link the USPTO-granted patents to the companies in DataGuide using Lee (2019)'s correspondence table.

We also use DataGuide's information on balance sheets to compute South Korea's private-sector return on equity. We use this metric to calculate the opportunity costs that G7P-supported firms incurred by investing in G7P-supported projects. The dataset includes information on 55,079 South Korean companies subject to external audits.[11]. We use data for the 1980–1991 period, before the G7P was implemented.

5 Empirical Strategy

Our design features (i) the use of targeted and almost-targeted (control) technological classes to estimate program effects and (ii) an event study that we estimate using a local projections difference-in-differences (LP-DID) approach (Jordà, 2005; Dube et al., 2024). As the G7P treated different technological classes over the years that it operated, we observe "cohorts" of technological classes targeted every year from 1992 to 2000.

The two features enable us to address concerns about identification that might threaten our analysis. The first relates to selection: Perhaps the selected technologies would have been ripe for success even in the absence of targeting. Though the overall technology selection process was indeed endogenous, the technologies that we use as controls were perceived as equally promising by the G7P experts, according to program records (KISTEP, 2002). The associated megaprojects were not implemented because of budget shocks and concerns about the ability of the program to sustain them over the long term. Our empirical analysis supports our claim of plausible exogeneity of treatment assignment, as treatment tells us nothing about exports, an explicitly targeted outcome, over the pre-G7P period.

The second feature of our design, the event study, is convenient because it enables the exploration of treatment dynamics and exploits the fact that the G7P targeted different classes over time. Formally, our identification assumption is that patenting and export performance in the targeted classes would have evolved similarly to their counterparts in the nontargeted classes had the G7P not been implemented. This assumption might take

[11] These are firms that met at least two of the following conditions: (i) assets over $11 million (2020 dollars), (ii) liabilities over $6.4 million, sales over $9.2 million, or more than 100 employees.

different forms. For example, they might relate to our previous discussion on selection. In Section 5, we never reject the null hypothesis that the pre-G7P treatment coefficients differ from zero at standard confidence levels. Again, pre-program exports and other variables targeted by the selection process have no correlation with assignment to G7P treatment.

We use the LP-DID approach to estimate our event study. This means that (i) we estimate regressions using ordinary least squares (OLS) separately for each year and (ii) we restrict the sample to comply with the clean-control condition. In practice, this means we keep only "newly treated" technological classes ($\Delta G7P_{s,g+h} = 1$) or clean controls ($G7P_{s,g+h} = 0$). We prefer LP-DID to other staggered DID estimators because it prevents us from making *forbidden comparisons*, whereby some treated observations are taken as controls for other treated observations. These might lead to contaminated coefficient estimates. Our choice of estimation method implies that we do not have to saturate our specification with pre-period coefficients to avoid contamination of the coefficients that we estimate.

The conditional independence assumption also relates to contemporary shocks to our explanatory variables that might bias our estimates. Given the multiple G7P cohorts, the coefficients that we estimate are not derived from single years and are, therefore, less likely to be driven by contemporary shocks. Moreover, we impose a relatively stringent set of controls to account for possibly correlated shocks.

5.1 Patenting

We estimate the effect of the G7P on future-citation-weighted innovation output and industry exports. Equation 1 is our baseline specification:

$$\Delta ihs(patents)_{s,g+h} = \alpha + \beta_{g+h}\Delta G7P_{s,g+h} + \delta_{c,t} + \sum_{j=1987}^{2015} X_s \gamma_j + \epsilon_{s,g+h} \tag{1}$$

$$\Delta ihs(patents)_{s,g+h} = ihs(patents)_{s,g+h} - ihs(patents)_{s,g-1} \tag{2}$$

$$\Delta G7P_{s,g+h} = G7P_{s,g+h} - G7P_{s,g-1} \tag{3}$$

$\Delta ihs(patents)_{s,g+h}$ is the change in (inverse hyperbolic sine) future-citation-weighted patents in an IPC 4-digit technological class s at h years after the G7P targeting relative to their number in the year $g-1$, the year before targeting. Our coefficient of interest on the right-hand side is β_{g+h}, which captures the average G7P effect on treated classes at different points in time. We include $\delta_{c,t}$, a calendar year–IPC 3-digit technological class c fixed effect to account for all shocks at this level. All our specifications include the interaction between technological class s's share of patenting output between 1987 and 1991, X_s, and calendar year dummies to account for potentially time-variant unobserved biases toward technologies in which South Korea had existing research capability. By specifying our model in a difference setting, we account for unobservable attributes at the IPC 4-digit level that do not change over time.

We allow h to be between -5 and +15—that is, we investigate our outcome in the period between the five years before a class was targeted and up to fifteen years after its targeting. Though we include more pre-treatment lags in the robustness checks, we choose this timeframe because planning exercises typically consider these time horizons. Our identification assumption is that, conditional on the fixed effects that we include and other variables on the right-hand side, the outcomes of the treated and control classes would have evolved similarly had the G7P not been implemented. We cluster standard errors at the IPC 4-digit level.

5.2 Exports

We study exports using our export sample, which is at the IPC 3-digit level. Equation 4 gives our baseline specification:

$$\Delta ihs(exports)_{c,g+h} = \alpha + \beta_{g+h}\Delta G7P_{c,g+h} + \delta_{d,t} + \sum_{j=1987}^{2015} X_c\gamma_j + \epsilon_{c,g+h} \tag{4}$$

$$\Delta ihs(exports)_{c,g+h} = ihs(exports)_{c,g+h} - ihs(exports)_{c,g-1} \tag{5}$$

$$\Delta G7P_{c,g+h} = G7P_{c,g+h} - G7P_{c,g-1} \tag{6}$$

$\Delta ihs(exports)_{c,g+h}$ is the change in (inverse hyperbolic sine) of exports in an IPC 3-digit technological class c at h years after the G7P targeting relative to their level in the year $g-1$, the year before targeting. The coefficient of interest is β_{g+h}, the G7P effect on treated classes at different points in time. We include a calendar year–IPC 1-digit technological class d fixed effect to account for shocks at this level. All our specifications include the interaction between technological class c's average share of exports between 1987 and 1991, X_s, and calendar year dummies to account for potentially time-variant unobserved biases toward technologies in which South Korea had existing export capacity. As we did for Equation 1, we set up Equation 4 in differences, which allows us to control for unobserved characteristics at the IPC 3-digit class level. Here, we also allow h to be between -5 and +15. We include more pre-treatment lags in the robustness checks in the appendix.

Our identification assumption is that, conditional on the fixed effects, the outcomes of the targeted and control classes would have evolved similarly had the G7P not been implemented. We cluster standard errors at the IPC 3-digit level, the level at which our explanatory variable changes in this case.

5.3 Cross-Country Evidence

How did G7P-targeted technological classes' patenting output and exports fare in comparison to those of the rest of the world? We study this question by taking our within-country estimations to cross-country samples in a triple-difference setting. Intuitively, we compare our baseline within–South Korea estimates to those for other countries—which effectively act as placebos. As for our South Korean samples, we use an LP-DID approach to estimate the relevant event studies.

5.4 Patenting

Equation 7 below shows the specification that we estimate, as above, using a standard local projection approach:

$$\begin{aligned}\Delta ihs(patents)_{s,g+h,k} = \alpha + \beta_{g+h}\Delta G7P_{s,g+h} \times I[\text{South Korea}] \\ + \delta_{c,t,k} + \sum_{j=1987}^{2015} X_{s,k}\gamma_j + \epsilon_{s,g+h,k} \quad (7)\end{aligned}$$

Note that it is identical to Equation 1 except for the inclusion of country subscript k and an indicator variable for South Korea. Here, $\Delta ihs(patents)_{s,g+h,k}$ is the change in (ihs) future-citation-weighted patents in an IPC 4-digit technological class s for country k at h years after the G7P targeting relative to their count in the year $g-1$, the year before targeting. We include $\delta_{c,t,k}$, a calendar year–country–IPC 3-digit technological class c fixed effect to account for all shocks at this level. We also include the interaction between technological class s's share of patenting output in country k between 1987 and 1991, $X_{s,k}$, and calendar year dummies. As in our baseline within-country specifications, we allow h to be between -5 and 15.

5.5 Exports

Equation 8 shows the specification that we estimate, as above, using a standard local projection approach:

$$\begin{aligned}\Delta ihs(exports)_{c,g+h,k} = \alpha + \beta_{g+h}\Delta G7P_{c,g+h} \times I[\text{South Korea}] \\ + \delta_{d,t,k} + \sum_{j=1987}^{2015} X_{c,k}\gamma_j + \epsilon_{c,g+h,k} \quad (8)\end{aligned}$$

Similarly to our patenting specification, Equation 8 is identical to Equation 4 except for its inclusion of the country k subscript and an indicator variable for South Korea. Here, $\Delta ihs(exports)_{s,g+h,k}$ is the change in (ihs) of exports in an IPC 3-digit technological class c in country k at h years after the G7P targeting relative to their level in $g-1$, the year before targeting. We include a country–calendar year–IPC 1-digit technological class d fixed effect to account for shocks at this level. All our specifications include the interaction between technological class c's average share of exports for country k between 1987 and 1991, X_s, k, and year calendar dummies. As in our baseline within-country specifications, we allow h to be between -5 and 15.

6 Results

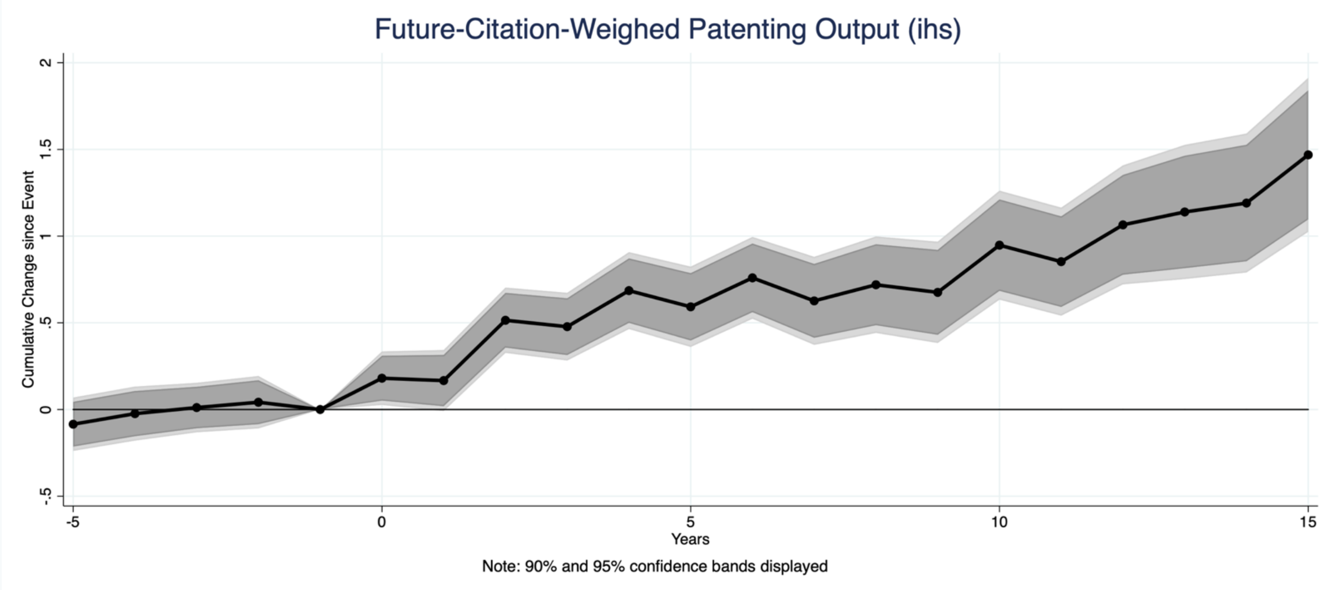
We find that future-citation-weighted patenting output and exports in G7P-targeted classes substantially increased relative to their counterparts in nontargeted classes over the long run. The dynamics of patenting and exports were, however, different: Whereas patenting output increased almost immediately following the targeting, exports started increasing only a few years after it. We first discuss our within-country results, including an expanded discussion on mechanisms and selection concerns, and then move to our cross-country findings.

6.1 Patenting

Figure 2 shows the result of estimating Equation 1 using the empirical strategy outlined above. We find that quality-adjusted (future-citation-weighted) patenting output in G7P-targeted classes increased relative to that in nontargeted classes. These effects varied over time. Our point estimates suggest that patenting output in the G7P-targeted classes increased by 16% the year after first receipt of G7P support relative to that of control classes in the year before treatment. This metric increases to 64% for the 5th year, 123% for the 10th year, and 232% for the 15th year. The evolution of the treatment effects over time suggests that the program spurred innovation relatively quickly and had an important long-term effect in the targeted classes. These effects are not linear: Our point estimates do not vary much between the third and the ninth years after targeting.

Figure 2 also shows no systematically different trends in outcomes in the targeted and control groups for the years before the G7P targeted a technological class. We cannot reject the null hypothesis that those coefficient estimates are equal to zero at standard confidence levels. Moreover, all the estimated coefficients are very close to zero in all cases.

Figure 2: South Korean Sample

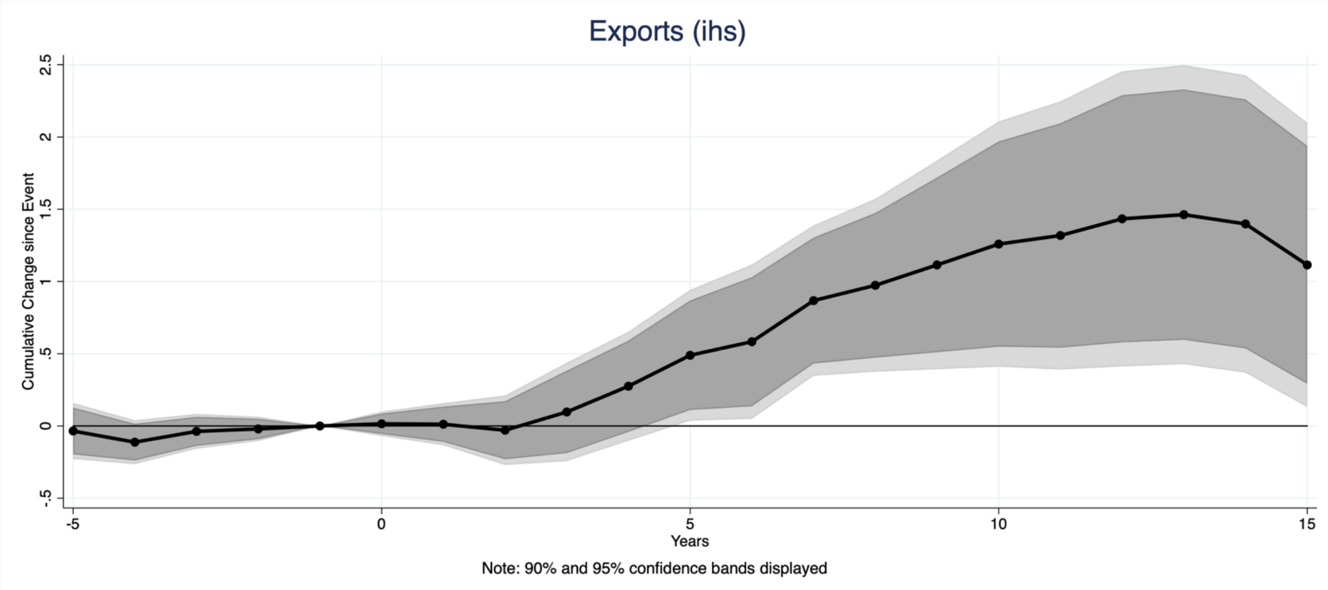


6.2 Exports

Figure 3 shows the result of estimating Equation 4 using the empirical strategy outlined above. We find that exports in the G7P-targeted classes increased relative to those in nontargeted classes over the long run. In contrast to patenting output, which responded almost immediately to the targeting, exports took some time to react. Our coefficients are essentially null for the first three years after treatment and become statistically different from zero at standard significance levels only for the 5th and subsequent years. These point estimates suggest that exports in the targeted classes had increased relative to those in the control sectors by 62% in the 5th year, 245% in the 10th year, and 204% in the 15th year. We note that these are real changes: Though we measure exports in nominal dollars, our fixed effects absorb price differentials over time.

Figure 3 also shows no differential trends in exports in the treated and control classes for the years in advance of treatment. We are unable to reject the null hypothesis that the pre-targeting coefficients are equal to zero in all cases. As we hinted before, we also interpret these results as a plausibility check for our research design in the patenting sample. It is widely acknowledged that exports played a central role in South Korea's economic miracle. Conversely, if there were any selection that our design does not account for, we should expect to observe it here, as external market potential was a variable that the megaproject selection process explicitly considered. As these results confirm, targeted and almost-targeted megaprojects passed the selection process. Thus, pre-G7P export performance tells us little about selection in our sample, given our design.

Figure 3: South Korean Sample



We implement several robustness checks to assess the extent to which decisions made while collecting data drive our findings. We assess the robustness of our findings to logarithmic transformation of the dependent variable and to the use of alternative definitions of patent nationality, alternative quality thresholds in our language model exercise, and longer pre-treatment lags. We refer the reader to the appendix while noting that our results are robust to these alternative choices.

6.3 Further Discussion on Selection

One concern is that our baseline results might reflect successful technology selection. However, while we discuss in detail above how our identification strategy deals with this matter and how the lack of differential trends in outcomes is informative about this issue, further discussion is warranted. One channel through which selection might operate is comparative advantage. Perhaps our estimates reflect that South Korean policymakers chose sectors that were already prone to success because they built on the underlying strengths of the South Korean economy. An example could be the electronics and home appliances sector, where South Korea was a relevant player even before the G7P.

We emphasize that our estimates control for such types of pre-existing strengths when we include pre-G7P shares of patenting and exports. We also estimate Equations 1 and 4 excluding the technological classes with pre-G7P outcomes (patenting or exports) at the 95th percentile or higher to further investigate the extent to which well-established sectors might drive our results. Figures 4 and 5 show that our findings remain unaltered. These results reflect that we appropriately controlled for those pre-existing strengths in our baseline specification.

Figure 4: South Korean Sample

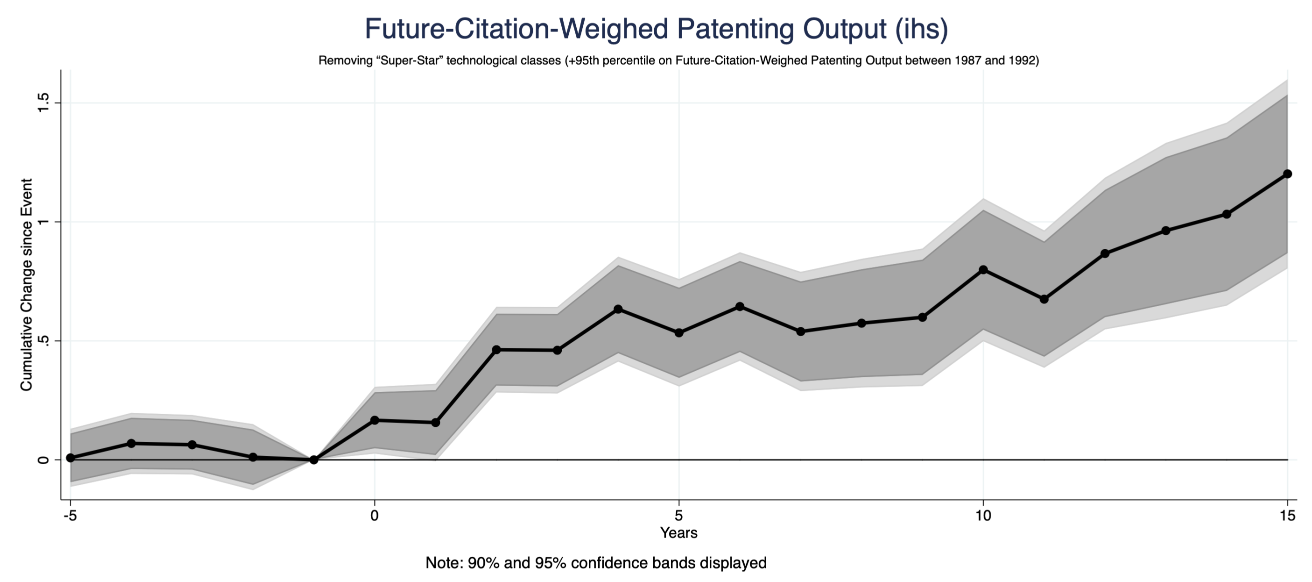
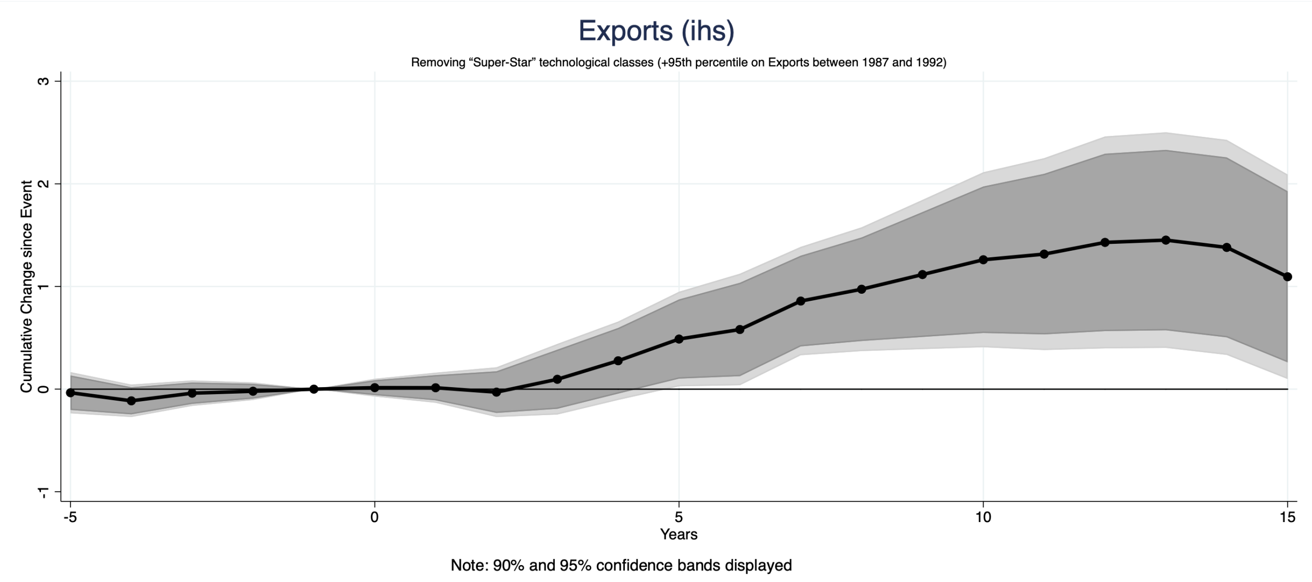


Figure 5: South Korean Sample



Most importantly, however, the G7P supported projects in which South Korea had well-known strengths—and others in which it had no development tradition, such as nuclear power and high-speed rail. In some of our control megaprojects, such as the high-speed ship project, South Korea had (and continues to have) global relevance. Such choices suggest that comparative advantage was not the sole driver of project selection.

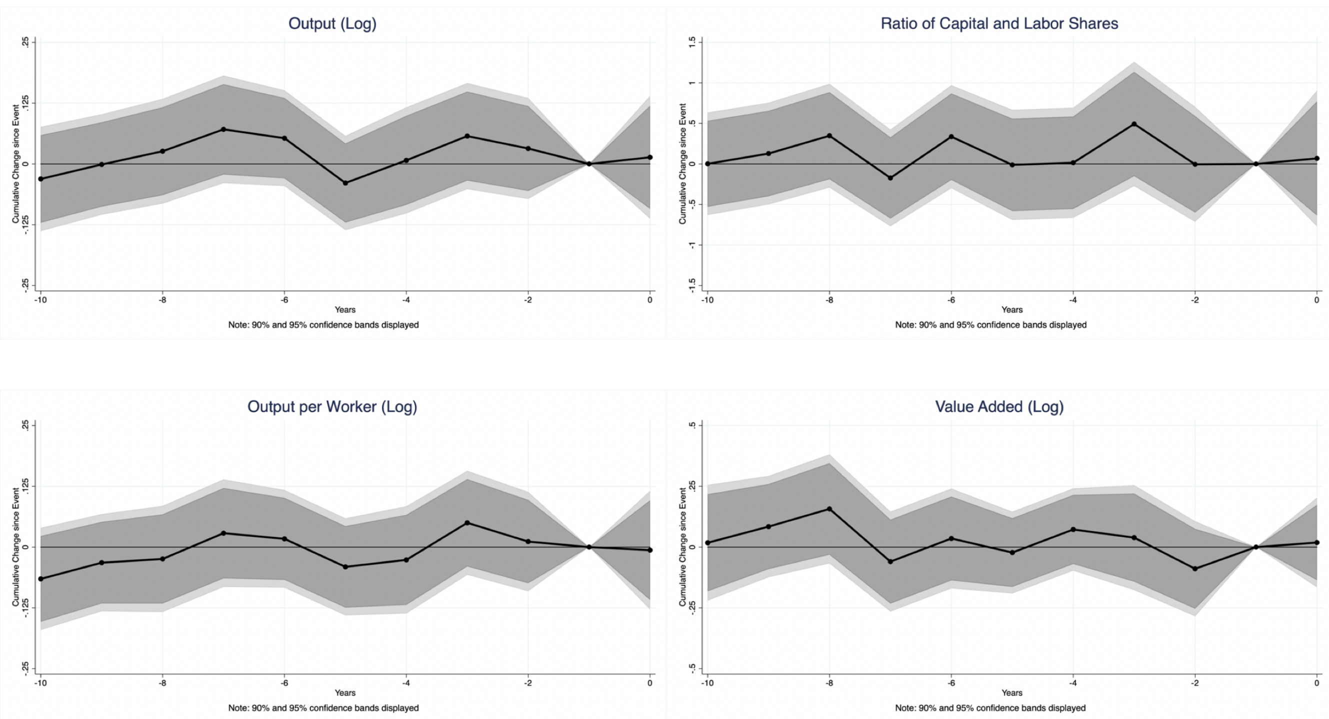
Though we are unable to rule out selection on unobservables, we show that selection on other observable economic variables was unlikely. This is what we would expect in view of the nature of the projects that were selected but not funded. To assess the validity of this expectation, we estimate Equation 9 below for our manufacturing sample using the standard LP-DID approach used throughout the paper:

$$\Delta Y_{f,c,g+h} = \alpha + \beta_{g+h}\Delta G7P_{g+h} + \delta_{c,t} + Age_f + \epsilon_{f,c,g+h} \tag{9}$$

where $\Delta Y_{f,c,g+h}$ is the change in variable Y for plant f in technological class c at h years after the G7P targeting relative to its level in $g-1$, the year before the G7P targeted the technological class in which the plant operates. We include $\delta_{c,t}$, a 3-digit technological class fixed effect, to account for shocks at this level. Age_f is plant f's age. We allow h to take values between -10 and 0. $\Delta G7P_{g+h}$ is defined as before in the paper. Our coefficient of interest is β_{g+h}, which is informative about differences between the targeted and almost-targeted groups in the variables that we look at.

We look at (log) output, (log) value added, (log) output per worker, and relative capital intensity—all variables that policymakers might have targeted while selecting the projects. Figure 6 shows the results of our estimating Equation 9. We find that targeting is not informative about these variables: Our coefficient estimates are typically very close to zero and are not statistically significant in any case for the years that we investigate. Though we are unable to rule out selection on unobservables, these findings alleviate remaining selection concerns.

Figure 6: South Korean Sample



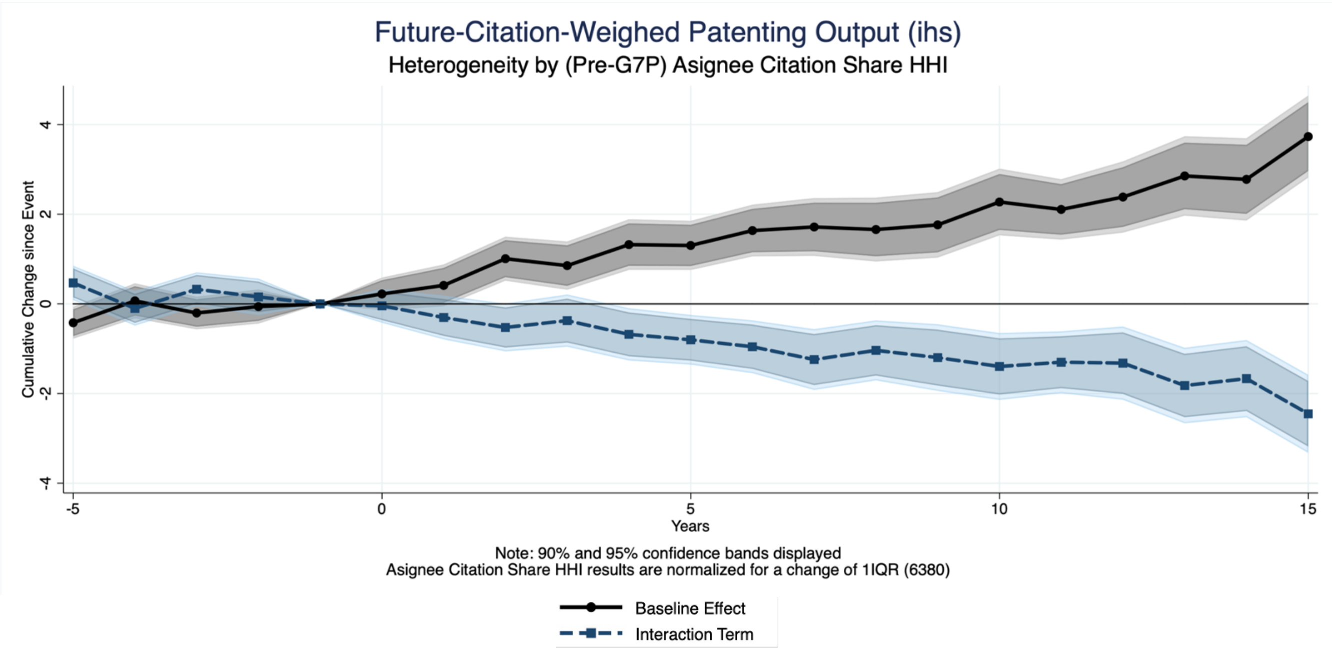
6.4 Mechanisms

To further understand the economics behind the G7P, we study the nature of the sectors that drive our baseline results. Different theories of economic growth (Romer, 1990) emphasize the role of knowledge spillovers in spurring innovation. Indeed, these spillovers often justify policy interventions to address market failures and align private incentives with societal goals. At the same time, industrial policy has in practice often (though not always) been associated with creating "national champions" able to exploit scale. However, one concern is that with scale and concentration might come more limited spillovers. What can we learn about this trade-off from the application of the G7P?

We measure the level of concentration of scientific output by computing the Hirschman-Herfindahl Index (HHI) for citation shares at the technological class level in the pre-G7P period. To do so, (i) we retrieve the number of citations that any South Korean assignee received for patents linked to a technological class over the pre-G7P period, (ii) we compute each assignee's share of citations for each class, and (iii) we compute the HHI for each technological class using those shares. We incorporate this measure by interacting it with G7P treatment–year dummies.

Figure 7 shows the results of this exercise. For ease of interpretation, we present the results for the HHI normalizing for a change of 6380, the difference between the 25th and 75th percentiles of the HHI distribution. Our findings suggest that sectors with lower concentration in scientific output performed much better than those with higher concentration. An HHI change of 6380 corresponds to a reduction in the baseline program effect of approximately three-quarters by the tenth year after receipt of program support. These effects are substantial and underline the relevance of spillovers in determining program success. Moreover, these might serve as a cautionary tale for pushes in industrial policy that end up increasing concentration.

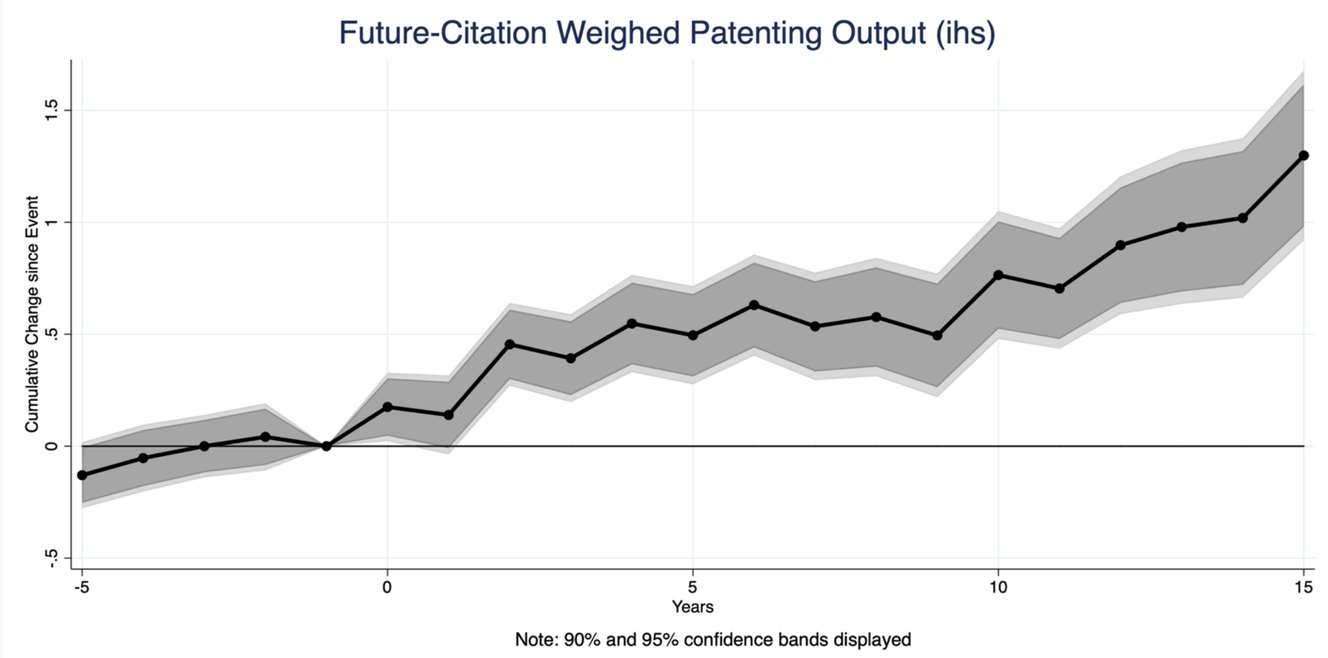
Figure 7: South Korean Sample



6.5 Cross-Country Evidence

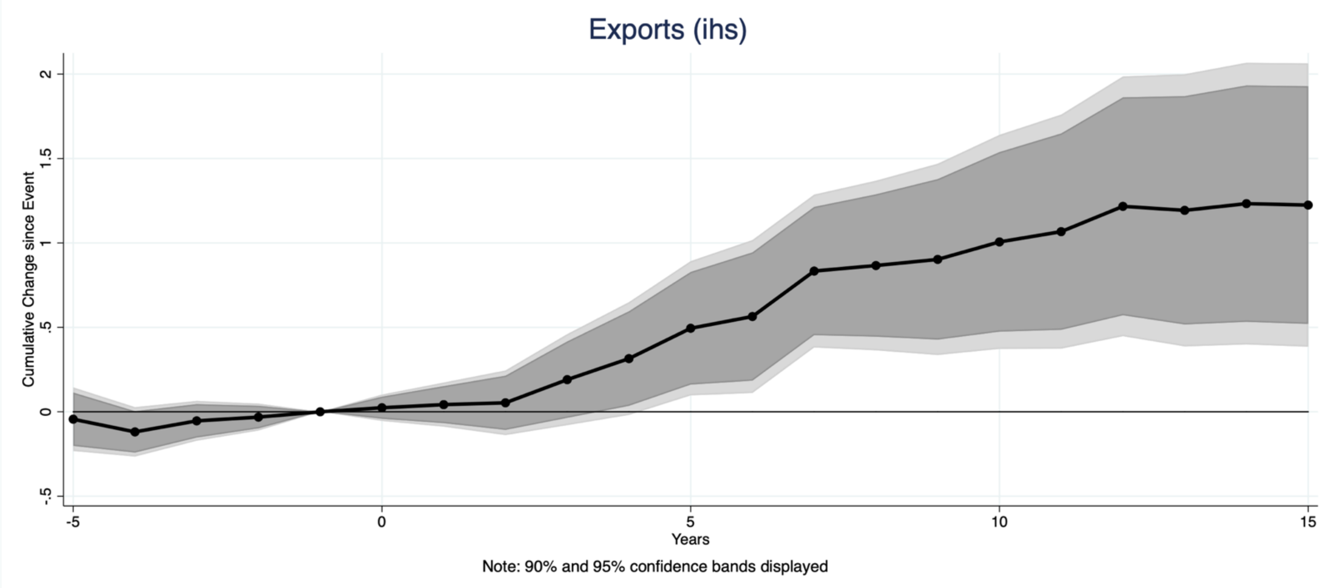
Patenting Figure 8 shows the result of our estimating Equation 7. Our results here are quite similar to those in Figure 2. We find that South Korea's patenting output in the G7P-targeted technological classes increased relative to other countries' following G7P support. The point estimates suggest that the increase was 64.1% by the 5th year, 147.8% by the 10th year, and 166.3% by the 15th year. Figure 8 also shows that the targeted technological classes in South Korea followed trends similar to those of their counterparts in other countries before receipt of program support.

Figure 8: Cross-Country Sample



Exports Figure 9 plots our findings from estimating Equation 8. Our results here are very similar, in both direction and magnitude, to those from our within-country regressions. We find that South Korea's exports in the G7P-targeted technological classes increased in comparison to other countries' after the extension of program support, though it took time for these effects to materialize. We find negligible effects for the first three years and detect a statistically significant increase in exports only by the 4th year. Our point estimates imply that exports increased relative to those in the year before receipt of program support by 64% by the 5th year, 173.5% by the 10th year, and 240.2% by the 15th year. We also fail to find differential trends in South Korean exports in G7P-targeted technological classes relative to those of other countries before receipt of program support.

Figure 9: Cross-Country Sample



6.6 Discussion

Our results highlight that the G7P shifted the direction in which the South Korean economy innovated. The quality-weighted patenting output of G7P-targeted technological classes grew substantially faster than that of control classes after treatment. These effects emerged quickly and persisted over time, suggesting that relative innovation levels in these classes changed permanently, even after the G7P ended operations in 2001. We find similar results when we estimate a triple-difference model in which we compare South Korean patenting output in the G7P-targeted technological classes to patenting output in other countries of the world, a finding that suggests important absolute-level effects. Overall, our analysis shows that the program successfully spurred high-quality innovation in the targeted technological classes. Our work helps rationalize how South Korea caught up to the technological frontier over the 1990s and 2000s (Kwon et al., 2017).

These shifts in the direction of innovation had important impacts on the real economy, even if they took time to materialize. We can detect a statistically significant effect on exports only by the 5th year after the program targeted a technological class. These results contrast with the finding of a relatively rapid impact on innovation output. Our findings are similar when we estimate a triple-difference model, suggesting that South Korea, already an export powerhouse in some sectors, improved its standing *relative* to other countries after its implementation of the G7P.

We find that the program was substantially more effective in spurring high-quality innovation in technological classes where scientific output was less concentrated at the advent of the G7P. We interpret this finding as indicative of the relevance of knowledge spillovers as determinants of the success of innovation policy. Moreover, it may also serve as a cautionary tale against excessive concentration, which might be a by-product of certain policies that, for instance, look to create "national champions" able to exploit economies of scale. A possible cost of higher concentration might be, for example, more limited spillovers.

Taken together, our results imply that the policy embedded in the G7P was relevant to South Korea's transition to a knowledge-intensive economy. This leap is one that countries often fail to accomplish. Most of the debate about East Asia's economic miracles has focused on the role that industrial policy played in enabling heavy industry over the 1970s. However, as countries such as Brazil and Mexico suggest, the obstacles to productive development do not stop with successful development of heavy industry. We show that technology policy played a role in increasing the sophistication of the South Korean economy *after* it developed a sizeable (heavy) manufacturing sector. Economic development, as Hirschman (1958) argues, is a complex process that necessitates a strategy, not a plan, and shifting policies to address the ever-changing character of the hurdles that developing economies face. Our findings are consistent with the story of a developmental state that opportunely shifted its industrial strategy to overcome those ever-changing hurdles.

7 Cost–Benefit Analysis

Policymakers often face budget constraints when deciding how to allocate public investment. These decisions are complicated by the fact that different policies might yield different returns on investment. This consideration is relevant to our study since the benefits that we find in our exercise might have materialized at an excessive cost, raising questions about the program's desirability. Was the G7P a cost-effective intervention?

Though we observe program investments in the G7P files, addressing the question of cost-effectiveness is challenging because we do not have a ready-to-use notion of the economic benefits yielded by the G7P. We overcome this challenge by implementing Kogan et al. (2017)'s patent valuation method in Korean data, which gives us the value in South Korean won of each individual innovation patent granted by the USPTO to publicly traded South Korean firms between 1992 and 2015. We combine this information with our reduced-form findings to identify the G7P's benefits.

7.1 Benefits

We use the reduced-form specification results to predict the number of patents that each G7P-supported technological class would have had in the absence of the program. We use the point estimates from our cross-country exercise. This approach amounts to assuming that patenting in the G7P-supported classes in Korea would have evolved in line with that in other countries—perhaps a less restrictive assumption than a parallel evolution of outcomes vis-à-vis those of the control technological classes in South Korea. In practice, however, the distinction is irrelevant since the point estimates from both exercises are very similar.

Given that we have point estimates for fifteen years, we compute the number of G7P-attributable patents using the point estimates for the first fifteen years after the G7P supported a technological class. Thus, if a technological class received support in 1992 (2000), we count the number of G7P-attributable patents between 1992 and 2007 (2000 and 2015).

Our empirical analysis is at the technological class level, implying that our measure already includes within–technological class, cross-firm spillovers. However, we ignore potential equilibrium effects across sectors. We identify a number of G7P-attributable patents granted within the first fifteen years after a technological class received program support. What is their economic value?

We follow Kogan et al. (2017) to infer the economic value of all USPTO patents granted to publicly traded South Korean firms between 1992 and 2015. We use daily stock market capitalizations from DataGuide and patent grant dates from USPTO to identify the dates when market participants learned about grant decisions. We use Lee (2019)'s correspondence table to match the USPTO assignees to the DataGuide data.

Intuitively, Kogan et al. (2017)'s method compares the market capitalization of patent assignees in the three days following a USPTO patent grant decision. The economic value of a patent results from adjustment of the changes in an assignee's market capitalization over the three-day post-grant window for broader market moves not related to the grant.[12]

To adjust for broader market moves, we isolate the assignee-specific return from the

[12]To provide a more conservative estimate of patent valuations, we depart from the original method which also adjusts for the possibility that market participants might anticipate to some extent the patent grant. Kogan et al. (2017) multiplies the inferred net-of-market-return patent values by the reciprocal of the unconditional probability that a filed patent is approved by the USPTO, which was 0.44 for the 1991-2001 period. If we follow the method exactly, the benefits we compute increase by a factor of 2.27.

broader market portfolio (South Korea's KOSPI index, in our case). Doing so requires us to impose assumptions about the distribution of the patent value and non–patent-related returns to back out the "true" signaling value of the patent grant. We follow Kogan et al. (2017) in assuming that patent values follow a normal distribution truncated at 0. Similarly, we assume that the non–patent-related returns follow a normal distribution. The method is robust to alternative distributional assumptions and provides a measure of the private valuation of each patent.

After implementing Kogan et al. (2017)'s technique, we have a valuation in Korean won for each patent granted by the USPTO to any publicly traded South Korean firm between 1992 and 2015. We use these valuations to compute the median value of a patent in each technological class every year. We multiply this measure by the number of G7P-attributable patents implied by our reduced-form exercise and come up with a current Korean won value of the G7P's benefits for each treated technological class. We convert the valuations to 1992 Korean won and compute their present value in 1992 using a 5% discount rate.

7.2 Costs

We include two costs in our analysis: R&D expenditures by program participants, including subsidies, and opportunity costs. We directly observe the R&D expenditures from the G7 program files. We compute the opportunity costs assuming that public investments would have invested the R&D expenditures in an alternative opportunity yielding a 7.5% per annum return, which would have been reinvested at the same rate. This choice comes from Kim (1996), who estimates that the social returns to investments in tertiary education in South Korea are between 7 and 7.5%. Private investments would accrue an annual return of 9.1%, which is the average sales-weighted Return On Equity (ROE) we compute between 1982 and 1991. As with the benefits, we discount these values to their present value in 1992 using a 5% discount rate.

7.3 Cost–Benefit Ratio and Internal Rate of Return

We find that the G7P was a cost-effective intervention. Under the assumptions described above, we calculate that the program yielded an internal rate of return of 20.9%, with its benefits amounting to 3.3 times its costs. We conduct sensitivity analyses to assess the robustness of our results to different assumptions about opportunity costs. Under the demanding assumption that alternative private and public investments would have garnered a 15% per annum return instead of the observed pre-G7P ROE and social returns to education, we find that the benefit-to-cost ratio falls to 1.58. Even under these conditions, the G7P would still be a cost-effective intervention.

8 Conclusion

We study how South Korea's first "mission-oriented" R&D program, the G7 Program, shaped the country's innovation and economic outcomes. We establish that the program shifted the direction in which the South Korean economy innovated over the 1990s and 2000s, when South Korea caught up to the knowledge frontier. Within ten years of receipt of program support, forward-citation-weighted patenting output in the targeted technological classes had doubled relative to that in control classes. The program effects were not limited to patenting: though the export effects emerged less immediately than did those on innovation activities, real exports in targeted sectors had tripled within ten years of the targeting relative to those in control classes. These results stand when we study cross-country evidence. Technological classes with less concentrated scientific output before the program drive our results. A cost–benefit analysis that uses market-based patent valuations suggests the G7P was highly cost-effective: its benefits exceeded its costs by over a factor of three. Our results point out that the G7P had an important role in transforming South Korea's industrial economy into an innovation-driven economy.

References

Acemoglu, D., Akcigit, U., Alp, H., Bloom, N., and Kerr, W. (2018). Innovation, reallocation, and growth. *American Economic Review*, 108(11):3450–3491.

Aghion, P. and Howitt, P. (1992). A model of growth through creative destruction. *Econometrica*, 60(2):323.

Akcigit, U., Hanley, D., and Serrano-Velarde, N. (2020). Back to basics: Basic research spillovers, innovation policy, and growth. *The Review of Economic Studies*, 88(1):1–43.

Amsden, A. H. (1992). *Asia's Next giant: South Korea and Late Industrialization*. Oxford University Press.

Barwick, P. J., Kalouptsidi, M., and Zahur, N. (2023). Industrial policy implementation: Empirical evidence from china's shipbuilding industry. *Review of Economic Studies*.

Bloom, N., Hassan, T. A., Kalyani, A., Lerner, J., and Tahoun, A. (2021). The diffusion of disruptive technologies.

Bloom, N., Van Reenen, J., and Williams, H. (2019). A toolkit of policies to promote innovation. *Journal of Economic Perspectives*, 33(3):163–184.

Chang, H.-J. (1993). The political economy of industrial policy in korea. *Cambridge Journal of Economics*, 17(2):131–157.

Chen, Z., Liu, Z., Suárez Serrato, J. C., and Xu, D. Y. (2021). Notching r&d investment with corporate income tax cuts in china. *American Economic Review*, 111(7):2065–2100.

Choi, J. and Shim, Y. (2023a). From adoption to innovation: State-dependent technology policy in developing countries. Manuscript.

Choi, J. and Shim, Y. (2023b). Technology adoption and late industrialization. Manuscript.

Criscuolo, C., Martin, R., Overman, H. G., and Van Reenen, J. (2019). Some causal effects of an industrial policy. *American Economic Review*, 109(1):48–85.

Dechezleprêtre, A., Einiö, E., Martin, R., Nguyen, K.-T., and Van Reenen, J. (2023). Do tax incentives increase firm innovation? an rd design for r&d, patents, and spillovers.

Dube, A., Girardi, D., Jorda, O., and Taylor, A. M. (2024). A local projections approach to difference-in-differences event studies.

Financial Times (2024). South korea pushes to export nuclear reactors to europe. *Financial Times*.

Giorcelli, M. (2019). The long-term effects of management and technology transfers. *American Economic Review*, 109(1):121–152.

Gollin, D., Parente, S., and Rogerson, R. (2002). The role of agriculture in development. *American Economic Review*, 92(2):160–164.

Gross, D. and Sampat, B. (2023). America, jump-started: World war ii r&d and the takeoff of the u.s. innovation system.

Gruber, J. and Johnson, S. (2023). *Jump-Starting America: How Breakthrough Science Can Revive Economic Growth and the American Dream*. PublicAffairs.

Hanlon, W. W. (2018). Skilled immigrants and american industrialization: Lessons from newport news shipyard. *Business History Review*, 92(4):605–632.

Hirschman, A. O. (1958). *The Strategy of Economic Development*. Yale Univ. Press.

Howell, S. T. (2017). Financing innovation: Evidence from r&d grants. *American Economic Review*, 107(4):1136–1164.

Johnson, C. (1982). *MITI and the Japanese Miracle*. Stanford University Press.

Jordà, (2005). Estimation and inference of impulse responses by local projections. *American Economic Review*, 95(1):161–182.

Juhász, R. (2018). Temporary protection and technology adoption: Evidence from the napoleonic blockade. *American Economic Review*, 108(11):3339–3376.

Juhász, R., Lane, N., and Rodrik, D. (2023). The new economics of industrial policy. *Annual Review of Economics*.

Kalouptsidi, M. (2017). Detection and impact of industrial subsidies: The case of chinese shipbuilding. *The Review of Economic Studies*, 85(2):1111–1158.

Kantor, S. and Whalley, A. (2023). Moonshot: Public r&d and growth.

Kim, H. C. (2020). National r&d program planning and management in terms of mop in korea. Report.

Kim, S. (1996). Education and economic development in south korea. *The Korean Journal of Policy Studies*, 11.

KISTEP (2002). A study on the highly advanced national (han) program evaluation i. Report.

KISTEP (2003). A study on the highly advanced national (han) program evaluation ii. Report.

Kogan, L., Papanikolaou, D., Seru, A., and Stoffman, N. (2017). Technological Innovation, Resource Allocation, and Growth. *The Quarterly Journal of Economics*, 132(2):665–712.

Krueger, A. O. (1995). East asian experience and endogenous growth theory. In *Growth Theories in Light of the East Asian Experience*. University of Chicago Press.

Kwon, H. (2021). *Changes by Competition: The Evolution of the South Korean Developmental State*. Oxford University Press.

Kwon, S., Lee, J., and Lee, S. (2017). International trends in technological progress: Evidence from patent citations, 1980–2011. *The Economic Journal*, 127(605).

Lane, N. (2023). Manufacturing revolutions: Industrial policy and industrialization in south korea.

Lee, J. (2019). Korea patent data project (kopdp).

Liu, E. and Ma, S. (2023). Innovation networks and r&d allocation.

Lybbert, T. J. and Zolas, N. J. (2014). Getting patents and economic data to speak to each other: An 'algorithmic links with probabilities approach for joint analyses of patenting and economic activity. *Research Policy*, 43(3):530–542.

Mazzucato, M. (2013). *The Entrepreneurial State*. Anthem Press.

Ministry of Science and Technology, Republic of Korea (1991). Highly advanced national project: A plan towards 21st century. Report.

Mitrunen, M. (2021). Industrial policy, structural change, and intergenerational mobility: Evidence from the finnish war reparations. Manuscript.

Moscona, J. and Sastry, K. (2023). Inappropriate technology: Evidence from global agriculture.

Noland, M. and Pack, H. (2003). *Industrial Policy in an Era of Globalization: Lessons from Asia*. Institute for International Economics.

Rodrik, D., Grossman, G., and Norman, V. (1995). Getting interventions right: How south korea and taiwan grew rich. *Economic Policy*, 10(20):53.

Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economy*, 98(5, Part 2).

Wade, R. (1990). *Governing the Market: Economic theory and the Role of Government in East Asian Industrialization*. Princeton University Press.

World Bank (2023). World development indicators. DataBank.

World International Patent Organization (2024). Ipc pulication help.

9 Appendix

9.1 Tables

Table A1: Countries Included in the Cross-Country Analysis (Patenting)

South Korea	Great Britain	Italy	Mexico
Taiwan	Japan	Israel	Portugal
United States of America	France	Malaysia	Rest of the World
Canada	Germany	Turkey	

Table A2: Countries Included in the Cross-Country Analysis (Exports)

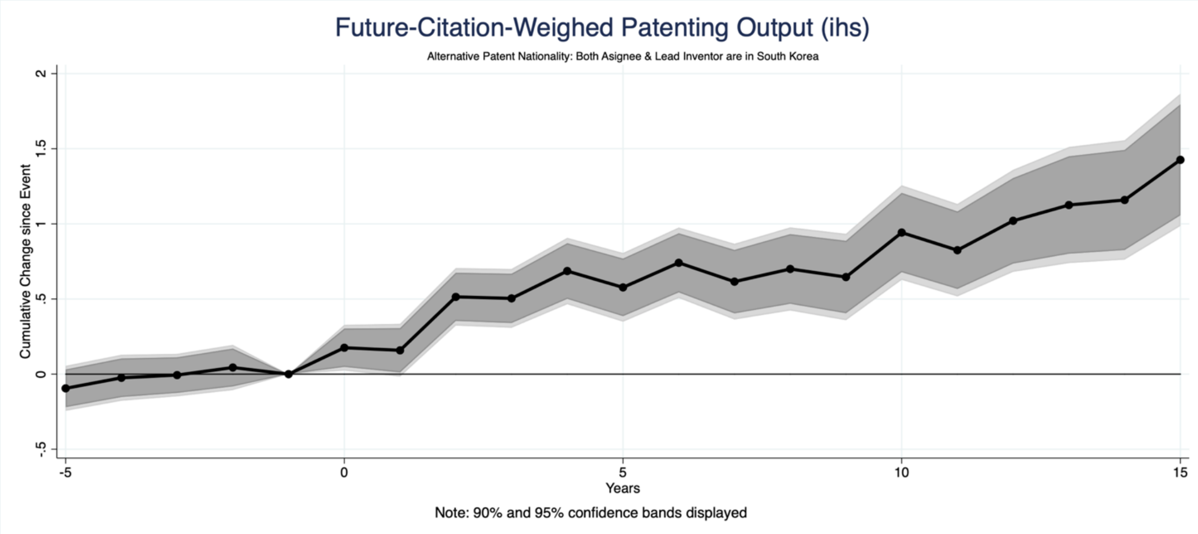
Algeria	Costa Rica	Greenland	Madagascar	Paraguay	Suriname
Angola	Cyprus	Grenada	Malawi	Peru	Sweden
Argentina	Czechoslovakia	Guadeloupe	Malaysia	Philippines	Switzerland
Australia	Democratic Republic of Yemen	Guatemala	Mali	Poland	Syrian Arab Republic
Austria	Denmark	Haiti	Malta	Portugal	Thailand
Bangladesh	Djibouti	Honduras	Martinique	Qatar	Togo
Barbados	Dominica	Hungary	Mauritius	Republic of Korea	Trinidad and Tobago
Belgium-Luxembourg	Ecuador	Iceland	Mexico	Réunion	Tunisia
Bhutan	Egypt	India	Morocco	Romania	Türkiye
Bolivia (Plurinational State of)	El Salvador	Indonesia	Nepal	Saint Kitts and Nevis	United Arab Emirates
Brazil	Ethiopia	Ireland	Netherlands (Kingdom of the)	Saint Lucia	United Kingdom of Great Britain and Northern Ireland
Brunei Darussalam	Faroe Islands	Israel	New Zealand	Saint Pierre and Miquelon	United States of America
Cameroon	Federal Republic of Germany	Italy	Nicaragua	Samoa	Uruguay
Canada	Fiji	Jamaica	Nigeria	Saudi Arabia	Vanuatu
Central African Republic	Finland	Japan	Norway	Senegal	Venezuela (Bolivarian Republic of)
Chile	France	Jordan	Oman	Seychelles	Yemen
China	French Guiana	Kenya	Other, Asia	Singapore	Yugoslavia
China, Hong Kong Special Administrative Region	French Polynesia	Kiribati	Pakistan	Solomon Islands	Zimbabwe
China, Macao Special Administrative Region	Germany	Kuwait	Panama	Spain	
Colombia	Greece	Libya	Papua New Guinea	Sri Lanka	

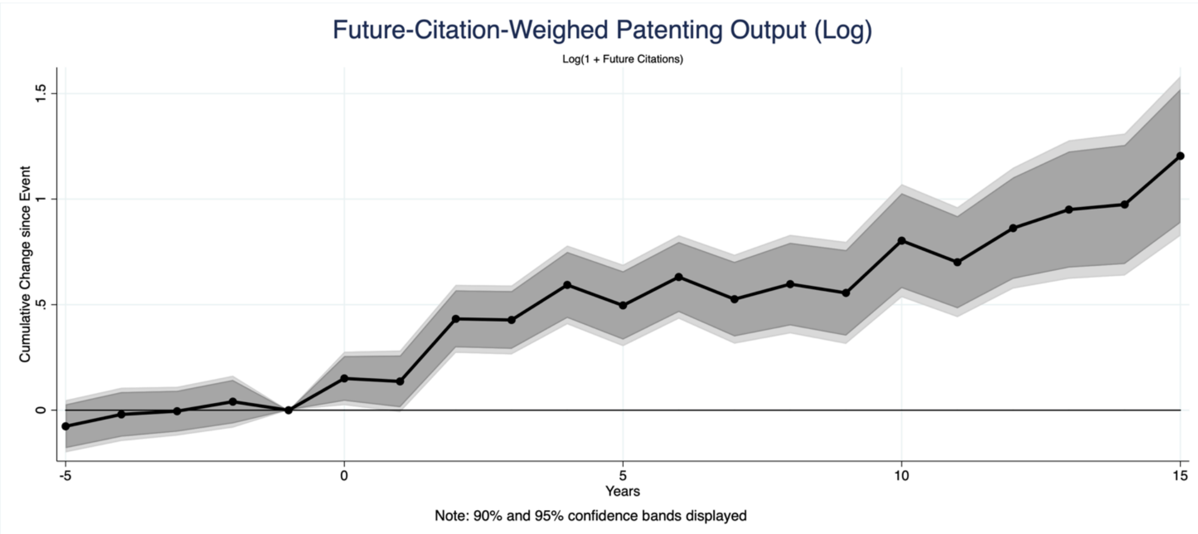
Table A3: Summary Statistics: Korean Future-Citation-Weighted Patents and Exports

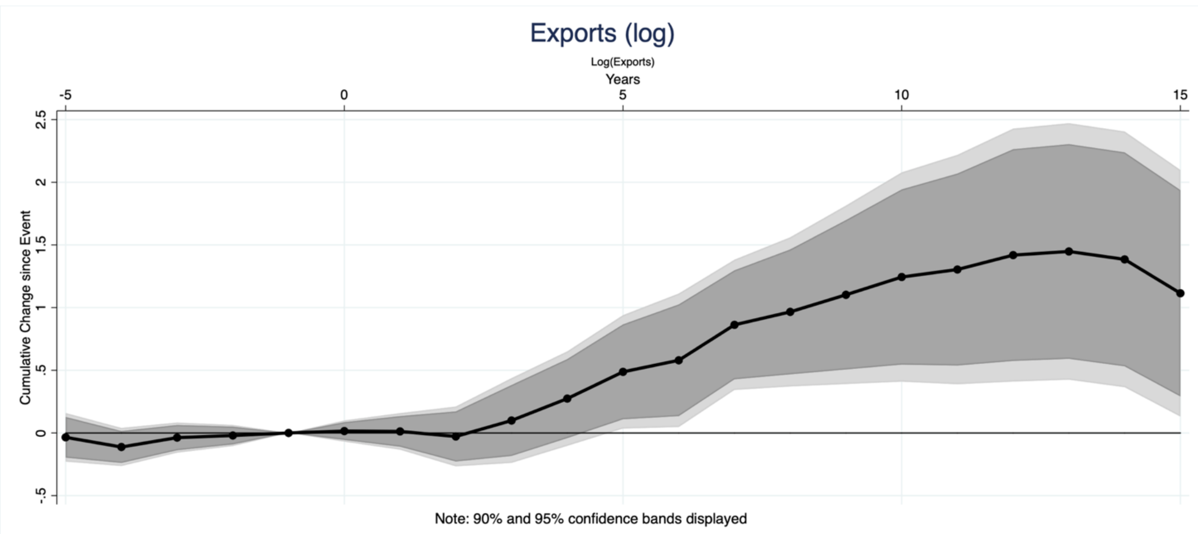
Variable	**Time Period**	**N**	**Mean**	**Standard Deviation**	**Min**	**Median**	**Max**
$ihs(patents)$	1980 - 2015	22,032	0.780	1.541	0.000	0.000	8.815
$ihs(patents)$	1980 - 1991	7,344	0.166	0.694	0.000	0.000	6.583
$ihs(patents)$	1992 - 2015	14,688	1.086	1.743	0.000	0.000	8.815
$ihs(exports)$	1980 - 2015	4,176	19.880	2.360	9.268	20.153	25.621
$ihs(exports)$	1980 - 1991	1,392	18.640	2.243	9.268	19.015	23.366
$ihs(exports)$	1992 - 2015	2,784	20.500	2.165	10.800	20.726	25.623

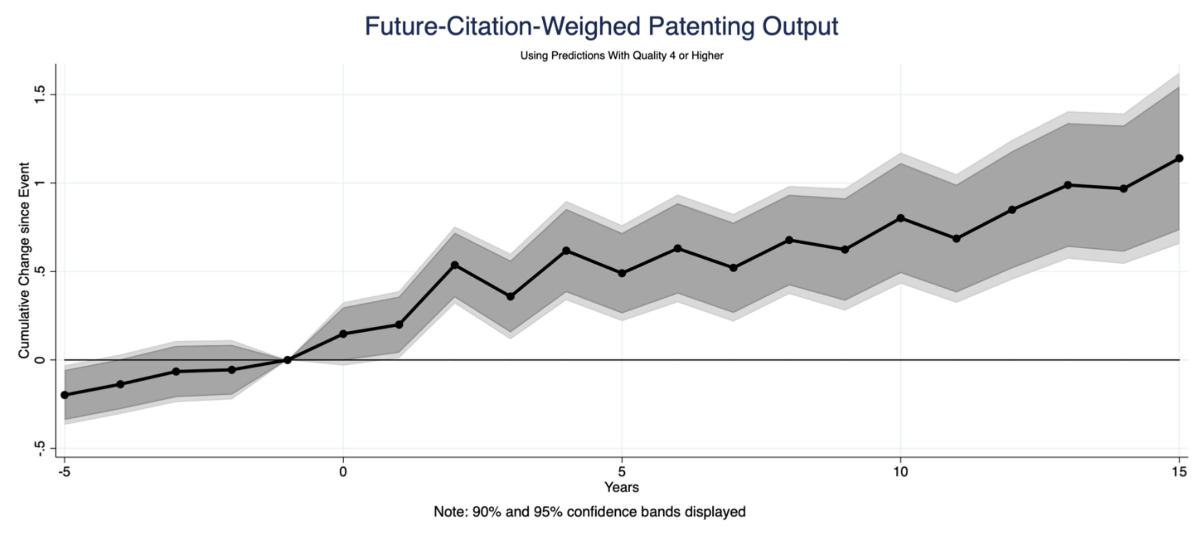
9.2 Robustness Checks

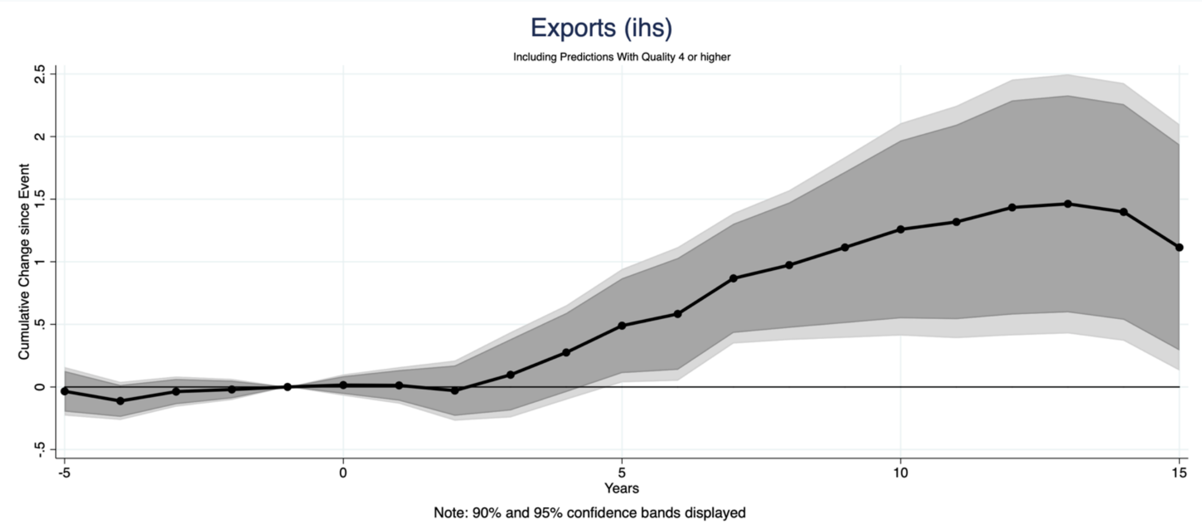
We assess the robustness of our findings to alternative definitions of patent nationality, logarithmic transformation of the dependent variable, and alternative quality thresholds in our language model exercise. All these are for the South Korean sample.

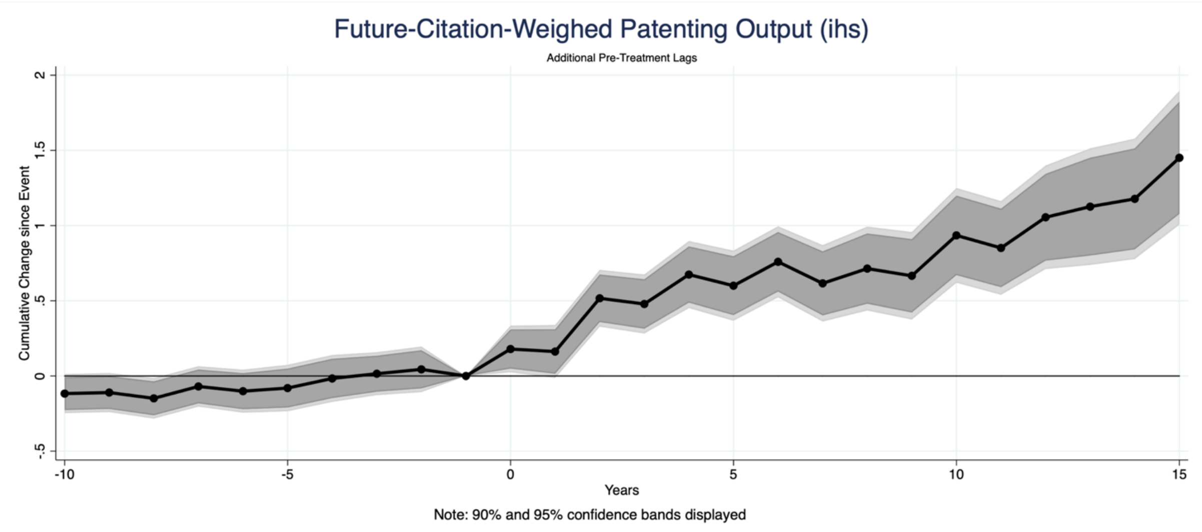

Future-Citation-Weighed Patenting Output (ihs)
Alternative Patent Nationality: Both Asignee & Lead Inventor are in South Korea
Cumulative Change since Event
-.5
0
.5
1
1.5
2
-5
0
5
10
15
Years
Note: 90% and 95% confidence bands displayed

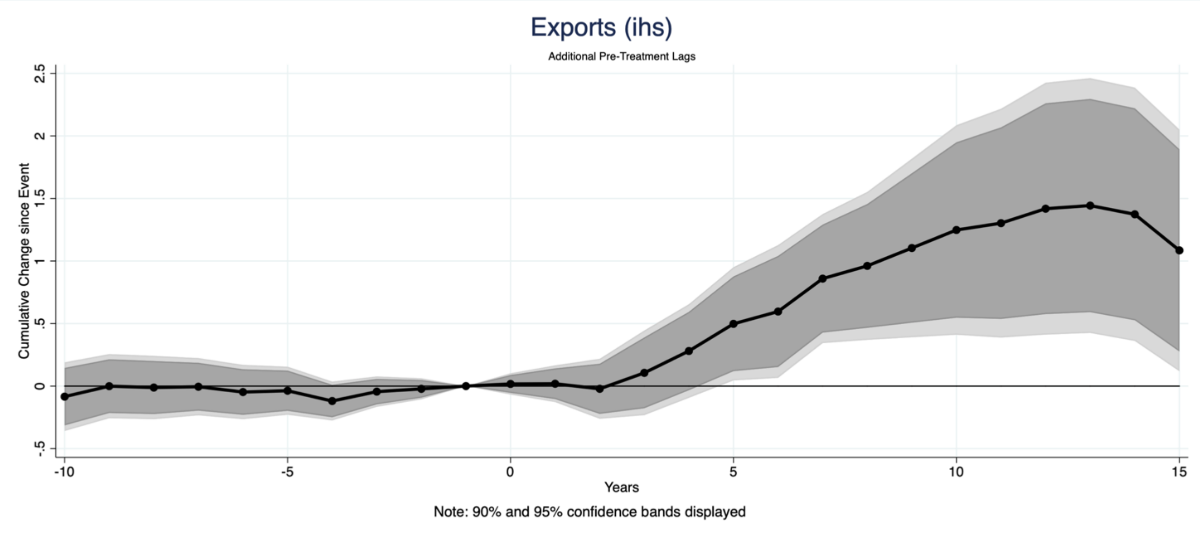

Future-Citation-Weighed Patenting Output (Log)
Log(1 + Future Citations)
Cumulative Change since Event
0
.5
1
1.5
-5
0
5
10
15
Years
Note: 90% and 95% confidence bands displayed


Exports (log)
Log(Exports)
Years
-5
0
5
10
15
Cumulative Change since Event
-.5
0
.5
1
1.5
2
2.5
Note: 90% and 95% confidence bands displayed


Future-Citation-Weighed Patenting Output
Using Predictions With Quality 4 or Higher
Cumulative Change since Event
Years
Note: 90% and 95% confidence bands displayed


Exports (ihs)
Including Predictions With Quality 4 or higher
Cumulative Change since Event
Years
Note: 90% and 95% confidence bands displayed


Future-Citation-Weighed Patenting Output (ihs)
Additional Pre-Treatment Lags
Cumulative Change since Event
Years
Note: 90% and 95% confidence bands displayed


Exports (ihs)
Additional Pre-Treatment Lags
Cumulative Change since Event
-.5
0
.5
1
1.5
2
2.5
-10
-5
0
5
10
15
Years
Note: 90% and 95% confidence bands displayed