International Students Bring Faster TFP Growth: an Extension to the Endogenous Growth Model

Yexuan Li*

2025-12-08

Abstract

Traditional endogenous growth models emphasize on the role of domestic R&D sectors, but do foreign technology absorption also contribute to domestic TFP growth? And which factors, such as international students, affect the foreign technology absorption? This paper addresses these questions through both theoretical framework and empirical analysis. We construct an extended endogenous growth model that incorporates foreign technology absorption and estimate the time series absorption rates for seven major countries from 2008 to 2017 using a rolling window approach. We then regress the estimated absorption rates on international student data and find that, for most countries, the students effect remains significant even after controlling for linear time trend. We further conduct a case study on China and show that there exists not only a significant linear time trend but also a strongly positive effect of returning international students, whereas imported books and copyrights imports are not statistically significant. Our study provides a useful paradigm for further empirical exploration of technology absorption and the causal identification of potential factors, laying foundation for future analyses.

*UC Berkeley, liyexuan25@berkeley.edu

1 Introduction

In this paper, we want to extend the classic endogenous growth model by adding an technology absorption term and estimate the time series absorption rate. Then we want to study what factors contribute to this rate. We would show that, international students lead to a positive and significant relationship with the absorption rate, while in other countries are not. However, other factors we tested such as imported books and copyrights, taking China as an example, do not show a significant result.

In the classical Solow growth model, technology progress is the only reason of long-run growth. Similarly, endogenous growth models emphasize the importance of knowledge production and human capital accumulation. Moreover, there are many factors in our lives, like international students, imported translated books and advanced foreign copyrights, which are also potential factors we want to test, influence how fast we are learning from frontier technology. This motivates us to study that, to what extent does these potential factors contribute to technological progress and, ultimately, to economic growth? The purpose of this study is to employ empirical data and econometric methods to estimate and examine the correlation between them.

Our approach is as follows. First, we combine the classical endogenous growth model with an extension that takes both domestic research and development sector (R&D) and foreign knowledge absorption into consideration. Next, we can directly implement regressions in discrete time using this equation and macroeconomic data. We use country-level panel regressions to estimate time series absorption rates of seven main countries we want to study (China, Germany, France, England, Japan, Russia and USA). Next, we try to use international students data of each country to explain the change in absorption rates, where we also add time trend to eliminate other potential time-related variables. Finally, we take China as an example to do an event study, use more detailed data such as international students who come back to work in China, imported books, software and copyrights, to test their influence.

The rest of this paper is structured as follows. There is a short chapter for literature review. In Chapter 3, I show and explain carefully what databases I use in my study. Next, I introduce our research design and show our regression functions in Chapter 4. After that I will show our regression results, including initial results, summary statistics, qualitative analyses and final results with some technical details and China event study. Finally, I give some remarks and conclusion on our study, and state our contribution to this field in Chapter 6. All supplementary tables and figures are in the appendix.

2 Literature Review

2.1 Theoretical Models about Endogenous Economic Growth

Romer (1990) gives a model about how research sectors create knowledge and increase Total Factor Productivity (TFP) and Jones (1995) expands this idea into a more general model.

Benhabib and Spiegel (1994) gives a model about technology difference driving TFP growth rate. Our model applies this idea.

Coe and Helpman (1995) studies the extent to which a country's total factor productivity (TFP) depends not only on domestic R&D capital but also on foreign R&D capital, which gives an intuition about the "spillover" effect of global innovation.

Cohen and Levinthal (1990) studies the importance of investment in R&D and absorptive capacity to firms.

In our study, we want to formalize these ideas and conduct an empirical research on "spillover" effect.

2.2 Historical Empirical Work

Abramitzky and Sin (2014) uses the collapse of Communism in Eastern Europe as a natural experiment to study "idea flows". They use book translations as a measure for it. They show that the removal of institutional barriers contributes to a sharp increase in translations from Western languages into former Communist countries, especially in applied and social sciences. They also show a convergence phenomenon that "Communist countries' translations of Western titles in the more scientific fields reached their levels in Western Europe post collapse", which also implies the effect of translation on economic growth. However, their research lacks further study about how translation contributes to technological growth, and that's what we want to explore.

Juhász et al. (2024) studies how translation helps Japan's economic development. Japan implemented many effective policies which leads to the boost of translated books. They quantify the amount of codified knowledge by industry and measure which they call "British patent relevance" (BPR) to capture the similarity of the text between industry technical manuals and British patents. It is what can explain why the Industrial Revolution first spread to Japan and not to any other non-Western country.

2.3 International Students and Technology Growth

Jonkers and Cruz-Castro (2013) reveals that returned scientists prefer co-publish internationally, contributing to the establishment of ties between the home and host systems. This study motivates us that, returned international students may also contribute to technology absorption in home country. Scellato et al. (2012, December) also shows that foreign-born researchers keep research links with colleagues in their country of origin.

However, there still lacks studies directly analyze the relationship between international students and technology growth.

2.4 Language, Translation and Trade

Melitz (2008) and Melitz and Toubal (2014) study how common native language (CNL) influence trade and use data set of series common native and spoken languages and linguistic proximity for 195 countries. Egger and Lassmann (2015) takes Switzerland as an example, doing a spatial RDD to confirm the effect of CNL on import.

Amano et al. (2016) showed how largely English dominates academic languages. They search Google Scholar in 16 languages and reveal that 35.6% of 75513 scientific documents on biodiversity conservation published in 2014 were not in English.

Yang et al. (2025) studies how localization of steam games driven by translation affects consumers' gains.

Merali (2024) shows that workers using Large Language Models (LLMs) for translation tasks significantly enhance their productivity, which is a huge shock to translation industry. This can be used as an IV to do causal inference.

However, few studies have directly examined the role of these factors we mentioned in explaining productivity or TFP growth. In this paper, we explicitly estimate time series absorption rate and test these factors. This quantitative work hasn't been fully studied before.

3 Data

In this chapter we not only introduce databases we use in our study, but also show some relative and useful data and explain why they are not so helpful in our study.

3.1 Macroeconomic Data

TFP Growth. Feenstra et al. (2025) gives us a full data set of country-level TFP and many other macroeconomic indicators like GDP per capita in Penn World Table (PWT). We mainly focus on the TFP data, and there are 185 countries ranging from 195 to 2023, totally 13690 observations. Here we have "ctfp" data calculated at current PPPs (USA=1) and "rtfpna" data for TFP at constant national prices (2021=1). Some data is lost but we only focus on seven main countries in the world. Actually, the base point chooses USA doesn't matter because, as we can see later, we can do some simple calculation to change base time point and base country. Besides TFP data, we also use the same database for population data.

Number of Researchers. From UNESCO UIS Statistics we can find researchers (in full-time equivalent) per million inhabitants. Same as above, this data ranges from 2000 to 2023 and lost in some years for some countries. This is important because researchers are important parts of technology growth. Some papers also explain this index as "Human Capital".

3.2 Data for International Students and Other Factors

UNESCO UIS Statistics. Here we can find the number of students studying abroad from 2000 to 2023, including total number from one particular country and origin and destination district. Unfortunately, this data set cannot provide more specific data about students from one country to another, which means it gathers all international students together. Moreover, the starting point is too late to fit other data sets. As a result, our data begins from 2000 and lack enough data points. This is a potential problem, however, as our results show, main coefficients are statistical significant.

Specific Data in China. We use data from the "China Statistical Yearbook" from 2009 to 2018, reporting returned international students and imported number of books and copyrights. In Appendix A we show three pictures in 2017 book. We use "Number of Returned Students" as an accurate estimation, since returned students contribute the most to technology growth in home country, which is a better index than what we get from UNESCO UIS Statistics. Then we use "Number" of "Imported" "Natural Science and S&T" books in figure 14, which directly relates to TFP growth. Here we choose to use the number of copies rather than kinds or monetary values because values are influenced by prices, while the number of kinds cannot capture spreading effects. For imported copyrights, we sum data of "Books", "Electronic Publications" and "Software" of USA, GBR, DEU, FRA, RUS and JPN as an estimation of imported technology materials. This is how we clean these data and apply to our study.

Index Translationum. UNESCO (n.d.) provides data about published translated books by countries and original languages from 1978 to 2012. There are 210 languages and less than 100 countries (number of countries varying every year). This data set is the only one I can find online directly related to the number of translated books, but some data is lost and seems not very reliable.

3.3 Other Useful Data

Industry Data in USA. From Fred we also get many databases. U.S. Bureau of Labor Statistics (2025a) provides data titled "Consumer Price Index for All Urban Consumers: Educational Books and Supplies in U.S. City Average", U.S. Bureau of Labor Statistics (2025b) for "Producer Price Index by Industry: Book Publishers: Book Publishing from Print Publishers", U.S. Bureau of Labor Statistics (2025c) for "Producer Price Index by Industry: Books Printing", U.S. Bureau of Labor Statistics (2025d) for "Producer Price Index by Industry: Printing Ink Manufacturing: Publication and Commercial Web Inks", and U.S. Census Bureau (2025) for "Total Revenue for Translation and Interpretation Services, Establishments Subject to Federal Income Tax, Employer Firms". What's more, IBISWorld provides data about translation services in the US which can be downloaded at their website. This database includes items like revenue, enterprises, employment and wages. However, these databases mentioned above are all about USA, thus can only be used to do event study. What's worse, these data are not a good estimation for absorption from foreign countries, so we do not use them.

Language Distance. Melitz and Toubal (2014) gives us a data set of language proximity. We can download the table from CEPII's website. We can use these data to measure the difficulty of absorbing knowledge from another country, which can be formalized as weights in different absorption terms from different countries. They gather common official, spoken and native language and adjusted value of linguistic proximity. There are 195 countries corresponding to 194 languages totally 37830 observations.

4 Research Design

The data sets available for this study are of limited quality. To address measurement problems, I plan to apply floating window method to estimate time series of technology absorption rates.

To use regression analysis to study the impact of technology absorption on TFP growth, we first estimate the technology absorption rate and then, in the second step, test its relationship with potential factors. Next, we will briefly introduce the regression functions.

4.1 Endogenous Growth Model

According to former studies about the form of technology growth function, we combine R&D sector and absorption from other countries effect like this:

$$\dot{A}_i = \underbrace{\phi_i H_i A_i}_{\text{domestic knowledge creation}} + \underbrace{H_i \sum_{j \neq i} \theta_{ij} \dot{A}_j}_{\text{foreign knowledge absorption}} , \tag{1}$$

where A_i denotes TFP level in country i, H_i for researchers, ϕ_i for domestic knowledge creation rate and θ_{ij} is the absorption rate we want.

The intuition for why we choose this function form is that, each country learn from foreign countries frontier knowledge. We assume that, we can decompose TFP into two parts, one for common and general knowledge and the other for specific and advanced knowledge, absorbed by other countries. We also tried other forms, for example, $\sum_{j \neq i} \theta_{ij}(A_j - A_i)\mathbb{1}\{A_j - A_i > 0\}$. However, this form cannot capture the idea that each country can develop their own technology and share with other countries, and its appearance in statistical significance is not satisfactory. Other form like directly sum on foreign TFP level, $\sum_{j \neq i} \theta_{ij} A_j$, is also excluded. For more details, please refer to Appendix B.

Following the theoretical framework, we can estimate the absorption rate of foreign knowledge using TFP data. The regression function can be rewritten as:

$$\Delta A_{i,t} = \beta_0 + \beta_{1i,t} H_{i,t} A_{i,t} + H_{i,t} \sum_{j \neq i} \beta_{2ij,t} \Delta A_{j,t} + \gamma_t + \varepsilon_{i,t}, \tag{2}$$

where:

- $A_{i,t}$: TFP level of country i at time t.
- $\Delta A_{j,t}$: Change of TFP level of country j from time t to $t+1$.
- $H_{i,t}$: Number of researchers of country i at time t.
- $\beta_{1i,t}$: Domestic knowledge creation rate of country i at time t.
- $\beta_{2ij,t}$: Technology absorption rate of country i from j at time t.
- γ_t: Time trend fixed effects.
- $\varepsilon_{i,t}$: Error term.

Since the definition of TFP data in PWT, we need to clean them according to the equation below:

$$A_{i,t} = \frac{rtfpna_{i,t} \times ctfp_{i,2021}}{rtfpna_{USA,2000}} = \frac{a_{i,t}}{a_{i,2021}} \times \frac{a_{i,2021}}{a_{USA,2021}} \times \frac{a_{USA,2021}}{a_{USA,2000}} = \frac{a_{i,t}}{a_{USA,2000}}, \tag{3}$$

where we use $a_{i,t}$ to denote the absolute TFP level of country i at time t. Thus, the $A_{i,t}$ we use is normalized by $A_{USA,2000}$. Our goal is to choose base country and base year for panel comparison. The reason for choosing 2000 instead of 2021 is that we do not want TFP to have a time convergence trend to 1.

There is a huge problem, that is each country only has 18 years observation, we are not able to test source country heterogeneity. For simplicity, we assume each source country has the same weight, that is:

$$\Delta A_{i,t} = \beta_0 + \beta_{1i,t} H_{i,t} A_{i,t} + \beta_{2i,t} H_{i,t} \sum_{j \neq i} \Delta A_{j,t} + \gamma_t + \varepsilon_{i,t}. \tag{4}$$

If we run regression on each country one by one, then we have 18 observations for 3 parameters. If we further keep the linear time trend, then would be 4 parameters.

4.2 Factors Regression

In the second step, we just run a simple linear regression to test the relationship between potential factors and the change in absorption rate. Taking international students as an example:

$$\beta_{2i,t} \;=\; \delta_0 + \delta_1 \mathit{stud}_6 + \gamma_t + \epsilon_i. \tag{5}$$

Here we also allow for linear time trend γ_t.

5 Initial Analysis

5.1 Summary Statistics

Figure 1 reports the summary statistics used in this study. The sample consists of seven countries, including China (CHN), Germany (DEU), France (FRA), the United Kingdom (GBR), Japan (JPN), Russia (RUS), and the United States (USA). This is a panel dataset covering the years from 2000 to 2017.

The variable *pop* denotes the population of each country, measured in millions. *rtfpna* represents the total factor productivity (TFP), defined as TFP at constant national prices (2021=1),

Variable	Obs	Mean	Std. Dev.	Min	Max
countrycode	0				
country	0				
year	126	2008.5	5.208839	2000	2017
pop	126	304.0283	432.2008	59.05734	1412.355
rtfpna	126	0.9289817	0.128159	0.5124558	1.078543
ctfp	126	0.8030526	0.2754428	0.2795744	1.277463
rsrch	126	3499.648	1250.595	549.5748	5352.831
H	126	0.6158425	0.4132705	0.1708751	1.745458
stud	126	122235.8	186448.4	21266	930743
stud_3	112	365158.2	543077.4	68806	2620123
stud_6	91	725566.4	1044940	139214	4790194
newtfp	126	0.8290753	0.2805231	0.2718622	1.161301
newdelta_A	126	0.0047743	0.0124231	-0.0493895	0.0355718
newdomes	126	0.4627651	0.3287506	0.1639725	1.41605
newS	126	0.0286457	0.0502369	-0.01792287	0.0927946
newspill	126	0.0150939	0.0430456	-0.02862944	0.1072934

Figure 1: Summary Statistics Table

while *ctfp* is TFP data calculated at current PPPs (USA=1). *rsrch* measures the share of researchers per million people. Thus, the product of *pop* and *rsrch* gives *H*, the total number of researchers measured in millions. Besides that, *stud* is the number of international students from each country, measured in individuals. On the purpose of capturing accumulative effect, we also construct two variables *stud_3* and *stud_6*, which means the sum of international students of the last three years and the last six years.

The last five variables are the main regression variables derived from the above statistics. *newtfp* is TFP data calculated by equation 3. *newdelta_A* denotes the year-to-year change in TFP (next year minus this year, to capture forward effect). *newdomes* captures the domestic technology level, defined as $H \times$ *newtfp* . *newS* is an aggregation parameter representing the sum of TFP differences of all foreign countries, regardless of sign. Finally, *newspill* measures the level of innovation spillovers from abroad, calculated as *newS* $\times$ H.

5.2 Qualitative Analysis

As a starting point, we conduct a simplified qualitative analysis. Here we remove the time subscript and treat each country's 18 observations from 2000 to 2017 as independent data points. Then we directly estimate the domestic innovation parameter β_{1i} and technology absorption parameter β_{2i}, assuming that these parameters remain constant over this period.

$$\Delta A_i \;=\; \beta_{0i} + \beta_{1i} H_i A_i \;+\; \beta_{2i} H_i \sum_{j \neq i} \Delta A_j + \epsilon_i. \tag{6}$$

Using our statistics in Figure 1, we can further simplify the equation as below and do regressions for each country:

$$\textit{newdelta\_A} \;=\; \beta_0 + \beta_1 \textit{newdomes} \;+\; \beta_2 \textit{newspill} + \gamma_t + \epsilon_i. \tag{7}$$

This qualitative analysis allows us to examine whether the two parameters are statistically significant, to observe their signs, and whether any time trend exists.

To test for a potential time trend, we include a linear time variable, denoted as *time_trend*, representing the annual index from 2000 to 2017, which only takes one degree of freedom:

$$\textit{delta\_A} \;=\; \beta_0 + \beta_1 \textit{domes} \;+\; \beta_2 \textit{spill} \;+\; \beta_3 \textit{time\_trend} + \epsilon_i. \tag{8}$$

Figure 2 presents the qualitative regression results corresponding to Equation 8. After viewing our results, we find that adding the *time_trend* variable can make β_2 extremely significant for almost all seven countries, while some countries like CHN and USA are not so significant. However, for USA the significance of *time_trend* substitudes the insignificance of β_2, and also show a very significant domestic knowledge creation.

ΔA	domestic		spill		time_trend	
	β_1	p>|t|	β_2	p>|t|	β_3	p>|t|
CHN	0.1340975*	0.068	0.0736589*	0.089	-0.0049847*	0.078
DEU	-0.1835221	0.226	1.055424***	0.000	0.0028485	0.167
FRA	-1.353142**	0.033	0.5674181***	0.000	0.0100564**	0.037
GBR	-0.01930771***	0.008	0.8468054***	0.000	0.0009699**	0.025
JPN	-0.2646052**	0.033	0.229047***	0.000	0.0006712**	0.030
RUS	0.0550646	0.085	0.5711799***	0.007	-0.0012357	0.257
USA	0.140723***	0.001	0.0149185	0.332	-0.0035587***	0.000

*Note: the statistical significance is defined as (i) *** if p-value is not larger than 0.01; (ii) ** if p-value is larger than 0.01 but not larger than 0.05; (iii) * if p-value is larger than 0.05 but not larger than 0.1.*

Figure 2: Qualitative Analysis of Absorption Rates, Controlling for Linear Time Trend

As shown, except for China and USA, the absorption coefficients are highly significant for all countries, and their signs are all positive, which can be interpreted that the more technology growth in other countries, the higher TFP growth home country has. Since we have recalculated the explanatory variables, it's hard to give a mathematical interpretation. The United States shows positive and significant β_0, which can be interpreted that the more researchers or higher TFP level, the higher TFP growth USA has.

However, the coefficients for the domestic innovation term are generally less significant. What's worse, some of them are negative. Possible explanations are, the data set is too small, and strong absorption effect dominate the technology growth. But this is still an interesting question, although it's not our key topic.

5.3 Time Series Estimation for Technology Absorption Rate

Next We employ a rolling-window method to estimate each country's time varying technology absorption rate. Specifically, we use six-year and nine-year rolling windows, meaning that for each estimation we rely on six or nine continuous annual observations to estimate the coefficients of domestic innovation and technology absorption. Both parameters are thus time series with country heterogeneity.

We here show the time series graph of nine-year rolling window, for example, using 2000-2008 observations to estimate $\beta_{2i,2008}$. Figure 3 displays the estimated technology absorption rates for China. In this figure, we plot both the estimated values and one standard deviation range above and below the estimates. For China, the absorption rate is significantly positive only in 2008, while shows a decreasing pattern till 2013. However, it's not so significant after 2008 except 2013.

In contrast, Figure 4 shows the results for Germany, where the estimated absorption rate remains significantly positive from 2008 to 2017 with estimation results larger than 0.6, and this rate remains a positive constant in a long time period through 2009 to 2016. We also show the results for all other countries.

However, a new question arise: why most countries show an abnormal point in 2017, the last year? Possible explanations are statistical edge effect by floating window methods or all the countries suffer from 2008 financial crisis, since nine-year estimation result in 2017 is reached

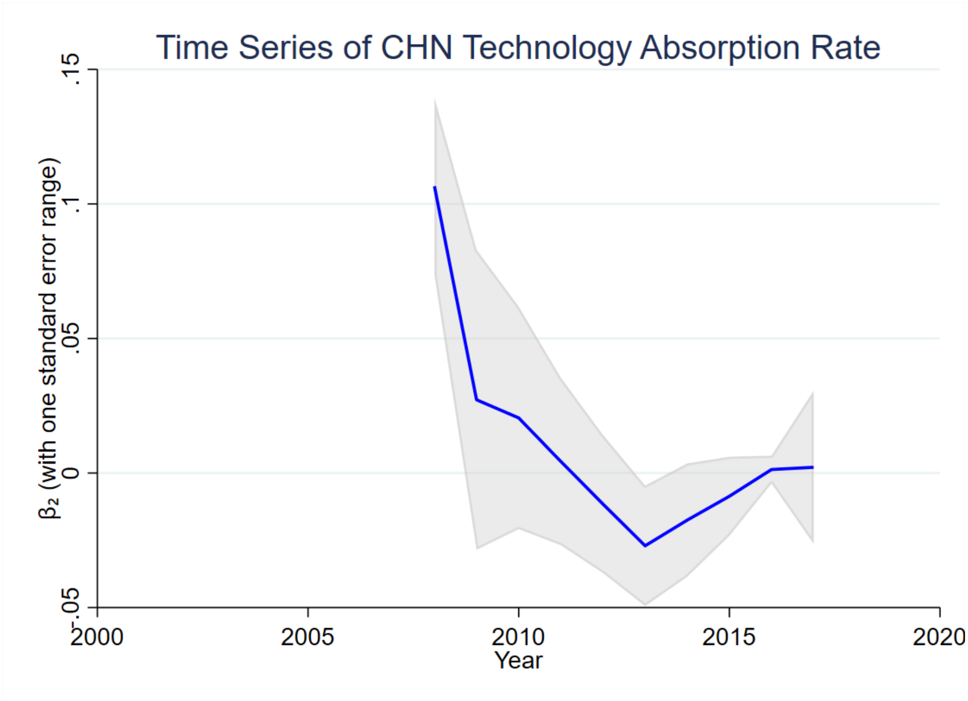

Figure 3: Time Series TFP Absorption Rate of CHN

using data from 2009 to 2017.

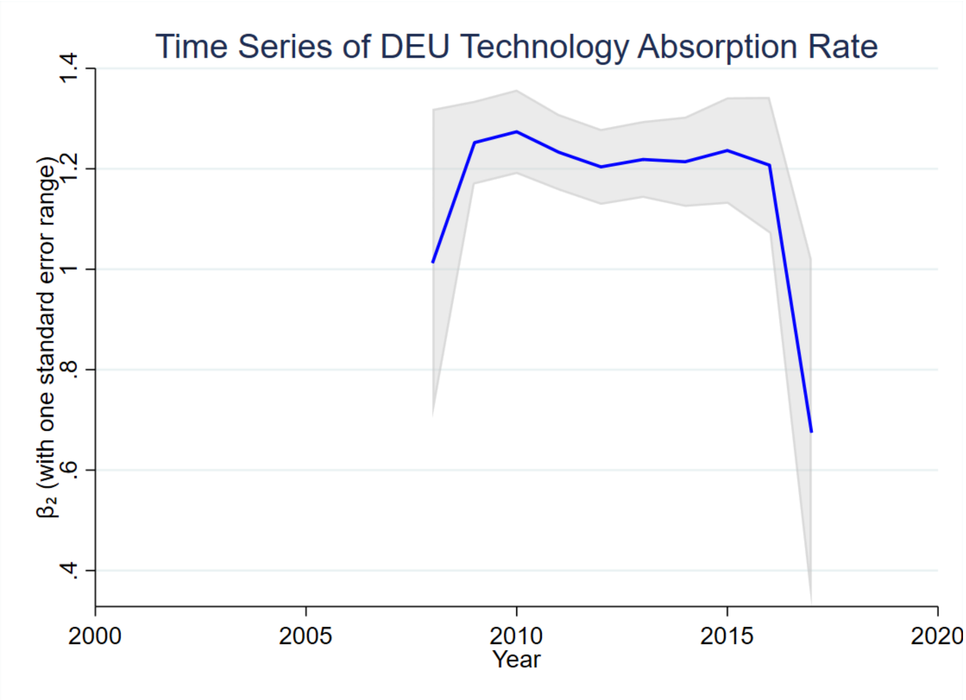

Figure 4: Time Series TFP Absorption Rate of DEU

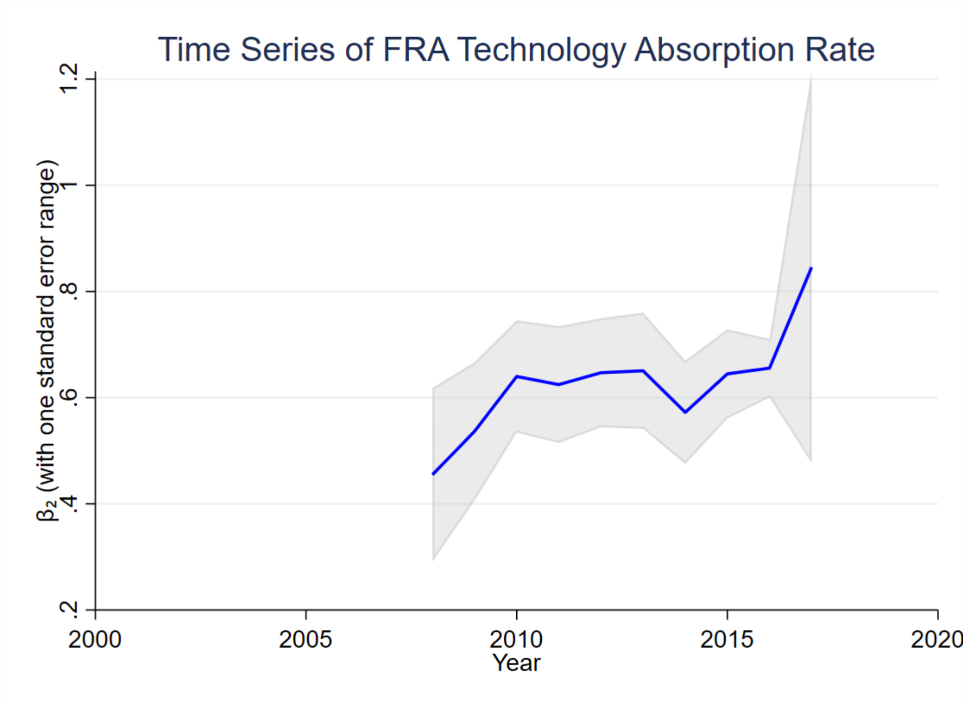

Figure 5: Time Series TFP Absorption Rate of FRA

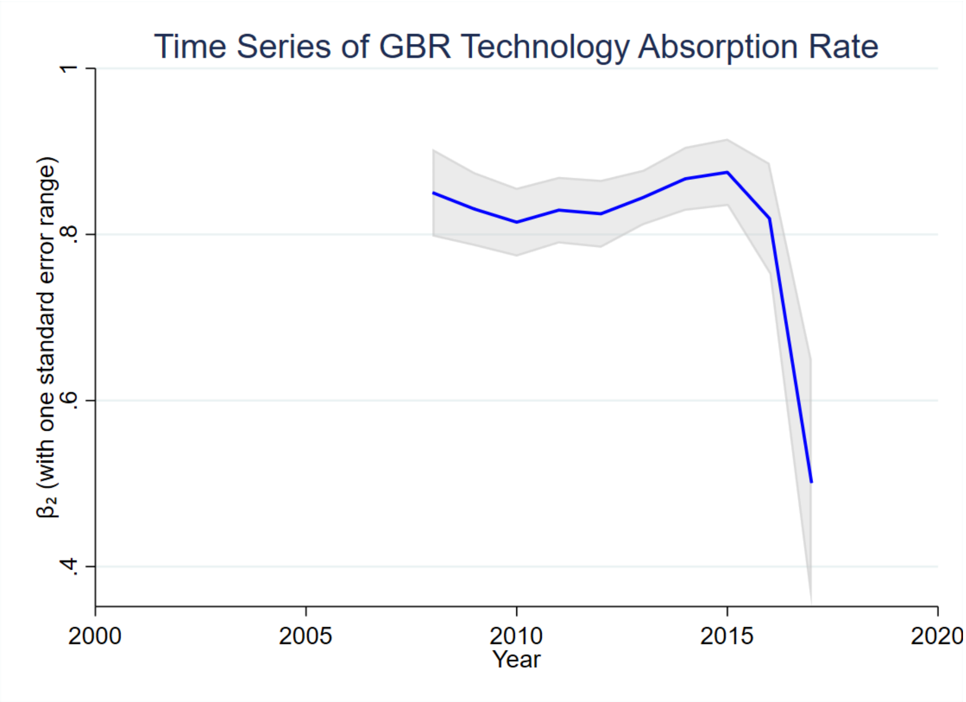

Figure 6: Time Series TFP Absorption Rate of GBR

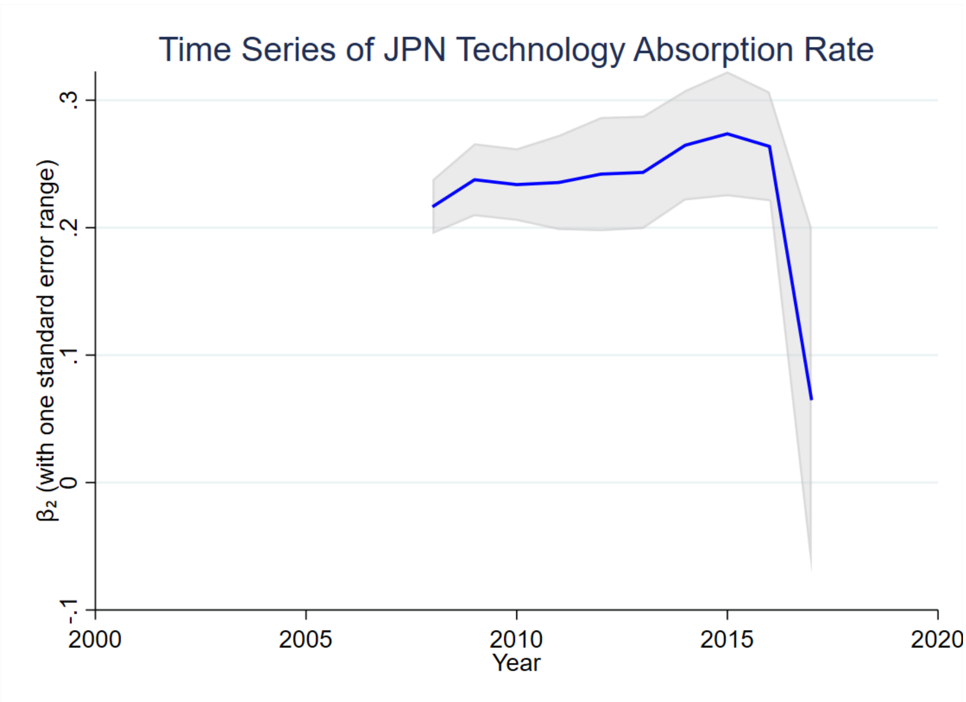

Figure 7: Time Series TFP Absorption Rate of JPN

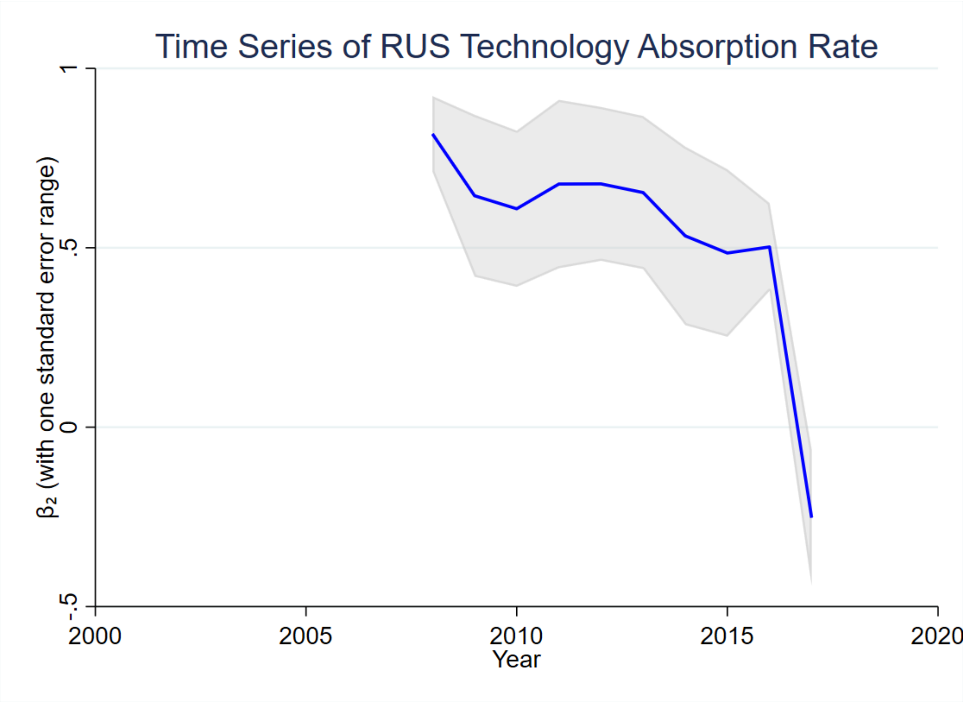

Figure 8: Time Series TFP Absorption Rate of RUS

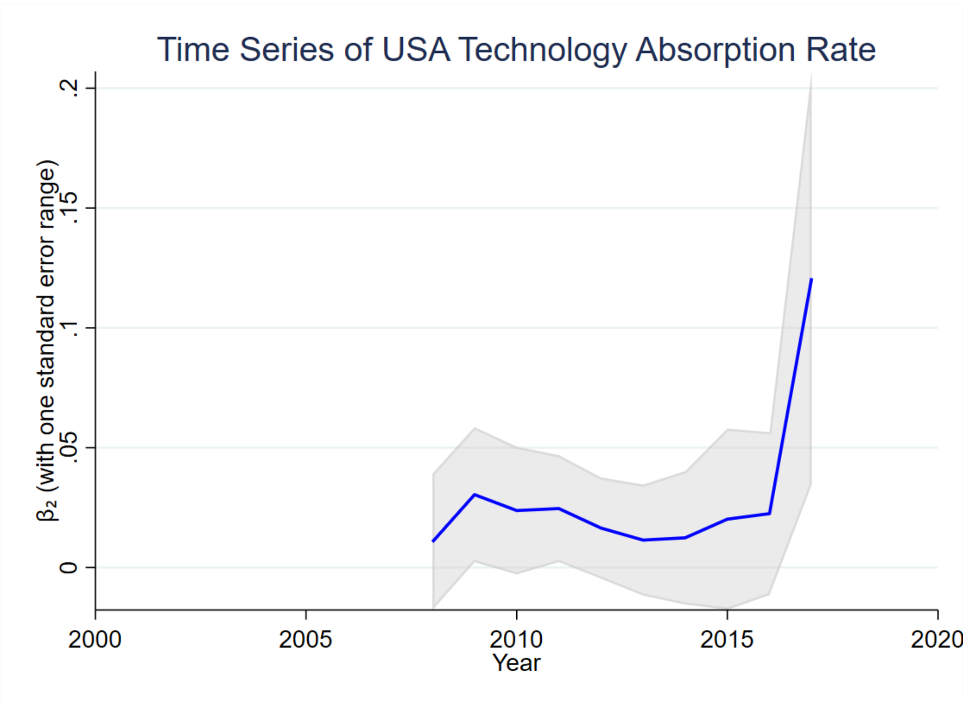

Figure 9: Time Series TFP Absorption Rate of USA

6 Results

6.1 Factors Statistics

Our initial explanatory variable was the national translation index. We employed the translation statistics provided by UNESCO and plot them into the figure shown below.

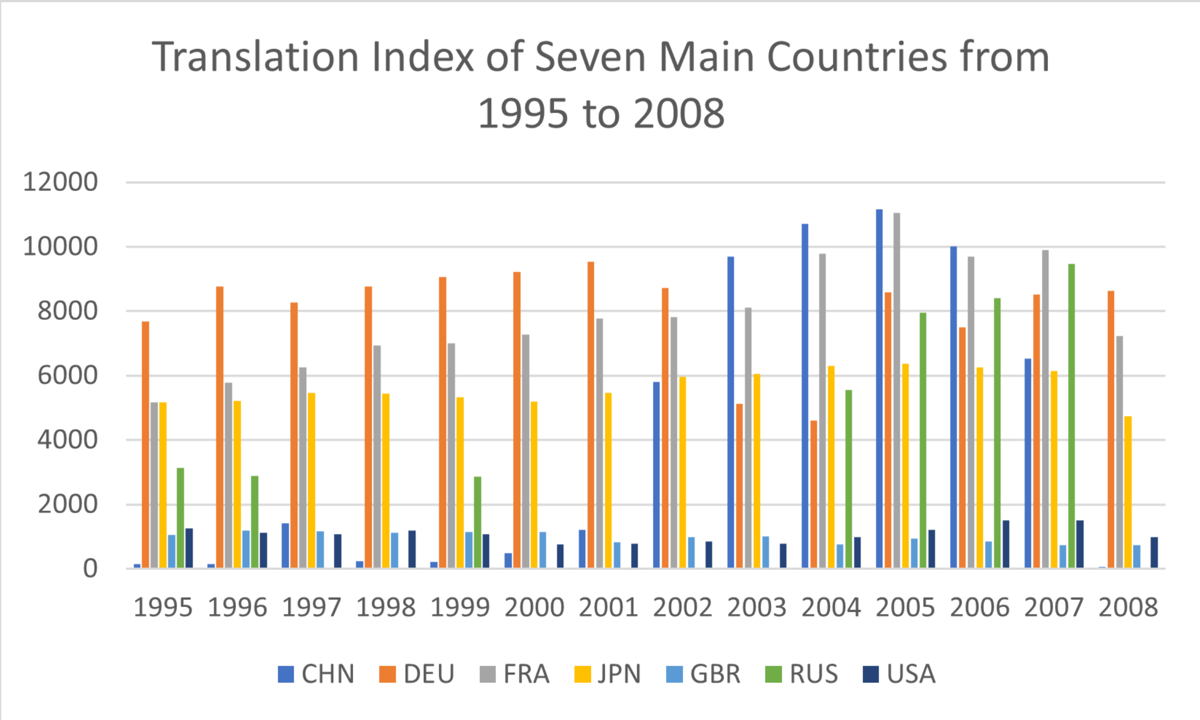

Figure 10: Translation Index

In constructing this data set, we excluded years with poor data quality and ultimately retained translation indexes for seven major countries from 1995 to 2008. Even within this period, however, some countries exhibit substantial fluctuations in the translation index, which we attribute primarily to measurement inaccuracies. In contrast, countries such as Germany, France, the United Kingdom, the United States, and Japan display relatively stable trends.

Nevertheless, since the translation data end in 2008, while our six-year parameter estimations of $\beta 2i, t$ begin in 2005, it is difficult to directly regress the absorption rates on the translation index with so few observations.

6.2 Estimation International Students Effect

As an alternative, we considered the number of international students as a second explanatory factor. This variable is available for the period 2000—2023, which aligns well with our estimated coefficients.

In Figure 11, the correlation coefficients for DEU and RUS are not statistically significant, and the coefficient for USA is only weakly significant. The other four countries exhibit relatively strong significance. Among them, the coefficients for CHN, FRA, RUS, and the USA are positive, which can be interpreted that controlling for other effects, in average we expect one more international student (measured as the six-year cumulative number) contributes to δ_1 increase in foreign technology absorption rate. We also find that the countries showing significant correlation coefficients also display significance in the linear time trend. Compared with the regression results without the time trend, including time trend indeed leads to better performance.

beta_2	stud_6		time trend	
	δ_1	p>|t|	δ_2	p>|t|
CHN	6.01e-07***	0.001	-0.1836888***	0.001
DEU	-2.36e-07	0.855	0.0307404	0.244
FRA	5.73e-06**	0.050	-0.2263085**	0.047
GBR	-7.59e-06	0.264	0.0227071	0.394
JPN	-5.60e-06**	0.038	-0.1083032**	0.044
RUS	0.0000439**	0.012	-0.5537361***	0.009
USA	6.75e-06*	0.067	-0.0850358*	0.070

*Note: the statistical significance is defined as (i) *** if p-value is not larger than 0.01; (ii) ** if p-value is larger than 0.01 but not larger than 0.05; (iii) * if p-value is larger than 0.05 but not larger than 0.1.*

Figure 11: Estimation of International Students Effect on Technology Absorption Rates

6.3 Event Study: China

Next, we conduct an event study using more detailed data from China. Because the data obtained from the UNESCO UIS Statistics are not of high quality, some results turned out to be insignificant or even negative, which goes beyond our expectations. Therefore, we use data from the "China Statistical Yearbook" instead. The data on returned international students are available for the years from 2000 to 2017, thus fully matching all observation points for the absorption rates. However, the data on imported books are available only from 2008 to 2017, and the data on imported copyrights only from 2010 to 2017, which may reduce the number of observations available for regression. Moreover, we cannot test cumulative effects for imported books and copyrights, as doing so would further decrease the number of observations. Nevertheless, to maintain consistency, we adopt a six-year cumulative measure for returned international students.

Moreover, given that both the technology absorption rate and the factors exhibit clear time trend patterns (as we can see in Figure 3–9), direct regression may lead to spurious correlation. To identify the true relationship between the two variables, we include a linear time trend to control for their common trend.

As in Figure 12, we run four sets of regressions. In each set we estimate with and without a linear time trend. In the first three sets, we include one factor at a time, while the final set includes all factors simultaneously.

beta_2	CHNstud_6		books		copyrights		time trend	
	β_1	p>|t|	β_2	p>|t|	β_3	p>|t|	β_3	p>|t|
CHN	-2.70e-08	0.143						
	2.06e-07***	0.001					-0.0578883***	0.001
			0.0000494	0.842				
			-0.0006471	0.117			-0.0123053**	0.045
					-3.56e-06	0.546		
					-4.01e-06	0.716	0.0002492	0.954
	2.48e-08	0.488	0.0002209	0.685	-0.0000129	0.364		
	2.42e-07**	0.025	-0.0001196	0.804	9.05e-06	0.535	-0.0740161*	0.053

*Note: the statistical significance is defined as (i) *** if p-value is not larger than 0.01; (ii) ** if p-value is larger than 0.01 but not larger than 0.05; (iii) * if p-value is larger than 0.05 but not larger than 0.1.*

Figure 12: Estimation of Abroad Students' Effect on Technology Absorption Rate

From the figure, we observe that when only returned international students are included, China exhibits a strong linear time trend, and the returned students display a significantly positive effect. This can be also interpreted as that controlling for other effects, in average we expect one more international student coming back to China (measured as the six-year cumulative number) contributes to 2.06e-07 increase in foreign technology absorption rate. However, when we regress only on imported books, there exists a relatively strong time trend, but the effect of imported books itself is not significant. The effect of imported copyrights is entirely insignificant, regardless of whether a time trend is included. When we include all three variables together, we still find a significantly positive effect of returned international students, along with a weakly significant time trend.

Our explanation is that returned international students indeed exhibit a significantly positive effect on domestic technology absorption, and this effect is even stronger than that of international students in the UNESCO UIS database since not all of them come back yet. In contrast, imported books and imported copyrights do not exhibit significant effects, possibly because this data set does not allow us to more precisely identify books or copyrights specifically related to technological absorption. A more important reason is that even if they are related to science and technology, these printed, approved, and imported books and copyrights mostly reflect existing technologies that are already outdated, making it difficult for them to capture the absorption of frontier technologies. In this sense, returned international students do bring back frontier knowledge and technology, which leads to faster TFP growth.

7 Conclusion

This study employs rich data and combines both theoretical extension and empirical analysis. We begin by extending the classical endogenous growth model, emphasizing the role of foreign technology absorption. We then transform the model into a regression framework and use data together with a floating window approach to estimate the time series of technology absorption rates for seven major countries. Next, we examine the impacts of potential factors on technology absorption rates, finding that international students exhibit significance and positive effects in many countries. We further conduct an event study using more detailed data from China and find that returned international students have a significantly positive effect on technology absorption rate, whereas imported books and copyrights do not show significant effects.

This paper makes several contributions. Firstly, we explicitly present an extended endogenous growth model and use it for empirical testing. Our results reveal a significant and positive relationship between foreign technology absorption and TFP growth, demonstrating the validity of the extended model. In addition, we provide rich evidence showing the significant relationship between international students and technology absorption rates. For data, we estimate the time series of technology absorption rates for seven countries from 2008 to 2017, showing their growth patterns during this period and providing an example for further empirical estimation.

Our study also leaves several potential issues for future analyses. Some arise from the statistical results：Why are the domestic knowledge creation rates negative for some countries? Why are the foreign absorption rates negative? Although these estimates are statistically significant, their negative signs require further interpretations. Moreover, our empirical strategy can be improved in several ways. For instance, the number of observations within our floating window estimation is so few, which may raise concerns about robustness, and we assume that all foreign countries share equal weights. Future research may differentiate the weights of different foreign technology absorption using measures such as language distance or absorption structure. Additionally, the seven major countries we select do not represent the whole world. For example, in the case of China’s imported books, Canada clearly outweighs Russia, yet Canada is not included in our regression though this factor is statistically insignificant. Last but not least, our study provides only estimation results and correlation analysis and still lacks causal identification, which requires further work.

References

Abramitzky, R., & Sin, I. (2014). Book translations as idea flows: The effects of the collapse of communism on the diffusion of knowledge. *Journal of the European Economic Association*, *12*(6), 1453–1520.

Amano, T., González-Varo, J. P., & Sutherland, W. J. (2016). Languages are still a major barrier to global science. *PLoS biology*, *14*(12), e2000933.

Benhabib, J., & Spiegel, M. M. (1994). The role of human capital in economic development evidence from aggregate cross-country data. *Journal of Monetary Economics*, *34*(2), 143–173. https://doi.org/https://doi.org/10.1016/0304-3932(94)90047-7

Coe, D. T., & Helpman, E. (1995). International R&D spillovers. *European Economic Review*, *39*(5), 859–887. https://doi.org/https://doi.org/10.1016/0014-2921(94)00100-E

Cohen, W. M., & Levinthal, D. A. (1990). Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly*, *35*(1), 128–152. Retrieved October 5, 2025, from http://www.jstor.org/stable/2393553

Egger, P. H., & Lassmann, A. (2015). The causal impact of common native language on international trade: Evidence from a spatial regression discontinuity design. *The Economic Journal*, *125*(584), 699–745. Retrieved October 4, 2025, from http://www.jstor.org/stable/24737103

Feenstra, R. C., Inklaar, R., Timmer, M. P., Roser, M., Rohenkohl, B., Arriagada, P., Hasell, J., Ritchie, H., & Ortiz-Ospina, E. (2025). *Total factor productivity – penn world table (processed by our world in data)* [Data adapted from Feenstra et al., "Penn World Table 11.0"]. Our World in Data; Groningen Growth; Development Centre. https://archive.ourworldindata.org/20251013-111729/grapher/tfp-at-constant-national-prices-20111.html

Jones, C. I. (1995). R & d-based models of economic growth. *Journal of political Economy*, *103*(4), 759–784.

Jonkers, K., & Cruz-Castro, L. (2013). Research upon return: The effect of international mobility on scientific ties, production and impact. *Research Policy*, *42*(8), 1366–1377. https://doi.org/https://doi.org/10.1016/j.respol.2013.05.005

Juhász, R., Sakabe, S., & Weinstein, D. (2024). *Codification, technology absorption, and the globalization of the industrial revolution* (tech. rep.). National Bureau of Economic Research.

Melitz, J. (2008). Language and foreign trade. *European Economic Review*, *52*(4), 667–699. https://doi.org/https://doi.org/10.1016/j.euroecorev.2007.05.002

Melitz, J., & Toubal, F. (2014). Native language, spoken language, translation and trade. *Journal of International Economics*, *93*(2), 351–363. https://doi.org/https://doi.org/10.1016/j.jinteco.2014.04.004

Merali, A. (2024). Scaling laws for economic productivity: Experimental evidence in llm-assisted translation. https://arxiv.org/abs/2409.02391

Romer, P. M. (1990). Endogenous technological change. *Journal of political Economy*, *98*(5, Part 2), S71–S102.

Scellato, G., Franzoni, C., & Stephan, P. (2012, December). *Mobile scientists and international networks* (Working Paper No. 18613). National Bureau of Economic Research. https://doi.org/10.3386/w18613

UNESCO. (n.d.). *Index translationum: Unesco database of book translations (1979–2012)*. https://www.unesco.org/xtrans/bsform.aspx?lg=0

U.S. Bureau of Labor Statistics. (2025a). *Consumer price index for all urban consumers: Educational books and supplies in u.s. city average* [Retrieved from FRED, Federal Reserve Bank of St. Louis]. Federal Reserve Bank of St. Louis. FRED. Retrieved October 21, 2025, from https://fred.stlouisfed.org/series/CUSR0000SEEA

U.S. Bureau of Labor Statistics. (2025b). *Producer price index by industry: Book publishers: Book publishing from print publishers* [Retrieved from FRED, Federal Reserve Bank of St. Louis]. Federal Reserve Bank of St. Louis. FRED. Retrieved October 21, 2025, from https://fred.stlouisfed.org/series/PCU511130511130P

U.S. Bureau of Labor Statistics. (2025c). *Producer price index by industry: Books printing* [Retrieved from FRED, Federal Reserve Bank of St. Louis]. Federal Reserve Bank of St. Louis. FRED. Retrieved October 21, 2025, from https://fred.stlouisfed.org/series/PCU323117323117

U.S. Bureau of Labor Statistics. (2025d). *Producer price index by industry: Printing ink manufacturing: Publication and commercial web inks* [Retrieved from FRED, Federal Reserve Bank of St. Louis]. Federal Reserve Bank of St. Louis. FRED. Retrieved October 21, 2025, from https://fred.stlouisfed.org/series/PCU32591032591062

U.S. Census Bureau. (2025). *Total revenue for translation and interpretation services, establishments subject to federal income tax, employer firms* [Retrieved from FRED, Federal Reserve Bank of St. Louis]. Federal Reserve Bank of St. Louis. FRED. Retrieved October 21, 2025, from https://fred.stlouisfed.org/series/REVEF54193TAXABL

Yang, H., Watabe, Y., & Kanasheuski, E. (2025). The tower of babel: Localization, translation, and international trade. *European Economic Review*, *179*, 105123. https://doi.org/https://doi.org/10.1016/j.euroecorev.2025.105123

Appendix

A Data from "China Statistical Yearbook"

In figure 13–15 we show some pictures from the 2017 "China Statistical Yearbook" as an example, including international students, imported books and copyrights data.

21-10 Statistics on Postgraduates and Students Studying Abroad

(person)

Year	Number of Postgraduates			Number of Students Studying Abroad	Number of Returned Students
	Graduates	Entrants	Enrolment		
1978	9	10708	10934	860	248
1980	476	3616	21604	2124	162
1985	17004	46871	87331	4888	1424
1990	35440	29649	93018	2950	1593
1995	31877	51053	145443	20381	5750
2000	58767	128484	301239	38989	9121
2001	67809	165197	393256	83973	12243
2002	80841	202611	500980	125179	17945
2003	111091	268925	651260	117307	20152
2004	150777	326286	819896	114682	24726
2005	189728	364831	978610	118515	34987
2006	255902	397925	1104653	134000	42000
2007	311839	418612	1195047	144000	44000
2008	344825	446422	1283046	179800	69300
2009	371273	510953	1404942	229300	108300
2010	383600	538177	1538416	284700	134800
2011	429994	560168	1645845	339700	186200
2012	486455	589673	1719818	399600	272900
2013	513626	611381	1793953	413900	353500
2014	535863	621323	1847689	459800	364800
2015	551522	645055	1911406	523700	409100
2016	563938	667064	1981051	544500	432500
2017	578045	806103	2639561	608400	480900

Figure 13: Number of Returned Students

23-6 Statistics on Imports and Exports of Books, Magazines and Newspapers (2017)

Item	Exports		Imports	
	Number (10 000 copies)	Value (10 000 USD)	Number (10 000 copies)	Value (10 000 USD)
National	**1870.72**	**6024.66**	**3255.60**	**31978.76**
Books Published	1232.71	5460.53	2033.59	17036.94
Philosophy, Social Science	177.09	1438.03	86.39	2549.01
Culture and Education	144.32	1001.70	436.38	3903.70
Literature and Art	198.92	1102.92	265.11	2108.04
Natural Science and S&T	46.81	291.03	66.38	2368.56
For Children	539.70	802.35	690.59	2371.98
General Books	125.87	824.50	488.74	3735.65
Magazines Published	335.19	504.37	311.74	13595.01
Newspapers Published	302.82	59.76	910.27	1346.81

a) Data are from national publication import and export units that have publication import business certificate. The same applies to the table following.

Figure 14: Imported Number of Books, Magazines and Newspapers

23-8 Basic Statistics on Registration of Copyright Contracts and Copyright Import and Export (2017)

(item)

Item	Total	Books	Audio Products	Video Products	Electronic Publica-tions	Software	Films	TV Pro-grams	Others
Total Number of Copyright Import During the Year	**18120**	**17154**	**147**	**364**	**372**	**12**	**10**	**61**	
United States	6645	6217	41	239	126	1	2	19	
United Kingdom	2991	2835	23	17	80			36	
Germany	951	933	2	11	3			2	
France	1164	1133	2	11	14		4		
Russia	93	90		1	2				
Canada	170	156		5	9				
Singapore	259	249	5		3	1		1	
Japan	2232	2101	25	54	46	4	1	1	
South Korea	183	168	1	4	10				
Hong Kong, China	165	139	15	3	6	2			
Macao, China									
Taiwan, China	946	917	18		9	2			
Others	2321	2216	15	19	64	2	3	2	

Figure 15: Imported Number of Copyrights

B Another Form of Absorption Term

Here we use another form of absorption term, that is, $\sum_{j \neq i} \theta_{ij}(A_j - A_i)\mathbb{1}\{A_j - A_i > 0\}$, and show some initial results. Figure 16 shows the summary statistics.

Variable	Obs	Mean	Std. Dev.	Min	Max
countrycode	0				
country	0				
year	126	2008.5	5.208839	2000	2017
pop	126	304.0283	432.2008	59.05734	1412.355
rtfpna	126	0.9289817	0.128159	0.5124558	1.078543
rsrch	126	3499.648	1250.595	549.5748	5352.831
stud	126	122235.8	186448.4	21266	930743
H	126	0.6158425	0.4132705	0.1708751	1.745458
delta_A	126	0.0080503	0.0175516	-0.0668102	0.0670522
domes	126	0.547174	0.3449089	0.1730319	1.568951
S_2	126	0.0483019	0.063474	-0.2119969	0.1320078
spill_2	126	0.0265654	0.0497186	-0.2451206	0.1340016

Figure 16: Summary Statistics Table, Another Form

We do not recalculate TFP data using equation 3. And we use backward type *delta_A* (this year minus the last year).

ΔA	domestic		spill		time trend	
	β₁	p>|t|	β₂	p>|t|	β₃	p>|t|
CHN	0.0998217*	0.051	0.0161119	0.646	-0.0068320*	0.064
DEU	0.0224634	0.873	0.6742410***	0.000	0.0006076	0.702
FRA	-0.7213124	0.467	0.5822999***	0.000	0.0050259	0.491
GBR	-0.2291494**	0.022	0.7504705***	0.000	0.0011261*	0.097
JPN	-0.2400447	0.192	0.3225586***	0.001	0.0012389	0.102
RUS	0.2613129*	0.070	0.9230331***	0.000	-0.0033507***	0.002
USA	0.0973703	0.178	0.0204282*	0.086	-0.0023302	0.124

*Note: the statistical significance is defined as (i) *** if p-value is not larger than 0.01; (ii) ** if p-value is larger than 0.01 but not larger than 0.05; (iii) * if p-value is larger than 0.05 but not larger than 0.1.*

Figure 17: Qualitative Analysis Results, Controlling for Time Trend, Another Form

In Figure 17 we report qualitative results. As we can see, most coefficients are not significant, even adding a linear time series.

We here show more time series graph of six-year rolling window, for example, using 2000-2005 observations to estimate $\beta_{2i,2005}$. Figure 18 displays the estimated technology absorption rates for China. The plotting program is the same as before.

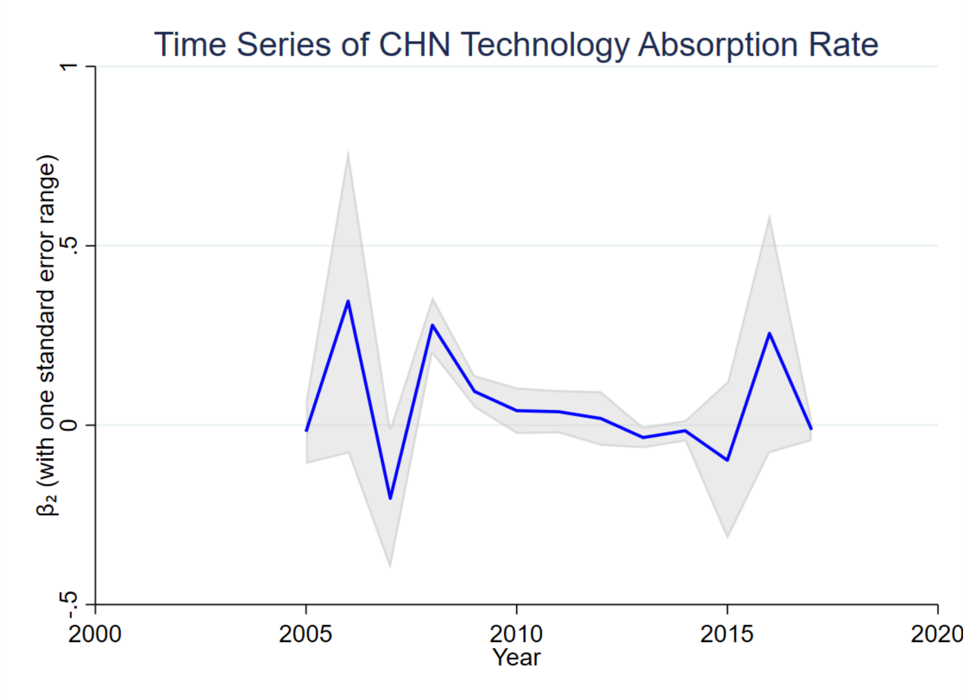

Figure 18: TFP Absorption Rate of CHN

Most of these estimation results are less significant than what we get in Chapter 5.3. And the point estimation results are much more smaller. The reason is, if we do not transform the base year to 2000, then all TFP data would converge to 1 in 2021, and *delta_A* also converges to 0. However, all factors show increasing pattern. Thus the coefficient will surely be negative.

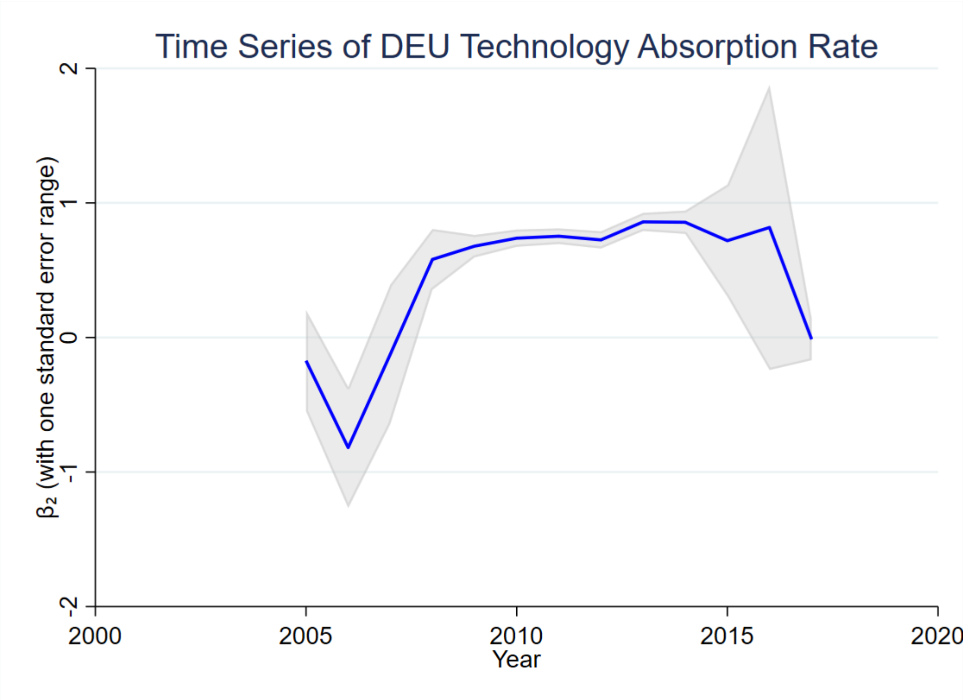

Figure 19: TFP Absorption Rate of DEU

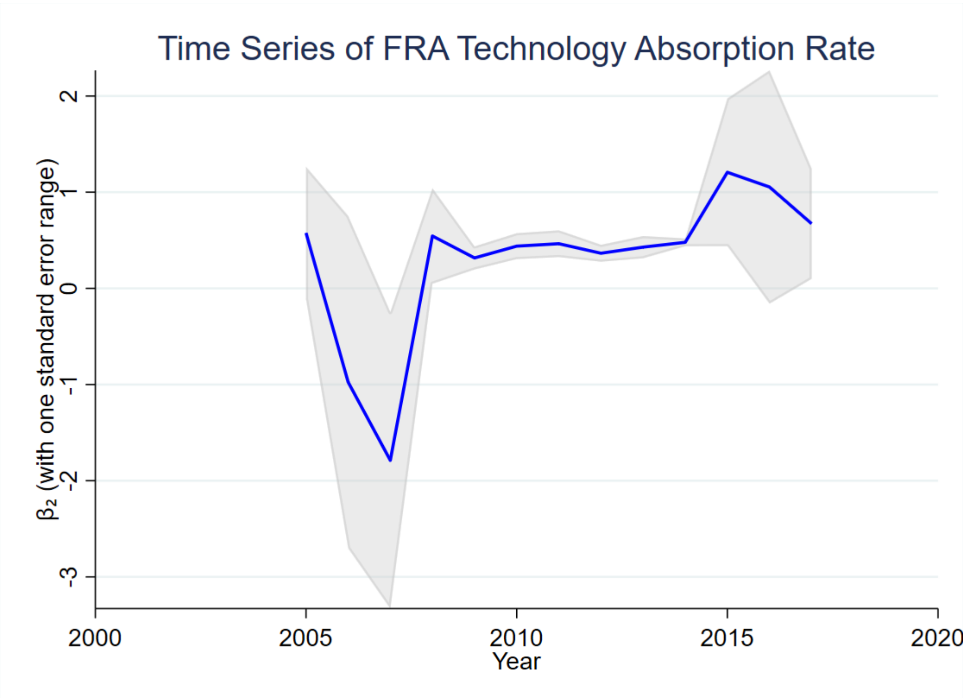

Figure 20: TFP Absorption Rate of FRA

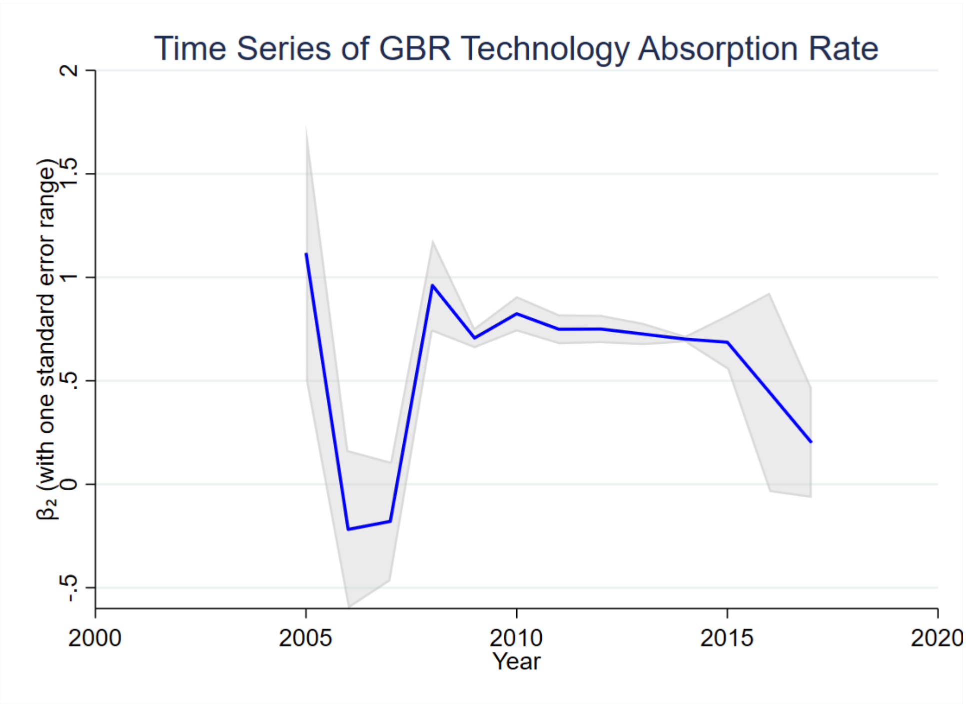

Figure 21: TFP Absorption Rate of GBR

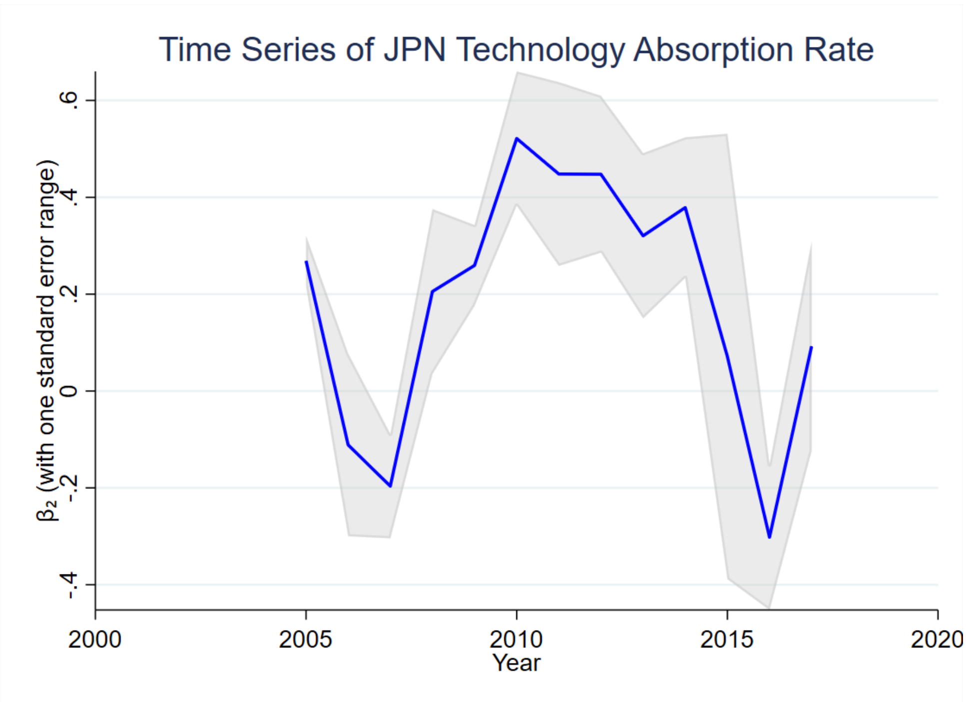

Figure 22: TFP Absorption Rate of JPN

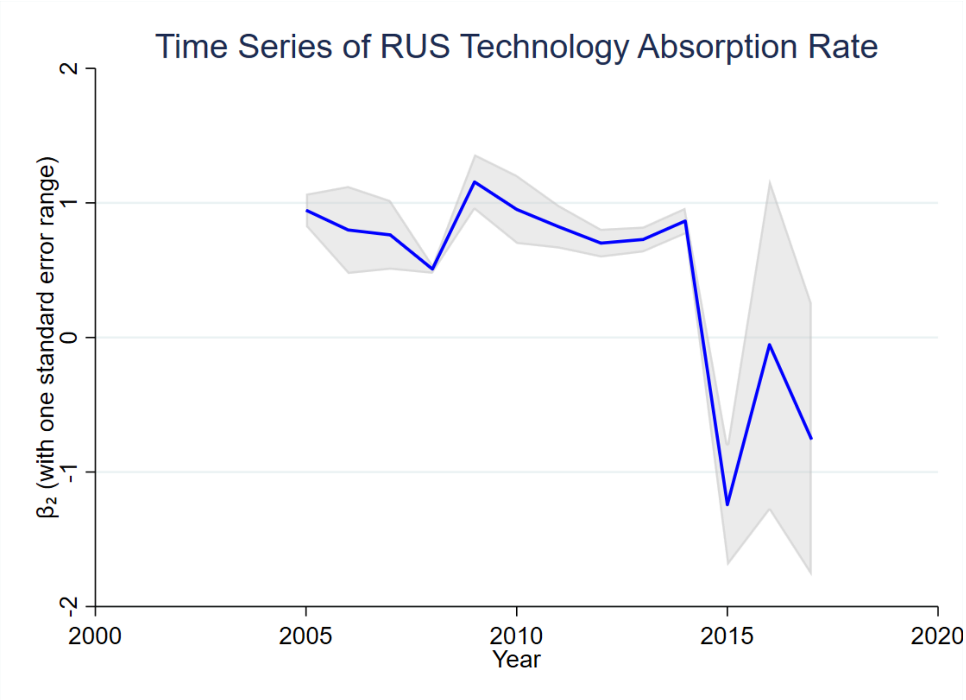

Figure 23: TFP Absorption Rate of RUS

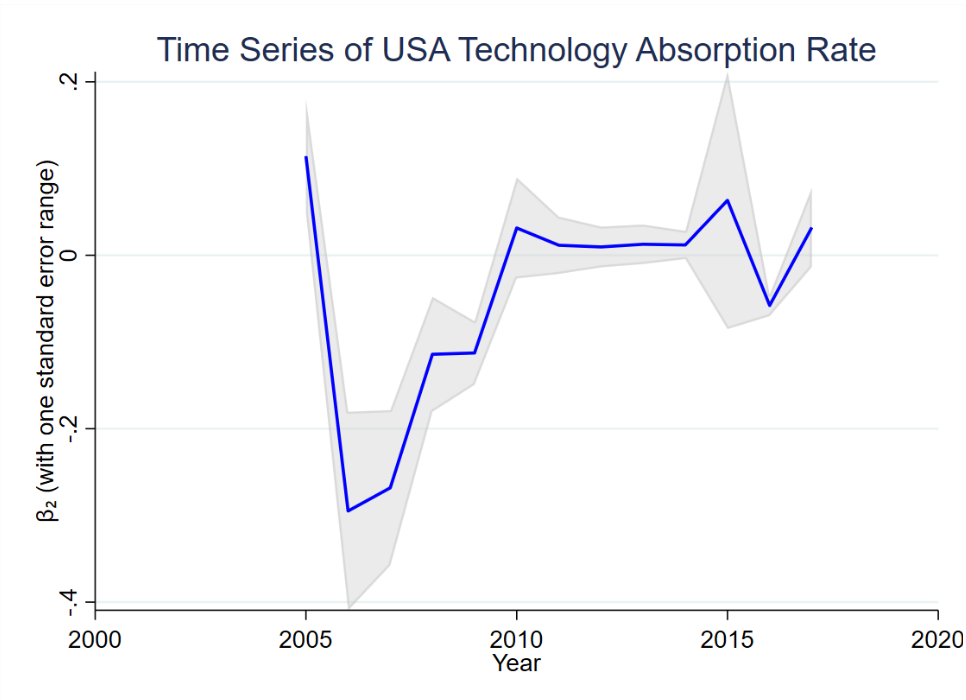

Figure 24: TFP Absorption Rate of USA