RESEARCH AND INNOVATION: FIGHTING THE PANDEMIC AND BOOSTING LONG-TERM GROWTH

Online Annex

Annex 3.1 Data Sources, Sample Coverage, and Variable Definitions

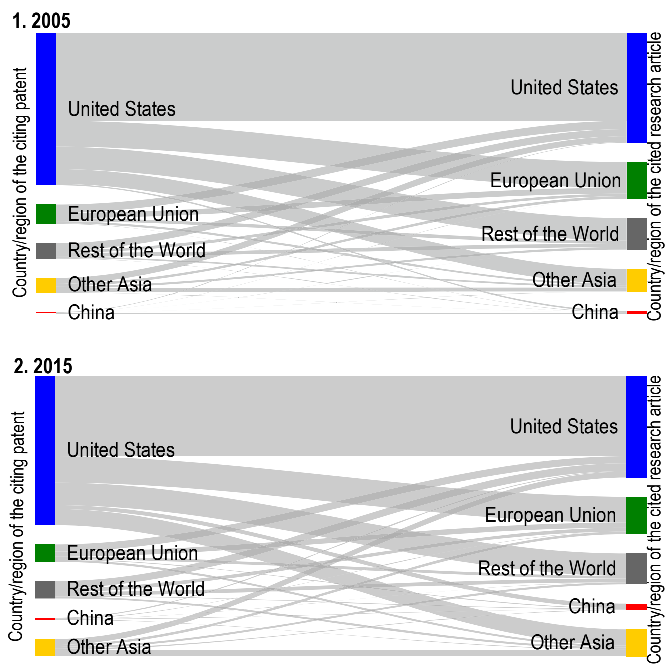
Data sources used in the chapter are listed in Annex Table 3.1.1.

The Reliance on Science (RoS) database, used for patent-to-(scientific)article citations, covers US and European patents. PATSTAT, which is used for patent-to-patent citations, provides global coverage from more than 190 patenting offices. The RoS is further matched to the USPTO database to obtain the country of residence of the inventor. This restricts the sample to patent-to-article citations in patents applied for in the USPTO; however, it still includes inventors from all over the world. To verify how representative the subsample is of cross-country citations, the critical element in our analysis of international spillovers, we compared patent-to-patent citations from the RoS-USPTO sample to patent-to-patent citations from the global universe of patents in PATSTAT. The correlation coefficient ranged from 0.97 to 0.98 using different citation lag windows (2, 4, 6, 8 and 10 years).

Each patent is assigned to a country based on the location of residence of the inventor. For patents with several inventors (or several countries for the same inventor), the main country is determined based on the largest share of countries. For patents with an equal split between countries, the patent is not assigned to any country. 2% of patents found in Reliance on Science and USPTO are not assigned to a country.

Each scientific article is assigned a country based on the academic affiliation of the author. Academic affiliation, provided by Reliance on Science, includes an institution name or geographic information that is matched to a country for 90% of unique affiliations.[1] For articles with several countries, the main country is determined based on the largest share of countries. Articles with an equal split between countries are not assigned to any country. This holds for

Annex Figure 3.1.1. The Geography of International Basic Knowledge Flows
(Citation share)



Sources: Reliance on Science; United States Patent and Trademark Office; and IMF staff calculations.
Note: The shaded area corresponds to the number of citations, where the left hand side represents 100% of the citing patents and the right hand side represents 100% of the cited articles.

[1] For the remaining 10 percent the used algorithm was not able to match affiliations to country.

1% of all articles. The flow of citations between patents and scientific articles is illustrated in Annex Figure 3.1.1.

Similarity in technological specialization between country pairs is calculated following Peri (2005). The calculation is based on patent classification into 131 technological categories as defined by the International Patent Classification codes. For each country, the vector of shares of patents falling into each category is used to calculate the uncentered correlation coefficient between the vectors of country *i* and country *j* . The resulting measure ranges from 0 (no overlap in technological classes) to 1 (complete overlap in technological classes). Technological distance is then defined as (1 – technological similarity). Similarity in scientific specialization and distance in scientific specialization are calculated analogously using the Reliance on Science database and the OECD classification of scientific fields (37 fields) for academic articles.

Technological development of each country is defined as the natural logarithm of the 5-year (2015-2019) average R&D expenditure per employed person. For each country pair, the difference between the technological development of the two countries was then calculated. R&D expenditure was taken from OECD Main Science and Technology Indicators; total employment was taken from Penn World Table 10.0.

Scientific development of each country is defined as the natural logarithm of the 5-year (2015-2019) average scientific article production per person employed in R&D. For each country pair, the difference between the scientific development of the two countries was then calculated. Scientific article production was calculated from Reliance on Science citations; number of R&D personnel was taken from OECD Main Science and Technology Indicators.

Patents are classified into clean energy technologies, dirty energy technologies, and emerging technologies based on USPTO's Cooperative Patent Classification (CPC) codes following Dechezleprêtre, Muckley, Neelakantan (2021).

The research expenditure on R&D is obtained from the OECD Main Science and Technology Indicators database, which is detailed by the type of expenditure. Following the OECD (2015) Frascati Manual, the data is collected following precise definitions for basic research, applied research and experimental development. Experimental development is combined with applied research into one category, and the data capture total expenditure (public and private) on R&D.

All regression results are produced using STATA 16.1.

Annex Table 3.1.1. Data Sources

Source	Indicators
PATSTAT Global 2020 Spring Edition	Patents flows; patent stock; patent-to-patent citations; International Patent Classification codes
Reliance on Science database; Marx and Fuegi (2020, 2020b)	Patent-to-article citations; OECD classification codes of scientific fields; year cited article was published; author affiliation
United States Patent and Trademark Office	Location of inventor; Cooperative Patent Classification codes; year patent application was filed
CEPII, GEO Dist; Mayer and Zignago (2011)	Common border dummy; common official language dummy; distance between countries' capital cities
World Bank, World Development Indicators; National Science Foundation, Science and Engineering Indicators	Scientific and technical journal articles (count)
OECD Product Market Regulation Statistics database	Indicator of regulation in product markets
OECD Science and Technology Indicators database	Basic R&D expenditure; basic R&D stock; non-basic R&D expenditure; non-basic R&D stock; total business enterprise R&D personnel; government total R&D personnel; higher education total R&D personnel; total employment
Penn World Table 10.0; Feenstra, Inklaar, and Timmer (2015)	Output-side real GDP at chained PPPs (in mil. 2017US$); number of persons engaged (in millions); average annual hours worked by persons engaged; TFP at constant national prices (2017=1); capital stock at current PPPs (in mil. 2017US$)
Barro-Lee Educational Attainment Dataset	Average years of schooling
Worldwide Governance Indicators	Control of corruption; government effectiveness; political stability and absence of violence/terrorism; rule of law; regulatory quality; voice and accountability (percentile rank)
World Economic Forum	Intellectual property protection; quality of overall infrastructure; quality of the education system
PRS Group, The International Country Risk Guide (ICRG)	Contract viability/expropriation
International Monetary Fund, Financial Development Index database	Financial development index
World Bank, Doing Business database	Ease of doing business score

Source: IMF staff compilation.

Annex Table 3.1.2. Economies Included in the Analysis

Figure / Exercise	List of Economies
Figure 3.1 (panel 1)	Australia; Austria; Belgium; Canada; Denmark; Finland; France; Germany; Greece; Hong Kong SAR; Ireland; Italy; Japan; Korea; Norway; Portugal; Singapore; Spain; Sweden; Switzerland; United Kingdom; United States
Figure 3.1 (panel 2)	Australia; Austria; Belgium; Canada; Denmark; Finland; France; Germany; Greece; Hong Kong SAR; Ireland; Italy; Japan; Korea; Norway; Portugal; Singapore; Spain; Sweden; Switzerland; United Kingdom; United States
Figure 3.1 (panel 3)	Argentina; Australia; Austria; Belgium; Chile; China; Czech Republic; Denmark; Estonia; France; Germany; Greece; Hungary; Iceland; Ireland; Israel; Italy ; Japan; Korea; Latvia; Lithuania; Luxembourg; Mexico; Netherlands; New Zealand; Norway; Poland; Portugal; Romania; Russia; Singapore; Slovak Republic; Slovenia; South Africa; Spain; Sweden; Switzerland; Taiwan Province of China; United Kingdom; United States
Figure 3.1 (panel 4)	Albania; Algeria; Antigua and Barbuda; Argentina; Armenia; Australia; Austria; Azerbaijan; Bahamas, The; Bahrain; Bangladesh; Barbados; Belarus; Belgium; Belize; Bolivia; Brazil; Brunei Darussalam; Bulgaria; Burkina Faso; Cameroon; Canada; Chile; China; Colombia; Costa Rica; Croatia; Cyprus; Czech Republic; Denmark; Dominican Republic; Ecuador; Egypt; El Salvador; Estonia; Eswatini; Fiji; Finland; France; Gabon; Georgia; Germany; Ghana; Greece; Grenada; Guatemala; Guinea; Honduras; Hong Kong SAR; Hungary; Iceland; India; Indonesia; Iran; Iraq; Ireland; Israel; Italy; Jamaica; Japan; Jordan; Kazakhstan; Kenya; Korea; Kuwait; Kyrgyz Republic; Latvia; Lebanon; Lithuania; Luxembourg; Macao SAR; Malaysia; Malta; Marshall Islands; Mauritius; Mexico; Moldova; Mongolia; Morocco; Nepal; Netherlands; New Zealand; Nicaragua; Niger; Nigeria; North Macedonia; Norway; Oman; Pakistan; Panama; Papua New Guinea; Paraguay; Peru; Philippines; Poland; Portugal; Puerto Rico; Qatar; Romania; Russia; Saudi Arabia; Serbia; Seychelles; Sierra Leone; Singapore; Slovak Republic; Slovenia; South Africa; Spain; Sri Lanka; St. Kitts and Nevis; Suriname; Sweden; Switzerland; Syria; Taiwan Province of China; Tanzania; Thailand; Trinidad and Tobago; Tunisia; Turkey; Uganda; Ukraine; United Arab emirates; United Kingdom; United States; Uruguay; Uzbekistan; Venezuela; Vietnam; Yemen; Zimbabwe
Figure 3.3	Albania; Algeria; Angola; Antigua and Barbuda; Argentina; Armenia; Australia; Austria; Azerbaijan; Bahrain; Bangladesh; Barbados; Belarus; Belgium; Benin; Bolivia; Bosnia and Herzegovina; Botswana; Brazil; Brunei Darussalam; Bulgaria; Cambodia; Cameroon; Canada; Central African Republic; Chad; Chile; China; Colombia; Costa Rica; Croatia; Cyprus; Czech Republic; Côte d'Ivoire; Denmark; Dominica; Dominican Republic; Ecuador; Egypt; El Salvador; Estonia; Ethiopia; Fiji; Finland; France; Gambia, The; Georgia; Germany; Ghana; Greece; Grenada; Guatemala; Guinea; Haiti; Hong Kong SAR; Hungary; Iceland; India; Indonesia; Iran; Iraq; Ireland; Israel; Italy; Jamaica; Japan; Jordan; Kazakhstan; Kenya; Korea; Kuwait; Lao P.D.R.; Latvia; Lebanon; Liberia; Libya; Lithuania; Luxembourg; Macao SAR; Malawi; Malaysia; Mali; Malta; Mauritius; Mexico; Micronesia; Moldova; Mongolia; Montenegro, Rep. of; Morocco; Mozambique; Myanmar; Namibia; Nepal; Netherlands; New Zealand; Nicaragua; Niger; Nigeria; North Macedonia; Norway; Oman; Pakistan; Panama; Papua New Guinea; Paraguay; Peru; Philippines; Poland; Portugal; Puerto Rico; Qatar; Romania; Russia; Saudi Arabia; Senegal; Serbia; Singapore; Slovak Republic; Slovenia; South Africa; Spain; Sri Lanka; Sudan; Sweden; Switzerland; Syria; Taiwan Province of China; Tajikistan; Tanzania; Thailand; Trinidad and Tobago; Tunisia; Turkey; Uganda; Ukraine; United Arab Emirates; United Kingdom; United States; Uruguay; Uzbekistan; Venezuela; Vietnam; West Bank and Gaza; Yemen; Zambia; Zimbabwe
Figure 3.4 (Panel 1)	Afghanistan; Albania; Algeria; Angola; Antigua and Barbuda; Argentina; Armenia; Aruba; Australia; Austria; Azerbaijan; The Bahamas; Bahrain; Bangladesh; Barbados; Belarus; Belgium; Belize; Benin; Bhutan; Bolivia; Bosnia and Herzegovina; Botswana; Brazil; Brunei; Darussalam; Bulgaria; Burkina Faso; Burundi; Cabo Verde; Cambodia; Cameroon; Canada; Central African Republic; Chad; Chile; China; Colombia; Republic of Congo; Costa Rica; Croatia; Cyprus; Czech Republic; Côte d'Ivoire; Denmark Djibouti; Dominica; Dominican Republic; Ecuador; Egypt; El Salvador; Eritrea; Estonia; Eswatini; Ethiopia; Fiji; Finland; France; Gabon; The Gambia; Georgia; Germany; Ghana; Greece; Grenada; Guatemala; Guinea; Guyana; Haiti; Honduras; Hong Kong SAR; Hungary; Iceland; India; Indonesia; Iran; Iraq; Ireland; Israel; Italy; Jamaica; Japan; Jordan; Kazakhstan; Kenya; Korea; Kuwait; Kyrgyz Republic; Lao P.D.R.; Latvia; Lebanon; Lesotho; Liberia; Libya; Lithuania; Luxembourg; Macao SAR; Madagascar; Malawi; Malaysia; Maldives; Mali; Malta; Marshall Islands; Mauritania; Mauritius; Mexico; Micronesia; Moldova; Mongolia; Morocco; Mozambique; Myanmar; Namibia; Nauru; Nepal; Netherlands; New Zealand; Nicaragua; Niger; Nigeria; North Macedonia; Norway; Oman; Pakistan; Palau; Panama; Papua New Guinea; Paraguay; Peru; Philippines; Poland; Portugal; Puerto Rico; Qatar; Romania; Russia; Rwanda; Samoa; San Marino; Saudi Arabia; Senegal; Seychelles; Sierra Leone; Singapore; Slovak Republic; Slovenia; Solomon Islands; Somalia; South Africa; Spain; Sri Lanka; St. Kitts and Nevis; Sudan; Suriname; Sweden; Switzerland; Syria; São Tomé and Príncipe; Taiwan Province of China; Tajikistan; Tanzania; Thailand; Togo; Tonga; Trinidad and Tobago; Tunisia; Turkey; Turkmenistan; Tuvalu; Uganda; Ukraine; United Arab Emirates; United Kingdom; United States; Uruguay; Uzbekistan; Vanuatu; Venezuela; Vietnam; Yemen; Zambia; Zimbabwe
Figure 3.5 (Panels 1 and 2)	Argentina; Australia; Austria; Belgium; Chile; China; Czech Republic; Denmark; Estonia; France; Germany; Greece; Hungary; Iceland; Ireland; Israel; Italy; Japan; Korea; Latvia; Lithuania; Luxembourg; Mexico; Netherlands; New Zealand; Norway; Poland; Portugal; Romania; Russia; Singapore; Slovak Republic; Slovenia; South Africa; Spain; Sweden; Switzerland; Taiwan Province of China; United Kingdom; United States
Figure 3.6	Albania; Algeria; Angola; Argentina; Armenia; Australia; Austria; Bahrain; Bangladesh; Barbados; Belgium; Belize; Benin; Bolivia; Brazil; Brunei Darussalam; Bulgaria; Burkina Faso; Burundi; Cambodia; Cameroon; Canada; Central African Republic; Chile; China; Colombia; Congo, Democratic Republic of the; Costa Rica; Croatia; Cyprus; Czech Republic; Côte d'Ivoire; Denmark; Dominican Republic; Ecuador; Egypt; El Salvador; Estonia; Eswatini; Ethiopia; Fiji; Finland; France; Gabon; Gambia, The; Germany; Ghana; Greece; Guatemala; Haiti; Honduras; Hong Kong SAR; Hungary; Iceland; India; Indonesia; Iran; Iraq; Ireland; Israel; Italy; Jamaica; Japan; Jordan; Kazakhstan; Kenya; Korea; Kuwait; Kyrgyz Republic; Lao P.D.R.; Latvia; Lesotho; Liberia; Lithuania; Luxembourg; Macao SAR; Madagascar; Malawi; Malaysia; Mali; Malta; Mauritania; Mauritius; Mexico; Moldova; Mongolia; Morocco; Myanmar; Namibia; Nepal; Netherlands; New Zealand; Nicaragua; Niger; Nigeria; Norway; Pakistan; Panama; Paraguay; Peru; Philippines; Poland; Portugal; Qatar; Romania; Russia; Saudi Arabia; Senegal; Serbia; Sierra Leone; Singapore; Slovak Republic; Slovenia; South Africa; Spain; Sri Lanka; Sudan; Sweden; Switzerland; Syria; Taiwan Province of China; Tajikistan; Tanzania; Thailand; Togo; Trinidad and Tobago; Tunisia; Turkey; Uganda; Ukraine; United Arab Emirates; United Kingdom; United States; Uruguay; Venezuela; Vietnam; Yemen; Zimbabwe

Source: IMF staff compilation.

Annex 3.2. The Diffusion of Basic and Applied Knowledge

This annex provides details and further robustness tests of the baseline results presented in the chapter's "The diffusion of basic and applied knowledge" subsection.

Gravity Model

The spatial diffusion of knowledge using patent data has been widely studied since Jaffe, Trajtenberg and Henderson (1993).[1] However, this literature has focused mostly on applied knowledge flows using patent-to-patent citations. The chapter extends this literature by studying *both* applied and basic knowledge spillovers. It does so by estimating a gravity-type model of international knowledge flows similar to Peri (2005) but for basic and applied knowledge flows distinctly.[2] In the model, bilateral country-to-country patent citations are regressed on measures of geographic and linguistic barriers, as well as scientific and technological distance.

Peri's (2005) model implies the following Poisson regression:

$$c_{ij} = \exp\{\alpha + \theta_i + \theta_j + \delta_1(diff\_bord) + \delta_2(diff\_lang) + \delta_3(spec\_dist) + \delta_4(geog\_dist) + e_{ij}\} \tag{3.1}$$

where c_{ij} is the bilateral citation count with country i citing country j. The model is fitted to patent-to-article citations to capture the determinants of basic knowledge flows, and to patent-to-patent citations to capture applied knowledge flows, where the citations are taken over a fixed 10-year window. The specific regressors are dummy variables taking the value 1 if i and j do not have a common border (*diff_bord*) or a common official language (*diff_lang*), a measure of specialization distance (*spec_dist*), and finally geographic distance in thousand kilometers (*geog_dist*).[3] The regressions include citing and cited country fixed effects to control for differences in the number of patent applications across countries as well as other factors that may influence a country's propensity to patent or to cite other patents. The model is estimated using the Pseudo-Poisson-Maximum Likelihood estimator, which is robust to the presence of significant heteroscedasticity in the data and the large number of dummies. It also allows for the inclusion of zero values for the dependent variable (Santos Silva and Tenreyro, 2006).

The baseline estimation, reported in column (1) of Annex Table 3.2.1, includes all countries for which bilateral citations are available in RoS (patent-to-article citations) and PATSTAT (patent-to-patent citations). The baseline results show that basic knowledge diffuses more strongly relative to applied knowledge: national borders only impede diffusion of applied knowledge (negative and significant coefficient); common language affects both types of flows, but has a marginally larger negative impact on applied knowledge diffusion; specialization distance matters more for applied knowledge than for basic knowledge pointing to the more generic nature of scientific discoveries and its potential application in diverse fields across countries. Note that the geographic distance variable is statistically significant for basic knowledge diffusion but not for applied knowledge. This is potentially explained by the legal requirement to cite earlier patents

[1] Jaffe, Trajtenberg and Henderson (1993) relied on the case-control matching method, which is more suited to micro-level analysis.

[2] See IMF (2018) for a recent application to applied knowledge spillovers.

[3] For the specialization distance variable, we use scientific specialization for patent-to-article citations and technological specialization for patent-to-patent citations; see Annex 3.1 for details on the construction of these variables.

when developing new marketable technologies – irrespective of the geographical distance – coupled with evidence of China and Korea becoming formidable competitors to the US and the EU in certain technology classes (e.g. 5G technology). Columns (2) and (3) confirm that these findings are unaffected by changing the length of the lag window for citations using 6 years and 2 years, respectively.

Annex Table 3.2.1. Gravity Model of Basic and Applied Knowledge Diffusion: Baseline Specification with Different Lag Lengths for Citation Window

	(1)		(2)		(3)	
	Patent-to-Article Citations (10Y)	Patent-to-Patent Citations (10Y)	Patent-to-Article Citations (6Y)	Patent-to-Patent Citations (6Y)	Patent-to-Article Citations (2Y)	Patent-to-Patent Citations (2Y)
No Common Border	0.147	-0.220***	0.130	-0.243***	0.065	-0.243***
	(0.102)	(0.056)	(0.097)	(0.058)	(0.071)	(0.063)
No Common Language	-0.141***	-0.156**	-0.132***	-0.162**	-0.099***	-0.172**
	(0.045)	(0.063)	(0.043)	(0.069)	(0.032)	(0.082)
Scientific Distance	-1.656***		-1.777***		-1.854***	
	(0.453)		(0.491)		(0.418)	
Technological Distance		-3.053***		-3.091***		-3.086***
		(0.232)		(0.241)		(0.270)
Geographic Distance (1,000 km)	-0.038***	-0.010	-0.038***	-0.008	-0.032***	-0.007
	(0.013)	(0.007)	(0.012)	(0.007)	(0.008)	(0.008)
Constant	11.083***	13.971***	10.588***	13.545***	9.438***	12.382***
	(0.046)	(0.040)	(0.045)	(0.039)	(0.047)	(0.040)
Number of Observations	19,506	30,104	18,552	29,589	16,359	26,252
Citing-Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Cited-Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

Source: IMF staff calculations.

Note: Robust standard errors (clustered by citing country) are reported in parentheses. 10Y = 10 years; 6Y = 6 years; 2Y = 2 years; km = kilometers. ***, **, and * indicate statistical significance at 1, 5, and 10 percent, respectively.

Controlling for productivity Differentials

In column (1) of Table 3.2.2, we include the productivity differential between country *i* and country *j* as an additional regressor in equation (3.1) while maintaining the distinction between basic and applied knowledge flows. This allows us to investigate whether countries with lower productivity tend to cite countries with higher productivity more often. The productivity differential measure is intended to proxy for the relative distance between countries in relation to the technological frontier. The productivity differential is measured as the (log) difference between output per worker averaged over the years 2015-2019. The results for the baseline variables are in line with the estimates in Table 3.2.1 with basic knowledge diffusing more strongly across most barriers. The output productivity differential shows a negative sign for patent-to-article citations indicating that countries cite science produced in countries with lower productivity less often. On the other hand, the coefficient for patent-to-patent citations is positive and significant indicating the opposite impact. This can again be rationalized by large competitive pressures from emerging markets in patenting activity, and legal requirements for advanced economies to cite patents from emerging economies competitors.

Annex Table 3.2.2. Gravity Model of Basic and Applied Knowledge Diffusion: Specifications Including the Difference in Levels of Output, Scientific and Technological Productivity (Top 5)

	(1)		(2)		(3)	
	Patent-to-Article Citations	Patent-to-Patent Citations	Patent-to-Article Citations	Patent-to-Patent Citations	Patent-to-Article Citations	Patent-to-Patent Citations
No Common Border	0.146	-0.221***	-0.007	-0.209***	-0.001	-0.346***
	(0.102)	(0.056)	(0.025)	(0.061)	(0.036)	(0.054)
No Common Language	-0.140***	-0.156**	-0.121***	-0.157**	-0.088***	-0.085**
	(0.045)	(0.063)	(0.028)	(0.068)	(0.019)	(0.041)
Scientific Distance	-1.682***		-2.048***		-1.248***	
	(0.450)		(0.518)		(0.267)	
Technological Distance		-3.054***		-3.023***		-2.996***
		(0.232)		(0.242)		(0.121)
Geographic Distance (1,000 km)	-0.038***	-0.010	-0.062***	-0.011	-0.015***	0.009
	(0.013)	(0.007)	(0.011)	(0.007)	(0.005)	(0.006)
Differences in Output Productivity	-2.074***	1.561***			-2.138***	1.588***
	(0.073)	(0.035)			(0.070)	(0.019)
Differences in Scientific Productivity			-0.769***			
			(0.055)			
Differences in Technological Productivity				1.319***		
				(0.002)		
Constant	-9.874***	-7.618***	-0.299	9.873***	-10.168***	-7.751***
	(0.286)	(0.830)	(0.226)	(0.047)	(0.258)	(0.797)
Number of Observations	17,669	24,806	1,560	1,892	17,139	24,026
Citing-Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Cited-Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

Source: IMF staff calculations.

Note: Robust standard errors (clustered by citing country) are reported in parentheses. km = kilometers. ***, **, and * indicate statistical significance at 1, 5, and 10 percent, respectively.

The productivity differential measure does not distinguish between the flows of basic and applied knowledge. Therefore, for robustness, we also include flow-specific measures of differences in scientific and technological productivity, respectively. To measure relative scientific (technological) productivity between country pairs, we use the log difference between scientific articles (patents) produced per total R&D personnel averaged over the years 2015-2019. Due to lack of data on R&D personnel in many countries, most notably the US, the sample is reduced to less than 2,000 observations (about a tenth of the original sample). This also allows to check the robustness of the results to the exclusion of the US. Estimates are reported in column (2). The coefficient stability in terms of magnitude and statistical significance is noteworthy, perhaps with the exception of scientific distance and geographic distance, which appear to have a stronger effect on basic knowledge flows under this specification. Interestingly, using these alternative productivity measures does not alter the main findings: differences in scientific (technological) productivity have a negative (positive) impact on the frequency of citations. In column (3), we use the same specification as column (1) but exclude China and Korea, two emerging markets that recently joined the ranks of the world's top innovators. The results in column (3) confirm that neither country is driving the results obtained in earlier specifications. Two differences are worth highlighting here: having no common border has a stronger impact on applied knowledge diffusion; geographic distance matters less for basic knowledge diffusion.

Using International Patent Families

The previous results used patent counts based on the "Top 5" definition, which included patents from the top 5 patenting offices (China, EU, Japan, Korea and US). In Table 3.2.3, we report robustness checks using the "international patent family" definition, which only counts patents that have been applied for in at least two distinct patenting offices. This restricts the count to patents with presumably higher commercial value to justify the higher cost to the inventors of applying in more than one patenting office. The column specifications are similar to those of Table 3.2.2. The results are largely consistent with those for the "Top 5". The only noteworthy difference is that geographic distance now seems to matter for applied knowledge diffusion in columns (2) and (3), while the impact of this same barrier on basic knowledge diffusion is smaller under this patent count.

Annex Table 3.2.3. Gravity Model of Basic and Applied Knowledge Diffusion: Specifications Including the Difference in Levels of Output, Scientific and Technological Productivity (International Patent Families)

	(1)		(2)		(3)	
	Patent-to-Article Citations	Patent-to-Patent Citations	Patent-to-Article Citations	Patent-to-Patent Citations	Patent-to-Article Citations	Patent-to-Patent Citations
No Common Border	0.023	-0.148***	-0.041	-0.134***	-0.046*	-0.253***
	(0.048)	(0.038)	(0.025)	(0.042)	(0.026)	(0.047)
No Common Language	-0.094***	-0.150***	-0.089***	-0.156**	-0.082***	-0.090**
	(0.023)	(0.058)	(0.032)	(0.063)	(0.023)	(0.038)
Scientific Distance	-1.436***		-3.000***		-0.883***	
	(0.441)		(0.746)		(0.289)	
Technological Distance		-2.978***		-3.011***		-2.927***
		(0.237)		(0.252)		(0.100)
Geographic Distance (1,000 km)	-0.019***	-0.016**	-0.040***	-0.015**	-0.009**	0.004
	(0.006)	(0.007)	(0.007)	(0.007)	(0.004)	(0.005)
Differences in Output Productivity	-1.578***	1.366***			-1.635***	1.392***
	(0.079)	(0.038)			(0.071)	(0.018)
Differences in Scientific Productivity			-1.210***			
			(0.053)			
Differences in Technological Productivity				1.253***		
				(0.003)		
Constant	0.826***	-7.239***	-4.058***	9.609***	-12.650***	-7.352***
	(0.048)	(0.738)	(0.509)	(0.046)	(0.285)	(0.665)
Number of Observations	11,622	23,559	1,560	1,892	11,186	22,949
Citing-Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Cited-Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

Source: IMF staff calculations.

Note: Robust standard errors (clustered by citing country) are reported in parentheses. km = kilometers. ***, **, and * indicate statistical significance at 1, 5, and 10 percent, respectively.

In addition to learning about the determinants of cross-border knowledge flows for basic and applied knowledge, the gravity model also gives the weights necessary to construct the foreign research stocks accessible to each country for the estimation of the ideas production function. Specifically, the frequency of bilateral citations between countries is taken into account in the construction of the foreign research stocks as it influences the intensity of international knowledge spillovers. This is described further in Online Annex 3.3.

Annex 3.3 Production Functions for Ideas

This Annex describes the methodology used to analyze the impact of own and foreign research inputs on innovation.

Specification

The empirical specification of the ideas production function yields the following regression:

$$ln(p_{i,t}) = \beta_1 ln(b_{i,t}) + \beta_2 ln(r_{i,t}) + \gamma_1 ln\left(\sum_{j \neq i} w_{ij}^{(b)} b_{j,t}\right) + \gamma_2 ln\left(\sum_{j \neq i} w_{ij}^{(r)} r_{j,t}\right) + \varepsilon_{i,t}, \quad (3.2)$$

where $p_{i,t}$ is the flow of new patent applications by an inventor resident in country i in year t, $b_{i,t}$ is the cumulative basic research stock by country i, and $r_{i,t}$ is the stock of applied research expenditure. The third and fourth terms in brackets are, respectively, the foreign basic and applied research stocks. The bilateral weights, $w_{ij}^{(b)}$ and $w_{ij}^{(r)}$, determine how accessible the foreign research stocks are to country i, and are derived from the gravity model discussed above. In particular, $w_{ij}^{(b)}$ ($w_{ij}^{(r)}$) are the fitted values, excluding the citing and cited country fixed effects, from the gravity Poisson regression using patent-to-article (patent-to-patent) citations. The research stocks are built from the annual R&D expenditure data for each country using the perpetual inventory method and assuming 10 percent depreciation (Peri 2005, IMF 2018).

Robustness

Table 3.3.1 shows the estimates of alternative specifications of the ideas production function. The estimates reveal a strong and significant relationship between basic research and innovation, a finding that is robust across all specifications. Column (1) reports panel OLS estimates with fixed effects and shows that the coefficient on the own basic research stock is positive and highly significant. To control for potential endogeneity of research stocks, column (2) estimates the same equation lagging each regressor by 1-year. Results are comparable. Columns (3) and (4) explore robustness to the assumed depreciation rate in the construction of the research stocks Column (3) assumes 5% depreciation for both basic and applied stocks, while column (4) assumes 5% depreciation for the basic stocks and 10% for the applied stocks. The positive and significant impact of own basic research persists, while the coefficient on foreign applied research increases in magnitude and now shows statistical significance. Column (5) applies the regression using non-overlapping 5-year averages of the observations to mitigate potential year-to-year volatility due to business cycle effects. The results yet again confirm the positive impact of own basic research. Panel unit root and cointegration tests reveal strong evidence of non-stationarity and cointegration among the variables. The dynamic OLS estimates are super consistent, and hence are robust to biases due to omitted variables and simultaneity. Columns (6) and (7) report specifications similar to column (1) but estimated using dynamic OLS, which efficiently utilizes the cointegration property of the data. Again, the relationship between basic research and innovation is positive and highly significant. The contribution of own applied research also appears positive and significant as is the contribution of foreign basic research.

Annex Table 3.3.1. Ideas Production Function: Alternative Specifications and Estimation Methods

	Panel OLS					Dynamic OLS	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log(Own BASIC Research Stock)	0.633***	0.570**	0.662***	0.554***	0.728***	0.514***	0.674***
	(0.226)	(0.217)	(0.216)	(0.196)	(0.265)	(0.103)	(0.126)
Log(Own APPLIED Research Stock)	0.275	0.222	0.180	0.352	0.256	0.674***	0.765***
	(0.246)	(0.241)	(0.239)	(0.257)	(0.244)	(0.113)	(0.144)
Log(Foreign BASIC Research Stock)	-0.106	-0.375	-0.974	-1.026	-0.741	0.559	1.358**
	(1.394)	(1.372)	(1.019)	(1.212)	(1.537)	(0.449)	(0.557)
Log(Foreign APPLIED Research Stock)	1.031	1.426	1.954**	1.944*	1.855	-0.055	-1.225*
	(1.382)	(1.385)	(0.832)	(1.019)	(1.541)	(0.513)	(0.663)
Constant	-2.655	-3.486	-4.546**	-3.395*	-4.831*		
	(2.275)	(2.444)	(1.749)	(1.924)	(2.502)		
Number of Observations	1,430	1,390	1,430	1,430	264	1,310	1,154
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Source: IMF staff calculations.

Note: Robust standard errors are reported in parentheses. Specification (1) is the baseline specification; specification (2) uses the first lag of the right-hand-side variables; specification (3) uses the research stocks with 5% depreciation in the construction of all the research stock; specification (4) uses the research stocks with 5% depreciation for the basic stocks and 10% depreciation for the applied stocks; specification (5) uses non-overlapping 5-year averages over the period 1980–2014; specification (6) is estimated using dynamic ordinary least squares (DOLS) with 1 lead and 1 lag; specification (7) is estimated using DOLS with 2 leads and 2 lags. Panel unit root tests (Levin, Lin and Chu 2002; Im, Pesaran and Shin 2003; Hadri 2000) and panel cointegration tests (Pedroni 2004; Westerlund 2007), not reported for brevity, confirm evidence of nonstationarity and cointegration in the data. Log = natural logarithm. ***, **, and * indicate statistical significance at 1, 5, and 10 percent, respectively.

Further Results for EMDEs

Table 3.3.2 shows the additional impact of the research stocks for EMDEs using an interaction dummy. The results in column (1) suggest that own basic research contributes more to innovation in EMDEs than in AEs; however, this result holds only if China is included in the sample. It may be reflective of the focus on some niche fields in which emerging markets are building specialized knowledge. Column (2) shows that the impact of own applied research is roughly the same across the two country groups.

Annex Table 3.3.2. Ideas Production Function: Further Results for EMDEs

	Panel OLS			
	(1)	(2)	(3)	(4)
Log(Own BASIC Research Stock)	0.528**	0.618***	0.641***	0.644***
	(0.225)	(0.222)	(0.196)	(0.197)
Log(Own APPLIED Research Stock)	0.267	0.260	0.409*	0.416*
	(0.250)	(0.255)	(0.231)	(0.227)
Log(Foreign BASIC Research Stock)	-0.328	-0.151	-0.721	-0.514
	(1.456)	(1.450)	(1.370)	(1.409)
Log(Foreign APPLIED Research Stock)	1.340	1.103	1.318	1.070
	(1.557)	(1.494)	(1.518)	(1.523)
Log(Own BASIC Research Stock) × EMDE	0.367***			
	(0.131)			
Log(Own APPLIED Research Stock) × EMDE		0.093		
		(0.183)		
Log(Foreign BASIC Research Stock) × EMDE			0.890***	
			(0.315)	
Log(Foreign APPLIED Research Stock) × EMDE				1.026***
				(0.354)
Constant	-3.468	-2.896	-2.603	-2.572
	(2.922)	(2.681)	(2.842)	(2.815)
Number of Observations	1,430	1,430	1,430	1,430
Country Fixed Effects	Yes	Yes	Yes	Yes

Source: IMF staff calculations.

Note: Robust standard errors are reported in parentheses. EMDE is a dummy variable where the reference group is advanced economies. EMDE = emerging market and developing economy; Log = natural logarithm. ***, **, and * indicate statistical significance at 1, 5, and 10 percent, respectively.

The estimation results in columns (3) and (4) show that knowledge diffusion is stronger to EMDEs, with both types of knowledge (basic and applied) being critical for innovation. This result corroborates the finding that EMDEs tend to cite foreign research more than home-grown research (see panel 3 of Figure 3.5 in the chapter). The robustness of these findings is tested by re-estimating column (1)-(4) while removing the included countries one-by-one. For column (2)-(4) the sign and significance is robust to this check, while the coefficient on basic research interacted with the EMDE dummy becomes insignificant when removing China from the sample.

Annex 3.4 Production Functions for Goods

This Annex describes the methodology used to analyze the relationship between innovation and productivity.

Empirical Framework

The analysis is based on the following panel specification:

$$lny_{i,t} = \beta lns_{i,t} + \delta lnc_{i,t} + \lambda lnh_{i,t} + \varepsilon_{i,t}, \tag{3.3}$$

where $y_{i,t}$ is output per worker (productivity) for country i in year t, $s_{i,t}$ is the stock of patents, $c_{i,t}$ is the stock of capital per worker, and $h_{i,t}$ is an index for human capital.[1] The coefficient of interest is β and captures the elasticity of productivity to the stock of patents. The stock of patents is computed using the annual flow of new patents along with a depreciation factor of 10 percent, while human capital is measured by the human capital index from the *Penn World Tables*. Country and time fixed effects are also included. Time fixed effects should capital elements common for all countries at each point in time, such as the global stock of patents.

Equation (3.3) is estimated using annual data for 138 advanced and emerging market and developing economies depending on data availability covering the period 1980 to 2017. Compared to Annex 3.3 a wider set of countries is included because the variables used in equation (3.3) are available for more countries. Robust standard errors are reported using the Huber/White/sandwich estimator.

Baseline Results

Table 3.4.1 reports the baseline results of equation 3.3. Column (1) reports the specification estimated with pooled ordinary least squares. In this specification the estimated elasticity of output per worker to the stock of patents is 0.057. This implies that a 1 percent increase in the stock of patents is associated with an 0.057 percent increase in output per worker. The estimated elasticity between output per worker, the capital stock per worker and the human capital index, respectively, is also positive albeit not significant for the human capital stock. In column (2), country fixed effects are added, which slightly increases the estimated elasticity on the patent stock to 0.063 percent. In column (3), time fixed effect is also added yielding an estimated elasticity on the patent stock of 0.055. In column (4), the coefficients on the capital stock per employee and the human capital index are restricted to sum to one which yields a coefficient on the patent stock of 0.048. Column (5) applies the regression using non-overlapping 5-year averages of the observations to mitigate potential year-to-year volatility due to business cycle effects. Overall, it is reassuring that the choice of estimation method does not affect the estimated elasticity much. In column (6) to (8), the dependent variable is changes from output per worker to TFP[2] Column (5) reports the results estimated without country or time fixed

[1] A potential concern about this specification is non-stationarity. However, tests for stationarity have been done for the linear combination between output per worker, the patent stock, capital stock per worker, and the human capital index. These point to stationarity.

[2] For these specifications, capital per worker is dropped as independent variable because TFP is derived as the part of output not explained by utilized labor or capital.

effects. In column (6) country fixed effects are added, and column (7) further adds time fixed effects. For all specifications the estimated elasticities are close to those achieved using output per worker (column 1-4), except for column (7) which yields an insignificant elasticity.

Annex Table 3.4.1. Productivity and Innovation

	Output per Worker					Total Factor Productivity		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log(Patent Stock)	0.057***	0.063***	0.055***	0.046***	0.048***	0.048***	0.041***	0.047***
	(0.014)	(0.015)	(0.015)	(0.005)	(0.009)	(0.009)	(0.011)	(0.013)
Log(Capital Stock per Employed)	0.514***	0.498***	0.465***	0.468***	0.479***	0.479***		
	(0.038)	(0.040)	(0.042)	(0.019)	(0.030)	(0.030)		
Log(Human Capital Index)	0.017	-0.031	-0.215	0.532***	0.521***	0.521***	-0.106	-0.121
	(0.188)	(0.202)	(0.263)	(0.019)	(0.030)	(0.030)	(0.112)	(0.132)
Constant	3.961***	4.170***	4.746***	4.332***	4.140***	4.140***	-0.207**	-0.230***
	(0.354)	(0.366)	(0.429)	(0.223)	(0.360)	(0.360)	(0.082)	(0.079)
Number of Observations	4680	4680	4680	4,680	886	886	3866	3866
Number of Countries	138	138	138	138	138	138	114	114
Country Fixed Effects	No	Yes	Yes	Yes	Yes	Yes	No	Yes
Time Fixed Effects	No	No	Yes	Yes	Yes	Yes	No	No
Coefficient Restriction	No	No	No	Yes	Yes	Yes	No	No
Sample	All	All	All	All	All	All	All	All

Source: IMF staff calculations.

Note: The estimations are based on a sample of countries from 1980 to 2017. Robust standard errors are reported in parentheses. Log = natural logarithm. ***, **, and * indicate statistical significance at 1, 5, and 10 percent, respectively.

Results with Institutional Interactions

Analysis reported in Table 3.4.2-3.4.3 is meant to investigate whether certain institutional features may strengthen the relationship between innovation and productivity. Specifically, columns (1) to (6) in Table 3.4.2 and 3.4.3 augment the baseline specification with additional institutional variables interacted with the patent stock. Following Coe and others (2009), the institutional variables are split into three groups (low, middle, and high) based on the average value for each country. These are used to create dummy variables for low and high level that are then interacted with the patent stock. In Table 3.4.2, the institutional variables include contract viability (column 1), financial development (column 2), intellectual property protection (column 3), quality of overall infrastructure (column 4), quality of the education system (column 5), and control of corruption (column 6). In Table 3.4.3, the institutional interactions include government effectiveness (column 1), political stability (column 2), rule of law (column 3), regulatory quality (column 4), voice and accountability (column 5), and years of schooling (column 6).

The interactions are only significant for financial development and years of schooling, where high levels are associated with a larger elasticity. The interaction is also significant for government effectiveness, where low levels are associated with a significantly lower elasticity.

Annex Table 3.4.2. Productivity and Innovation with Institutional Interactions

	X = Contract Viability	X = Financial Development	X = Intellectual Property Protection	X = Quality of Overall Infrastructure	X = Quality of the Education System	X = Control of Corruption
	(1)	(2)	(3)	(4)	(5)	(6)
Log(Patent Stock)	0.047**	0.025	0.055***	0.064***	0.069***	0.058***
	(0.021)	(0.019)	(0.015)	(0.017)	(0.014)	(0.015)
Log(Human Capital Index)	-0.206	-0.196	-0.204	-0.201	-0.173	-0.200
	(0.273)	(0.261)	(0.264)	(0.265)	(0.272)	(0.282)
Log(Capital Stock per Employed)	0.465***	0.466***	0.465***	0.465***	0.464***	0.465***
	(0.042)	(0.041)	(0.042)	(0.040)	(0.041)	(0.042)
Log(Patent Stock) × High X	0.012	0.038**	0.001	-0.011	-0.024	-0.003
	(0.020)	(0.017)	(0.018)	(0.018)	(0.017)	(0.019)
Log(Patent Stock) × Low X	0.016	-0.005	-0.008	-0.060	-0.052	-0.022
	(0.029)	(0.093)	(0.043)	(0.040)	(0.039)	(0.039)
Constant	4.735***	4.751***	4.739***	4.761***	4.763***	4.750***
	(0.409)	(0.426)	(0.422)	(0.410)	(0.422)	(0.432)
Number of Observations	4,680	4,680	4,680	4,680	4,680	4,680
Number of Countries	138	138	138	138	138	138
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Coefficient Restriction	No	No	No	No	No	No
Sample	All	All	All	All	All	All

Source: IMF staff calculations.

Note: The estimations are based on a sample of countries from 1980 to 2017. Robust standard errors are reported in parentheses. Log = natural logarithm. ***, **, and * indicate statistical significance at 1, 5, and 10 percent, respectively.

Annex Table 3.4.3. Productivity and Innovation with Institutional Interactions, Continued

	X = Government Effectiveness	X = Political Stability	X = Rule of Law	X = Regulatory Quality	X = Voice and Accountability	X = Years of Schooling
	(1)	(2)	(3)	(4)	(5)	(6)
Log(Patent Stock)	0.065***	0.075***	0.060***	0.057***	0.040*	0.050***
	(0.016)	(0.020)	(0.016)	(0.016)	(0.020)	(0.018)
Log(Human Capital Index)	-0.214	-0.224	-0.177	-0.184	-0.194	0.026
	(0.274)	(0.271)	(0.278)	(0.275)	(0.280)	(0.297)
Log(Capital Stock per Employed)	0.460***	0.464***	0.462***	0.465***	0.469***	0.473***
	(0.040)	(0.042)	(0.042)	(0.041)	(0.042)	(0.041)
Log(Patent Stock) × High X	-0.017	-0.023	-0.006	-0.002	0.021	0.044**
	(0.018)	(0.022)	(0.018)	(0.018)	(0.024)	(0.022)
Log(Patent Stock) × Low X	-0.134**	-0.028	-0.048	-0.037	0.029	-0.043
	(0.056)	(0.021)	(0.043)	(0.057)	(0.022)	(0.029)
Constant	4.872***	4.745***	4.776***	4.738***	4.683***	4.458***
	(0.419)	(0.418)	(0.436)	(0.420)	(0.442)	(0.413)
Number of Observations	4,680	4,680	4,680	4,680	4,680	4,680
Number of Countries	138	138	138	138	138	138
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Coefficient Restriction	No	No	No	No	No	No
Sample	All	All	All	All	All	All

Source: IMF staff calculations.

Note: The estimations are based on a sample of countries from 1980 to 2017. Robust standard errors are reported in parentheses. Log = natural logarithm. ***, **, and * indicate statistical significance at 1, 5, and 10 percent, respectively.

Annex 3.5 Policy Experiments

This annex makes use of estimates in Annex 3.3 and 3.4 to conduct two policy experiments. First, it computes the effect on productivity of increasing own and foreign research stocks (Figure 3.7.1 in the main chapter). Second, it calculates the effect on global innovation and productivity of scientific decoupling between the United States and China (Figure 3.7.2 in the main chapter).

The Effect of Higher Research Stocks on Productivity

The effect on productivity of increasing research stocks is investigated by combining the estimated elasticities (i) from research stocks to patenting activity (Annex 3.3), and (ii) from patent stocks to productivity (Annex 3.4). Specifically, the estimated effect of an X percent increase in the stock of basic research is

$$X \times \eta_{\{b^i,p\}} \times \eta_{\{p,y\}}, \qquad i \in \{o, f\}$$

where $\eta_{\{b^i,p\}}$ is the elasticity from own (*i*=o) or foreign (*i*=f) basic research to patenting activity, while $\eta_{\{p,y\}}$ is the estimated elasticity from the domestic stock of patents to output per worker. It is assumed that the research stock is permanently increased, which permanently increase the flow and thus also stock of patents. Baseline parameters from Section 3.3 and Section 3.4 are used for parametrization of $\eta_{\{b^i,p\}}$, $\eta_{\{b^o,p\}}$, and $\eta_{\{p,y\}}$;they are set to 0.674, 1.358, and 0.046. A 10 percent increase in the stock of own and foreign basic research thus yields a productivity increase of 0.31 and 0.62, respectively.

Scientific Decoupling

The potential effects on innovation and productivity of scientific decoupling are investigated through construction of counter-factual stocks of foreign research. Specifically, the intensity of scientific citations between the two countries is reduced gradually, which reduces the stock of foreign basic knowledge for both countries (Table 3.5.1). In the empirical framework above, this reduces annual innovation through equation (3.2), and in turn reduces productivity through equation (3.3). Specifically, the effect on the annual flow of patents from a reduction in the stock of foreign basic knowledge can be written as

$$\Delta ln(p_{i,t}) = \gamma_1 \Delta ln\left(\sum_{j \neq i} w_{ij}^{(b)} b_{j,t}\right)$$

while the change in productivity can be expressed as

$$\Delta lny_{i,t} = \beta \Delta lns_{i,t}.$$

To translate the reduction in the annual flow of patents ($p_{i,t}$) to the stock of patents ($s_{i,t}$), it is assumed that the change in the stock of foreign basic knowledge is permanent. Finally, the country-specific results are weighted by global patent and employment shares to translate into global results.

Table 3.5.1 illustrates these effects for various degrees of decoupling. Columns (1) and (2) show the change in the stock of basic foreign knowledge for China and the United States, respectively.

Columns (3) and (4) convert this drop in foreign basic research into a drop in the annual flow of patents using the estimated elasticity of patents with respect to foreign basic research.[3] Column (5) computes the global effect by applying the sample shares of annual patenting activity. Columns (6) and (7) compute the resulting effect on productivity by applying by applying the estimated elasticity of productivity with respect to patent stocks. Finally, column (8) denotes the computed effect on global productivity by applying the employment shares for China and the United States.

Annex Table 3.5.1. Estimated Effect of Scientific Decoupling on Global Innovation and Productivity

	Foreign Basic Knowledge (Percentage change)		Annual Flow of Patents (Percentage change)			Output per Worker (Percentage change)		
	China	USA	China	USA	Global	China	USA	Global
Decoupling (Percent)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
10	-3.21	-0.55	-4.37	-0.74	-0.41	-0.20	-0.03	-0.06
20	-6.54	-1.09	-8.88	-1.49	-0.83	-0.41	-0.07	-0.12
30	-9.97	-1.65	-13.54	-2.24	-1.25	-0.62	-0.10	-0.18
40	-13.53	-2.20	-18.37	-2.99	-1.69	-0.85	-0.14	-0.25
50	-17.22	-2.76	-23.38	-3.75	-2.13	-1.08	-0.17	-0.32
60	-21.05	-3.32	-28.58	-4.51	-2.58	-1.31	-0.21	-0.39
70	-25.03	-3.89	-33.99	-5.28	-3.04	-1.56	-0.24	-0.46
80	-29.18	-4.45	-39.62	-6.05	-3.51	-1.82	-0.28	-0.54
90	-33.51	-5.02	-45.50	-6.82	-3.99	-2.09	-0.31	-0.62
100	-38.03	-5.60	-51.64	-7.60	-4.48	-2.38	-0.35	-0.70

Source: IMF staff calculations.

Note: The table shows the estimated effect on global innovation (measured as the flow of new patents) and productivity of a given reduction (in percent) of citations between the United States and China. To construct the “Global” effect on annual patents, patent shares of 32 percent and 4 percent are used for USA and China (sample shares in 2016). To construct the “Global” effect on output per worker, employment shares of 6 percent and 29 percent are used for USA and China (sample shares in 2016). An estimated elasticity of 1.358 is used for patents with respect to the stock of foreign basic knowledge, and an estimated elasticity of 0.046 of productivity with respect to the stock of patents.

[3] Here it is assumed that a given drop in patent application translates into a proportional drop in granted patents.

Annex 3.6 Policy Analysis

The model used in the chapter is a version of Akcigit et al. 2021, re-estimated to comparable data across three advanced economies. This Annex outlines the basics of the model and covers more detail on the re-estimation process discussed in the main body of the chapter.

The Model

For an in-depth presentation of the model, the reader is referred to Akcigit et al. 2021. Here, the annex focusses on the microeconomic structure of firms as this determines their incentives for distinct basic and applied research.

The framework is a multi-firm, multi-industry, multi-product model. Each firm sells potentially many products. These products are grouped by industry, of which there are a fixed number, M. Industries, in turn, are differentiated from each other in two ways. First, products within an industry are closer mutual substitutes than products in different industries. Second, the effects of basic and applied research vary across industries. As a result, cross-market presence is the key determinant of individual firms' incentives for applied and basic research – firms selling a wider range of different types of product have greater incentives for basic research (more on this below).

To be more concrete, we say that firm j sells Pj products in Mj industries, with $Mj<M$. and $Pj \geq Mj$. For example, a simple economy might have 5 industries in total, and a firm j might operate in two industries, with 5 products in the first and 2 in the second. In which case, $Mj = 2$ and $Pj = 7$.

Firms can conduct both basic and applied research within a given industry. Both types of research have increasing marginal cost and the within-industry effects are similar: both types of research increase the probability that firms will become more productive in that industry, although the particular product line where this occurs is random. However, only basic research has a cross-industry impact. Basic research in one industry raises the probability that productivity will increase in another industry, so long as the firm is already operating in that industry. As a result, large multi-industry firms can better capture the gains of basic research. This induces a correlation between firm size and basic research effort, in line with the data.

To return to the example of firm j above, either applied or basic research in either of the industries that they operate in will have a direct within-industry effect. For instance, if the firm sells products in industries numbered 1 and 2, applied research in industry 1 only raises the probability of productivity gains in industry 1. This can take the form either of an improvement to product they already sell (so Pj remains the same) or of supplanting a competitor's product (so Pj increases). Basic research in industry 1 has the same effect plus an additional cross-industry spillover, increasing the chance of improved productivity also in industries 2 through 5 (each with equal probability). Of these, the only place that the firm can make use of these gains is in industry 2. And so the firm benefits more from basic research if it operates in more industries.

Note that although basic research has a cross-industry spillover that applied research lacks, increasing marginal costs mean that firms will always do some of both types of research, increasing their research effort until marginal costs are equal.

Because the scope of the firm's activities determines the incentives for (the type of) innovation, the aggregate effect of pro-research policies depends on the distribution of firm sizes. More large firms mean a greater role for basic research all else equal.

The preceding is a somewhat simplified version of the main model mechanisms. There are several other further wrinkles in the final version (large "breakthrough" innovations, depreciation rates of product line efficiency varying with type of research, entry and exit of firms, expansion into new industries, and the like). However, the basic mechanism is governed by the trade-offs and incentives outlined here.

Estimation

Estimation involves matching some 30 moments in order to infer values for the 18 different parameters that govern the model solution. Given that the number of model moments exceeds the number of parameters, the model is over-identified. This means that there is a tension between matching the different aspects of the data. This tension is resolved by using an efficient weighting regime within the context of a simulated method of moments. For a given parameter combination, the model is solved and the steady-state density of firms simulated. The estimation routine then minimizes the divergence of the simulated results from the data, weighting them in proportion to their information about the model parameters. This weighting scheme means that deviations from the target moments are largest for moments which have the least information about the model's parameters

In their original paper, Akcigit and co-authors estimate the model using French data. This is appealing because French firms survey data is unique in its inclusion of questions about the explicit division of firms' research effort into basic and applied components. This provides information on the relationship between firm size and research effort and intensity, and pins down 16 of the data moments. In the absence of analogous data for other countries, the estimation procedure employed here assumes that this relationship is broadly representative of that elsewhere.

The country-specific variation in the estimation instead comes from key macroeconomic moments, summarized below. These conditions pin down a range of parameters relevant to the policy recommendations of the model including households' time preference and risk aversion, firms' entry and exit rates, and the probability of mergers.

The data for these aggregate moments come from three sources.:

- *Aggregate macroeconomic data* allow the calculation of three moments the return on sales, firm exit rate, and average aggregate growth. These pin down dynamics of the corporate sector, and so influence expected future profits and thus returns to (and hence incentives for) research.

- *National firm surveys* inform two further moments: the ages of small and large firms. These pin down the speed at which firms grow on average, and also affect firms' incentives to engage in research.
- *PATSTAT data* on individual citations to patents are aggregated to provide the last four moments: the average and standard deviation of the number of citations to public and private patents. These determine the public benefits of research, as they proxy for the positive spillovers from research.

Note that throughout, estimation is conditional on the average values observed for the size of public research and research subsidies. Of course, these moments are not fixed in the policy experiments presented in the chapter but instead vary when alternate policy scenarios are considered. Generally, the model matches the targeted moments well, similar to Akcigit et al. 2021, with a median absolute average error on growth rates of less than one half of one percent.