Mitigating technology gaps' contribution to international income inequality

David Angenendt
TUM School of Management, Technical University of Munich
and Centre for Business
Research, University of Cambridge

Franco Mariuzzo
School of Economics and Centre for Competition Policy, University of East Anglia

Junjun Zhang
School of Economics and Centre for Competition Policy, University of East Anglia

CCP Working Paper 25-01
This version: 06 January 2025

Abstract: Sampson's article, 'Technology Gaps, Trade, and Income,' examines the impact of innovation efficiency gaps on income, wages, and trade dynamics. Our replication, which involves utilizing additional patent metrics, broadening the country selection, extending the time frame, widening the range of the trade elasticity, and excluding outliers, reinforces the significant role of technology gaps in shaping economic inequality. However, our findings indicate that the strength of this effect varies depending on country heterogeneity and the measures of innovation used.

Contact Details:

David Angenendt	david.angenendt@tum.de
Franco Mariuzzo	F.Mariuzzo@uea.ac.uk
Junjun Zhang	junjun.zhang@uea.ac.uk

ISSN 1745-9648

Mitigating technology gaps' contribution to international income inequality *

David Angenendt† Franco Mariuzzo‡ Junjun Zhang§

January 6, 2025

Abstract

Sampson's article, 'Technology Gaps, Trade, and Income,' examines the impact of innovation efficiency gaps on income, wages, and trade dynamics. Our replication, which involves utilizing additional patent metrics, broadening the country selection, extending the time frame, widening the range of the trade elasticity, and excluding outliers, reinforces the significant role of technology gaps in shaping economic inequality. However, our findings indicate that the strength of this effect varies depending on country heterogeneity and the measures of innovation used.

JEL classifiers: D31, F14, O31, O33, O47.

*We thank the University of East Anglia for hosting a Replication Games session in July 2023, and we would like to express our appreciation to Abel Brodeur for sharing valuable insights during the event. We also extend our gratitude to Farasat Bokhari for providing valuable feedback on this work.

†TUM School of Management, Technical University of Munich (Germany) and Centre for Business Research, University of Cambridge (UK). Email: david.angenendt@tum.de.

‡School of Economics and Centre for Competition Policy, University of East Anglia (UK). Email: f.mariuzzo@uea.ac.uk.

§School of Economics and Centre for Competition Policy, University of East Anglia (UK). Email: junjun.zhang@uea.ac.uk.

1 Introduction

Sampson (2023) develops and tests a theory of technology gaps and their implications for inequality, laying the foundations for a new branch of literature. Alongside Buera and Oberfield (2020) and Cai et al. (2022), Sampson contributes to an emerging body of literature investigating the interrelations between international trade, innovation, and technology diffusion. By constructing an endogenous growth model that considers factors at both the industry and country levels, their study captures variations in innovation efficiency among countries and differences in innovation levels and adoption choices across industries. These factors collectively contribute to shaping the equilibrium conditions that influence technology gaps, trade dynamics, and income and wage inequality.

The model by Sampson demonstrates that countries with higher innovation efficiency—measured by changes in either R&D intensity or patent intensity relative to the U.S.—tend to exhibit a greater comparative advantage in industries characterized by a higher degree of innovation dependence. The calibration of innovation efficiency at the country and industry levels is based on various extensive OECD datasets, including bilateral trading data, R&D expenditures, and patent statistics, spanning the years 2010-2014. Additionally, they conduct a counterfactual analysis by assuming uniform innovation efficiency across all sample countries. One of their key findings is that technological disparities account for approximately 25% to 33% of the observed nominal wage variation between the considered economies.

Our primary contribution in this article is to assess the external validity of Sampson's main result. We demonstrate that technology gaps explain a larger share of between-country differences when the country sample is more heterogeneous. We illustrate heterogeneity along two dimensions: the innovation measure and the level of economic development. In more homogeneous samples, the technology gaps explain a lower share in some instances, a negligible portion of wage and income differences. This finding complements Sampson's model by highlighting how representativeness and determinant selection influence empirical results, offering further nuance to the role of technology gaps in economic outcomes.

Sampson investigates innovation efficiency using two distinct data sources: R&D expenditure (OECD 2023b) and counts of 'triadic patent families' (TPFs) (OECD 2023a), both by industry and country. While both sources are commonly employed in the literature, they have limitations in fully capturing innovation. We suggest that R&D expenditure data likely overestimates truly novel research effort, while counts of TPFs likely underestimate such efforts. As discussed later, innovation as captured by the expenditure data also encompasses activities closer in spirit to diffusion than novel innovation. TPFs, on the other hand, are restricted to patent applications jointly filed in the U.S., Europe, and Japan. This introduces two types of biases: one towards highly valuable innovations and a second towards firms from countries that host a 'triadic patent office'.

Building on a suggestion by De Rassenfosse et al. (2013), we adopt an alternative approach by counting patent applications filed at any patent office worldwide, aggregating them into patent families to accommodate differences in national patent laws and procedures. We find a stronger correlation of this measure with both of Sampson's metrics (changes in R&D intensity or patent intensity relative to the U.S.) than the correlation between these metrics themselves. Additionally, while the correlation with the TPF measure is especially strong, the shape of the distribution more closely mirrors that of R&D expenditure. When this broader patent application measure is used to calibrate the model, we observe a more uniform distribution of innovation efficiency across countries and innovation dependence across industries. Broader innovation measures yield attenuated results for wage and income dispersion compared to the original metrics, offering a conservative lower bound on the impact of technology gaps. This highlights the need to consider measure-specific sensitivities while affirming the broader relevance of technology gaps in explaining economic inequality.

Moreover, we explore variations in country composition, excluding six developing and six developed nations identified as outliers in the original paper's Figure 6. Unlike Sampson, we preclude these countries from the dataset before equilibrium calculation. With 24% of the dataset excluded, we scrutinize potential influences on the results. This analysis contributes to evaluating the model's sensitivity to variations in country composition and their implications for the original study's conclusions. We find a significant reduction in the effect of technology gaps on inequality when the six developing countries are removed.

The geographical distribution of innovation has been extensively studied, revealing patterns of both concentration and fragmentation across regions. Studies consistently show that innovation activities tend to cluster in specific locations, driven by factors such as agglomeration economies, access to skilled labor, and R&D spillovers. However, the methods used to measure innovation can significantly affect the interpretation of these patterns, particularly when alternative metrics highlight less concentrated innovation distributions.

R&D surveys are a common method for quantifying innovation efforts but are not without biases. As highlighted by Kleinknecht et al. (2002) and Faber and Hesen (2004), such surveys may underestimate innovation activities, especially in small firms whose inventive work is harder to capture. This issue is especially pronounced in developing and emerging economies, where small firms dominate the economy (Bogliacino et al. 2012). Conversely, overestimation may occur when firms include broader innovation-related expenditures that go beyond the OECD's precise R&D definitions (Kleinknecht et al. 2002). These biases complicate cross-country comparisons and may obscure a more dispersed pattern of innovation, particularly in underrepresented regions.

The literature consistently identifies metropolitan regions as key innovation hubs, benefiting from agglomeration effects, access to skilled labor, and localized knowledge spillovers. For instance, Crescenzi et al. (2023) describe innovation as concentrated in globally interconnected metropolitan hotspots, while Boschma (2005) highlights the role of various forms of proximity—geographic, social, and organizational—in driving innovation. Glaeser (2008) and Carlino and Kerr (2015) further emphasize the importance of spatial clustering and face-to-face interactions in fostering innovation. Such concentrated patterns align with findings that R&D productivity is often enhanced in densely populated innovation clusters (Carlino and Kerr 2015).

Alternative measures of innovation paint a more fragmented picture. Broader metrics, which capture dispersed activities, attenuate the apparent dominance of innovation hubs. This is particularly relevant when accounting for regions where small firms or informal innovation efforts play a significant role, as these are often underrepresented in traditional metrics. Studies such as Dominguez Lacasa et al. (2019) and Lee et al. (2021) suggest that patent data, while useful, may not fully capture the diversity of innovation activities in emerging economies. Patent-based measures face particular challenges in cross-country comparisons, including 'home bias,' wherein innovation activities are disproportionately associated with a single patent office (Hinze and Schmoch 2005). To address this, researchers have developed measures such as triadic patent families (Grupp et al. 1996), which count patents filed in multiple jurisdictions to reduce bias and enhance comparability among developed nations. However, this approach is less suitable for comparisons involving developing economies due to cost barriers and differing motivations for patent filings (Frietsch and Schmoch 2009). For example, Khan and Dernis (2006) demonstrate that more inclusive patent measures, such as Patent Cooperation Treaty applications, yield higher shares of patents from developing countries. Yet, even these approaches struggle to fully capture innovation in less-developed regions, where firms may lack the resources to file patents internationally. Recent work by Bruns and Kalthaus (2020) underscores the importance of selecting appropriate patent measures, demonstrating that restrictive counts (e.g., triadic patents) can decrease effect sizes in some studies while increasing them in others. This variability highlights the nuanced impact of alternative patent measures on observed geographical innovation patterns.

Our findings contribute to this ongoing discourse by illustrating how alternative measures of innovation—those that reflect broader, less concentrated patterns—yield attenuated results compared to traditional metrics. These findings align with existing literature, which demonstrates that while innovation tends to concentrate in metropolitan hotspots, dispersed regions also meaningfully contribute to global innovation, albeit with weaker outcomes. By connecting these insights with the methodological challenges of measuring R&D and patent data, our analysis underscores the importance of inclusive approaches that better represent innovation across diverse geographic and economic contexts.

Additionally, we explore six modifications to the calibration decisions made by the author in the original study. (i) We incorporate an analysis of additional patent metrics from the OECD and a measure of patent citations based on our more extensive patent measure. (ii) By broadening the scope, we expand the number of countries from 26 to 29. This expansion includes country-years with observations for at least 10 industries, relaxing the original requirement of 14 industries. (iii) By broadening the temporal scope, we incorporate an additional two years of data into the original four-year dataset. Subsequently, we explore various partitions of time periods to assess their impact. (iv) The original paper includes a robustness check in which the author varies the trade elasticity within the range of 2.5 to 8.5. In our analysis, we extend this examination by adjusting the trade elasticity to a lower value, 1, and a higher value, 10.5, to further assess its robustness. (v) We exclude two outliers, specifically, the Paper and Paper Products (17) industry and the Agriculture, Forestry, and Fishing (0103) industry. These outliers, identified using information from Figure 4 in Sampson's paper, are examined for their impact on the results. (vi) We rerun the regression Equation (33) from Sampson and examine how outliers affect the regression. The outcomes of this comprehensive set of robustness checks affirm the reliability and consistency of the results observed in the original paper. Some variation in the results is observed when different patent metrics are used. Specifically, Sampson's findings on wage or income inequality sometimes become more pronounced and other times more confined. The key takeaway is that when the metric captures patent concentration, the results tend to align with or exceed those in the original paper. Conversely, when the patent measure is more evenly distributed across countries, the observed inequality decreases.

The structure of the paper is the following. In Section 2, we explore the inherent challenges of Sampson's research, while Section 3 evaluates the robustness of their methodology. This analysis is organized into various subsections, each addressing different aspects of Sampson's work. Through this critical lens, we examine the employed methodology, assess data reliability, and evaluate the validity of the conclusions drawn. Additionally, we highlight the strengths and limitations within Sampson's research, offering a nuanced perspective on its overall quality and contribution to the field. Section 4 concludes.

2 Unveiling challenges

2.1 Using a more representative measure of innovation

Sampson relies on two distinct data sources to obtain a proxy for innovation efficiency, captured by variations in R&D intensity (the ratio of R&D expenditure to value added) or patenting intensity (the ratio of patents count to value added) relative to the U.S. The first is the OECD's Analysis of Business and Economic Research and Development (ANBERD) database, providing aggregated business R&D expenditure data by industry and country (OECD 2023b). The second source involves counts of 'triadic patent families' (TPF) by technology class, also sourced from the OECD (OECD 2023a). A triadic patent family encompasses patent applications containing the same invention filed simultaneously at the U.S., European, and Japanese patent offices. These TPF counts are then mapped to industries using a widely-accepted technology-industry correspondence table (Lybbert and Zolas 2014). While both OECD data sources are commonly employed in the literature, they have limitations in fully capturing innovative efforts.

In Section I.A of their article, Sampson defines R&D as an investment 'to create new ideas and technologies through innovation' (p. 477). Alternatively, firms can pursue an 'adoption' strategy, oriented towards 'learning about and implementing existing production techniques'—often referred to as the imitation or diffusion of existing technology in other contexts. If R&D expenditure data was reliably restricted to effort spent on 'innovation' as per the above definition, it would be the preferred data source to proxy for innovation efficiency. However, the OECD's ANBERD R&D expenditure data, collected following the definitions in the Frascati Manual (OECD 2015), encompasses activities that could be appropriately categorized as adoption efforts. This includes tasks such as identifying discrepancies when replicating existing results and incorporating additional material into the maintenance manual of a complex system (OECD 2015, p. 46). Given the challenges of measuring R&D in general, and its comparability across countries in particular, as highlighted in Sampson (2023, p. 495), this potential mismeasurement makes it essential to validate the results using an alternative measure.

The most desirable alternative measure is counts of all patent filings submitted by nationals of a country. In contrast, the triadic patent families used by Sampson are often considered a proxy for particularly valuable technologies, due to the costs of seeking protection in several patent offices (Criscuolo 2006, Nagaoka et al. 2010, van Zeebroeck 2011, De Rassenfosse et al. 2013).

Unlike commercial outcomes, Sampson's study specifically focuses on technological progress. This distinction is crucial because technological advancements do not always align with commercial profitability; significant inventions may yield little financial gain for the inventor, even if they are valuable to others (Shankar et al. 1998, Hoppe 2000). This constrains the applicability of TPFs as an innovation proxy. Moreover, the 'home advantage' of firms from countries hosting triadic patent offices, along with their increased financial capacity to file foreign patents, can introduce biases in international comparisons, potentially disadvantaging emerging economies and exacerbating disparities in measured innovation.[1] Recognizing these considerations is vital for studies aiming to compare countries at diverse developmental stages. Finally, the importance of patent protection varies across industries, as does the propensity to seek patent protection internationally. As a result, differences in the composition of a country's industrial base can significantly impact triadic patent counts, even among nations with similar levels of economic development Sternitzke (2009).

We employ a more representative metric by counting *all* patent application families, regardless of the patent offices where they were submitted. To achieve this, we rely on the European Patent Office (EPO)'s Patent Statistics (PATSTAT) database (version autumn 2021), which consolidates data from 90 global patent offices and is widely used in research (Kang and Tarasconi 2016). Notably, the OECD also relies on this database for their

[1] For instance, De Rassenfosse et al. (2022) find that obtaining a patent in an export destination significantly increases the patent owner's exports to that market. This suggests that patenting abroad is often motivated by the intention to export to those markets, a strategy that may not always be economically feasible for firms from emerging economies. Consequently, triadic patents may not fully capture the innovation capacity of those countries.

TPF counts.[2] This alternative approach alleviates 'geographic bias' and refrains from imposing a 'filter on patent value' (De Rassenfosse et al. 2013).

While we advocate for a more inclusive patent count, we concur with the established use of *patent families*—defined by OECD (2009) as sets of related patents filed in multiple countries to protect the same invention—instead of raw counts of patent applications to facilitate international comparisons across diverse national patent systems (Dernis et al. 2002, Nagaoka et al. 2010, De Rassenfosse et al. 2013). For instance, countries like Japan traditionally require the filing of a greater number of separate patent applications for the same invention compared to the U.S. and Europe. Adopting the concept of patent families, each linked to a unique International Patent Documentation (INPADOC) patent family ID (`INPADOC_FAMILY_ID`) in PATSTAT Table 201, effectively addresses and mitigates this issue (Park and Hingley 2009).

For each `INPADOC_FAMILY_ID`, we determine the earliest filing date to assign a unique invention year. Inventor countries are extracted from PATSTAT Tables 206 and 207. However, owing to significant missing data in these tables (De Rassenfosse et al. 2013, 2019), we additionally utilize imputed location data provided by De Rassenfosse and Seliger (2021). Patent families are linked to countries based on the relative share of each country among all inventor countries within the family (Dernis and Khan 2004). For example, if a family has 'Italy' assigned twice, 'Germany' once, and 'United Kingdom' once as the inventor countries, the family is accounted for as 0.50 patent families for Italy, and 0.25 for both Germany and the UK, respectively.[3] Only inventor countries that are part of the original dataset and patent families with a first filing between 2010-2014 are considered.

Table 1 provides summary statistics for the various innovation measures analyzed in

[2]For further details, refer to the OECD Triadic Patent Families database usage instructions, July 2020, available at https://www.oecd.org/sti/intellectual-property-statistics-and-analysis.htm#ip-data.

[3]While the OECD, relying solely on U.S. patent office data, counts each inventor only once per patent family, our data does not allow tracking the same inventor across different applications. Consequently, we count an inventor country as many times as it occurs within the same patent family. The impact of this difference in approach on national patent counts is minimal, as demonstrated in Table 1. Unreported robustness checks, involving calculating country shares first at the application level before aggregating them at the family level, yield nearly identical results.

Table 1. Summary statistics of different innovation measures.

Innovation	R&D exp. (bn.)	count of patent application families					
Data source:	OECD	OECD	PS	PS	PS	PS	PS
Patent family definition:		triadic	triadic	triadic	triadic	ID	ID
Location data:			US data only		imputed		imputed
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Full sample ($N = 2,454$)							
mean	112	91	90	76	39	1304	1676
SD	1100	367	371	318	219	4908	7135
median	0.184	4.8	6	5	1	124	101
mean/SD	*0.10*	*0.25*	*0.24*	*0.24*	*0.18*	*0.27*	*0.23*
mean/median	*608.70*	*18.96*	*14.70*	*15.20*	*39.00*	*10.52*	*16.59*
Country level ($S = 25$)							
mean	97.6	90	89	76	39	1300	1663
SD	383	209	207	177	117	2609	3999
median	0.398	13.3	14	11.8	3.3	205	188
mean/SD	*0.25*	*0.43*	*0.43*	*0.43*	*0.33*	*0.50*	*0.42*
mean/median	*245.23*	*6.77*	*6.21*	*6.44*	*11.82*	*6.34*	*8.85*
Industry level ($J = 20$)							
mean	109	91	90	76	39	1308	1682
SD	247	120	121	104	63	1831	2354
median	26.5	56	59	50	26	853	1027
mean/SD	*0.44*	*0.76*	*0.74*	*0.73*	*0.62*	*0.71*	*0.71*
mean/median	*4.11*	*1.63*	*1.52*	*1.52*	*1.50*	*1.53*	*1.64*

Notes: PS=PATSTAT, the EPO's worldwide statistical database; ID=INPADOC patent family definition. Each observation is a country-industry-year tuple, where the industry (J) is one of the 20 two-digit ISIC manufacturing sectors for which Sampson obtained R&D and patent data, and the country (S) is one of the 25 OECD countries used in their analysis.Using $T = 5$ years of data for each country, the total number of observations therefore would be $S \times J \times T = 2,500$. The actual sample size of 2,454 indicates 46 cases of missing innovation data Columns (1) and (2) report, respectively, statistics for the R&D intensity and patent intensity used in the original paper. Columns (3) through (7) use patent data independently obtained from PATSTAT, with variations in aggregation level and selection of patent applications. Specifically: Column (3) utilizes the OECD's triadic patent family definition and use inventor country locations only from patent applications filed at the US Patent & Trademark Office, aiming to replicate Column (2). Columns (4) and (5) utilize the OECD's triadic patent family definition but include inventor country locations from all available patent applications . Columns (6) and (7) employ the `INPADOC` patent family definition that comes with PATSTAT, including patent applications not filed simultaneously in all 'top-3' patent offices. Additionally, Columns (3), (4) and (6) only use inventor location data from PATSTAT, (5) and (7) incorporate additional inventor location data from De Rassenfosse and Seliger (2021).

our replication. While our objective is to compare the author's results with those obtained using our preferred innovation measure, we also report several intermediate measures to systematically investigate the influence of different choices in data construction on the results. In addition to reporting the author's R&D expenditure and patent count (Columns 1 and 2), we replicate the OECD's triadic patent measure using multiple sources. Data on patent applications and inventor locations are sourced from the EPO's PATSTAT database. Individual patent applications are grouped into families using two approaches: one with data from the OECD's Science, Technology, and Innovation (STI) department, which is directly compatible with PATSTAT (Columns 3–5), and the other using the alternative 'INPADOC' patent family (IPF) definition provided by the EPO as part of PATSTAT (columns 6–7). We also explore the impact of different sources of inventor location data. Column (3) follows the OECD's approach, using inventor location data exclusively from U.S. Patent & Trademark Office filings. Columns (4) and (6) use inventor locations data from all available patent offices in PATSTAT. Lastly, Columns (5) and (7) incorporate additional inventor locations imputed by De Rassenfosse and Seliger (2021) for applications with missing data in PATSTAT.

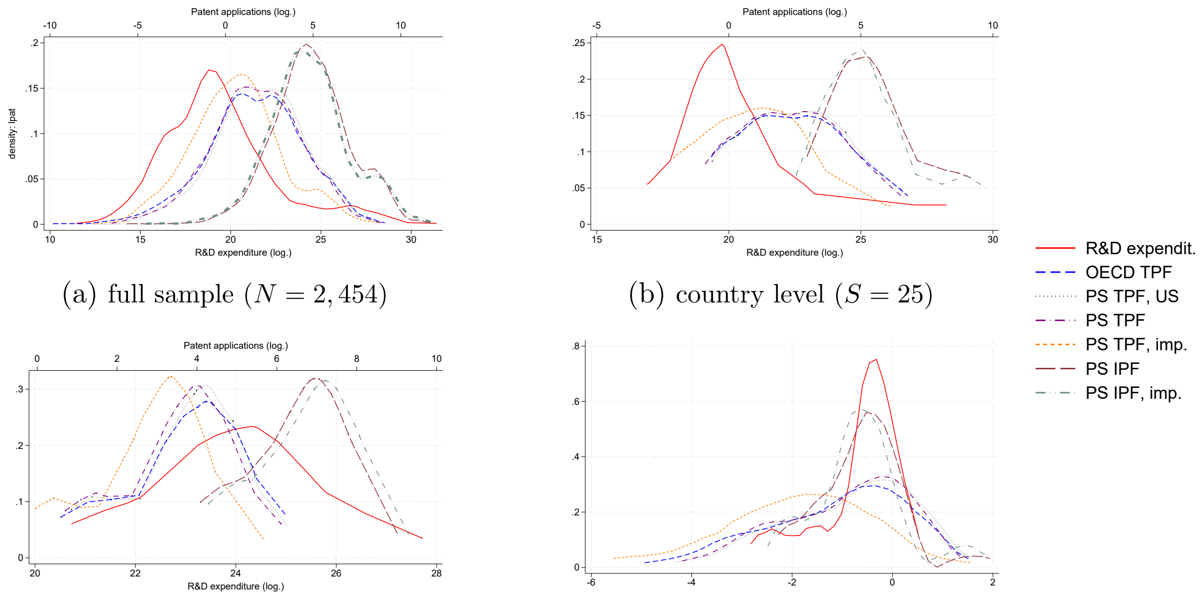
The PATSTAT-based replication yields overall comparable values. Using a broader range of inventor location sources reduces the mean for the triadic patent family definition but increases it for the worldwide definition.[4] Panels a-c in Figure 1 reveal comparable distribution shapes across R&D measures, with R&D expenditures exhibiting a notably longer right tail. The utilization of imputed inventor locations and the worldwide patent family definition slightly extends the right tail of the patent count distribution. Particularly when considering country-level data (Figure 1b), the distribution of IPF counts aligns more closely with the R&D expenditure distribution than the TPF counts. It is only at the industry level that the shape of the distributions of the patent counts more closely mirror each other compared to the distribution of R&D-expenditure (Figure 1c).

To investigate the extent to which the different innovation measures capture the same definition of innovation, we calculated correlation coefficients, reported in Table A1 in the appendix.[6] It is noteworthy that the IPF count is more strongly correlated with R&D expenditure (0.53 without and 0.43 with imputation of inventor locations) than the TPF count used by Sampson (0.20 without and 0.13 with imputation). At the same time, the correlation between the two patent-based counts remains very high (0.80 without and

[4]This is likely due to the combination of family definition and data sources. For TPFs, PATSTAT contains at least *some* inventor location. However, certain offices, like the Japanese Patent Office, do not provide location data to PATSTAT (De Rassenfosse and Seliger 2021). Thus, using more complete location coverage distributes the same number of TPFs across a larger number of countries, some of which are outside of the sample, reducing the sample mean. Conversely, many IPFs entirely lack location data in PATSTAT. (Remember that an IPF may comprise as little as a single patent application.) Expanding location coverage hence allows more IPFs to be assigned a country and therefore used in analysis, increasing the mean.

[5]Notes: 'imp' in the legend denotes the use of imputed inventor locations.

[6]To further validate the IPF count, we examined its correlation with R&D expenditure values obtained from external survey data. Due to the small sample size from various country-year combinations, these results are presented only in Appendix A of the appendix, and were unable to use the survey-based innovation measures to calibrate Sampson's model. However, the correlations suggest that in the survey sample, the IPF count performs as a stronger proxy for business R&D expenditure than both the TPF count and the OECD's aggregate R&D expenditure.



(a) full sample ($N = 2,454$)

(b) country level ($S = 25$)

(c) industry level ($J = 20$)

(d) country-level innovation efficiency

Figure 1. Density plots comparing different innovation measures.[5]

0.91 with imputation).

It is well known that the ability and propensity to patent inventions differs between industries. We therefore also calculate industry-specific correlation coefficients. Table A2 presents the averages for these 20 industry-specific correlations. The average correlations between the patent-based measures and R&D spending are all larger than the correlations calculated using the full sample above, but still considerably greater for our suggested measures (between 0.27 and 0.31 for the TPF counts and between 0.70 and 0.72 for the IPF counts).

To illustrate the range of heterogeneity between industries, the maximum correlation between the two measures used by the author is 0.73, found in industry 13 (textiles), whereas the correlations between those two and our preferred measure are 0.96 for R&D expenditure and 0.87 for the TPF count (using 97 jointly non-missing observations). At the other end of the spectrum is industry 14 (wearing apparel) with a correlation between the author's measures of only 0.05. In the same industry, our IPF count is correlated with R&D expenditure at 0.54 and with the TPF count at 0.80 (using 91 observations).

All this corroborates our interpretation of the IPF measure capturing a definition of R&D that is somewhat 'in the middle' between the author's two original measures. In fact, the observed mean correlation between IPF and R&D expenditure of 0.891 is only 0.001 smaller than the maximum possible correlation. In other words, given its mean correlation with R&D expenditure, our suggested patent count is, on average across the sample industries, as strongly correlated with the TPF count as is technically possible, and vice versa.[7]

Figure 1d presents the calculation of countries' innovation efficiency, which is R&D intensity or patent intensity normalized to the U.S. We primarily have data on patent intensity, precisely defined as the ratio of the average patent family count to the average value added over time for each industry and country. Following Sampson, this ratio is then logged, and the log value for Germany is subtracted. Next, the median across industries is calculated for each country, and finally, the value for the U.S. is subtracted to obtain the patent measure of innovation efficiency. The distributions obtained using R&D expenditure and IPF counts exhibit similar shapes, contrasting with the flatter distribution observed when using TPF counts. The distribution based on expenditures features a longer left tail, while the IPF-based distribution contains an upper outlier. Details on the underlying values for the density curves in Figure 1d are provided in Table 2.

The values in Table 2 show that when employing worldwide patent counts, notable shifts in efficiency are observed for certain countries. Belgium and France experience a decrease, while Japan and Korea witness an increase. Particularly, countries at the lower end in the author's original values (Chile, Czechia, Hungary, Mexico, Poland, and Turkey) show enhanced innovation efficiency relative to the U.S. when using worldwide instead of triadic patent counts. Ireland and Poland also display an improvement compared to expenditure-based estimates. Across all columns, worldwide patent counts yield the highest mean estimates, aligning with the objective of mitigating the 'disadvantaging' of less-developed economies in the patent-based measure.

[7] We calculate the maximum possible correlation by replacing one of the correlations in the matrix by ρ, calculating the determinant of the matrix, and then exploiting the fact that matrix must be positive semi-definite. This is a well-known property of a covariance matrix that can be shown to apply also a matrix of mean covariances. The resulting interval of solutions to the quadratic equation describes the range of feasible values of ρ given the other covariances.

Table 2. Innovation efficiency by country — including alternative measures of patent counts

Innovation efficiency measure:	R&D intensity	Patenting intensity					
Data source:	OECD	OECD	PS	PS	PS	PS	PS
Patent family definition:		triadic	triadic	triadic	triadic	ID	ID
Location data:			US only		imputed		imputed
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Australia (AUS)	-0.25	-0.99	-0.74	-0.71	-2.01	-1.15	-0.49
Austria (AUT)	**0.00**	-0.15	**0.07**	-0.16	-0.95	-0.34	-0.60
Belgium (BEL)	**0.27**	-0.08	**0.22**	**0.04**	-0.44	-0.76	-1.14
Canada (CAN)	-0.57	-1.10	-0.90	-0.74	-1.46	-0.49	-0.70
Chile (CHL)	-2.50	-3.00	-2.42	-2.55	-3.07	-1.92	-2.21
Czechia (CZE)	-1.30	-2.55	-2.28	-2.46	-2.79	-1.06	-1.02
Germany (DEU)	-0.31	**0.05**	**0.10**	-0.03	-0.96	**0.14**	**0.09**
Denmark (DNK)	-0.38	-0.02	**0.13**	**0.06**	-0.63	-0.19	-0.62
Spain (ESP)	-0.84	-1.79	-1.40	-1.62	-2.43	-1.03	-1.06
Finland (FIN)	-0.09	-0.30	**0.09**	**0.02**	-1.03	-0.17	-0.16
France (FRA)	**0.30**	**0.11**	**0.30**	0.16	-1.39	-0.12	-0.05
United Kingdom (GBR)	-0.45	-0.05	-0.29	-0.18	-1.21	-0.38	-0.69
Hungary (HUN)	-1.38	-2.29	-1.96	-2.01	-2.89	-0.86	-0.87
Ireland (IRL)	-0.72	-0.85	-0.34	-0.48	-0.61	-0.11	-0.26
Italy (ITA)	-0.66	-0.97	-0.77	-0.90	-2.15	-1.16	-1.23
Japan (JPN)	0.14	0.91	**0.96**	**0.84**	**0.91**	-0.08	**1.19**
Korea (KOR)	-0.24	-0.50	-0.40	-0.38	-2.64	**1.62**	**1.70**
Mexico (MEX)	-2.61	-4.30	-3.74	-3.62	-4.82	-1.95	-1.98
Netherlands (NLD)	-0.52	**0.51**	**0.59**	**0.48**	-0.84	**0.03**	-0.38
Norway (NOR)	-0.43	-0.87	-0.65	-0.69	-2.07	-0.72	-0.97
Poland (POL)	-2.05	-2.60	-2.26	-2.16	-3.14	-0.55	-0.43
Portugal (PRT)	-0.52	-2.52	-2.04	-2.13	-3.52	-1.85	-1.92
Slovenia (SVN)	-0.31	-1.41	-1.51	-1.71	-1.98	-0.61	-0.54
Turkey (TUR)	-1.78	-3.39	-2.97	-3.04	-4.91	-1.86	-1.95
USA	0.00	0.00	0.00	0.00	0.00	0.00	0.00
mean	-0.69	-1.13	-0.88	-0.96	-1.88	-0.62	-0.65
SD	0.81	1.34	1.22	1.20	1.40	0.80	0.90

Notes: The innovation measure used in each column is identical to the one described in the Notes below Table 1. Innovation efficiency is expressed relative to that of the United States, which is normalized to one. For this reason, the table shows a value of zero in each column for the USA. To aid interpretation, non-negative values are highlighted in bold. R&D intensity is calculated as the ratio of R&D expenditure to industry value-added. What we term "patenting intensity" is that same ratio but with the corresponding patent count in the numerator.

Table 3 and Table 4 replicate Tables 1 and 3 from the original paper. The first set of results is derived from Sampson's Equation (33), whereby the dependent variable is $\log\left(\frac{EX_{js\tilde{s}}}{EX_{j\tilde{s}\tilde{s}}}\right) - (\sigma-1)\log\left(\frac{w_{\tilde{s}}}{w_s}\right)$, where j denotes the industry, s represents the exporting country and $\widetilde{s}$ the destination country. So that, $EX_{s\widetilde{s}}$ signifies the value of trade from country s to country $\widetilde{s}$. Additionally, σ represents the Armington elasticity of demand, and w_s the wage in country s. Moreover, we describe the key independent variable of interest, b_s, the log of normalized innovation efficiency of country s, specifically in the form of medians calculated across industries, as described earlier, and denoted as $\log\left(\frac{RD_{js}}{RD_{j\tilde{s}}}\right)$. The term $(1-\sigma)b_s$ on the right-hand side of Equation (33) is interacted with country dummies. This adjustment provides coefficients for innovation dependence at the country level, with the U.S. serving as the reference country. In Table 3, Columns (5) and (6) depict a slightly reduced estimated average innovation dependence. Notably, the industry 'Computer, Electronic, and Optical Products (26)' is no longer a pronounced outlier among industries with the highest innovation dependence. Meanwhile, 'Mining and Quarrying' maintains its position as the industry with the lowest innovation dependence. The null hypothesis of equal innovation dependence across industries is now rejected below the 5% level, as opposed to the previous 1% level.

In Table 4, countries exhibit increased similarity when our alternative patent count is employed. The model now accounts for only 17% of nominal wage dispersion, roughly half of the explanatory power achieved with the author's original innovation measures. The model's ability to explain real income dispersion is diminished by the same relative amount. The anticipated average change in both outcomes from eliminating differences in innovation efficiency between countries is now reduced to approximately one-third of the original estimates.[8]

While we acknowledge, as stated in the introduction, that R&D expenditure likely overcounts true innovation efforts and triadic patent families likely undercount them,

[8]While the dispersion of innovation intensity is reduced, as indicated by the SID values in Table 3, the overall impact of narrowing innovation gaps on income and wage dispersion is not known in advance. Wage or income dispersion is determined by the ratio of counterfactual to observed standard deviations. Although the counterfactual standard deviation (the numerator of the equation on p. 505 of Sampson) decreases as innovation intensity becomes more uniform, the observed standard deviation of wages or income also declines, making unpredictable the effect on the ratio.

Table 3. Innovation dependence by industry — alternative patent counts

Innovation efficiency measure:	R&D intensity	Patenting intensity				
Data source:	OECD	OECD	OECD	PS	PS	PS
Patent family definition:		triadic	triadic	triadic	ID	ID
Location data:			US only	imputed		imputed
	(1)	(2)	(3)	(4)	(5)	(6)
Agriculture, forestry and fishing (0103)	0.17	0.01	0.01	0.01	0.02	0.06
	(0.09)	(0.06)	(0.06)	(0.05)	(0.07)	(0.05)
Mining and quarrying (0508)	-0.11	-0.14	-0.16	-0.05	-0.21	-0.22
	(0.13)	(0.08)	(0.09)	(0.1)	(0.07)	(0.07)
Food products, beverages and tobacco (1012)	0.21	0.06	0.07	0.06	0.04	0.07
	(0.08)	(0.06)	(0.06)	(0.04)	(0.06)	(0.05)
Textiles (13)	0.29	0.12	0.12	0.03	0.11	0.12
	(0.06)	(0.05)	(0.05)	(0.03)	(0.05)	(0.03)
Wearing apparel (14)	0.33	0.13	0.12	0.01	0.09	0.11
	(0.05)	(0.04)	(0.04)	(0.02)	(0.03)	(0.02)
Leather and related products (15)	0.34	0.12	0.10	0.02	0.07	0.13
	(0.08)	(0.07)	(0.07)	(0.04)	(0.05)	(0.04)
Wood and products of wood and cork, except furniture (16)	0.20	0.03	0.03	0.01	0.02	0.06
	(0.07)	(0.05)	(0.05)	(0.03)	(0.05)	(0.04)
Paper and paper products (17)	0.34	0.13	0.14	0.09	0.10	0.13
	(0.07)	(0.05)	(0.05)	(0.04)	(0.06)	(0.05)
Printing and reproduction of recorded media (18)	0.27	0.11	0.12	0.10	0.09	0.10
	(0.06)	(0.04)	(0.04)	(0.03)	(0.06)	(0.05)
Coke and refined petroleum products (19)	0.14	0.05	0.05	0.06	0.02	0.05
	(0.08)	(0.04)	(0.04)	(0.04)	(0.07)	(0.06)
Chemicals and chemical products (20)	0.38	0.19	0.20	0.14	0.14	0.15
	(0.09)	(0.06)	(0.06)	(0.06)	(0.08)	(0.05)
Basic pharmaceutical products and pharmaceutical preparations (21)	0.22	0.17	0.18	0.13	0.13	0.14
	(0.14)	(0.10)	(0.10)	(0.08)	(0.09)	(0.07)
Rubber and plastics products (22)	0.38	0.18	0.19	0.11	0.15	0.17
	(0.05)	(0.04)	(0.04)	(0.03)	(0.06)	(0.04)
Other non-metallic mineral products (23)	0.30	0.12	0.13	0.07	0.12	0.14
	(0.06)	(0.04)	(0.04)	(0.03)	(0.05)	(0.04)
Basic metals (24)	0.27	0.18	0.20	0.14	0.09	0.12
	(0.07)	(0.03)	(0.04)	(0.04)	(0.06)	(0.05)
Fabricated metal products, except machinery and equipment (25)	0.33	0.14	0.14	0.10	0.12	0.14
	(0.06)	(0.04)	(0.04)	(0.03)	(0.06)	(0.05)
Computer, electronic and optical products (26)	0.60	0.30	0.30	0.13	0.12	0.17
	(0.12)	(0.05)	(0.06)	(0.03)	(0.05)	(0.06)
Electrical equipment (27)	0.37	0.19	0.19	0.12	0.10	0.15
	(0.10)	(0.04)	(0.04)	(0.04)	(0.05)	(0.05)
Machinery and equipment n.e.c. (28)	0.38	0.21	0.23	0.16	0.14	0.18
	(0.11)	(0.06)	(0.06)	(0.05)	(0.07)	(0.06)
Motor vehicles, trailers and semi-trailers (29)	0.27	0.19	0.21	0.17	0.11	0.15
	(0.08)	(0.03)	(0.04)	(0.03)	(0.06)	(0.06)
Other transport equipment (30)	0.26	0.00	0.00	0.01	-0.01	0.05
	(0.13)	(0.05)	(0.06)	(0.06)	(0.06)	(0.06)
Furniture, other manufacturing (3133)	0.25	0.10	0.12	0.07	0.11	0.12
	(0.07)	(0.05)	(0.04)	(0.04)	(0.06)	(0.04)
Observations	171K	171K	171K	171K	171K	171K
R-squared	0.70	0.69	0.69	0.69	0.68	0.69
TCC†	Yes	Yes	Yes	Yes	Yes	Yes
PLC†	Yes	Yes	Yes	Yes	Yes	Yes
CAC†	Yes	Yes	Yes	Yes	Yes	Yes
AID†	0.28	0.12	0.12	0.08	0.08	0.10
SID	0.13	0.09	0.10	0.06	0.08	0.08
F test	0.07	0.00	0.00	0.00	0.08	0.02

Notes: †TCC=Trade-cost controls; PLC=Productivity-level controls; CAC=Comparative advantage controls; AID=Average innovation-dependence; SID=Standard deviation of innovation-dependence. The F-test assesses whether innovation-dependence exhibits equal levels of significance across various industries. The standard errors are clustered by importer-industry, and they are presented within parentheses. We incorporate exporter-industry fixed effects, industry dummy variable interactions with six bilateral distance intervals, and with a dummy variable indicating whether the nations share a border, a common language, or a free trade agreement—all examples of trade cost restrictions. Productivity is significantly influenced by rule of law, corruption prevention, political stability, regulatory quality, voice and accountability, ease of doing business, and private credit as a percentage of GDP. Comparative advantage controls include interactions of industry dummy variables with the importer's rule of law, log private credit as a proportion of GDP, log physical capital per employee, and human capital. The F-test equalizes innovation-dependence across industries (p-value reported). The innovation measure used in each column is identical to that one described in the Notes below Table 1, except that here we do not report results obtained with the count of triadic patent families with inventor locations based on PATSTAT data only (column (4) in Tables 1 and 3). Columns (1) and (2) are identical to columns (3) and (4), respectively, in Sampson's Table 1. For Columns 3-6, the only distinction lies in the innovation measure used to calculate innovation intensity; all computations follow the same methodology as those used by Sampson.

both measures may introduce bias in the same direction when comparing country pairs. Imagine a scenario where each country's total R&D effort comprises (1) inventions with high immediate commercial value, (2) inventions with low immediate commercial value, and (3) imitation efforts (that are still novel enough to be considered R&D by the OECD's definition). Triadic patent counts would approximate (1), while R&D expenditure proxies the sum of all three parts. Our proposed measure aims to proxy the sum of (1) and (2). If a reference country exhibits both higher total R&D expenditure and a larger share of triadic patent applications among all patented inventions, countries will appear more disparate using the author's R&D measures than with our measure. Section 6 of De Rassenfosse et al. (2013) suggests the presence of such differences between patent indicators across countries, but a more in-depth investigation would necessitate additional data, surpassing the scope of this article.

2.2 Exploring variations in the number of countries

In this section, we refine the composition of countries by initially excluding six developing nations identified as outliers in Figure 6 of the original paper, followed by the exclusion of six developed countries that appear on the top north-east side of the same figure. It is noteworthy that the exclusion of six developing countries, identified as outliers in Figure 6, has also been carried out in Sampson (2023). However, a key distinction lies in the approach: while Sampson removed the countries after calculating the equilibrium, we preclude these countries from the dataset before performing equilibrium calculation.

Our rationale for this choice arises from data limitations, as Sampson's list of countries

Table 4. Counterfactual results — alternative patent counts

		Innovation efficiency measure:	R&D intensity	Patenting intensity				
		Data source: Patent family definition: Location data:	OECD	OECD triadic	PS triadic US only	PS triadic imputed	PS ID	PS ID imputed
			(1)	(2)	(3)	(4)	(5)	(6)
1.	Nominal wage	Average change relative to US	0.18	0.14	0.11	0.15	0.05	0.06
		Dispersion ratio	0.32	0.27	0.26	0.2	0.11	0.17
2.	Real income per capita	Average change relative to US	0.06	0.04	0.04	0.05	0.02	0.02
		Dispersion ratio	0.17	0.13	0.13	0.12	0.05	0.08

Notes: Column definitions are equivalent to those in Table 3.

notably represents an incomplete subset of the 38 OECD countries and, naturally, only a fraction of the 195 countries worldwide. Our objective is to evaluate the impact of this incomplete representation on the equilibrium and the overarching message conveyed in the paper. By investigating whether the removal of 6 out of 25 countries, constituting 24% of the dataset, influences the results, we aim to shed light on potential implications for the external validity of the findings. Caution may be warranted if such exclusions significantly impact the outcomes.

Table 5. Innovation intensity summary statistics

b_s	obs	Mean	Std. dev	Min	Max
		R&D intensity			
Model 1	25	-0.69	0.81	-2.61	0.30
Model 2	19	-0.29	0.33	-0.84	0.30
Model 3	19	-0.90	0.81	-2.61	0.00
		Patenting intensity			
Model 1	25	-1.13	1.34	-4.30	0.91
Model 2	19	-0.53	0.83	-2.52	0.91
Model 3	19	-1.55	1.25	-4.30	0.05

Notes: Model 1 displays the original results derived from the original paper's innovation efficiency. Model 2 presents results after excluding the six low-income countries before calculating the equilibrium—a departure from the original methodology where these countries were excluded after equilibrium calculations. Finally, in Model 3, we omit the six high-income countries, as identified in Figure 6 of the original paper.

In this section, each table is structured to offer varied perspectives on the results. Tables with columns (or rows) labeled with a prefix of '1' (model 1) present the original findings. Similarly, those labeled with a prefix of '2' (model 2) display results after excluding six developing countries. Finally, tables with a prefix of '3' (model 3) document outcomes following the removal of six developed countries.

In Table 5, we observe that the average value R&D intensity relative to the U.S. in the case of removing developing countries, -0.29, is considerably lower in absolute value than the original value of -0.69. Conversely, with the exclusion of developed countries, the value of -0.90 is higher in absolute value than the original. The rationale behind this discrepancy lies in the characteristics of model 2, where the data sample exclusively comprises high-income countries. These high-income countries exhibit a relatively similar

Table 6. Innovation dependence by industry

Innovation efficiency measure	R&D intensity			R&D intensity			R&D intensity			Patenting intensity		
Industries	(1A)	(2A)	(3A)	(1B)	(2B)	(3B)	(1C)	(2C)	(3C)	(1D)	(2D)	(3D)
0103 (Agriculture)	0.45	0.05	0.50	0.33	0.24	0.39	0.17	0.25	0.18	0.01	-0.04	0.09
	(0.06)	(0.14)	(0.07)	(0.05)	(0.15)	(0.05)	(0.09)	(0.14)	(0.10)	(0.06)	(0.08)	(0.07)
0508 (Mining)	0.37	-0.21	0.42	0.25	-0.01	0.31	-0.11	-0.04	-0.09	-0.14	-0.20	-0.15
	(0.09)	(0.25)	(0.11)	(0.07)	(0.21)	(0.08)	(0.13)	(0.18)	(0.19)	(0.08)	(0.06)	(0.13)
1012 (Food)	0.48	0.09	0.50	0.36	0.28	0.39	0.21	0.28	0.16	0.06	0.00	0.13
	(0.05)	(0.12)	(0.07)	(0.04)	(0.14)	(0.05)	(0.08)	(0.10)	(0.10)	(0.06)	(0.07)	(0.06)
13 (Textiles)	0.51	0.05	0.53	0.42	0.24	0.46	0.29	0.33	0.28	0.12	0.10	0.13
	(0.05)	(0.13)	(0.06)	(0.05)	(0.13)	(0.06)	(0.06)	(0.06)	(0.10)	(0.05)	(0.05)	(0.11)
14 (Apparel)	0.47	-0.09	0.49	0.37	0.21	0.42	0.33	0.18	0.38	0.13	0.04	0.24
	(0.06)	(0.11)	(0.06)	(0.06)	(0.14)	(0.06)	(0.05)	(0.15)	(0.08)	(0.04)	(0.05)	(0.07)
15 (Leather)	0.48	-0.19	0.50	0.39	0.01	0.43	0.34	-0.27	0.39	0.12	0.01	0.25
	(0.06)	(0.14)	(0.07)	(0.07)	(0.14)	(0.07)	(0.08)	(0.18)	(0.08)	(0.07)	(0.07)	(0.09)
16 (Wood)	0.52	0.08	0.57	0.40	0.27	0.45	0.20	0.20	0.21	0.03	-0.04	0.09
	(0.06)	(0.09)	(0.08)	(0.04)	(0.10)	(0.05)	(0.07)	(0.09)	(0.09)	(0.05)	(0.05)	(0.05)
17 (Paper)	0.58	0.13	0.61	0.45	0.31	0.49	0.34	0.37	0.32	0.13	0.05	0.20
	(0.05)	(0.09)	(0.07)	(0.04)	(0.11)	(0.05)	(0.07)	(0.09)	(0.09)	(0.05)	(0.07)	(0.05)
18 (Printing)	0.58	0.06	0.60	0.46	0.25	0.49	0.27	0.24	0.25	0.11	0.04	0.14
	(0.06)	(0.11)	(0.08)	(0.04)	(0.13)	(0.05)	(0.06)	(0.06)	(0.09)	(0.04)	(0.04)	(0.05)
19 (Petrol)	0.48	0.17	0.46	0.36	0.32	0.35	0.14	0.15	0.09	0.05	0.08	-0.01
	(0.05)	(0.11)	(0.07)	(0.04)	(0.10)	(0.05)	(0.08)	(0.08)	(0.10)	(0.04)	(0.04)	(0.04)
20 (Chemicals)	0.59	0.15	0.62	0.47	0.33	0.51	0.38	0.28	0.36	0.19	0.08	0.21
	(0.05)	(0.10)	(0.06)	(0.05)	(0.12)	(0.05)	(0.09)	(0.10)	(0.11)	(0.06)	(0.04)	(0.06)
21 (Pharma)	0.62	-0.16	0.65	0.50	0.05	0.53	0.22	0.09	0.38	0.17	0.05	0.24
	(0.07)	(0.21)	(0.09)	(0.06)	(0.22)	(0.06)	(0.14)	(0.22)	(0.13)	(0.10)	(0.09)	(0.14)
22 (Plastics)	0.60	0.13	0.63	0.48	0.29	0.52	0.38	0.36	0.33	0.18	0.13	0.19
	(0.05)	(0.10)	(0.06)	(0.04)	(0.13)	(0.05)	(0.05)	(0.08)	(0.08)	(0.04)	(0.05)	(0.05)
23 (Minerals)	0.57	0.07	0.67	0.45	0.25	0.52	0.30	0.23	0.43	0.12	0.05	0.31
	(0.05)	(0.09)	(0.07)	(0.04)	(0.11)	(0.04)	(0.06)	(0.06)	(0.10)	(0.04)	(0.05)	(0.08)
24 (Basic metals)	0.58	0.14	0.77	0.43	0.29	0.66	0.27	0.30	0.43	0.18	0.19	0.31
	(0.05)	(0.13)	(0.07)	(0.05)	(0.14)	(0.04)	(0.07)	(0.09)	(0.12)	(0.03)	(0.03)	(0.10)
25 (Fabric. metals)	0.60	0.10	0.57	0.48	0.28	0.40	0.33	0.28	0.16	0.14	0.06	0.18
	(0.05)	(0.08)	(0.08)	(0.04)	(0.11)	(0.05)	(0.06)	(0.06)	(0.14)	(0.04)	(0.04)	(0.06)
26 (Computers)	0.65	0.31	0.67	0.49	0.49	0.58	0.60	0.73	0.59	0.30	0.27	0.44
	(0.06)	(0.14)	(0.08)	(0.04)	(0.13)	(0.05)	(0.12)	(0.16)	(0.13)	(0.05)	(0.06)	(0.05)
27 (Electrical)	0.61	0.08	0.67	0.53	0.25	0.58	0.37	0.29	0.38	0.19	0.13	0.22
	(0.09)	(0.11)	(0.08)	(0.07)	(0.15)	(0.06)	(0.10)	(0.11)	(0.13)	(0.04)	(0.04)	(0.07)
28 (Machinery)	0.71	0.10	0.77	0.60	0.29	0.66	0.38	0.26	0.43	0.21	0.13	0.31
	(0.08)	(0.12)	(0.07)	(0.05)	(0.15)	(0.05)	(0.11)	(0.13)	(0.12)	(0.06)	(0.05)	(0.10)
29 (Vehicles)	0.55	0.15	0.57	0.39	0.36	0.40	0.27	0.37	0.16	0.19	0.24	0.18
	(0.05)	(0.11)	(0.06)	(0.04)	(0.12)	(0.05)	(0.08)	(0.08)	(0.14)	(0.03)	(0.06)	(0.06)
30 (Other trans.)	0.56	0.00	0.67	0.38	0.24	0.48	0.26	0.21	0.49	0.00	-0.10	0.13
	(0.10)	(0.13)	(0.09)	(0.06)	(0.17)	(0.06)	(0.13)	(0.13)	(0.18)	(0.05)	(0.06)	(0.06)
3133 (Furniture)	0.55	-0.06	0.57	0.42	0.14	0.45	0.25	0.22	0.19	0.10	0.03	0.16
	(0.07)	(0.15)	(0.08)	(0.04)	(0.17)	(0.05)	(0.07)	(0.08)	(0.09)	(0.05)	(0.06)	(0.05)
Observations	171K	137K	136K	171K	137K	136K	171K	137K	136K	171K	137K	136K
R-squared	0.52	0.11	0.55	0.65	0.34	0.68	0.70	0.43	0.73	0.69	0.43	0.73
TCC†	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
PLC†	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
CAC†	No	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
AID†	0.55	0.05	0.58	0.43	0.24	0.47	0.28	0.24	0.28	0.12	0.06	0.17
SID	0.08	0.13	0.08	0.07	0.11	0.08	0.13	0.18	0.15	0.09	0.11	0.11
F test	0.13	0.82	0.28	0.00	0.93	0.00	0.07	0.08	0.62	0.00	0.00	0.00

Notes: Row definitions are equivalent to those in Table 3. Columns 1A-1D present the original results from Table 1, page 498. Columns 2A-2D showcase results after excluding the six low-income countries before calculating the equilibrium; in the original results, these countries were excluded after equilibrium calculations. In columns 3A-3D, we exclude the six high-income countries, identified using Figure 6 in the original paper. The standard errors are clustered by importer-industry, and they are presented within brackets. We incorporate exporter-industry fixed effects, industry dummy variable interactions with six bilateral distance intervals, and with a dummy variable indicating whether the nations share a border, a common language, or a free trade agreement—all examples of trade cost restrictions. Productivity is significantly influenced by rule of law, corruption prevention, political stability, regulatory quality, voice and accountability, ease of doing business, and private credit as a percentage of GDP. Comparative advantage controls include interactions of industry dummy variables with the importer's rule of law, log private credit as a proportion of GDP, log physical capital per employee, and human capital. The F-test equalizes innovation-dependence across industries (p-value reported).

level of innovation, contributing to a more uniform innovation landscape relative to the U.S. A similar pattern is observed with patent intensity. Excluding developing countries results in a mean patent intensity of -0.53, which is smaller in absolute terms than the original value of -1.13. Conversely, excluding developed countries increases the value to -1.55, which is larger in absolute terms than the original.

In Table 6, an expanded version of Table 1 from the original paper, we present the computed estimates of innovation dependence. Following the original paper's insights, the gradual integration of trade costs, productivity levels, and comparative advantage shows a systematic reduction in the average estimated innovation dependence, as seen in AID values in columns 1A-1D.

A noteworthy observation emerges when we exclude developing countries from the analysis. In column 2A, where only trade cost controls are considered, the mean estimated innovation dependence sharply diminishes to 0.05, a substantial reduction from the original figure of 0.55. This suggests a minimal impact of innovation when solely incorporating trade cost controls. However, with the subsequent introduction of controls for productivity levels and comparative advantage (columns 2B-2D), the estimated innovation dependence aligns more closely with the original result, although still slightly lower in most scenarios.

Columns 3A to 3D exhibit mean innovation dependence estimates similar to the original results when the top six developed countries are excluded from the analysis. The variation pattern remains consistent even after the inclusion of additional control variables. Although columns 3A-3D exhibit quite similar results, suggesting minimal changes when developed countries are excluded, they imply that the presence or absence of developed countries in the dataset may not significantly alter the outcomes.

The observations above suggest the robustness of the results when a cluster comprising both developed and developing countries is present in the data sample. The similarity in results implies that the dynamics within this mixed cluster contribute to consistent findings. Conversely, when the data sample exclusively includes developed countries, a distinct and homogeneous development pattern emerges. This shared pattern among developed nations can potentially exert a notable influence on the results. In essence, the observed similarity in outcomes may be attributed to the shared development trajectories among developed countries within the dataset.

In Model 2, the focus on developed countries results in a convergence of innovation efficiency, as measured by nominal wages and real income. This convergence among developed nations reduces the gap in innovation efficiency and innovation dependence, as shown in Table 5 and Table 6, respectively.

Table 7. Counterfactual Results

Innovation efficiency measure (1)	Outcome (2)	R & D intensity (3)	Patenting intensity (4)	R & D intensity generalized model (5)
1A Nominal wage	Average change relative to US	0.18	0.14	0.18
2A Nominal wage	Average change relative to US	0.09	0.05	0.03
3A Nominal wage	Average change relative to US	0.23	0.26	0.22
1B Nominal wage	Dispersion ratio	0.32	0.27	0.31
2B Nominal wage	Dispersion ratio	0.24	0.22	0.07
3B Nominal wage	Dispersion ratio	0.27	0.30	0.26
1C Real income	Average change relative to US	0.06	0.04	0.06
2C Real income	Average change relative to US	0.02	0.01	0.01
3C Real income	Average change relative to US	0.07	0.08	0.07
1D Real income	Dispersion ratio	0.17	0.13	0.16
2D Real income	Dispersion ratio	0.09	0.07	0.03
3D Real income	Dispersion ratio	0.14	0.15	0.13

Notes: For detailed descriptions of Models 1-3, refer to the notes in Table 5. The top panels (1 and 2) display the average log wage change and its standard deviation relative to the United States, comparing the counterfactual economy with the calibrated model. In the bottom panels (3 and 4), similar statistics are shown for real GDP per capita, which is measured as GDP per working-age individual. The model in Column 3 is calibrated using R&D data, while Column 4 is calibrated with patent data. Column 5 presents the calibration of the generalized model from Section IVA, using R&D data.

This trend is mirrored in Table 7, where the values in rows 2A-2D also suggest a relatively low magnitude of inequality compared to the original results. The coherence arises from the homogeneity in innovation efficiency among developed nations in Model 2, highlighting the impact of exclusively including developed countries on innovation dependence metrics. The results obtained through calibrations of the generalized model are particularly striking, indicating a near-complete disappearance of inequality.

Sampson (2023) conducted a robustness check by excluding outlier countries as depicted in Figure 6 of their paper. However, they removed these outliers after computing the equilibrium and reported only the more favourable wage dispersion ratio of 0.25 and the income dispersion ratio of 0.10. In our analysis, we present all results in rows 2A-2D of Table 8 following their approach. The results in column 3 are consistent between the two models. However, in column 4, a substantial difference emerges when the model is generalized. For instance, the wage dispersion ratio is 0.07 in Model 1 compared to 0.25 in Model 2, and the income dispersion ratio is 0.03 in Model 1 compared to 0.10 in Model 2. The generalized model is more influenced by the outliers, specifically the six poorest countries.

3 Demonstrating robustness

We subject the original results to various checks to test their robustness. We start the section by comparing additional patent metrics obtained from OECD and PATSTAT data. The transparency and quality of Sampson's code and data-sharing practices greatly facilitated our replication of the original work

Table 8. Counterfactual Results

Innovation efficiency measure (1)	Outcome (2)	R & D intensity (3)	Patenting intensity (4)	R & D intensity generalized model (5)
1A Nominal wage	Average change relative to US	0.09	0.05	0.03
2A Nominal wage	Average change relative to US	0.08	0.06	0.08
1B Nominal wage	Dispersion ratio	0.24	0.22	0.07
2B Nominal wage	Dispersion ratio	0.25	0.31	0.25
1C Real income	Average change relative to US	0.02	0.01	0.01
2C Real income	Average change relative to US	0.03	0.02	0.03
1D Real income	Dispersion ratio	0.09	0.07	0.03
2D Real income	Dispersion ratio	0.10	0.11	0.10

Notes: Model 1 presents results after excluding the six low-income countries before calculating the equilibrium—a departure from the original methodology where these countries were excluded after equilibrium calculations. Model 2 shows the result after excluding the six low-income countries and computing the equilibrium. We used the same method as described on page 507 of the original paper. For detailed descriptions of row and column definitions, refer to the notes in Table 7.

3.1 Additional patent innovation measures

This section further examines the impact of using different patent metrics. One set of metrics restricts attention to a single patent office or filing route, intentionally introducing additional geographical bias into the analysis. Where possible, we consider patent grants instead of patent applications, as granted patents have passed examination by the patent office and are more likely to capture novel technological contributions. This set includes European Patent Office Grants (EPO_G); Patent Cooperation Treaty Applications (PCT_A), which cover international patent applications filed through the PCT system; and United States Patent and Trademark Office Grants (USPTO_G).[9] A second set of metrics involves patent families, which group together patents filed in multiple countries to protect a single invention. Here, we expand the family definition to include filings at all five of the world's largest patent offices (IP5) instead of the three used by Sampson.[10] Additionally, we complement the IPF measure introduced in subsection 2.1 by citations received from other patent families. [11] Such citations-weighted patent counts are often used to capture the 'quality' of an invention, which can reflect both technological impact and commercial value (Ejermo 2009, Jaffe and de Rassenfosse 2017). Therefore, we expect results similar to those obtained with triadic patent families.

As shown in Table B1, TRIADIC patents and USPTO_G patents both capture high-quality innovations, but they exhibit notable differences in their mean innovation efficiency values. TRIADIC patents, which cover a selective international scope, show a mean innovation efficiency gap of -1.13, indicating a larger average innovation gap compared to most other metrics. USPTO_G patents, with a mean value of -1.95, and CIT patents at -1.92, suggest an even greater innovation efficiency gap, likely skewed by their

[9] We construct all but the CIT measure using the same patent data as Sampson, only modifying the filtering criterion which in the original is `keep if kindpatent=='FAMILIES'`.

[10] The five largest patent offices are the United States Patent and Trademark Office (USPTO), the European Patent Office (EPO), the Japan Patent Office (JPO), the Korean Intellectual Property Office (KIPO), and the China National Intellectual Property Administration (CNIPA).

[11] Citations data is sourced from PATSTAT table 209. Citations are counted at the level of the `INPADOC_FAMILY_ID`, meaning that a citation by another patent family is counted as a single citation, regardless of how many individual patent applications within that family make the citation. Intra-family citations are excluded. For our primary measure, we consider citations received within 5 years of the first application filing to prevent earlier filings from accumulating higher citations counts simply due to longer exposure (OECD 2009), but results are unchanged without this time cutoff.

focus on the U.S. market (USPTO_G) and higher citation rates (CIT). However, when broader geographical data is considered, as seen with the EPO_G, IP5, and PCT_A metrics, the mean innovation efficiency gap decreases to -0.32, -0.60, and -0.62, respectively, reflecting more balanced innovation levels compared to the U.S.

Table B2 presents the innovation efficiency gap by country. The TRIADIC metric provides a balanced global view of innovation efficiency. Alternative metrics show regional biases. EPO_G assigns higher weights to European countries, reflecting the European focus of the EPO. Conversely, USPTO_G assigns higher weights to the U.S. and high-tech countries closely linked to the U.S. market, highlighting the concentration of high-tech innovation in these regions. IP5 and PCT_A metrics provide a more balanced international perspective but still show variations, with IP5 favoring broader international collaboration and PCT_A emphasizing early-stage global patenting. Interestingly, the citations metric exhibits an almost equally strong U.S. bias as the USPTO-based metric. This bias may be partly due to the stricter citations requirements at the USPTO compared to other offices, such as the EPO (Harhoff et al. 1999). As a result, U.S. patents tend to generate more citations on average than European patents (Michel and Bettels 2001), and the citations recipients are predominantly domestic (Bacchiocchi and Montobbio 2010).

Table B3 compares innovation dependence by industry using these alternative patent metrics. The first column, labeled TRIADIC, replicates Sampson's original findings, showing significant dependence in the chemical, pharmaceutical, and electronic industries, among others. IP5 is broadly comparable. PCT_A reveals higher innovation dependence in high-tech industries such as computers, electronics, and pharmaceuticals, underscoring the strategic importance of early and extensive international patent protection in these sectors. It also has the highest average innovation dependence across all metrics, with statistically significant differences in the wearing apparel and electrical equipment sectors, while EPO_G has the lowest average score. In general, the EPO_G and USPTO_G metrics yield industry estimates that closely reflect the respective domestic industrial base. EPO_G scores higher than TRIADIC in coke & refined petroleum products, while USPTO_G produces higher values for textiles, chemicals, and transport equipment. However, all metrics consistently identify the computer industry as the most innovation-dependent by a significant margin, clearly distinguishing them from the IPF count introduced in subsection 2.1.

Our main focus is on the counterfactuals and the role of innovation in explaining income and wage inequality. These counterfactual results shown in Table 9 demonstrate that the choice of patent metric influences the estimated impact of innovation efficiency on wage and income dispersion, complementing Table 4. TRIADIC metrics show a strong link between innovation efficiency and wage dispersion, confirming Sampson's findings that innovation efficiency differences account for nearly one-quarter of nominal wage dispersion within the OECD countries. USPTO_G and CIT metrics reveal a higher impact, emphasizing the role of U.S.-centric high-tech innovation and the role of highly cited patents. In contrast, EPO_G and IP5 metrics suggest a slightly more modest impact, reflecting regional differences. The same principle applies to the real income per capita, although innovation efficiency differences have a smaller impact on this variable compared to nominal wages.

The choice of patent metric significantly influences the analysis of innovation dependence and its economic implications. Broader geographical data tend to dilute the intensity of innovation dependence, which explains less of the technology gap. Patent citations, often used as a proxy for patent quality, play an important role. For this measure of patent intensity, the results align more closely with USPTO_G. It is important to note that patent citations should be interpreted with caution,[12] as they may not fully reflect patent quality. This is due to factors such as self-citations by applicants, variation in examination procedures between patent offices (Jaffe and de Rassenfosse 2017), and geographic biases that favor large, advanced economies —biases that differ from those found in the TPF count. For instance, knowledge spillovers, and thus patent citations, are influenced by national borders (Thompson and Fox-Kean 2005, Thompson 2006). Merely being in a country with many other patentees can therefore inflate citation counts beyond what innovation quality alone would merit (Harhoff et al. 1999). Lastly, the

[12] Bruns and Kalthaus (2020) argue for a similar level of caution when interpreting results obtained with counts of patent grants instead of patent applications.

role of the innovation measure here is to identify R&D effort, which is then related to value added. This way, the eventual quality or impact of inventions is already captured in the denominator of innovation efficiency and should therefore not be included in the numerator.

Table 9. Counterfactual results

			TRIADIC	EPO_G	IP5	PCT_A	USPTO_G	CIT
		Innovation efficiency measure	Patenting intensity	Patenting intensity	Patenting intensity	Patenting intensity	Patenting intensity	Patenting intensity
1.	Nominal wage	Average change relative to US	0.14	0.02	0.07	0.11	0.32	0.28
		Dispersion ratio	0.27	0.22	0.24	0.26	0.30	0.30
2.	Real income	Average change relative to US	0.04	0.01	0.02	0.04	0.11	0.09
	per capita	Dispersion ratio	0.13	0.10	0.11	0.12	0.16	0.15

Notes: EPO_G (European Patent Office Grants): Patents granted by the EPO, reflecting successful innovations that have passed rigorous examination processes in Europe. FAMILIES: Patent families, which indicate the same invention filed in multiple patent offices, showcasing the global reach and perceived importance of the innovation. IP5 (IP5 Offices): Patents filed in the five major patent offices (EPO, JPO, KIPO, SIPO, USPTO), representing a broad international scope. PCT_A (Patent Cooperation Treaty Applications): Applications filed under the PCT, offering a pathway to seek patent protection in multiple countries with a single application. USPTO_G (United States Patent and Trademark Office Grants): Patents granted by the USPTO, highlighting innovations recognized and protected in the United States. CIT (Citations to IPF patent families): Number of citations received by INPADOC patent families within five years after the first filing date.

3.2 Refining country inclusion

Our initial step involves adjusting the filter set presented in the original paper, which excludes countries with over two-thirds of industries featuring missing values. By extending our analysis to include country-years with observations available for at least 10 industries, we increase the total number of OECD countries from 25 to 28. The supplementary three countries incorporated into the study are Estonia, Slovakia, and Sweden. The outcomes closely resemble the original findings and are reported, benchmarked against those in the original paper, in Appendix C.[13]

3.3 Modifying the time frame

We initiate the comparison by adjusting the time period. First, we extend the original time period from 2010-2014 to 2010-2016 and then partition the intervals into two sub-

[13] In addition to our primary analysis, we conducted further investigations: (i) We refined the original sample by excluding any country with more than 1/2 industries with missing values, resulting in a sample of 17 countries (comprising six developing and 11 developed countries). (ii) We applied a more stringent criterion by excluding any country with more than 8/10 industries with missing values, resulting in a sample of 28 countries (six developing and 22 developed countries). (iii) To maintain balance, we kept an equal number of developed and developing countries. Specifically, we randomly selected six developed countries to match the number of six observed developing countries, totaling 12 countries. (iv) We conducted analyses with three countries exhibiting a high level of innovation removed (excluding the U.S., which serves as the reference country). Importantly, the results remained consistent with the original findings.

groups 2010-2012 and 2014-2016. The results of those three groups closely align with those in the original paper and can be found in Appendix D.

3.4 Changing the values of trade elasticity

The original paper incorporates a robustness check where the author varies the trade elasticity within the range of 2.5 to 8.5. In our analysis, we extend this examination by adjusting the trade elasticity to 1 and 10.5 to further assess its robustness. Overall, the findings closely align with the original results, and the new results are provided in Appendix E.

3.5 Dropping outlier industries

In the original paper (Figure 4), the author identified two outlier industries: Agriculture, Forestry, and Fishing (0103), and Paper and Paper Products (17). The results excluding these two industries, remain similar to the original ones and are detailed in Appendix F.

3.6 Excluding outliers in the regression analysis

Verardi and Croux (2009) highlight the vulnerability of OLS to outliers, which can distort parameter estimation by assigning excessive weight to observations with large residuals. To address this issue, they emphasize robust regression estimators that are less sensitive to outliers. In our analysis, we utilize the Stata module developed by Jann (2022), providing a package for robust fixed-effect regression and outlier detection. Subsequently, we exclude these identified outliers and reapply the fixed-effect regression, following the approach in Sampson (2023). The outcomes of excluding outliers and employing robust regression from Jann (2022) align consistently with the original results, as presented in Appendix G.

4 Conclusions

Sampson's work has made a significant contribution to the development of a new strand of economics literature. We successfully replicated the original paper's results and subjected them to various robustness checks, all of which they withstood. Our contribution lies in the choice of the innovation measure and questioning the selection of developing countries. Our findings affirm the robustness of the original paper's framework across diverse methodological extensions. The results provide nuanced insights into the sensitivity of inequality measures to alternative innovation metrics, demonstrating that technology gaps remain a key factor in economic disparities while highlighting areas for future research.

While the group of countries under examination constitutes a relatively homogeneous subset, representing only a fraction of the 195 countries globally, it remains noteworthy that within this narrowed scope, a discernible ranking becomes evident (see Figure 6 in the original article). The inherent challenge in comparing these countries lies in their diverse industry structures and varying levels of economic development —a complexity further compounded by limitations in available data.

Moreover, no single innovation indicator is flawless. Therefore, results should be cross-verified using a range of innovation indicators that capture different facets of innovation measurement (Hagedoorn and Cloodt 2003, Lanjouw and Schankerman 2004, Martínez 2011). The triadic patent count favors countries with affiliated patent offices and economically prosperous firms. Introducing a more technologically oriented indicator helps mitigate some of this bias. We have incorporated additional patent metrics with different geographical focus, and for the most part, these metrics reinforced Sampson's main findings.

Given the crucial policy implications of this paper, further research is warranted, potentially employing a more comprehensive list of developing countries and utilizing alternative sources of innovation data.

References

Bacchiocchi, E. and Montobbio, F.: 2010, International Knowledge Diffusion and Home-bias Effect: Do USPTO and EPO Patent Citations Tell the Same Story?, *The Scandinavian Journal of Economics* **112**(3), 441–470.

Bogliacino, F., Perani, G., Pianta, M. and Supino, S.: 2012, Innovation and Development:

The Evidence From Innovation Surveys, *Latin American Business Review* **13**(3), 219–261.

Boschma, R.: 2005, Proximity and innovation: A critical assessment, *Regional Studies* **39**(1), 61–74.
URL: *https://doi.org/10.1080/0034340052000320887*

Bruns, S. B. and Kalthaus, M.: 2020, Flexibility in the selection of patent counts: Implications for p-hacking and evidence-based policymaking, *Research Policy* **49**(1), 103877.

Buera, F. J. and Oberfield, E.: 2020, The Global Diffusion of Ideas, *Econometrica* **88**(1), 83–114.

Cai, J., Li, N. and Santacreu, A. M.: 2022, Knowledge Diffusion, Trade, and Innovation across Countries and Sectors, *American Economic Journal: Macroeconomics* **14**(1), 104–145.

Carlino, G. and Kerr, W. R.: 2015, Chapter 6 - agglomeration and innovation, *in* G. Duranton, J. V. Henderson and W. C. Strange (eds), *Handbook of Regional and Urban Economics*, Vol. 5 of *Handbook of Regional and Urban Economics*, Elsevier, pp. 349–404.
URL: *https://www.sciencedirect.com/science/article/pii/B9780444595171000064*

Cox, N. J.: 2008, Speaking Stata: Correlation with Confidence, or Fisher's z revisited, *The Stata Journal: Promoting communications on statistics and Stata* **8**(3), 413–439.

Crescenzi, R., Iammarino, S., Ioramashvili, C., Rodríguez-Pose, A. and Storper, M.: 2023, The Geography of Innovation: Local Hotspots and Global Innovation Networks, *SSRN Electronic Journal* .
URL: *https://www.ssrn.com/abstract=4424262*

Criscuolo, P.: 2006, The 'home advantage' effect and patent families. A comparison of OECD triadic patents, the USPTO and the EPO, *Scientometrics* **66**(1), 23–41.

De Rassenfosse, G., Dernis, H., Guellec, D., Picci, L. and van Pottelsberghe de la Potterie, B.: 2013, The worldwide count of priority patents: A new indicator of inventive activity, *Research Policy* **42**(3), 720–737.

De Rassenfosse, G., Grazzi, M., Moschella, D. and Pellegrino, G.: 2022, International patent protection and trade: Transaction-level evidence, *European Economic Review* **147**, 104160.

De Rassenfosse, G., Kozak, J. and Seliger, F.: 2019, Geocoding of worldwide patent data, *Scientific Data* **6**(1), 260.

De Rassenfosse, G. and Seliger, F.: 2021, Imputation of missing information in worldwide patent data, *Data in Brief* **34**, 106615.

Dernis, H., Guellec, D. and Van Pottelsberghe, B.: 2002, Using Patent Counts for Cross-Country Comparisons of Technology Output, *OECD STI Review* **27**, 129–146.

Dernis, H. and Khan, M.: 2004, *Triadic Patent Families Methodology*, OECD Science, Technology and Industry Working Papers, No. 2004/02, OECD Publishing, Paris.

Dominguez Lacasa, I., Jindra, B., Radosevic, S. and Shubbak, M.: 2019, Paths of technology upgrading in the BRICS economies, *Research Policy* **48**(1), 262–280.

Ejermo, O.: 2009, Regional Innovation Measured by Patent Data—Does Quality Matter?, *Industry and Innovation* **16**(2), 141–165.

Faber, J. and Hesen, A. B.: 2004, Innovation capabilities of European nations: Cross-national analyses of patents and sales of product innovations, *Research Policy* **33**(2), 193–207.

Frietsch, R. and Schmoch, U.: 2009, Transnational patents and international markets, *Scientometrics* **82**(1), 185–200.

Glaeser, E. L.: 2008, *Cities, Agglomeration, and Spatial Equilibrium*, Oxford University Press, Oxford, U.K.

Grupp, H., Münt, G. and Schmoch, U.: 1996, Assessing different types of patent data for describing high-technology performance, *in* OECD (ed.), *Innovation, Patents and Technological Strategies*, Innovation, Patents and Technological Strategies. - Paris : Organisation for Economic Co-operation and Development, ISBN 92-64-14661-x. - 1996, p. 17-56, OECD Publishing, Paris, pp. 271–284.

Hagedoorn, J. and Cloodt, M.: 2003, Measuring innovative performance: Is there an advantage in using multiple indicators?, *Research Policy* **32**(8), 1365–1379.

Harhoff, D., Narin, F., Scherer, F. M. and Vopel, K.: 1999, Citation Frequency and the Value of Patented Inventions, *The Review of Economics and Statistics* **81**(3), 511–515.

Hinze, S. and Schmoch, U.: 2005, Opening the Black Box: Analytical Approaches and their Impact on the Outcome of Statistical Patent Analyses, *in* H. F. Moed, W. Glänzel and U. Schmoch (eds), *Handbook of Quantitative Science and Technology Research: The Use of Publication and Patent Statistics in Studies of S&T Systems*, Springer Netherlands, Dordrecht, pp. 215–235.

Hoppe, H. C.: 2000, Second-mover advantages in the strategic adoption of new technology under uncertainty, *International Journal of Industrial Organization* **18**(2), 315–338.

Jaffe, A. B. and de Rassenfosse, G.: 2017, Patent citation data in social science research: Overview and best practices, *Journal of the Association for Information Science and Technology* **68**(6), 1360–1374.

Jann, B.: 2022, Robreg: Stata module providing robust regression estimators.
URL: *https://EconPapers.repec.org/RePEc:boc:bocode:s458931*

Kang, B. and Tarasconi, G.: 2016, PATSTAT revisited: Suggestions for better usage, *World Patent Information* **46**, 56–63.

Khan, M. and Dernis, H.: 2006, Global Overview of Innovative Activities from the Patent Indicators Perspective, *OECD Science, Technology and Industry Working Papers 2006/03* .

Kleinknecht, A., Van Montfort, K. and Brouwer, E.: 2002, The Non-Trivial Choice between Innovation Indicators, *Economics of Innovation and New Technology* **11**(2), 109–121.

Lanjouw, J. O. and Schankerman, M.: 2004, Patent Quality and Research Productivity: Measuring Innovation with Multiple Indicators, *The Economic Journal* **114**(495), 441–465.

Lee, J.-D., Lee, K., Meissner, D., Radosevic, S. and Vonortas, N. (eds): 2021, *The Challenges of Technology and Economic Catch-up in Emerging Economies*, Oxford University Press.

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

Martínez, C.: 2011, Patent families: When do different definitions really matter?, *Scientometrics* **86**(1), 39–63.

Michel, J. and Bettels, B.: 2001, Patent citation analysis. A closer look at the basic input data from patent search reports, *Scientometrics* **51**(1), 185–201.

Nagaoka, S., Motohashi, K. and Goto, A.: 2010, Chapter 25 - patent statistics as an innovation indicator, *in* B. H. Hall and N. Rosenberg (eds), *Handbook of the Economics of Innovation*, Vol. 2 of *Handbook of the Economics of Innovation*, North-Holland, pp. 1083–1127.

OECD: 2009, *OECD Patent Statistics Manual*, OECD Publishing, Paris.

OECD: 2015, *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development*, The Measurement of Scientific, Technological and Innovation Activities, OECD Publishing, Paris.

OECD: 2023a, Patents by main technology and by International Patent Classification

(IPC), OECD Patent Statistics (database), doi: 10.1787/data-00508-en. Last accessed 2023-07-31.

OECD: 2023b, The OECD Analytical Business Enterprise Research and Development (ANBERD) Database, url: `http://oe.cd/anberd`. Last accessed 2023-07-31.

Park, W. G. and Hingley, P.: 2009, Patent family data and statistics at the European Patent Office, *American University Department of Economics Working Paper Series, No. 2009-08* .

Sampson, T.: 2023, Technology gaps, trade, and income, *American Economic Review* **113**(2), 472–513.

Shankar, V., Carpenter, G. S. and Krishnamurthi, L.: 1998, Late Mover Advantage: How Innovative Late Entrants Outsell Pioneers, *Journal of Marketing Research* **35**(1), 54–70.

Sternitzke, C.: 2009, Defining triadic patent families as a measure of technological strength, *Scientometrics* **81**(1), 91–109.

Thompson, P.: 2006, Patent Citations and the Geography of Knowledge Spillovers: Evidence from Inventor- and Examiner-added Citations, *The Review of Economics and Statistics* **88**(2), 383–388.

Thompson, P. and Fox-Kean, M.: 2005, Patent Citations and the Geography of Knowledge Spillovers: A Reassessment, *The American Economic Review* **95**(1), 450–460.

van Zeebroeck, N.: 2011, The puzzle of patent value indicators, *Economics of Innovation and New Technology* **20**(1), 33–62.

Verardi, V. and Croux, C.: 2009, Robust regression in stata, *The Stata Journal* **9**(3), 439–453.

URL: *https://doi.org/10.1177/1536867X0900900306*

Appendices

A Correlations between innovation measures

A.1 Correlations between innovation measures used to calibrate the model

Table A1. Correlations between innovation measures in the full sample

Innovation measure:	R&D exp.	Patent application family counts					
Data source: Patent family definition: Location data:	OECD	OECD triadic	OECD triadic US only	PS triadic	PS triadic imputed	PS ID	PS ID imputed
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1)	1						
(2)	0.1997	1					
(3)	0.1877	0.9907	1				
(4)	0.1870	0.9848	0.9781	1			
(5)	0.1289	0.9558	0.9328	0.9616	1		
(6)	0.5263	0.8008	0.7938	0.7588	0.7705	1	
(7)	0.4280	0.9085	0.8763	0.9178	0.9102	0.8636	1

Notes: Correlation coefficients are calculated for the set of jointly non-missing values, including 2,181 observations.

Table A2. Average within-industry correlation between innovation measures

Innovation measure:	R&D exp.	Patent application family counts					
Data source: Patent family definition: Location data:	OECD	OECD triadic	PS triadic US only	PS triadic	PS triadic imputed	PS ID	PS ID imputed
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1)	1						
(2)	0.3072	1					
(3)	0.2932	0.7599	1				
(4)	0.3083	0.7586	0.7587	1			
(5)	0.2741	0.9696	0.9702	0.9725	1		
(6)	0.7213	0.6884	0.7073	0.6723	0.6083	1	
(7)	0.7048	0.8907	0.8764	0.8876	0.8574	0.8212	1

Notes: Values are calculated by applying Fisher's z transformation (Cox 2008) to each industry-specific correlation coefficient, averaging over these transformed values, and then applying the inverse transformation to each average.

A.2 Correlations with R&D expenditure from external survey data

Table A3. Correlations between innovation measures and survey R&D expenditure

Innovation measure:		R&D exp.	TPF	IPF	R&D expenditure		
Data source:		OECD		PS	World Bank Enterprise Surveys		
		(1)	(2)	(3)	(4)	(5)	(6)
	(1)	1					
	(2)	0.4187	1				
	(3)	0.6748	0.6487	1			
all firms	(4)	0.3861	0.0238	0.5204	1		
recent innovators	(5)	0.3747	0.0132	0.4912	0.9629	1	
new-to-the-market	(6)	0.3135	0.0121	0.4026	0.9338	0.9856	1

Notes: The table reports pairwise correlation coefficients. Column (1) contains aggregate R&D expenditure from the OECD's ANBERD database as used by Sampson as the primary innovation measure. Column (2) uses counts of triadic patent families as used by Sampson as the alternative innovation measure. Column (3) uses counts of INPADOC patent families as introduced in subsection 2.1. Column (4) uses R&D expenditure from the World Bank's Enterprise Surveys, weighted and summed at the industry level. Columns (5) and (6) follow the same approach but in summing at the industry level only consider R&D expenditure by firms that recently introduced an innovation (5) or that recently introduced an innovation that was at least new to their market (6).

Table A4. Regressions between innovation measures and survey R&D expenditure

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OECD R&D expend.	8.506	4.115	0.044	0.423***	149.276	0.276	0.298***	0.298
	(9.222)	(3.621)	(0.028)	(0.049)	(99.351)	(0.184)	(0.051)	(0.214)
TPF count	-81.996	-37.068	-0.449**	-0.993*	-63.113**	-0.117**	-1.551***	-0.143**
	(65.157)	(38.112)	(0.195)	(0.517)	(26.881)	(0.050)	(0.535)	(0.058)
IPF count	114.802**	54.215**	0.840***	1.306***	465.428***	0.860***	1.619***	0.990***
	(44.661)	(21.855)	(0.134)	(0.297)	(175.974)	(0.325)	(0.307)	(0.379)
Constant	-399.532**	-138.967**	12.394***	-0.351	174.913***	0.130*	-1.542*	0.189**
	(183.343)	(64.617)	(0.548)	(0.877)	(40.198)	(0.074)	(0.906)	(0.087)
Dependent variable:								
values:	lvl	lvl	log	log	lvl	std	log	std
missing values = 0?		yes		yes	yes	yes	yes	yes
new-to-market only?							yes	yes
Independent variables:	log	log	log	log	std	std	log	std
Observations	197	371	197	371	371	371	371	371
R^2	0.051	0.037	0.241	0.324	0.041	0.041	0.241	0.038

Standard errors in parentheses. The variables are expressed as follows: "lvl" = million USD; "log" = natural logarithm of the 'lvl' value; "std" = standardized 'lvl' value (mean = 0, std. dev. = 1). In columns that combine "log" with missing values replaced by zero, the dependent variable is calculated as the logarithm of (expenditure +1). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In this subsection, we compare the three innovation measures used in the main text to a measure of R&D expenditure obtained from the World Bank's Enterprise Surveys.[14] The R&D expenditure values are weighted to account for the representativeness of each responding firm and summed up at the country-year-industry level, the level of analysis of Sampson's paper. The goal is to validate the IPF count against a measure of R&D effort by businesses from an external, independent data source. In addition, this measure

[14]Source: World Bank Enterprise Surveys, www.enterprisesurveys.org. We thank the Enterprise Analysis Unit of the Development Economics Global Indicators Department of the World Bank for the data.

is based on micro-data, allowing for the construction of customized innovation metrics similar to those derived from patent data, which—as discussed in the main text—is not possible with aggregate R&D statistics.

Table A3 shows that the IPF count is the strongest predictor of firm-level R&D effort obtained from the survey data, with a correlation coefficient just over 0.5, outperforming the OECD's ANBERD R&D expenditure data, which has a coefficient just below 0.4. While the lower correlation for the OECD's R&D measure may be due to chance in the specific sample of country-industry-year combinations, it is clear that the IPF count is as a robust proxy for business R&D across a diverse range of countries. In contrast, the TPF count is nearly uncorrelated with business R&D in this sample. Correlations with R&D spending only by innovators or new-to-the-market innovators are generally weaker, but the relative pattern remains. Table A4 further supports this finding with a simple linear regression of aggregated survey R&D values on the three other innovation measures. Only the IPF count shows a statistically significant relationship with the dependent variable in all columns, whereas the TPF count, when controlling for the other two measures, is negatively associated.

Data construction. The World Bank provides both single-country and multi-country datasets via its survey website. Multi-country datasets are restricted to a core set of of questions that are comparable across countries, while single-country datasets contain the full set of responses to questionnaires that are tailored to reflect each country's specific characteristics and concerns. We start by using the most recent 'standardized dataset' (`New_Comprehensive_July_5_2024.dta`), listed as `StandardizedNew-2006-2023-core4.zip` in the download list) to develop a micro-data-based measure of business R&D spending. This dataset includes firm-level data identified by a unique ID and associated with two-digit ISIC (rev. 4) industries. Key variables include dummies for product (`h1`) and process (`h5`) innovation, for new-to-the-market product innovation (`h2`), and the occurrence of R&D expenditure (`h8`) during the preceding fiscal year. By merging with individual country datasets using unique IDs, we also obtain R&D expenditure values (`h9`). We restrict the dataset to manufacturing industries (ISIC codes 10-33) and the list of countries analyzed by Sampson plus Argentina (due to data availability). We convert firms' expenditure data to 2015 US dollars, apply sampling weights (`wt`), and sum it at the industry level for each country (or country-year in the case of Argentina). This process yields expenditure data for the following 26 countries (with year and number of industries covered): Argentina (2006, 2010, 2017; 22), Austria (2021; 18), Belgium (2020; 20); Bulgaria (2019; 17); Croatia (2019; 9); Czechia (2019; 16); Denmark (2020; 21); Estonia (2019; 15); Finland (2020; 19); France (2021; 19); Germany (2021; 20); Greece (2018; 15); Hungary (2019; 15); Ireland (2020; 17); Italy (2019; 14); Latvia (2019; 11); Luxembourg (2020; 7); Netherlands (2020; 17); Poland (2019; 10); Portugal (2019; 14); Romania (2019; 12); Slovak Republic (2019; 13); Slovenia (2019; 20); Spain (2021; 21); Sweden (2020; 19); and Turkey (2019; 14).

The other three innovation measures are calculated as described in the main paper. However, we use more recent versions of the relevant datasets: the 2024 version of the OECD's ANBERD dataset, the 2023 version of the OECD's Patents by Technology dataset, and the Spring 2024 version of PATSTAT.

B Supplementary tables of patent intensity measures

Table B1. Patent intensity summary statistics

	Obs	Mean	Std. Dev.	Min	Max
TRIADIC	25	-1.13	1.34	-4.30	0.91
EPO_G	25	-0.32	1.17	-3.59	1.01
IP5	25	-0.60	1.06	-2.95	0.63
PCT_A	25	-0.62	0.85	-2.77	0.53
USPTO_G	25	-1.95	1.19	-4.44	0
CIT	25	-1.92	1.27	-4.30	0.24

Notes: Row definitions are equivalent to the corresponding column definitions in Table 9.

Table B2. Patent intensity by country

Country	Patenting intensity					
	TRIADIC	EPO_G	IP5	PCT_A	USPTO_G	CIT
Australia (AUS)	-0.99	-1.21	-0.82	-0.52	-1.89	-1.56
Austria (AUT)	-0.15	**0.93**	**0.22**	-0.16	-1.25	-1.56
Belgium (BEL)	-0.08	**0.63**	-0.08	-0.22	-1.37	-1.59
Canada (CAN)	-1.10	-0.91	-0.24	-0.58	-0.97	-1.17
Chile (CHL)	-3.00	-2.33	-2.28	-1.63	-3.69	-4.03
Czech Republic (CZE)	-2.55	-0.96	-1.41	-1.65	-3.03	-2.64
Germany (DEU)	**0.05**	**1.07**	**0.47**	**0.05**	-1.17	-0.93
Denmark (DNK)	-0.02	**1.10**	**0.31**	**0.26**	-0.84	-1.17
Spain (ESP)	-1.79	-0.62	-1.17	-0.89	-2.76	-2.55
Finland (FIN)	-0.30	**0.68**	**0.25**	**0.03**	-1.14	-0.86
France (FRA)	**0.11**	**0.90**	**0.24**	-0.03	-1.14	-1.11
United Kingdom (GBR)	-0.15	**0.23**	-0.14	-0.06	-1.27	-1.32
Hungary (HUN)	-2.29	-0.80	-1.04	-1.02	-2.53	-2.64
Ireland (IRL)	-0.85	**0.14**	-0.22	**0.14**	-1.49	-1.54
Italy (ITA)	-0.97	**0.42**	-0.40	-0.70	-1.97	-2.20
Japan (JPN)	**0.91**	**0.27**	**0.63**	**0.23**	-0.78	-0.22
South Korea (KOR)	-0.50	-0.68	**0.52**	-0.10	-0.95	**0.18**
Mexico (MEX)	-4.30	-3.59	-2.95	-2.77	-3.86	-4.18
Netherlands (NLD)	**0.51**	**0.93**	**0.55**	**0.53**	-0.82	-1.29
Norway (NOR)	-0.87	**0.00**	-0.27	-0.27	-1.53	-1.88
Poland (POL)	-2.60	-1.11	-1.81	-1.86	-3.60	-3.37
Portugal (PRT)	-2.52	-1.29	-1.93	-1.54	-3.32	-3.53
Slovenia (SVN)	-1.41	**0.05**	-0.67	-0.72	-2.84	-4.22
Turkey (TUR)	-3.39	-1.86	-2.75	-1.92	-4.44	-4.22
United States (USA)	0.00	0.00	0.00	0.00	0.00	0.00

Notes: Column definitions are equivalent to those in Table 9. Positive values are highlighted in bold.

Table B3. Innovation dependence by industry

	Patenting intensity					
VARIABLES	TRIADIC	EPO_G	IP5	PCT_A	USPTO_G	CIT
Agriculture, forestry and fishing (0103)	0.01	-0.01	0.02	0.04	0.10	0.10
	(0.06)	(0.07)	(0.08)	(0.12)	(0.07)	(0.05)
Mining and quarrying (0508)	-0.14	-0.11	-0.22	-0.23	0.00	-0.11
	(0.08)	(0.09)	(0.11)	(0.15)	(0.09)	(0.08)
Food products, beverages and tobacco (1012)	0.06	0.05	0.07	0.10	0.13	0.12
	(0.06)	(0.06)	(0.08)	(0.12)	(0.06)	(0.05)
Textiles (13)	0.12	0.06	0.13	0.21	0.20	0.18
	(0.05)	(0.07)	(0.07)	(0.09)	(0.05)	(0.04)
Wearing apparel (14)	0.13	0.08	0.13	**0.28**	0.15	0.17
	(0.04)	(0.06)	(0.04)	(0.08)	(0.08)	(0.04)
Leather and related products (15)	0.12	0.01	0.13	0.33	0.15	0.17
	(0.07)	(0.09)	(0.10)	(0.14)	(0.06)	(0.05)
Wood and products of wood and cork, except furniture (16)	0.03	0.04	0.03	0.00	0.07	0.08
	(0.05)	(0.06)	(0.06)	(0.09)	(0.05)	(0.04)
Paper and paper products (17)	0.13	0.12	0.15	0.13	0.19	0.19
	(0.05)	(0.06)	(0.07)	(0.10)	(0.06)	(0.05)
Printing and reproduction of recorded media (18)	0.11	0.11	0.11	0.14	0.15	0.14
	(0.04)	(0.05)	(0.06)	(0.08)	(0.05)	(0.04)
Coke and refined petroleum products (19)	0.05	0.13	0.06	0.04	0.08	0.05
	(0.04)	(0.05)	(0.06)	(0.07)	(0.05)	(0.05)
Chemicals and chemical products (20)	0.19	0.14	0.21	0.33	0.26	0.21
	(0.06)	(0.07)	(0.08)	(0.10)	(0.06)	(0.06)
Basic pharmaceutical products and pharmaceutical preparations (21)	0.17	0.10	0.20	0.34	0.28	0.19
	(0.10)	(0.08)	(0.12)	(0.14)	(0.10)	(0.10)
Rubber and plastics products (22)	0.18	0.15	0.20	0.22	0.20	0.21
	(0.04)	(0.05)	(0.05)	(0.07)	(0.04)	(0.03)
Other non-metallic mineral products (23)	0.12	0.12	0.15	0.16	0.14	0.15
	(0.04)	(0.05)	(0.05)	(0.07)	(0.05)	(0.04)
Basic metals (24)	0.18	0.15	0.19	0.29	0.21	0.18
	(0.03)	(0.05)	(0.06)	(0.07)	(0.04)	(0.05)
Fabricated metal products, except machinery and equipment (25)	0.14	0.15	0.16	0.18	0.16	0.17
	(0.04)	(0.05)	(0.06)	(0.08)	(0.05)	(0.04)
Computer, electronic and optical products (26)	0.30	0.30	0.26	0.45	0.32	0.28
	(0.05)	(0.06)	(0.08)	(0.10)	(0.08)	(0.05)
Electrical equipment (27)	0.19	0.17	0.22	**0.35**	0.18	0.18
	(0.04)	(0.05)	(0.06)	(0.07)	(0.06)	(0.04)
Machinery and equipment n.e.c. (28)	0.21	0.17	0.26	0.37	0.25	0.23
	(0.06)	(0.05)	(0.08)	(0.11)	(0.08)	(0.07)
Motor vehicles, trailers and semi-trailers (29)	0.19	0.16	0.19	0.26	0.23	0.22
	(0.03)	(0.04)	(0.06)	(0.07)	(0.05)	(0.05)
Other transport equipment (30)	0.00	-0.08	0.01	-0.02	0.09	0.11
	(0.05)	(0.07)	(0.07)	(0.10)	(0.08)	(0.07)
Furniture, other manufacturing (3133)	0.10	0.12	0.13	0.14	0.14	0.14
	(0.05)	(0.06)	(0.06)	(0.09)	(0.05)	(0.04)
Observations	171,152	171,152	171,152	171,152	171,152	171,152
R-squared	0.69	0.69	0.69	0.69	0.69	0.70
TCC†	Yes	Yes	Yes	Yes	Yes	Yes
PLC†	Yes	Yes	Yes	Yes	Yes	Yes
CAC†	Yes	Yes	Yes	Yes	Yes	Yes
AID†	0.12	0.10	0.13	0.19	0.17	0.15
SID	0.09	0.09	0.11	0.16	0.07	0.08
F test	0.00	0.02	0.04	0.00	0.24	0.05

Notes: Row definitions are equivalent to those in Table 3. Column definitions are equivalent to those in Table 9. Values are shown in bold when the new metric is significantly higher than the TRIADIC at the 10% level. The F-test equalizes innovation-dependence across industries (p-value reported).

C Refining country inclusion

Table C1. Innovation intensity summary statistics

b_s	obs	Mean	Std. dev	Min	Max
		R&D intensity			
Model 1	25	-0.69	0.81	-2.61	0.30
Model 2	28	-0.75	0.82	-2.61	0.30
		Patenting intensity			
Model 1	25	-1.13	1.34	-4.30	0.91
Model 2	28	-1.17	1.33	-4.30	0.91

Notes: Model 1 encompasses the original findings as established in the paper. Expanding our analysis to encompass country-years with data available for a minimum of 10 industries in model 2 has resulted in an augmentation of the overall country count to 28.

Table C2. Innovation dependence by industry

Innovation efficiency measure	R&D intensity		R&D intensity		R&D intensity		Patenting intensity	
Industries	(1A)	(2A)	(1B)	(2B)	(1C)	(2C)	(1D)	(2D)
0103 (Agriculture)	0.45	0.46	0.33	0.33	0.17	0.16	0.01	0.01
	(0.06)	(0.05)	(0.05)	(0.04)	(0.09)	(0.09)	(0.06)	(0.06)
0508 (Mining)	0.37	0.39	0.25	0.25	-0.11	-0.11	-0.14	-0.14
	(0.09)	(0.08)	(0.07)	(0.06)	(0.13)	(0.13)	(0.08)	(0.08)
1012 (Food)	0.48	0.51	0.36	0.37	0.21	0.23	0.06	0.07
	(0.05)	(0.04)	(0.04)	(0.05)	(0.08)	(0.08)	(0.06)	(0.05)
13 (Textiles)	0.51	0.50	0.42	0.38	0.29	0.22	0.12	0.08
	(0.05)	(0.05)	(0.05)	(0.05)	(0.06)	(0.07)	(0.05)	(0.05)
14 (Apparel)	0.47	0.47	0.37	0.35	0.33	0.29	0.13	0.11
	(0.06)	(0.06)	(0.06)	(0.06)	(0.05)	(0.06)	(0.04)	(0.04)
15 (Leather)	0.48	0.48	0.39	0.37	0.34	0.30	0.12	0.11
	(0.06)	(0.06)	(0.07)	(0.07)	(0.08)	(0.08)	(0.07)	(0.07)
16 (Wood)	0.52	0.51	0.40	0.37	0.20	0.16	0.03	0.02
	(0.06)	(0.06)	(0.04)	(0.04)	(0.07)	(0.07)	(0.05)	(0.05)
17 (Paper)	0.58	0.58	0.45	0.43	0.34	0.30	0.13	0.12
	(0.05)	(0.05)	(0.04)	(0.03)	(0.07)	(0.06)	(0.05)	(0.05)
18 (Printing)	0.58	0.58	0.46	0.44	0.27	0.24	0.11	0.10
	(0.06)	(0.05)	(0.04)	(0.04)	(0.06)	(0.06)	(0.04)	(0.04)
19 (Petrol)	0.48	0.46	0.36	0.32	0.14	0.09	0.05	0.03
	(0.05)	(0.05)	(0.04)	(0.04)	(0.08)	(0.08)	(0.04)	(0.04)
20 (Chemicals)	0.59	0.58	0.47	0.44	0.38	0.33	0.19	0.16
	(0.05)	(0.05)	(0.05)	(0.05)	(0.09)	(0.09)	(0.06)	(0.06)
21 (Pharma)	0.62	0.65	0.50	0.50	0.22	0.23	0.17	0.18
	(0.07)	(0.07)	(0.06)	(0.06)	(0.14)	(0.13)	(0.1)	(0.09)
22 (Plastics)	0.60	0.59	0.48	0.44	0.38	0.31	0.18	0.16
	(0.05)	(0.05)	(0.04)	(0.04)	(0.05)	(0.06)	(0.04)	(0.04)
23 (Minerals)	0.57	0.57	0.45	0.44	0.30	0.27	0.12	0.11
	(0.05)	(0.04)	(0.04)	(0.04)	(0.06)	(0.06)	(0.04)	(0.04)
24 (Basic metals)	0.58	0.59	0.43	0.41	0.27	0.24	0.18	0.17
	(0.05)	(0.05)	(0.05)	(0.05)	(0.07)	(0.07)	(0.03)	(0.04)
25 (Fabric. metals)	0.60	0.59	0.48	0.45	0.33	0.30	0.14	0.13
	(0.05)	(0.05)	(0.04)	(0.04)	(0.06)	(0.06)	(0.04)	(0.04)
26 (Computers)	0.65	0.65	0.49	0.45	0.60	0.51	0.30	0.23
	(0.06)	(0.06)	(0.04)	(0.06)	(0.12)	(0.13)	(0.05)	(0.06)
27 (Electrical)	0.61	0.59	0.53	0.48	0.37	0.31	0.19	0.14
	(0.09)	(0.09)	(0.07)	(0.08)	(0.1)	(0.09)	(0.04)	(0.05)
28 (Machinery)	0.71	0.71	0.60	0.57	0.38	0.33	0.21	0.19
	(0.08)	(0.07)	(0.05)	(0.05)	(0.11)	(0.1)	(0.06)	(0.06)
29 (Vehicles)	0.55	0.54	0.39	0.35	0.27	0.21	0.19	0.16
	(0.05)	(0.05)	(0.04)	(0.04)	(0.08)	(0.09)	(0.03)	(0.03)
30 (Other trans.)	0.56	0.58	0.38	0.38	0.26	0.24	0.00	0.00
	(0.1)	(0.08)	(0.06)	(0.05)	(0.13)	(0.12)	(0.05)	(0.06)
3133 (Furniture)	0.55	0.55	0.42	0.41	0.25	0.21	0.10	0.09
	(0.07)	(0.06)	(0.04)	(0.04)	(0.07)	(0.07)	(0.05)	(0.05)
Observations	171K	185K	171K	185K	171K	185K	171K	185K
R-squared	0.52	0.54	0.65	0.65	0.70	0.70	0.69	0.69
TCC†	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
PLC†	No	No	Yes	Yes	Yes	Yes	Yes	Yes
CAC†	No	No	No	No	Yes	Yes	Yes	Yes
AID†	0.55	0.55	0.43	0.41	0.28	0.24	0.12	0.10
SID	0.08	0.07	0.07	0.07	0.13	0.11	0.09	0.08
F test	0.13	0.09	0.00	0.01	0.07	0.24	0.00	0.01

Notes: Row and column definitions are equivalent to those in Table 3. Columns 1A-1D encompass the original findings as established in the paper. Expanding our analysis to encompass country-years with data available for a minimum of 10 industries in columns 2A-2D has resulted in an augmentation of the overall country count to 28. The standard errors are clustered by importer-industry, and they are presented within brackets.

Table C3. Counterfactual Results

Innovation efficiency measure (1)	Outcome (2)	R & D intensity (3)	Patenting intensity (4)	R & D intensity generalized model (5)
1A Nominal wage	Average change relative to US	0.18	0.14	0.18
2A Nominal wage	Average change relative to US	0.18	0.13	0.17
1B Nominal wage	Dispersion ratio	0.32	0.27	0.31
2B Nominal wage	Dispersion ratio	0.33	0.25	0.29
1C Real income	Average change relative to US	0.06	0.04	0.06
2C Real income	Average change relative to US	0.06	0.04	0.05
1D Real income	Dispersion ratio	0.17	0.13	0.16
2D Real income	Dispersion ratio	0.16	0.11	0.14

Notes: For detailed descriptions of Models 1-3, refer to the notes in Table C1. For detailed descriptions of row and column definitions, refer to the notes in Table 7.

Table D1. Innovation intensity summary statistics

b_s	obs	Mean	Std. dev	Min	Max
R&D intensity					
Model 1	25	-0.69	0.81	-2.61	0.30
Model 2	25	-0.74	0.86	-3.03	0.17
Model 3	24	-0.60	0.73	-2.47	0.33
Model 4	23	-0.72	0.89	-3.20	0.13
Patenting intensity					
Model 1	25	-1.13	1.34	-4.30	0.91
Model 2	25	-1.12	1.35	-4.27	0.85
Model 3	24	-0.94	1.22	-4.04	0.88
Model 4	23	-0.84	1.27	-3.69	1.15

Notes: Model 1 incorporates the initial discoveries outlined in the paper. Model 2 extends the original timeframe from 2010-2014 to 2010-2016. Model 3 divides the timeframe into 2010-2012, while Model 4 focuses on the interval from 2014-2016.

D Modifying time frame

Table D2. Innovation dependence by industry

Innovation efficiency measure	R&D intensity				R&D intensity				R&D intensity				Patenting intensity			
Industries	2010-2014 (1A)	2010-2016 (2A)	2010-2012 (3A)	2014-2016 (4A)	2010-2014 (1B)	2010-2016 (2B)	2010-2012 (3B)	2014-2016 (4B)	2010-2014 (1C)	2010-2016 (2C)	2010-2012 (3C)	2014-2016 (4C)	2010-2014 (1D)	2010-2016 (2D)	2010-2012 (3D)	2014-2016 (4D)
0103 (Agriculture)	0.45	0.43	0.47	0.39	0.33	0.31	0.31	0.2	0.17	0.09	0.16	0.15	0.01	-0.02	0	-0.02
	(0.06)	(0.06)	(0.06)	(0.07)	(0.05)	(0.05)	(0.05)	(0.06)	(0.09)	(0.06)	(0.09)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
0508 (Mining)	0.37	0.33	0.38	0.33	0.25	0.23	0.2	0.14	-0.11	-0.25	-0.27	-0.18	-0.14	-0.17	-0.19	-0.18
	(0.09)	(0.09)	(0.1)	(0.1)	(0.07)	(0.07)	(0.07)	(0.09)	(0.13)	(0.13)	(0.14)	(0.15)	(0.08)	(0.08)	(0.08)	(0.08)
1012 (Food)	0.48	0.45	0.5	0.43	0.36	0.33	0.33	0.24	0.21	0.12	0.19	0.17	0.06	0.03	0.05	0.01
	(0.05)	(0.06)	(0.05)	(0.07)	(0.04)	(0.05)	(0.05)	(0.07)	(0.08)	(0.06)	(0.08)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
13 (Textiles)	0.51	0.47	0.51	0.45	0.42	0.39	0.34	0.3	0.29	0.18	0.23	0.22	0.12	0.06	0.09	0.07
	(0.05)	(0.05)	(0.06)	(0.06)	(0.05)	(0.05)	(0.06)	(0.07)	(0.06)	(0.04)	(0.07)	(0.05)	(0.05)	(0.04)	(0.05)	(0.04)
14 (Apparel)	0.47	0.4	0.49	0.37	0.37	0.33	0.25	0.23	0.33	0.27	0.31	0.37	0.13	0.07	0.13	0.11
	(0.06)	(0.04)	(0.09)	(0.03)	(0.06)	(0.06)	(0.07)	(0.06)	(0.05)	(0.08)	(0.09)	(0.06)	(0.04)	(0.03)	(0.05)	(0.03)
15 (Leather)	0.48	0.43	0.49	0.41	0.39	0.36	0.3	0.3	0.34	0.26	0.31	0.17	0.12	0.06	0.12	0.01
	(0.06)	(0.04)	(0.05)	(0.06)	(0.07)	(0.06)	(0.07)	(0.09)	(0.08)	(0.08)	(0.1)	(0.07)	(0.07)	(0.08)	(0.07)	(0.03)
16 (Wood)	0.52	0.5	0.53	0.47	0.4	0.39	0.37	0.28	0.2	0.19	0.17	0.25	0.03	-0.01	0.01	-0.03
	(0.06)	(0.06)	(0.07)	(0.07)	(0.04)	(0.04)	(0.04)	(0.05)	(0.07)	(0.05)	(0.07)	(0.06)	(0.05)	(0.04)	(0.05)	(0.05)
17 (Paper)	0.58	0.55	0.6	0.55	0.45	0.44	0.41	0.34	0.34	0.3	0.32	0.34	0.13	0.09	0.1	0.09
	(0.05)	(0.05)	(0.06)	(0.07)	(0.04)	(0.04)	(0.04)	(0.06)	(0.07)	(0.05)	(0.08)	(0.05)	(0.05)	(0.05)	(0.06)	(0.05)
18 (Printing)	0.58	0.55	0.59	0.53	0.46	0.43	0.43	0.33	0.27	0.23	0.23	0.26	0.11	0.08	0.09	0.04
	(0.06)	(0.06)	(0.07)	(0.06)	(0.04)	(0.04)	(0.04)	(0.05)	(0.06)	(0.06)	(0.06)	(0.06)	(0.04)	(0.04)	(0.04)	(0.05)
19 (Petrol)	0.48	0.49	0.49	0.45	0.36	0.38	0.33	0.26	0.14	0.15	0.04	0.17	0.05	0.04	0.04	0.05
	(0.05)	(0.05)	(0.06)	(0.07)	(0.04)	(0.04)	(0.05)	(0.06)	(0.08)	(0.06)	(0.08)	(0.07)	(0.04)	(0.03)	(0.03)	(0.04)
20 (Chemicals)	0.59	0.56	0.6	0.55	0.47	0.46	0.42	0.38	0.38	0.25	0.3	0.28	0.19	0.13	0.14	0.16
	(0.05)	(0.05)	(0.05)	(0.07)	(0.05)	(0.06)	(0.05)	(0.08)	(0.09)	(0.1)	(0.1)	(0.09)	(0.06)	(0.07)	(0.06)	(0.08)
21 (Pharma)	0.62	0.53	0.63	0.46	0.5	0.44	0.42	0.31	0.22	0.12	0.06	0.21	0.17	0.12	0.11	0.09
	(0.07)	(0.06)	(0.08)	(0.08)	(0.06)	(0.06)	(0.07)	(0.08)	(0.14)	(0.12)	(0.14)	(0.15)	(0.1)	(0.06)	(0.1)	(0.04)
22 (Plastics)	0.6	0.56	0.61	0.55	0.48	0.45	0.42	0.36	0.38	0.28	0.36	0.33	0.18	0.14	0.15	0.15
	(0.05)	(0.04)	(0.07)	(0.05)	(0.04)	(0.03)	(0.04)	(0.04)	(0.05)	(0.06)	(0.07)	(0.04)	(0.04)	(0.03)	(0.04)	(0.03)
23 (Minerals)	0.57	0.54	0.57	0.53	0.45	0.43	0.4	0.33	0.3	0.24	0.26	0.28	0.12	0.08	0.1	0.05
	(0.05)	(0.05)	(0.07)	(0.06)	(0.04)	(0.04)	(0.04)	(0.05)	(0.06)	(0.05)	(0.06)	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)
24 (Basic metals)	0.58	0.54	0.6	0.5	0.43	0.39	0.42	0.28	0.27	0.19	0.31	0.18	0.18	0.13	0.16	0.15
	(0.05)	(0.06)	(0.04)	(0.07)	(0.05)	(0.05)	(0.05)	(0.07)	(0.07)	(0.06)	(0.08)	(0.05)	(0.03)	(0.02)	(0.04)	(0.03)
25 (Fabric. metals)	0.6	0.58	0.6	0.57	0.48	0.47	0.43	0.36	0.33	0.31	0.31	0.33	0.14	0.12	0.12	0.1
	(0.05)	(0.05)	(0.07)	(0.06)	(0.04)	(0.05)	(0.04)	(0.06)	(0.06)	(0.06)	(0.05)	(0.06)	(0.04)	(0.04)	(0.04)	(0.05)
26 (Computers)	0.65	0.54	0.68	0.47	0.49	0.44	0.45	0.29	0.6	0.41	0.66	0.37	0.3	0.29	0.31	0.26
	(0.06)	(0.06)	(0.06)	(0.07)	(0.04)	(0.05)	(0.05)	(0.06)	(0.12)	(0.11)	(0.16)	(0.09)	(0.05)	(0.05)	(0.06)	(0.11)
27 (Electrical)	0.61	0.57	0.63	0.6	0.53	0.53	0.49	0.52	0.37	0.35	0.34	0.32	0.19	0.16	0.17	0.2
	(0.09)	(0.05)	(0.09)	(0.12)	(0.07)	(0.03)	(0.06)	(0.09)	(0.1)	(0.06)	(0.1)	(0.09)	(0.04)	(0.03)	(0.05)	(0.03)
28 (Machinery)	0.71	0.66	0.68	0.7	0.6	0.57	0.51	0.55	0.38	0.31	0.29	0.37	0.21	0.15	0.16	0.15
	(0.08)	(0.06)	(0.1)	(0.12)	(0.05)	(0.04)	(0.06)	(0.07)	(0.11)	(0.05)	(0.12)	(0.07)	(0.06)	(0.04)	(0.07)	(0.04)
29 (Vehicles)	0.55	0.53	0.55	0.51	0.39	0.41	0.36	0.29	0.27	0.27	0.27	0.2	0.19	0.17	0.19	0.1
	(0.05)	(0.04)	(0.06)	(0.05)	(0.04)	(0.05)	(0.04)	(0.05)	(0.08)	(0.09)	(0.1)	(0.04)	(0.03)	(0.03)	(0.04)	(0.03)
30 (Other trans.)	0.56	0.57	0.59	0.54	0.38	0.47	0.35	0.37	0.26	0.24	0.25	0.3	0	0.05	-0.02	0.01
	(0.1)	(0.07)	(0.12)	(0.08)	(0.06)	(0.08)	(0.09)	(0.08)	(0.13)	(0.09)	(0.19)	(0.07)	(0.05)	(0.05)	(0.06)	(0.05)
3133 (Furniture)	0.55	0.5	0.56	0.51	0.42	0.4	0.37	0.32	0.25	0.18	0.2	0.26	0.1	0.06	0.08	0.03
	(0.07)	(0.07)	(0.08)	(0.07)	(0.04)	(0.04)	(0.05)	(0.06)	(0.07)	(0.07)	(0.07)	(0.06)	(0.05)	(0.06)	(0.05)	(0.06)
Observations	171K	235K	99K	31K	171K	235K	99K	31K	171K	235K	99K	31K	171K	235K	99K	31K
R-squared	0.52	0.54	0.50	0.53	0.65	0.66	0.65	0.67	0.70	0.71	0.70	0.75	0.69	0.71	0.70	0.74
TCC†	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
PLC†	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
CAC†	No	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
AID†	0.55	0.51	0.56	0.49	0.43	0.41	0.38	0.32	0.28	0.21	0.24	0.24	0.12	0.08	0.10	0.07
SID	0.08	0.07	0.07	0.08	0.07	0.07	0.07	0.09	0.13	0.13	0.17	0.12	0.09	0.09	0.10	0.09
F test	0.13	0.04	0.36	0.03	0.00	0.00	0.01	0.02	0.07	0.00	0.01	0.01	0.00	0.00	0.00	0.00

Notes: Row and column definitions are equivalent to those in Table 3. Columns 1A-1D present the original results from Table 1, page 498. In columns 2A-2D, we extend the period by two years, spanning from 2010-2014 to 2010-2016 (following Sampson's Readme file). In columns 3A-3D, we divide the timeframe into 2010-2012. In column 4A-4D, we focus on the interval from 2014-2016. The standard errors are clustered by importer-industry, and they are presented within brackets.

Table D3. Counterfactual Results

Innovation efficiency measure (1)	Outcome (2)	R & D intensity (3)	Patenting intensity (4)	R & D intensity generalized model (5)
1A Nominal wage	Average change relative to US	0.18	0.14	0.18
2A Nominal wage	Average change relative to US	0.16	0.15	0.18
3A Nominal wage	Average change relative to US	0.14	0.10	0.09
4A Nominal wage	Average change relative to US	0.10	0.05	0.14
1B Nominal wage	Dispersion ratio	0.32	0.27	0.31
2B Nominal wage	Dispersion ratio	0.31	0.27	0.32
3B Nominal wage	Dispersion ratio	0.28	0.24	0.21
4B Nominal wage	Dispersion ratio	0.33	0.26	0.35
1C Real income	Average change relative to US	0.06	0.04	0.06
2C Real income	Average change relative to US	0.05	0.04	0.06
3C Real income	Average change relative to US	0.05	0.03	0.03
4C Real income	Average change relative to US	0.03	0.02	0.04
1D Real income	Dispersion ratio	0.17	0.13	0.16
2D Real income	Dispersion ratio	0.15	0.12	0.16
3D Real income	Dispersion ratio	0.14	0.10	0.09
4D Real income	Dispersion ratio	0.15	0.10	0.16

Notes: For detailed descriptions of Models 1-3, refer to the notes in Table D1. For detailed descriptions of row and column definitions, refer to the notes in Table 7.

E Changing values of trade elasticity

Table E1. Innovation intensity summary statistics

b_s	obs	Mean	Std. dev	Min	Max
R&D intensity					
Model 1	25	-0.69	0.81	-2.61	0.30
Model 2	25	-0.69	0.81	-2.61	0.30
Model 3	25	-0.69	0.81	-2.61	0.30
Patenting intensity					
Model 1	25	-1.13	1.34	-4.30	0.91
Model 2	25	-1.13	1.34	-4.30	0.91
Model 3	25	-1.13	1.34	-4.30	0.91

Notes: Model 1 encompasses the original findings as established in the paper. In Model 2, we have adjusted the preferred trade elasticity to 1, deviating from the original value of 6.53. In Model 3, we have adjusted the preferred trade elasticity to 10.5, deviating from the original value of 6.53.

Table E2. Innovation dependence by industry

Innovation efficiency measure	R&D intensity			R&D intensity			R&D intensity			Patenting intensity		
	6.53	1	10.5	6.53	1	10.5	6.53	1	10.5	6.53	1	10.5
Industries	(1A)	(2A)	(3A)	(1B)	(2B)	(3B)	(1C)	(2C)	(3C)	(1D)	(2D)	(3D)
0103 (Agriculture)	0.45	-0.17	0.50	0.33	-0.32	0.38	0.17	-0.50	0.22	0.01	-0.55	0.05
	(0.06)	(0.20)	(0.05)	(0.05)	(0.21)	(0.04)	(0.09)	(0.41)	(0.07)	(0.06)	(0.25)	(0.05)
0508 (Mining)	0.37	-0.66	0.44	0.25	-0.83	0.32	-0.11	-2.36	0.05	-0.14	-1.47	-0.05
	(0.09)	(0.30)	(0.07)	(0.07)	(0.32)	(0.05)	(0.13)	(0.68)	(0.10)	(0.08)	(0.36)	(0.06)
1012 (Food)	0.48	0.00	0.51	0.36	-0.15	0.39	0.21	-0.23	0.24	0.06	-0.19	0.08
	(0.05)	(0.12)	(0.05)	(0.04)	(0.14)	(0.04)	(0.08)	(0.28)	(0.07)	(0.06)	(0.21)	(0.05)
13 (Textiles)	0.51	0.42	0.51	0.42	0.27	0.43	0.29	0.50	0.27	0.12	0.27	0.11
	(0.05)	(0.14)	(0.05)	(0.05)	(0.15)	(0.05)	(0.06)	(0.24)	(0.06)	(0.05)	(0.18)	(0.04)
14 (Apparel)	0.47	0.13	0.50	0.37	-0.07	0.40	0.33	0.86	0.30	0.13	0.56	0.10
	(0.06)	(0.22)	(0.06)	(0.06)	(0.23)	(0.06)	(0.05)	(0.36)	(0.06)	(0.04)	(0.21)	(0.04)
15 (Leather)	0.48	0.33	0.49	0.39	0.12	0.40	0.34	1.03	0.29	0.12	0.25	0.11
	(0.06)	(0.31)	(0.06)	(0.07)	(0.30)	(0.06)	(0.08)	(0.59)	(0.07)	(0.07)	(0.48)	(0.05)
16 (Wood)	0.52	0.25	0.54	0.40	0.11	0.42	0.20	-0.31	0.24	0.03	-0.40	0.06
	(0.06)	(0.19)	(0.06)	(0.04)	(0.21)	(0.03)	(0.07)	(0.26)	(0.06)	(0.05)	(0.13)	(0.04)
17 (Paper)	0.58	0.59	0.58	0.45	0.44	0.45	0.34	0.57	0.33	0.13	0.24	0.12
	(0.05)	(0.15)	(0.05)	(0.04)	(0.15)	(0.04)	(0.07)	(0.26)	(0.06)	(0.05)	(0.18)	(0.05)
18 (Printing)	0.58	0.64	0.58	0.46	0.51	0.46	0.27	0.18	0.28	0.11	0.15	0.11
	(0.06)	(0.14)	(0.06)	(0.04)	(0.15)	(0.04)	(0.06)	(0.18)	(0.06)	(0.04)	(0.09)	(0.04)
19 (Petrol)	0.48	-0.01	0.51	0.36	-0.15	0.39	0.14	-0.98	0.22	0.05	-0.43	0.09
	(0.05)	(0.19)	(0.04)	(0.04)	(0.21)	(0.04)	(0.08)	(0.31)	(0.07)	(0.04)	(0.20)	(0.03)
20 (Chemicals)	0.59	0.72	0.58	0.47	0.57	0.47	0.38	0.67	0.36	0.19	0.52	0.17
	(0.05)	(0.21)	(0.05)	(0.05)	(0.19)	(0.05)	(0.09)	(0.40)	(0.08)	(0.06)	(0.26)	(0.05)
21 (Pharma)	0.62	0.87	0.60	0.50	0.65	0.48	0.22	-0.11	0.25	0.17	0.47	0.15
	(0.07)	(0.23)	(0.07)	(0.06)	(0.22)	(0.05)	(0.14)	(0.70)	(0.11)	(0.10)	(0.50)	(0.07)
22 (Plastics)	0.60	0.77	0.59	0.48	0.60	0.47	0.38	0.79	0.35	0.18	0.48	0.16
	(0.05)	(0.17)	(0.05)	(0.04)	(0.18)	(0.03)	(0.05)	(0.22)	(0.05)	(0.04)	(0.17)	(0.03)
23 (Minerals)	0.57	0.55	0.57	0.45	0.41	0.45	0.29	0.28	0.30	0.12	0.17	0.11
	(0.05)	(0.15)	(0.05)	(0.04)	(0.15)	(0.04)	(0.06)	(0.19)	(0.05)	(0.04)	(0.10)	(0.04)
24 (Basic metals)	0.58	0.56	0.58	0.43	0.41	0.43	0.26	0.51	0.25	0.18	0.61	0.15
	(0.05)	(0.19)	(0.05)	(0.05)	(0.21)	(0.04)	(0.07)	(0.38)	(0.06)	(0.03)	(0.19)	(0.03)
25 (Fabric. metals)	0.60	0.74	0.59	0.47	0.61	0.47	0.33	0.56	0.32	0.14	0.30	0.13
	(0.05)	(0.16)	(0.05)	(0.04)	(0.17)	(0.03)	(0.06)	(0.16)	(0.05)	(0.04)	(0.10)	(0.04)
26 (Computers)	0.65	1.23	0.61	0.49	1.05	0.45	0.60	3.67	0.39	0.29	1.95	0.18
	(0.06)	(0.25)	(0.07)	(0.04)	(0.24)	(0.04)	(0.12)	(0.75)	(0.09)	(0.05)	(0.34)	(0.04)
27 (Electrical)	0.61	1.16	0.57	0.53	1.04	0.49	0.37	1.22	0.31	0.19	0.88	0.14
	(0.09)	(0.34)	(0.08)	(0.07)	(0.32)	(0.06)	(0.10)	(0.57)	(0.07)	(0.04)	(0.30)	(0.03)
28 (Machinery)	0.71	1.54	0.66	0.60	1.40	0.54	0.38	0.88	0.35	0.21	0.84	0.17
	(0.08)	(0.31)	(0.07)	(0.05)	(0.29)	(0.04)	(0.11)	(0.58)	(0.08)	(0.06)	(0.39)	(0.04)
29 (Vehicles)	0.55	0.39	0.56	0.39	0.23	0.40	0.27	0.53	0.26	0.19	0.66	0.16
	(0.05)	(0.19)	(0.05)	(0.04)	(0.19)	(0.03)	(0.08)	(0.52)	(0.06)	(0.03)	(0.31)	(0.02)
30 (Other trans.)	0.56	0.39	0.57	0.38	0.23	0.39	0.26	0.02	0.27	-0.00	-0.28	0.02
	(0.10)	(0.38)	(0.09)	(0.06)	(0.35)	(0.05)	(0.13)	(0.74)	(0.09)	(0.05)	(0.23)	(0.05)
3133 (Furniture)	0.55	0.42	0.56	0.42	0.25	0.44	0.25	-0.03	0.27	0.10	0.09	0.10
	(0.07)	(0.18)	(0.06)	(0.04)	(0.21)	(0.04)	(0.07)	(0.22)	(0.06)	(0.05)	(0.15)	(0.05)
Observations	171K	171K	171K	171K	171K	171K	171K	171K	171K	171K	171K	171K
R-squared	0.52	0.27	0.59	0.65	0.28	0.77	0.70	0.32	0.81	0.69	0.33	0.80
TCC†	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
PLC†	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
CAC†	No	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
AID†	0.55	0.49	0.55	0.43	0.34	0.43	0.28	0.35	0.28	0.12	0.23	0.11
SID	0.08	0.49	0.05	0.07	0.49	0.05	0.13	1.08	0.07	0.09	0.66	0.06
F test	0.13	0	0.89	0	0	0.13	0.07	0	0.8	0	0	0.11

Notes: Row and column definitions are equivalent to those in Table 3. Model 1 encompasses the original findings as established in the paper. Columns 1A-1D encompass the original findings as established in the paper. In Columns 2A-2D, we have adjusted the preferred trade elasticity to 1, deviating from the original value of 6.53. In Columns 3A-3D, we have adjusted the preferred trade elasticity to 10.5, deviating from the original value of 6.53. The standard errors are clustered by importer-industry, and they are presented within brackets.

Table E3. Counterfactual Results

Innovation efficiency measure (1)	Outcome (2)	R & D intensity (3)	Patenting intensity (4)	R & D intensity generalized model (5)
1A Nominal wage	Average change relative to US	0.18	0.14	0.18
2A Nominal wage	Average change relative to US	0.33	0.35	0.36
3A Nominal wage	Average change relative to US	0.17	0.12	0.17
1B Nominal wage	Dispersion ratio	0.32	0.27	0.31
2B Nominal wage	Dispersion ratio	0.50	0.65	0.52
3B Nominal wage	Dispersion ratio	0.32	0.23	0.30
1C Real income	Average change relative to US	0.06	0.04	0.06
2C Real income	Average change relative to US	0.09	0.10	0.09
3C Real income	Average change relative to US	0.06	0.04	0.06
1D Real income	Dispersion ratio	0.17	0.13	0.16
2D Real income	Dispersion ratio	0.23	0.31	0.23
3D Real income	Dispersion ratio	0.18	0.12	0.17

Notes: For detailed descriptions of Models 1-3, refer to the notes in Table E1. For detailed descriptions of row and column definitions, refer to the notes in Table 7.

F Dropping outlier industries

Table F1. Innovation intensity summary statistics

b_s	obs	Mean	Std. dev	Min	Max
R&D intensity					
Model 1	25	-0.69	0.81	-2.61	0.30
Model 2	25	-0.69	0.85	-2.93	0.33
Patenting intensity					
Model 1	25	-1.13	1.34	-4.30	0.91
Model 2	25	-1.18	1.36	-4.54	0.85

Notes: Model 1 encompasses the original findings as established in the paper. In Model 2, we have excluded two outliers, namely the Paper and Paper Products (17) industry and the Agriculture, Forestry, and Fishing (0103) industry. These outliers were identified based on information from Figure 4 in Sampson (2023).

Table F2. Innovation dependence by industry

Innovation efficiency measure	R&D intensity		R&D intensity		R&D intensity		Patenting intensity	
Industries	(1A)	(2A)	(1B)	(2B)	(1C)	(2C)	(1D)	(2D)
0103 (Agriculture)	0.45		0.33		0.17		0.01	
	(0.06)		(0.05)		(0.09)		(0.06)	
0508 (Mining)	0.37	0.34	0.25	0.23	-0.11	-0.15	-0.14	-0.15
	(0.09)	(0.08)	(0.07)	(0.06)	(0.13)	(0.14)	(0.08)	(0.08)
1012 (Food)	0.48	0.44	0.36	0.33	0.21	0.18	0.06	0.05
	(0.05)	(0.04)	(0.04)	(0.04)	(0.08)	(0.08)	(0.06)	(0.06)
13 (Textiles)	0.51	0.48	0.42	0.39	0.29	0.27	0.12	0.11
	(0.05)	(0.04)	(0.05)	(0.05)	(0.06)	(0.06)	(0.05)	(0.05)
14 (Apparel)	0.47	0.44	0.37	0.35	0.33	0.32	0.13	0.13
	(0.06)	(0.05)	(0.06)	(0.05)	(0.05)	(0.06)	(0.04)	(0.04)
15 (Leather)	0.48	0.45	0.39	0.36	0.34	0.32	0.12	0.12
	(0.06)	(0.06)	(0.07)	(0.07)	(0.08)	(0.08)	(0.07)	(0.07)
16 (Wood)	0.52	0.49	0.40	0.37	0.20	0.18	0.03	0.03
	(0.06)	(0.05)	(0.04)	(0.04)	(0.07)	(0.07)	(0.05)	(0.04)
17 (Paper)	0.58		0.45		0.34		0.13	
	(0.05)		(0.04)		(0.07)		(0.05)	
18 (Printing)	0.58	0.54	0.46	0.43	0.27	0.26	0.11	0.10
	(0.06)	(0.05)	(0.04)	(0.04)	(0.06)	(0.06)	(0.04)	(0.04)
19 (Petrol)	0.48	0.45	0.36	0.34	0.14	0.12	0.05	0.05
	(0.05)	(0.04)	(0.04)	(0.04)	(0.08)	(0.08)	(0.04)	(0.04)
20 (Chemicals)	0.59	0.54	0.47	0.44	0.38	0.34	0.19	0.18
	(0.05)	(0.05)	(0.05)	(0.05)	(0.09)	(0.09)	(0.06)	(0.06)
21 (Pharma)	0.62	0.57	0.50	0.45	0.22	0.15	0.17	0.15
	(0.07)	(0.07)	(0.06)	(0.06)	(0.14)	(0.13)	(0.1)	(0.1)
22 (Plastics)	0.60	0.57	0.48	0.45	0.38	0.37	0.18	0.18
	(0.05)	(0.04)	(0.04)	(0.03)	(0.05)	(0.05)	(0.04)	(0.04)
23 (Minerals)	0.57	0.53	0.45	0.42	0.30	0.28	0.12	0.11
	(0.05)	(0.04)	(0.04)	(0.04)	(0.06)	(0.06)	(0.04)	(0.04)
24 (Basic metals)	0.58	0.53	0.43	0.39	0.27	0.23	0.18	0.17
	(0.05)	(0.05)	(0.05)	(0.05)	(0.07)	(0.07)	(0.03)	(0.03)
25 (Fabric. metals)	0.60	0.56	0.48	0.44	0.33	0.31	0.14	0.14
	(0.05)	(0.04)	(0.04)	(0.03)	(0.06)	(0.06)	(0.04)	(0.04)
26 (Computers)	0.65	0.59	0.49	0.45	0.60	0.54	0.30	0.30
	(0.06)	(0.04)	(0.04)	(0.03)	(0.12)	(0.1)	(0.05)	(0.05)
27 (Electrical)	0.61	0.58	0.53	0.51	0.37	0.36	0.19	0.19
	(0.09)	(0.08)	(0.07)	(0.07)	(0.1)	(0.09)	(0.04)	(0.04)
28 (Machinery)	0.71	0.66	0.60	0.56	0.38	0.34	0.21	0.21
	(0.08)	(0.07)	(0.05)	(0.05)	(0.11)	(0.09)	(0.06)	(0.06)
29 (Vehicles)	0.55	0.51	0.39	0.36	0.27	0.27	0.19	0.19
	(0.05)	(0.04)	(0.04)	(0.03)	(0.08)	(0.08)	(0.03)	(0.03)
30 (Other trans.)	0.56	0.50	0.38	0.34	0.26	0.19	0.00	0.00
	(0.1)	(0.08)	(0.06)	(0.06)	(0.13)	(0.11)	(0.05)	(0.05)
3133 (Furniture)	0.55	0.51	0.42	0.40	0.25	0.24	0.10	0.10
	(0.07)	(0.06)	(0.04)	(0.04)	(0.07)	(0.06)	(0.05)	(0.05)
Observations	171K	151K	171K	151K	171K	151K	171K	151K
R-squared	0.52	0.52	0.65	0.65	0.70	0.70	0.69	0.70
TCC†	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
PLC†	No	No	Yes	Yes	Yes	Yes	Yes	Yes
CAC†	No	No	No	No	Yes	Yes	Yes	Yes
AID†	0.55	0.51	0.43	0.40	0.28	0.26	0.12	0.12
SID	0.08	0.07	0.07	0.07	0.13	0.13	0.09	0.09
F test	0.13	0.08	0.00	0.00	0.07	0.03	0.00	0.00

Notes: Row and column definitions are equivalent to those in Table 3. Columns 1A-1D encompass the original findings as established in the paper. In columns 2A-2D, we have excluded two outliers, namely the Paper and Paper Products (17) industry and the Agriculture, Forestry, and Fishing (0103) industry. These outliers were identified based on information from Figure 4 in Sampson (2023). The standard errors are clustered by importer-industry, and they are presented within brackets.

Table F3. Counterfactual Results

Innovation efficiency measure (1)	Outcome (2)	R & D intensity (3)	Patenting intensity (4)	R & D intensity generalized model (5)
1A Nominal wage	Average change relative to US	0.18	0.14	0.18
2A Nominal wage	Average change relative to US	0.16	0.15	0.18
1B Nominal wage	Dispersion ratio	0.32	0.27	0.31
2B Nominal wage	Dispersion ratio	0.31	0.27	0.32
1C Real income	Average change relative to US	0.06	0.04	0.06
2C Real income	Average change relative to US	0.05	0.04	0.06
1D Real income	Dispersion ratio	0.17	0.13	0.16
2D Real income	Dispersion ratio	0.15	0.12	0.16

Notes: For detailed descriptions of Models 1-3, refer to the notes in Table F1. For detailed descriptions of row and column definitions, refer to the notes in Table 7.

G Dropping outlier in the regression

Table G1. Innovation dependence by industry

Innovation efficiency measure	R&D intensity			R&D intensity			R&D intensity			Patenting intensity		
Industries	(1A)	(2A)	(3A)	(1B)	(2B)	(3B)	(1C)	(2C)	(3C)	(1D)	(2D)	(3D)
0103 (Agriculture)	0.45	0.45	0.46	0.33	0.36	0.33	0.17	0.29	0.17	0.01	0.13	0.01
	(0.06)	(0.05)	(0.05)	(0.05)	(0.04)	(0.05)	(0.09)	(0.07)	(0.09)	(0.06)	(0.06)	(0.06)
0508 (Mining)	0.37	0.39	0.38	0.25	0.31	0.26	-0.11	-0.05	-0.10	-0.14	-0.06	-0.13
	(0.09)	(0.08)	(0.08)	(0.07)	(0.06)	(0.06)	(0.13)	(0.14)	(0.13)	(0.08)	(0.08)	(0.08)
1012 (Food)	0.48	0.48	0.48	0.36	0.38	0.36	0.21	0.21	0.21	0.06	0.12	0.06
	(0.05)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.08)	(0.09)	(0.08)	(0.06)	(0.06)	(0.06)
13 (Textiles)	0.51	0.50	0.51	0.42	0.45	0.42	0.29	0.31	0.29	0.12	0.18	0.12
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.06)	(0.07)	(0.06)	(0.05)	(0.06)	(0.05)
14 (Apparel)	0.47	0.45	0.47	0.37	0.40	0.37	0.33	0.32	0.33	0.13	0.18	0.14
	(0.06)	(0.05)	(0.06)	(0.06)	(0.06)	(0.06)	(0.05)	(0.08)	(0.05)	(0.04)	(0.05)	(0.04)
15 (Leather)	0.48	0.44	0.48	0.39	0.41	0.39	0.34	0.34	0.34	0.12	0.22	0.13
	(0.06)	(0.05)	(0.06)	(0.07)	(0.07)	(0.07)	(0.08)	(0.07)	(0.08)	(0.07)	(0.06)	(0.07)
16 (Wood)	0.52	0.54	0.52	0.40	0.44	0.40	0.20	0.21	0.20	0.03	0.12	0.03
	(0.06)	(0.07)	(0.06)	(0.04)	(0.05)	(0.04)	(0.07)	(0.08)	(0.07)	(0.05)	(0.04)	(0.04)
17 (Paper)	0.58	0.56	0.58	0.45	0.46	0.45	0.34	0.32	0.34	0.13	0.19	0.13
	(0.05)	(0.06)	(0.05)	(0.04)	(0.04)	(0.04)	(0.07)	(0.09)	(0.07)	(0.05)	(0.06)	(0.05)
18 (Printing)	0.58	0.64	0.58	0.46	0.54	0.46	0.27	0.34	0.28	0.11	0.23	0.11
	(0.06)	(0.06)	(0.06)	(0.04)	(0.04)	(0.04)	(0.06)	(0.08)	(0.06)	(0.04)	(0.05)	(0.04)
19 (Petrol)	0.48	0.52	0.48	0.36	0.43	0.36	0.14	0.31	0.15	0.05	0.17	0.06
	(0.05)	(0.05)	(0.04)	(0.04)	(0.04)	(0.04)	(0.08)	(0.07)	(0.08)	(0.04)	(0.04)	(0.03)
20 (Chemicals)	0.59	0.56	0.59	0.47	0.49	0.48	0.38	0.45	0.38	0.19	0.27	0.19
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.09)	(0.10)	(0.09)	(0.06)	(0.07)	(0.06)
21 (Pharma)	0.62	0.59	0.62	0.50	0.53	0.50	0.22	0.11	0.22	0.17	0.20	0.17
	(0.07)	(0.08)	(0.07)	(0.06)	(0.06)	(0.06)	(0.14)	(0.18)	(0.14)	(0.10)	(0.12)	(0.09)
22 (Plastics)	0.60	0.54	0.60	0.48	0.46	0.48	0.38	0.36	0.38	0.18	0.22	0.19
	(0.05)	(0.06)	(0.05)	(0.04)	(0.05)	(0.03)	(0.05)	(0.08)	(0.05)	(0.04)	(0.05)	(0.04)
23 (Minerals)	0.57	0.59	0.57	0.45	0.49	0.45	0.29	0.29	0.29	0.12	0.19	0.12
	(0.05)	(0.06)	(0.05)	(0.04)	(0.04)	(0.04)	(0.06)	(0.07)	(0.06)	(0.04)	(0.04)	(0.04)
24 (Basic metals)	0.58	0.54	0.58	0.43	0.44	0.43	0.26	0.37	0.27	0.18	0.28	0.19
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.07)	(0.10)	(0.07)	(0.03)	(0.04)	(0.03)
25 (Fabric. metals)	0.60	0.58	0.60	0.47	0.48	0.48	0.33	0.31	0.33	0.14	0.20	0.14
	(0.05)	(0.06)	(0.05)	(0.04)	(0.05)	(0.04)	(0.06)	(0.08)	(0.06)	(0.04)	(0.05)	(0.04)
26 (Computers)	0.65	0.54	0.65	0.49	0.46	0.49	0.60	0.61	0.60	0.29	0.34	0.30
	(0.06)	(0.09)	(0.06)	(0.04)	(0.07)	(0.04)	(0.12)	(0.15)	(0.12)	(0.05)	(0.06)	(0.05)
27 (Electrical)	0.61	0.52	0.61	0.53	0.50	0.53	0.37	0.33	0.37	0.19	0.20	0.19
	(0.09)	(0.08)	(0.09)	(0.07)	(0.06)	(0.07)	(0.10)	(0.08)	(0.10)	(0.04)	(0.04)	(0.04)
28 (Machinery)	0.71	0.68	0.71	0.60	0.61	0.60	0.38	0.30	0.38	0.21	0.25	0.22
	(0.08)	(0.07)	(0.08)	(0.05)	(0.05)	(0.05)	(0.11)	(0.09)	(0.11)	(0.06)	(0.06)	(0.06)
29 (Vehicles)	0.55	0.49	0.56	0.39	0.39	0.39	0.27	0.33	0.28	0.19	0.26	0.19
	(0.05)	(0.05)	(0.05)	(0.04)	(0.05)	(0.04)	(0.08)	(0.11)	(0.08)	(0.03)	(0.04)	(0.03)
30 (Other trans.)	0.56	0.58	0.56	0.38	0.47	0.38	0.26	0.16	0.25	-0.00	0.04	0.00
	(0.10)	(0.12)	(0.09)	(0.06)	(0.08)	(0.06)	(0.13)	(0.08)	(0.12)	(0.05)	(0.04)	(0.05)
3133 (Furniture)	0.55	0.56	0.55	0.42	0.48	0.43	0.25	0.26	0.25	0.10	0.20	0.11
	(0.07)	(0.07)	(0.06)	(0.04)	(0.04)	(0.04)	(0.07)	(0.08)	(0.06)	(0.05)	(0.05)	(0.05)
Observations	171K	171K	168K	171K	171K	168K	171K	171K	168K	171K	171K	168K
R-squared	0.52	0.31	0.54	0.65	0.34	0.66	0.70	0.35	0.71	0.69	0.35	0.70
TCC†	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
PLC†	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
CAC†	No	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
AID†	0.55	0.53	0.55	0.43	0.45	0.43	0.28	0.29	0.28	0.17	0.19	0.12
SID	0.08	0.07	0.07	0.07	0.06	0.07	0.13	0.13	0.13	0.09	0.08	0.09
F test	0.13	0.25	0.12	0.00	0.01	0.00	0.07	0.36	0.06	0.00	0.00	0.00

Notes: Row and column definitions are equivalent to those in Table 3. We replicated the original results in Columns 1A-1D as reported in the paper. We then applied the robustness check method from Jann (2022), which uses a robust fixed effect regression to identify outliers. We found that the proportions of outliers in Columns 2A-2D were 36.57%, 36.55%, 36.59%, and 36.63%, respectively. In Columns 3A-3D, we excluded these outliers and reran the original regression as in Columns 1A-1D.