Federal Reserve Bank of Dallas

The Effects of U.S. Public R&D on Global Growth

Gustavo De Souza, Andrew J. Fieldhouse, Karel Mertens, Ishan B. Nath and Valerie A. Ramey

Working Paper 2622 **August 2026**

Research Department

https://doi.org/10.24149/wp2622

The Effects of U.S. Public R&D on Global Growth[*]

Gustavo De Souza[†], Andrew J. Fieldhouse[‡], Karel Mertens[§],
Ishan B. Nath[±] and Valerie A. Ramey[°]

July 15, 2026

Abstract

This paper provides the first causal estimates of the global social returns to U.S. public R&D. We use a narrative identification strategy to quantify the effects of exogenous shocks to U.S. R&D appropriations on the dynamic TFP response of 69 foreign economies from 1980–2019. A U.S. R&D appropriations shock equal to 1 percent of the federal R&D capital stock raises foreign TFP by approximately 1 percent after 12 years. This response is driven primarily by nondefense rather than defense R&D and is concentrated in non-OECD economies. These patterns are most consistent with diffusion occurring through openly accessible scientific content, capital-embodied trade, and technological leapfrogging by economies further from the global frontier. A back-of-the-envelope calculation suggests that the global social returns to U.S. public nondefense R&D are roughly twice as large as the domestic returns, meaning the U.S. captures about half of these productivity benefits.

JEL Classification: E62, F62, O33, O38, O47.

Keywords: Public R&D, International R&D Spillovers, Social Returns to R&D, Productivity, Technological Diffusion.

---

[*] We are grateful to Robert Inklaar and Pete Klenow for helpful comments and discussion. We also thank Phyo Ba Kyu with the United Nations Statistics Division for providing data. Eric Chang and Yuan Zhou provided excellent research assistance. The views in this paper are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Chicago, the Federal Reserve Bank of Dallas, or the Federal Reserve System.

[†] Gustavo De Souza, Federal Reserve Bank of Chicago.

[‡] Andrew J. Fieldhouse, Mays Business School, Texas A&M University.

[§] Karel Mertens, Federal Reserve Bank of Dallas and CEPR, mertens.karel@gmail.com.

[±] Ishan B. Nath, Harvard Kennedy School and NBER.

[°] Valerie A. Ramey, Hoover Institution, Stanford University, NBER and CEPR.

1 Introduction

The nonrivalry of ideas is a core foundation of modern growth theory (Romer, 1990). Because knowledge produced anywhere can be applied everywhere, any one country's investment in research and development (R&D) can in principle raise productivity around the world—a positive externality with first-order implications for the social returns to R&D.

A broad class of growth models embeds such cross-country knowledge spillovers as the central force tying together patterns of long-run growth across the world. These models can explain how large, persistent differences in productivity levels across countries can coexist with the absence of permanent divergence in growth rates, why frontier countries with very different levels of domestic innovation tend to grow together, and why episodes of growth by economies further from the technological frontier can be especially rapid at first and taper off as they catch up.[1]

Yet the critical parameter in these canonical models of global growth—the magnitude of cross-country knowledge spillovers—has generally been calibrated indirectly to match descriptive moments from cross-country patterns of total factor productivity (TFP), R&D capital, trade flows, and patent data rather than using direct empirical estimates. Recent work has produced credible causal evidence on the *domestic* social returns to public R&D (Fieldhouse and Mertens, 2026a), but corresponding estimates of the *global* social returns to public R&D do not yet exist. Identifying the magnitude of global productivity spillovers from public R&D in a frontier country matters for a wide range of policy questions, including cost-benefit analysis of R&D funding, the case for international coordination to internalize positive R&D externalities, and the potential productivity costs of de-globalization policies that risk impeding technological diffusion.

This paper provides the first causal estimates of the global social returns to U.S. public R&D, meaning R&D work funded by the government but performed by a mix of federal agencies, national laboratories, private firms, universities, and other nonprofits. We extend the time-series narrative identification of Fieldhouse and Mertens (2026a), which establishes a strong causal link between exogenous changes in U.S. federal R&D appropriations and U.S. business-sector productivity growth, to an international panel local

[1] See Aghion and Howitt (1992); Coe and Helpman (1995); Howitt (2000); Klenow and Rodríguez-Clare (2005); Acemoglu et al. (2006); Lucas (2009); Buera and Oberfield (2020); Hsieh et al. (2023), among others.

projections framework covering 69 foreign countries over 1980–2019. Our identifying variation is based on the Fieldhouse and Mertens (2023) narrative shocks to defense and nondefense R&D appropriations—orthogonalized to lagged U.S. controls—for the five major federal agencies funding U.S. R&D: the Department of Defense, Department of Energy, National Aeronautics and Space Administration, National Institutes of Health, and National Science Foundation. With respect to policy endogeneity, the identifying assumption is even weaker than for the U.S., since U.S. R&D policy is unlikely to respond to cyclical economic conditions abroad. The longer-run endogeneity concern is not attenuated in the same way, however: common global technology news that predicts future U.S. TFP can also predict future foreign TFP, simultaneously influencing U.S. R&D appropriations and foreign productivity. We address this residual concern using lagged controls, as detailed in Section 3. We estimate impulse responses of TFP in other countries (jointly and separately by subgroups) to defense, nondefense, and combined (defense and nondefense) U.S. R&D shocks over a 15-year horizon.

We find that foreign productivity responds significantly and persistently to U.S. nondefense R&D shocks. In the full sample of countries excluding the U.S., a 1 percent shock to combined U.S. defense and nondefense R&D appropriations relative to the federal R&D capital stock gradually raises average foreign TFP, with the effect peaking at a 1 percent increase approximately 12 years after the shock and remaining significant through the end of the 15-year horizon. When separating the effects by type of shock, we find that the dynamic pattern of gradual diffusion over the initial decade after a shock is similar for both defense and nondefense R&D shocks, but the magnitude of the effect peaks slightly higher for nondefense shocks (0.6 percent versus 0.4 percent) and the TFP response to nondefense shocks is far more persistent at longer horizons. These headline results survive a battery of additional tests: the responses for both OECD and non-OECD countries are robust to omitting any individual country from the panel, so that no outlier country unduly drives the estimated TFP responses, and rolling regressions show that the results are relatively stable across estimation windows within the 1980–2019 sample.

The foreign TFP responses to U.S. R&D shocks display a particular form of heterogeneity: countries with smaller domestic stocks of R&D and intellectual-property capital, relative to the U.S., see relatively larger TFP responses. This pattern holds both within the OECD—when splitting the sample by above- and below-median R&D capital stocks—

and more broadly, when separating original OECD members from countries that never joined the group. The TFP response to a 1 percent combined R&D shock in non-OECD countries peaks at roughly 1.8 percent and persists to the end of the 15-year horizon. This is more than twice as large as the 0.7 percent peak effect in OECD countries (excluding the U.S.), which reverts to zero at longer horizons. Importantly, we show that the larger non-OECD response is not an artifact of differential TFP measurement across these groups. The effects of nondefense R&D shocks are also larger and more persistent in both OECD and non-OECD countries than the effects of defense R&D.

The heterogeneous patterns in the transmission of U.S. public R&D shocks to foreign TFP help narrow the set of candidate mechanisms driving global technological diffusion over the last half century, supporting some channels while casting doubt on others. The disproportionate effects of nondefense relative to defense R&D support the hypothesis that diffusion benefits from the free flow of ideas, goods, and technologies. While many defense innovations are initially classified and intended only for domestic use (or, with time, shared only with close allies), nondefense R&D produces scientific publications and tradable products that become widely available across the world. U.S. defense R&D funding also largely skews toward *development*, whereas nondefense funding is concentrated in fundamental *research*, where spillovers are thought to be larger (Akcigit et al., 2020; Fieldhouse and Mertens, 2026a).

On the other hand, if diffusion operated primarily through the direct flow of knowledge, we might expect countries with more domestic R&D capacity—and thus a greater ability to absorb and apply new ideas—to benefit most (Cohen and Levinthal, 1990; Howitt, 2000). Our results, however, show the opposite: benefits flow disproportionately to countries with *less* domestic R&D capacity, pointing away from direct knowledge transfer and toward channels that do not require domestic absorptive capacity. A second hypothesis we can distinguish is the classic catch-up story: economies further behind the frontier have more to gain from the larger *stock* of unused technologies already available at the frontier (Gerschenkron, 1962; Abramovitz, 1986). But that story turns on the existing stock of frontier technology, whereas the nondefense R&D shock we study delivers a *pulse* of new technology that is equally available to every country regardless of distance to the frontier. Our larger estimated TFP responses for non-OECD countries and those with smaller domestic R&D capital stocks must instead reflect differences in how readily countries adopt new frontier technologies.

Two other mechanisms can account for this adoption margin, consistent with the heterogeneous TFP spillovers we observe. The first is diffusion through capital-embodied trade (Eaton and Kortum, 2001; Caselli and Wilson, 2004; Ayerst, Ibrahim, MacKenzie, and Rachapalli, 2023): U.S. nondefense R&D reaches foreign economies through the direct imports of machinery, pharmaceuticals, software, and agricultural technologies (Farrokhi and Pellegrina, 2023), all of which can raise TFP in those countries by providing more advanced productive inputs. Moreover, lower-income countries tend to have higher import penetration of capital goods used in domestic production and thus can gain more from improvements in U.S. technology, consistent with our larger estimated TFP responses to U.S. public R&D in these countries.

The second plausible candidate is leapfrogging (Brezis, Krugman, and Tsiddon, 1993): economies further from the frontier have less locked-in legacy capital and thus face lower transition costs to incorporating newly invented technology. This dynamic is illustrated by the roll-out of mobile telephones in countries that largely skipped the landline stage and the deployment of solar power with battery storage in places lacking established electricity grids. A related variant of this mechanism is that countries further from the frontier grow more through imitation rather than innovation (Acemoglu, Aghion, and Zilibotti, 2006), and are thus better placed to incorporate advances made by frontier innovators. Openly accessible scientific content, capital-embodied trade, and leapfrogging are all consistent with both the broader literature and the new causal evidence in this paper.

Beyond shedding light on how public R&D diffuses abroad, our findings speak to two central questions for the international coordination of innovation policy: Is the rest of the world free-riding by scaling back its own R&D as the U.S. increases public R&D? Or is the global supply of R&D resources sufficiently inelastic that the U.S. bids them away from other countries when it expands its own R&D? Either channel would show up as U.S. R&D crowding out foreign R&D, so to test for them, we estimate how foreign R&D activity—proxied by R&D capital and R&D researchers—responds to the U.S. R&D appropriations shocks. We find that increases in U.S. public R&D modestly crowd *in* foreign private-sector and higher-education R&D investment; we also find positive albeit imprecisely estimated effects of U.S. R&D appropriations shocks on the number of R&D researchers employed in other OECD countries. Taken together, we find no evidence that

U.S. public R&D crowds out foreign R&D investment or reallocates R&D researchers from foreign countries to the U.S., although data limitations preclude such analysis for the non-OECD countries capturing the largest TFP spillovers. These results rule out displacement of foreign R&D, whether through free-riding or competition for scarce resources, but they do not speak to free-riding in the broader public-good sense, i.e., other countries benefiting from U.S. R&D investments they did not fund.

This paper relates most directly to a small empirical literature on causal identification of international R&D spillovers. The closest previous work is by Moretti, Steinwender, and Van Reenen (2025), who use a shift-share instrumental variable (IV) strategy based on predicted defense R&D spending to identify within-country crowd-in effects of government defense R&D on private R&D investment and productivity growth. Their paper also documents cross-border crowd-in within the same industry, with private-sector R&D responding positively to public R&D in other OECD countries. Our finding of foreign R&D crowd-in is consistent with this cross-border result, but our central contributions are quantifying the dynamic productivity spillovers over longer horizons, documenting their concentration in non-OECD economies, and showing their greater strength for non-defense than defense R&D. In particular, the stronger spillovers we find for *nondefense* R&D and non-OECD countries speak directly to margins their design cannot identify.

Other empirical work related to this paper includes within-firm evidence on R&D spillovers to multinational affiliates' productivity (Bilir and Morales, 2020), patent-citation evidence on cross-border knowledge flows (Branstetter, 2001), quasi-experimental evidence on international scientific knowledge flows (Iaria, Schwarz, and Waldinger, 2018), correlational evidence linking trade-weighted foreign R&D capital to domestic TFP (Coe and Helpman, 1995; Coe, Helpman, and Hoffmaister, 1997, 2009), and more recent evidence on technological diffusion through international trade (Ayerst, Ibrahim, MacKenzie, and Rachapalli, 2023). On the theoretical and quantitative side, our magnitudes bear on a growing literature studying cross-country idea flows and the role of trade as a diffusion mechanism (Alvarez, Buera, and Lucas, 2013; Cai, Li, and Santacreu, 2022; Hsieh, Klenow, and Nath, 2023; Akcigit, Ates, and Impullitti, 2026), in which the magnitude of cross-country spillovers is typically calibrated to indirect moments rather than estimated directly.

Our paper makes two broad contributions. First, to our knowledge, we provide the first causally identified estimates of international TFP spillovers and global social returns to U.S. public R&D. Weighting our TFP impulse responses for each country group by real GDP and discounting future productivity benefits, we quantify a lower bound for the wedge between global and domestic social returns to public nondefense R&D; these calculations imply that the global social returns are roughly twice as large as the domestic returns, so that the U.S. captures at most about half of the total productivity benefits from its own public R&D investment. This builds on existing evidence of domestic firm-to-firm spillovers from private R&D (Bloom, Schankerman, and Van Reenen, 2013) and domestic social returns to public R&D (Fieldhouse and Mertens, 2026a). Second, our cross-country heterogeneity analysis is the first to use the response to a quasi-experimental frontier R&D pulse to narrow the set of plausible diffusion channels, pointing toward channels that operate through the openly accessible outputs of nondefense R&D and disproportionately benefit economies further from the technological frontier.

The remainder of the paper proceeds as follows. Section 2 describes our data and measurement choices. Section 3 lays out our empirical strategy. Section 4 presents the main results on cross-country TFP responses and heterogeneity across countries. Section 5 explores mechanisms and documents foreign crowd-in of R&D activity in response to increases in U.S. public R&D. Section 6 quantifies global social returns to public R&D relative to domestic returns, and Section 7 concludes.

2 Data and Measurement

To study international productivity spillovers from the U.S. technology frontier to the rest of the world and the mechanisms of global technological diffusion, we first assemble an international panel dataset of productivity measures and R&D activity.

2.1 International Productivity Data

Our primary outcome of interest is TFP, which, under common assumptions, is interpreted as a measure of technological progress or know-how. Consistent TFP measurement across countries is notoriously difficult, and the underlying measurement choices turn out to matter substantively for interpreting our results and the transmission mechanisms behind our cross-country spillover estimates. There are real tradeoffs among coverage breadth, sample length, and measurement quality in international TFP data;

we navigate these tradeoffs explicitly and document our choices throughout.

Our international panel dataset of TFP and related production function inputs and output is primarily compiled from the Penn World Table (PWT) version 11.0 (Feenstra et al., 2015), which offers the broadest panel of TFP data for both advanced and emerging market economies across regions. PWT measures TFP as a Solow residual from a constant-returns translog production function in labor and an aggregate capital stock, with a capital share that varies by country and time. The aggregate capital stock is constructed using the perpetual inventory method (PIM) from United Nations (UN) National Accounts data, aggregating up to nine asset classes including software and other intellectual property products. However, the country coverage of the intellectual-property components varies across countries' System of National Accounts (SNA) vintages. For countries reporting under SNA 2008, R&D investment is capitalized in the underlying national accounts and is included in PWT's aggregate capital stock construction. Conversely, countries still reporting under SNA 1968 or 1993 do not have R&D capitalized, since the pre-2008 SNA frameworks treated R&D expenditure as a current expense rather than as capital formation. For these countries, R&D investment is therefore absent from the aggregate capital stock and instead shows up in the PWT TFP residual. This measurement-regime distinction turns out to be central to interpreting our spillover estimates, as revisited in Section 4.

Fieldhouse and Mertens (2026a) document significant longer-run U.S. TFP responses to exogenous R&D appropriations shocks, but there are two important methodological wedges between the U.S. TFP series used in that paper and the PWT TFP series we use here. Fieldhouse and Mertens (2026a) primarily study the response of the utilization-adjusted TFP for the U.S. business sector originally produced by Fernald (2012) and maintained by the Federal Reserve Bank of San Francisco. The PWT TFP series we use, by contrast, are constructed for the entire economy rather than just the business sector, and are not adjusted for cyclical utilization rates of factor inputs. Critically, the PWT TFP series net out both private R&D and government R&D investment from the Solow residual—at least for countries reporting under SNA 2008—whereas only private R&D capital and other business-sector inputs are netted out of the Fernald measure. Even when foreign R&D capital responds endogenously to U.S. R&D appropriations shocks—as our impulse response functions in Section 5.1 will show—capitalization under SNA 2008 nets the measured capital-service contribution of that induced R&D out of the TFP residual. It does not, however, remove the mediated effects from increased R&D, such as local knowledge spillovers, input complementarities, and social returns in excess of its measured user cost; so for SNA 2008 countries, the PWT residual captures these alongside direct spillovers from the U.S. frontier, not direct spillovers alone. For countries still using SNA 1968 or 1993 frameworks, any R&D capital response instead flows into the TFP residual.

2.2 Baseline Sample and Panel of Countries

We compile and construct data for as many countries over as much of the post-WWII era as possible, while taking care to avoid major confounding influences on measured productivity that would not be absorbed by country fixed effects in our regression models. In particular, we structure the panel to avoid four potential sources of contamination from structural breaks and known measurement problems in TFP data that would not reflect technological diffusion.

First, to avoid contamination of measured TFP from post-WWII reconstruction dynamics and rapid capital re-accumulation in advanced economies (e.g., France, Germany, Japan), we use a delayed 1980–2019 sample window for our baseline estimation.[2] A parallel concern motivates the delayed start for the non-OECD economies in our panel, where early post-WWII productivity dynamics also reflect pre-liberalization trade and industrial regimes and, in many countries, post-independence institutional transitions, as opposed to cleanly capturing technological diffusion. Moreover, we prefer a uniform sample window for all countries, which allows cleaner comparison of our results across the full panel and country subgroups. Second, we exclude petro states and other heavy resource-export-oriented economies, where TFP is likely biased by extraction and exchange rate fluctuations (e.g., Saudi Arabia, Qatar), as PWT does not provide resource-adjusted TFP measures for these countries.[3] Third, we exclude former Soviet-bloc transition economies, which suffer from well-known data quality issues during central planning and from institutional regime changes that do not reflect technological progress

[2]The start date is chosen based on diagnostic tests of relative output, capital, and labor productivity growth in these post-WWII reconstruction economies. We deliberately end our sample in 2019 to avoid COVID-era productivity measurement problems, which are of particular concern for TFP measures that are not adjusted for cyclical factor utilization rates such as the PWT data we use.

[3]For example, PWT TFP for Saudi Arabia is nearly monotonically decreasing from its first observation in 1974, with the 1973–74 and 1979–81 oil shocks inflating measured TFP.

per se (e.g., Poland, Romania). Finally, we exclude countries with one-off political or economic collapses or prolonged civil wars that appear to be reflected in measured TFP (e.g., Sudan, Zimbabwe).

After these exclusions, we are left with a panel of 70 countries with TFP data consistently available over 1978–2019, i.e., with two years of pre-sample data for the lagged controls in our regressions. Our baseline panel of countries includes 30 modern OECD countries—24 of which are 'original' OECD countries (defined as membership at the start of our sample in 1980)—and 40 non-OECD countries.[4] Using the World Bank's regional classification, our final panel provides reasonably broad coverage for countries across the global regions of East Asia and the Pacific (12), Europe and Central Asia (21), Latin America and the Caribbean (17), and North America (USA and Canada); coverage is relatively more limited for the Middle East and North Africa (5), South Asia (3), and Sub-Saharan Africa (10).[5] See Appendix Table A.1 for a list of all countries in our baseline sample.

Beyond studying TFP responses by country group or region, we also code each country-year by the UN Statistics Division's own classification of SNA vintage from the National Accounts Official Country Data Database (1968, 1993, or 2008). Doing so allows us to separate countries where PWT TFP nets out both public and private R&D capital accumulation (SNA 2008) from those where R&D investment effects instead flow into the TFP residual (SNA 1968 or 1993). This distinction is also central to interpreting the responses of foreign R&D capital stocks we study in Section 5, particularly whether they are reflected in or netted out of measured TFP. Appendix Table A.1 also reports the SNA system in use for each country and year in our baseline sample, since countries can transition between SNA systems over time as national accounts are revised.

2.3 International R&D Activity Data

The international panel of R&D activity comes from two OECD Main Science and Technology Indicators (MSTI) datasets, again balancing tradeoffs among coverage breadth, sample length, and measurement quality. These data serve two roles in the analysis. First,

[4]Note that the Soviet-bloc exclusion in particular drops eight current OECD member states (Czechia, Estonia, Hungary, Latvia, Lithuania, Poland, Slovakia, and Slovenia), which is why our panel includes only 30 of the 38 current OECD members.

[5]https://datatopics.worldbank.org/sdgatlas/archive/2017/the-world-by-region.html.

they are direct outcomes of interest, allowing us to quantify the mechanisms through which spillovers manifest—for example, whether foreign R&D investment is crowded in (or out) by U.S. R&D appropriations shocks. Second, their consistent cross-country definitions and construction provide a complement to the PWT series, identifying channel-specific responses that PWT's aggregate residual treatment cannot separate (e.g., private-sector R&D capital versus government R&D capital).

We extend the methodology of Coe et al. (2009) to construct our international panel of R&D capital stocks for the 24 advanced economies with requisite data availability. Following their PIM approach, we use OECD MSTI R&D expenditure data, converted to constant purchasing power parity (PPP) dollars using World Bank deflators, to construct capital stock series for three of the four major sectors performing R&D reported in MSTI: private-sector, government, and higher-education; we omit the private non-profit sector (the final and smallest performing sector in the MSTI data) because those data are too intermittently reported to construct reliable capital stocks for many of the countries in our panel.[6] Our one departure from Coe et al. (2009) is the assumed depreciation rate for R&D capital, δ. Whereas they assume $\delta = 5\%$, our baseline uses $\delta = 16\%$—the rate the Bureau of Economic Analysis applies to government R&D, higher-education R&D, and scientific R&D services, and the rate adopted by Fieldhouse and Mertens (2026a) for government R&D capital. Since government and higher education are two of the three performing sectors in our panel, 16% is a natural rate for these stocks; it also lies within the range the BEA applies to business R&D across industries, so we use it uniformly across all three sectors. That said, we use the 5% rate of Coe et al. (2009) for robustness checks on this assumption.[7] We also construct a "total" R&D capital series for each country by summing the constructed private-sector, government, and higher education R&D capital stocks, again omitting the private non-profit sector. The MSTI historical R&D expenditure data go back as far as 1963 for some countries, but coverage is inconsistent before the 1990s for most countries, which limits our panel of R&D capital stocks to the same 24 advanced economies in the Coe et al. (2009) database (all current OECD members).

Lastly, we construct a panel of total R&D personnel across all sectors, based on full-

[6]Sparse data reporting for the private non-profit sector even affects major G7 economies, e.g., Germany (reported in only 12 of 43 years) and Italy (no data before 2002).

[7]Estimated responses of R&D capital and returns to government R&D are sensitive to the assumed depreciation rate (Fieldhouse and Mertens, 2026a); Appendices A and B report the $\delta = 5\%$ construction and the corresponding results.

time equivalent (FTE) employment, from the OECD. Our panel consists of 14 countries for which data are largely consistently available over 1978–2023, and any missing observations are linearly interpolated; see Appendix A for details.[8] We focus on countries excluding the U.S. to isolate international spillovers in R&D intensity from the domestic U.S. response; the two notable omissions from the remaining countries are the United Kingdom and France, whose R&D personnel series are unavailable or inconsistent over this period.

3 Empirical Strategy

We extend the time-series narrative identification strategy of Fieldhouse and Mertens (2026a) to an international panel local projections framework to estimate the causal effects of U.S. public R&D on international productivity growth and innovation spillovers. Fieldhouse and Mertens (2026a) identify a strong causal link between government-funded R&D and U.S. business-sector productivity growth, using exogenous changes in federal R&D appropriations from a companion narrative analysis (Fieldhouse and Mertens, 2023). There are two overarching endogeneity concerns relating future TFP growth to present choices about federal R&D funding: 1) a short-run concern about policy endogeneity, e.g., policymakers increasing R&D spending in response to a recession or oil shock that will decrease measured TFP; and 2) a longer-run endogeneity concern about policymakers responding to news about promising new technologies correlated with future productivity growth, e.g., by increasing R&D appropriations today. These endogeneity concerns look different in an international panel context, but they remain the two sources of endogeneity our identification strategy seeks to guard against.

To address the potential threat of policy endogeneity, the Fieldhouse and Mertens (2023) companion narrative analysis examines all significant annual changes in federal R&D appropriations for five major agencies over 1947–2019 and classifies a subset as conditionally exogenous to current macroeconomic conditions (and the rest as endogenous to those conditions).[9] Fieldhouse and Mertens (2026a) aggregate these agency-specific exogenous policy changes into 'defense' and 'nondefense' R&D appropriations

[8] The 14 countries by ISO3 codes are as follows: AUT, BEL, CAN, DEU, DNK, ESP, FIN, IRL, ITA, JPN, NLD, NOR, PRT, and SWE. We use national R&D personnel totals, so differences in how the historical (1963–1980) and modern (1981–2023) series classify personnel are immaterial; see Appendix A.2 for details.

[9] The five major R&D funding agencies are the Department of Defense, Department of Energy, National Aeronautics and Space Administration, National Institutes of Health, and National Science Foundation.

shocks and use them as external instruments to estimate TFP responses and domestic rates of return to government R&D. The narrative classification is intended to remove any policy variation motivated by short-run cyclical conditions, since those conditions might independently influence short- or medium-run TFP growth and thus bias estimates of the returns to government-funded R&D. Policy endogeneity is much less of a concern in our international panel of countries (which excludes the U.S.) because U.S. R&D policy periodically responds to domestic cyclical conditions but does not appear to respond endogenously to foreign cyclical conditions (Fieldhouse and Mertens, 2023).

A residual identification concern, distinct from U.S. R&D policy endogeneity, is that foreign TFP growth may contain a cyclical component correlated with global business conditions through demand channels. This cyclical variation could bias short- to medium-run impulse responses if unabsorbed, since global recessions that coincide with periods of elevated U.S. R&D spending and share a common geopolitical origin may independently depress foreign TFP through channels unrelated to technology spillovers. Fieldhouse and Mertens (2026a) address the analogous domestic concern by using utilization-adjusted TFP measures (Fernald, 2012), which remove cyclical factor utilization fluctuations from the U.S. TFP residual, and by controlling for lagged utilization rates. As noted in Section 2.1, such utilization-adjusted TFP series and utilization rates are not available for most countries in our international panel. We therefore include country-specific real GDP per capita growth as lagged controls in the panel local projections, which should absorb any short-run demand-driven TFP fluctuations, thus isolating the longer-run technology spillovers.

To address longer-run potential sources of endogeneity, Fieldhouse and Mertens (2026a) embed the R&D appropriations shocks in regression models that also use lagged macroeconomic controls to capture predictable variation in future U.S. TFP growth from recent R&D investments and technology news. The benchmark set of controls includes the government R&D capital stock, private R&D capital stock, (log) cumulated changes in real federal R&D appropriations (capturing all changes in R&D appropriations for these agencies, not just those classified as exogenous), U.S. stock returns for R&D-intensive sectors, the Ramey and Zubairy (2018) defense news shocks, and the utilization-adjusted TFP and factor utilization rates of Fernald (2012). By contrast, longer-run endogeneity concerns are not attenuated in our international panel setting, and we therefore adopt a similar set of controls.

Rather than including all these lagged controls in every panel regression, we first orthogonalize the narrative U.S. R&D appropriations shocks to the baseline Fieldhouse and Mertens (2026a) lagged controls and use the residuals as external instruments. We estimate three separate regressions for the 'nondefense' (n) and 'defense' (d) shocks as well as a 'combined' (c) shock series, all on quarterly data over 1947Q1–2019Q4. The nondefense and defense shocks follow Fieldhouse and Mertens (2026a) exactly: the exogenous changes in R&D appropriations are scaled to the one-year lagged government nondefense and defense R&D capital stocks, respectively. The combined shock is constructed separately as the sum of the nondefense and defense appropriations innovations scaled together relative to the one-year lagged total federal R&D capital stock, then orthogonalized using the same lag-augmented regression structure; it captures the aggregate federal R&D impulse without imposing a decomposition by nondefense/defense function.[10] The residual instruments are a slightly narrower object than Fieldhouse and Mertens's (2026a) shocks, additionally partialling out variation predictable from U.S. macroeconomic conditions, but serve the same identification role in our panel setting. The construction of these aggregated shocks also helps mitigate the concern that contemporaneous global shocks—geopolitical shifts, synchronized scientific breakthroughs, or technology news—could simultaneously drive U.S. R&D appropriations and independently shape future foreign TFP. The narrative appropriations shocks are orthogonalized to lagged macroeconomic controls at quarterly frequency before aggregation, so they reflect discrete legislative decisions tied to specific quarters: global technology news shocks materializing in subsequent quarters of the same year would not contaminate them.

Using these orthogonalized residuals as our shock measure, our baseline regression framework for estimating the response of log TFP (or other outcome variable of interest)

[10]The Fieldhouse and Mertens (2023) exogenous R&D appropriations shocks and U.S.-specific control variables are all from the Fieldhouse and Mertens (2026a) replication files with the exception of updated vintages of utilization-adjusted TFP and utilization rates (from the Federal Reserve Bank of San Francisco). For the combined shock, the numerator is the explicit sum of the nondefense and defense appropriations innovations rather than Fieldhouse and Mertens (2026a)'s aggregate series, which differs in years where the Department of Energy threshold for a "significant" appropriations change is met by the total but not by each component separately.

$y_{i,t}$ in country i over horizon h is:

$$y_{i,t+h} = \beta^0_{k,h}\hat{z}^k_t + \sum_{j=1}^{p_z} \beta^j_{k,h}\hat{z}^k_{t-j} + \sum_{j=1}^{p_x} \zeta^{j\prime}_h \boldsymbol{x}_{i,t-j} + \gamma_i + \varepsilon_{i,t+h} \tag{1}$$

where $\hat{z}^k_t$ is the residual federal R&D appropriations shock for one category of spending, with $k \in \{n, d, c\}$ indexing the nondefense, defense, and combined shock series, and $\beta^0_{k,h}$ captures the impulse response of shock k at horizon h. Because the R&D appropriations shocks have long-lasting effects on U.S. productivity growth, we include five annual lags of the orthogonalized shock series $\hat{z}^k$. We also include two lags of a country-specific vector of controls, $\boldsymbol{x}_{i,t}$, which always includes country i's log TFP (in levels) and real GDP per capita growth as well as utilization-adjusted TFP for the U.S. business sector (Fernald, 2012). Equation (1) is estimated on a pooled panel of countries, either the full sample or country-group subsamples; γ_i is a country fixed effect and standard errors are clustered at the country level throughout, allowing for arbitrary within-country serial correlation of the regression residuals across time.[11] Given the gradual response of U.S. TFP to public R&D in Fieldhouse and Mertens (2026a), our baseline specification iteratively estimates (1) over a 15-year horizon: $h = 0, 1, ..., 15$. Equation (1) is identified up to a scale factor, and all results reported below are scaled to imply a shock equal to one percentage point of the lagged nondefense, defense, or total federal R&D capital stock, depending on the shock, i.e., a one-unit increase in $\hat{z}^k_t$.[12]

4 Effects of U.S. Public R&D on Foreign Productivity

This section reports our main results for international TFP spillovers from U.S. public R&D appropriations shocks. Section 4.1 documents statistically significant TFP spillovers that build gradually over horizons of 5–15 years, then studies heterogeneity across and within country groups to identify where those spillovers are concentrated. Section 4.2 documents that our headline results of significant international TFP spillovers are robust to individual country exclusions from the panel, and that results are also relatively sta-

[11]Two additional inference concerns are left for future work: these standard errors do not account for estimation uncertainty in the generated instruments, and country clustering does not absorb contemporaneous cross-country residual correlation.

[12]Note that Fieldhouse and Mertens (2026a) instead report results scaled to induce a peak 1 percent response of the U.S. government R&D capital stock.

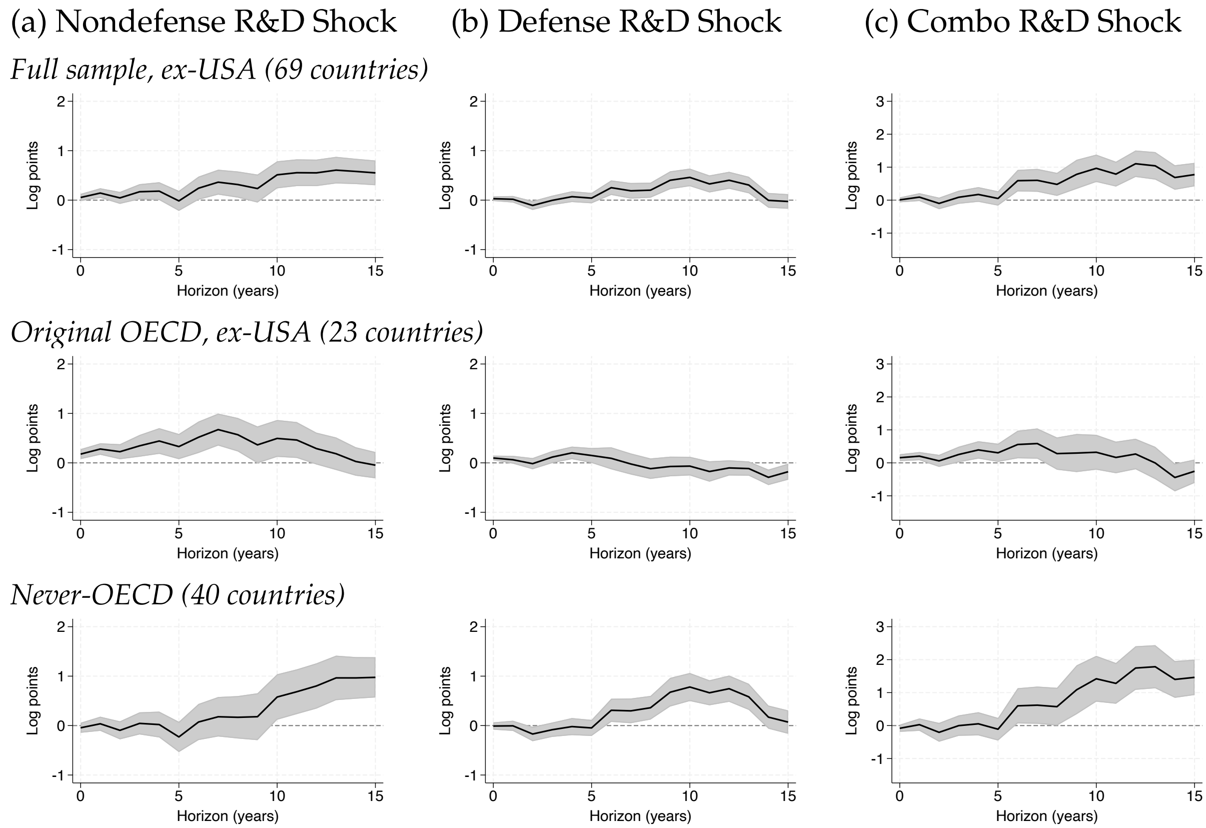
ble over rolling subsample windows within 1980–2019. Section 4.3 studies the role of country-varying TFP measurement for our headline results, specifically the treatment of intellectual property as a factor input to PWT TFP measurement across SNA vintages. Lastly, Section 4.4 compares U.S. TFP responses across the utilization-adjusted and unadjusted Fernald (2012) business-sector measures and the PWT economy-wide series—the U.S. being the only country for which all three are available over our full estimation sample—which suggests that our PWT-based headline estimates are, if anything, conservative. Throughout, we find that international TFP spillovers are concentrated in countries further from the U.S. technological frontier and with smaller domestic R&D capital stocks, pointing against a Cohen and Levinthal (1990) absorptive capacity mechanism and toward other channels such as capital-embodied trade, leapfrogging in economies with less lock-in of legacy technologies, and imitation of frontier technologies as a mechanism for catch-up growth.

4.1 Main Productivity Spillover Results

Figure 1 reports impulse response functions estimated from (1) of foreign PWT TFP to the federal nondefense (left panel), defense (middle panel), and combined (right panel) U.S. R&D appropriations shocks across three sample subgroups: the top row for the full panel excluding the U.S., the middle row for the 'original' OECD countries excluding the U.S., and the bottom row for the non-OECD countries. Shaded bands are 95 percent confidence intervals based on country-clustered standard errors. For our headline results we focus on the original OECD countries rather than the full modern OECD panel, in part because of more consistent TFP measurement across the former: Chile, Colombia, and Costa Rica (later entrants) used SNA 1968 vintage accounts for much of our sample; see Appendix Table A.1. Appendix Figure B.1 reports analogous results for the modern OECD (both with and without the U.S.).

For the full international sample, U.S. nondefense R&D shocks cause a delayed but eventually highly significant response of foreign TFP, which rises roughly 0.5 percent by the end of the 15-year horizon. This mirrors the finding in Fieldhouse and Mertens (2026a) that nondefense R&D appropriations shocks cause a highly significant but delayed response of U.S. business-sector TFP, which becomes persistently significant after roughly eight years—slightly faster than the foreign response in Figure 1(a), which builds more

Figure 1: Foreign TFP Response to an Increase in U.S. R&D Appropriations



Notes: Estimates based on (1) using the narrative measure of changes in federal nondefense (left panel), defense (middle panel), and combined (right panel) U.S. R&D appropriations shocks. Shaded bands are 95 percent confidence intervals based on country-clustered standard errors. Impulses scaled to imply a 1 percent increase in U.S. R&D appropriations as a share of the respective federal R&D capital stock. *Sample*: 1980–2019.

gradually. The TFP response to defense R&D shocks becomes statistically significant after only six years for the full international panel, and is smaller and less persistent than that of the nondefense shock; the response peaks at roughly 0.4 percent at 10 years and decays back to zero after 13 years. The response to the combined shock is substantially larger than either component shock, building gradually to a peak of roughly 1.0 percent around 12 to 13 years and remaining elevated through the 15-year horizon. Because the combined shock is normalized to total federal R&D capital—averaging approximately 1.6 times the nondefense stock and 2.8 times the defense stock over our sample—a one-unit combined shock represents a larger dollar impulse than either component shock individually; the proportionally larger combined response is thus to be expected.

TFP responses for the original OECD countries excluding the U.S. (middle row of Figure 1) differ markedly from the full-panel responses above. We see a more immediate but transitory TFP response to the nondefense R&D shock, which peaks at roughly 0.7 percent after seven years for the original OECD countries. Conversely, the TFP response to the defense R&D shock is essentially flat for the first 10 years and then turns modestly negative at the longest horizons, mirroring Fieldhouse and Mertens's (2026a) negative long-run TFP point estimates for the defense R&D shock in their baseline specification, albeit more precisely estimated here in the panel of original OECD countries. The TFP response to the combined shock peaks at roughly 0.5 percent after six to seven years—similar in timing and broadly comparable in magnitude to the nondefense response—before gradually declining and becoming statistically indistinguishable from zero at longer horizons. Appendix Figure B.1 shows that the TFP impulse responses for the full modern OECD and modern OECD excluding the U.S. are virtually identical to those for the original OECD countries excluding the U.S., despite more varied TFP measurement regimes among the later OECD entrants.

The bottom row of Figure 1 shows that the delayed, significant TFP response for the full international sample is entirely driven by the non-OECD countries, as would be expected given the more transitory OECD responses that fade to zero at long horizons (middle row above). The nondefense R&D shock has no economically or statistically significant effect for the first nine years, but then TFP in the non-OECD countries rises sharply with a peak increase of roughly 1.0 percent at the end of the 15-year horizon. Their TFP response to defense R&D shocks is less delayed and more transitory, with a response that is significant at the 95 percent level between horizons of 6 and 14 years and a peak response of roughly 0.8 percent after 10 years. The response to the combined shock is qualitatively similar to the nondefense response in timing—slow buildup over the full horizon—but more precisely estimated and somewhat attenuated at longer horizons; this attenuation reflects the contribution of the defense response, which decays after its mid-horizon peak.

The response to the combined shock is substantially larger than the nondefense response throughout, peaking at roughly 1.8 percent around 12 to 13 years and remaining above 1.4 percent at the 15-year horizon. The timing of the response—a gradual buildup over the full horizon—broadly mirrors the nondefense response, whereas the larger response to the combined shock reflects its broader normalization, consistent with the full-sample results.

The non-OECD countries driving the significant aggregate international productivity responses in Figure 1 differ from the OECD countries in myriad ways. Most pertinent to international productivity spillovers from the U.S., they tend to be less affluent and further from the technological frontier (as measured by real GDP per capita relative to the U.S., or by relative TFP levels) and less R&D intensive. They are also more likely to use SNA 1968 or 1993 vintages and hence to have intellectual property investment reflected in the PWT TFP residual rather than netted out via the aggregate capital stock. To analyze heterogeneity in TFP responses across key country characteristics, we next re-estimate our baseline TFP specification from (1) splitting the original OECD (excluding the U.S.) and non-OECD subgroups by three characteristics—real GDP per capita, R&D capital, and population—each measured as a 1977–79 pre-sample log average relative to the U.S. and split at the within-subgroup median. Table A.2 reports below-median indicators for each split by country and subgroup membership.

Figure 2 replicates Figure 1 for the original OECD countries excluding the U.S. (top row) and non-OECD countries (bottom row), splitting each subgroup by relative GDP per capita as a proxy for distance from the U.S. technological frontier. In the top row, the impulse responses are qualitatively similar for the above- and below-median OECD countries; the most notable difference is that the IRFs are more precisely estimated for the wealthier (above-median) OECD countries, which may reflect differences in macroeconomic volatility or data quality across OECD national income levels. The bottom row, by contrast, shows clearer within-cohort heterogeneity for the non-OECD countries: the TFP response is modestly stronger and more persistent at longer horizons for the countries further from the U.S. productivity frontier (below-median).

Taken together, Figure 2 provides evidence consistent with the possibility of technological "leapfrogging" (Brezis et al., 1993)—the hypothesis that countries further from the productivity frontier catch up more quickly because they can adopt mature technologies without sunk costs in earlier generations, e.g., the adoption of mobile phones in sub-Saharan Africa without first building landlines.

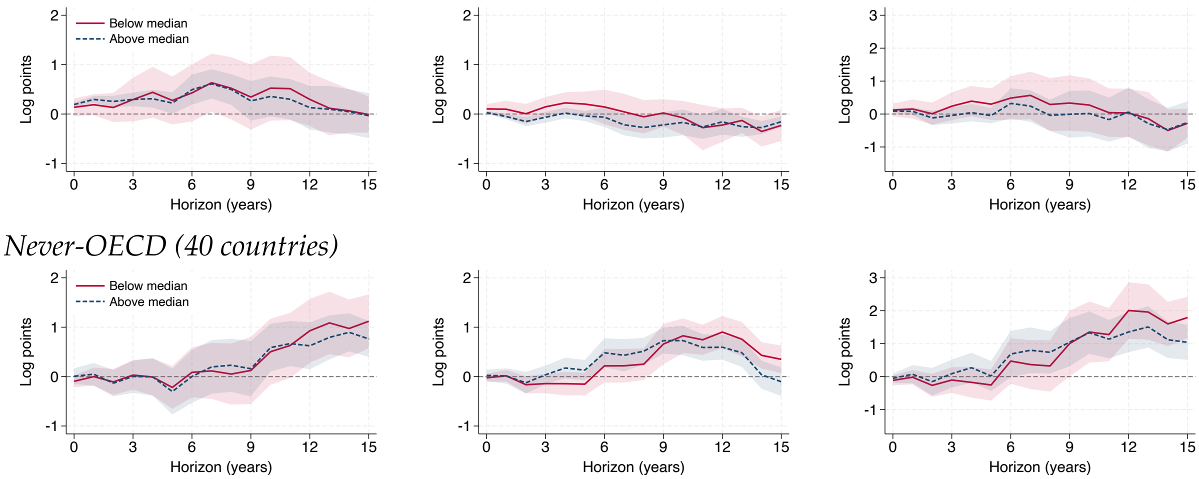
Whereas Figure 2 split countries by distance to the U.S. technological frontier, Figure 3 replicates the exercise splitting each subgroup by IPP and R&D capital stock (PWT

Figure 2: Foreign TFP Response to an Increase in U.S. R&D Appropriations: Heterogeneous Responses by Distance from the Productivity Frontier

(a) Nondefense R&D Shock (b) Defense R&D Shock (c) Combo R&D Shock

Original OECD, ex-USA (23 countries)

Never-OECD (40 countries)



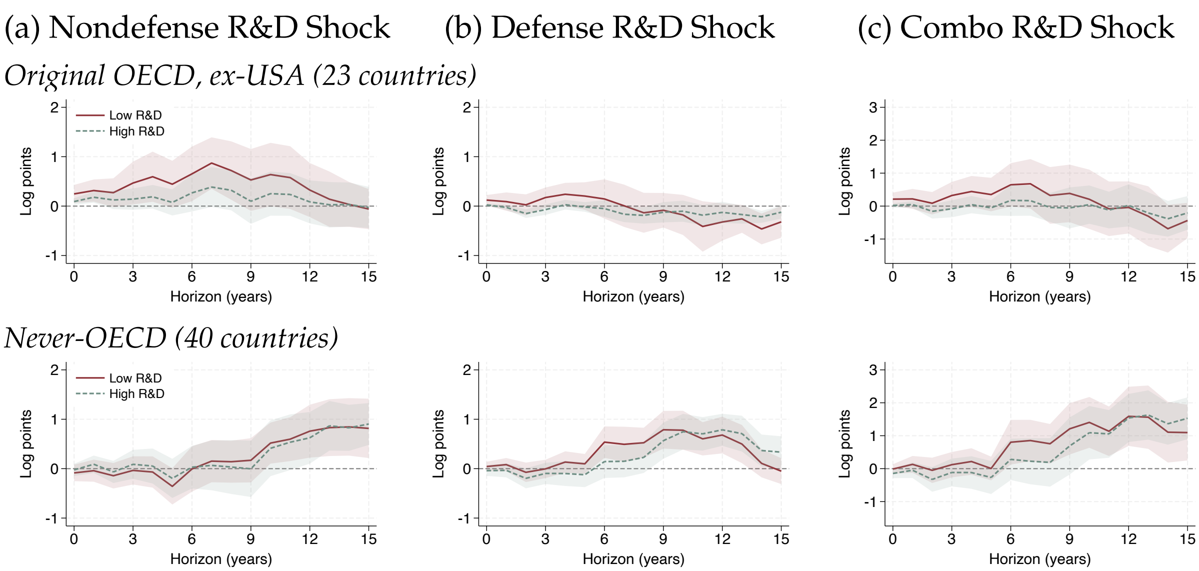
Notes: Estimates based on (1) using the narrative measure of changes in federal nondefense (left panel), defense (middle panel), and combined (right panel) U.S. R&D appropriations shocks. Shaded bands are 95 percent confidence intervals based on country-clustered standard errors. Impulses scaled to imply a 1 percent increase in U.S. R&D appropriations as a share of the respective federal R&D capital stock. Solid cranberry (dashed navy) lines correspond to countries below (above) the sample median of log GDP per capita relative to the United States, averaged over 1977–1979 within that country group; below-median countries lie further from the productivity frontier. *Sample*: 1980–2019.

rnn_other), providing a more direct proxy for countries' capacity to absorb R&D-driven technology spillovers from abroad.[13] Most strikingly, the TFP response of the original OECD countries excluding the U.S. to the nondefense R&D appropriations shock is driven by countries with a below-median initial R&D capital stock. For the above-median countries, the response is neither statistically significant nor far from zero at medium or long-run horizons. By contrast, the below-median OECD response is statistically significant and peaks at roughly 0.8 percent—somewhat higher than the peak response for all original OECD countries in Figure 1—though with wider confidence intervals for the below-median estimates at longer horizons. Overall, the heterogeneity across OECD

[13]Because PWT's *rnn_other* captures R&D investment as part of the aggregate capital stock only for countries reporting under SNA 2008, this proxy more cleanly captures the R&D capital stock for OECD countries (almost all of which were on SNA 2008 by the mid-1990s) than for non-OECD countries (many of which remained on SNA 1968 or 1993, where R&D investment instead enters the PWT TFP residual). We revisit this measurement question explicitly in Section 4.3.

countries and the comparison between OECD and non-OECD countries point toward mechanisms that are not consistent with the "absorptive capacity" mechanism where countries with more domestic R&D capacity can better adopt new technologies (Cohen and Levinthal, 1990; Howitt, 2000).

Figure 3: Foreign TFP Response to an Increase in U.S. R&D Appropriations: Heterogeneous Responses by Existing R&D Capital Stock



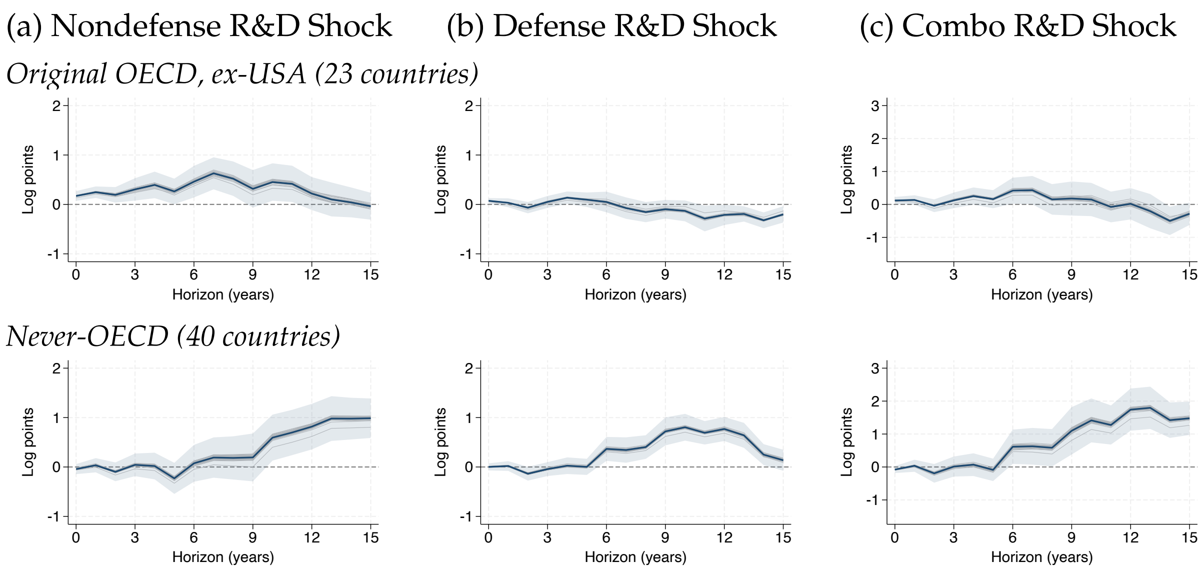
Notes: Estimates based on (1) using the narrative measure of changes in federal nondefense (left panel), defense (middle panel), and combined (right panel) U.S. R&D appropriations shocks. Shaded bands are 95 percent confidence intervals based on country-clustered standard errors. Impulses scaled to imply a 1 percent increase in U.S. R&D appropriations as a share of the respective federal R&D capital stock. Solid maroon (dashed teal) lines correspond to countries below (above) the sample median of each country's log IPP and R&D capital stock relative to the United States (PWT *rnn_other*), averaged over 1977–1979 within that country group. *Sample*: 1980–2019.

The bottom row of Figure 3 shows a more nuanced pattern: peak TFP responses of the non-OECD countries are more similar in magnitude across the R&D capital stock split than was the case for OECD countries above. However, the response becomes significantly different from zero earlier, with the peak also reached sooner, for the below-median R&D capital stock countries. Taken together, Figure 3 suggests that the R&D capital stock gradient governs response magnitude within the OECD cohort but appears to govern response timing within the non-OECD cohort.

As a final test for heterogeneous responses within the OECD and non-OECD country groups, Appendix Figure B.2 reports TFP responses for each subsample split by population.[14] The key takeaway is that the TFP spillovers for OECD countries are entirely concentrated in relatively less populous countries, whereas the TFP spillovers in non-OECD countries are qualitatively similar but slightly stronger in more populous countries.

4.2 Subsample Stability and Robustness

Figure 4: Foreign TFP Response to an Increase in U.S. R&D Appropriations: Robustness to Individual Country Exclusions



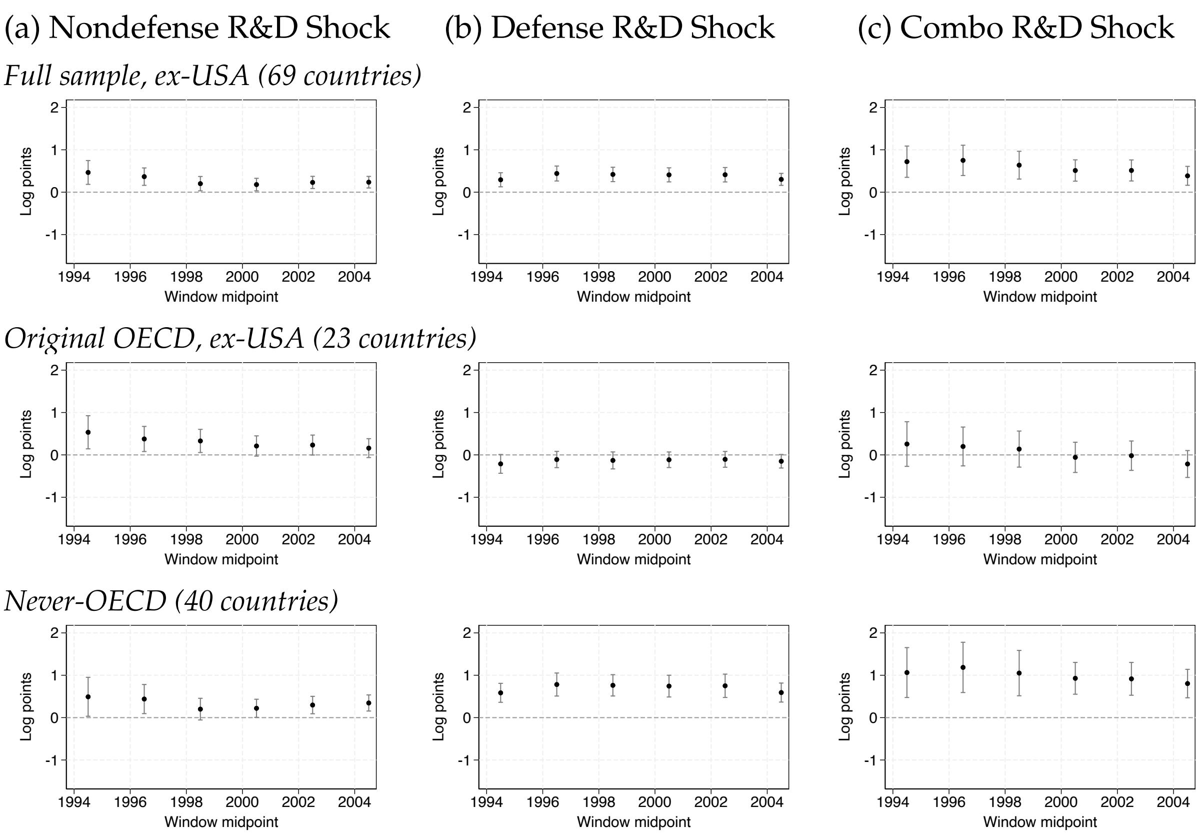
Notes: Estimates based on (1) using the narrative measure of changes in federal nondefense (left panel), defense (middle panel), and combined (right panel) U.S. R&D appropriations shocks. Grey lines show estimates after sequentially excluding each country from the sample; navy lines show the full-cohort baseline estimates. Shaded bands are 95 percent confidence intervals for the full-cohort baseline estimates based on country-clustered standard errors. Impulses scaled to imply a 1 percent increase in U.S. R&D appropriations as a share of the respective federal R&D capital stock. *Sample*: 1980–2019.

To test the sensitivity of our pooled panel estimates to the potential influence of outlier countries, Figure 4 replicates the main results for the original OECD countries excluding the U.S. (top row) and non-OECD countries (bottom row) when leaving one country out of each sample subgroup. The thicker blue lines and associated shaded 95 percent confidence bands replicate the results from Figure 1 and each finer line shows how the

[14]This split by raw population has less obvious implications from growth theory for productivity spillovers than real GDP per capita or R&D capital intensity, so we omit it from the body of the paper for brevity.

point estimates change when re-estimating (1) while dropping one OECD country (top) or non-OECD country (bottom) from the panel. The point estimates are all quite similar both qualitatively and quantitatively, indicating that the headline results are robust and no single outlier country is unduly influencing the baseline results for either country subgroup or any of the U.S. R&D appropriations shocks.

Figure 5: Foreign TFP Response to an Increase in U.S. R&D Appropriations: Stability of h=10 Estimates from 30-year Rolling Regressions



Notes: Estimates for $h = 10$ from rolling regression versions of (1) using the narrative measure of changes in federal nondefense (left panel), defense (middle panel), and combined (right panel) U.S. R&D appropriations shocks. Each rolling regression sample is 30 years with forecast horizon $h = 0, ..., 10$, staggered by two years across six windows spanning 1980–2009 to 1990–2019. The x-axis denotes the midpoint year for each sample window. Circles are point estimates and bands are 95 percent confidence intervals based on country-clustered standard errors. Impulses scaled to imply a 1 percent increase in U.S. R&D appropriations as a share of the respective federal R&D capital stock.

To study the stability of our headline results over time, we re-estimate (1) as rolling regressions with 30-year subsample windows staggered in two-year increments from 1980–

2009 through 1990–2019. Given the truncated sample window, we reduce the forecast horizon from 15 years to 10 years, leaving 20 years of effective sample per rolling window. Figure 5 reports the $h = 10$ TFP response of foreign PWT TFP to the federal nondefense (left panel), defense (middle), and combined (right) U.S. R&D appropriations shocks, for the full panel excluding the U.S. (top row), original OECD countries excluding the U.S. (middle), and non-OECD countries (bottom). The x-axis denotes the midpoint year for each sample window, i.e., estimates plotted at 1995 correspond to results for the first 1980–2009 window. Circles are the $h = 10$ point estimates for each rolling regression and bands are the associated 95 percent confidence intervals based on country-clustered standard errors.

For the full sample, the $h = 10$ impulse response estimate is always significant at the 95 percent confidence level for all six rolling regressions and across each of the three U.S. R&D appropriations shocks, suggesting that the headline finding of significant international productivity spillovers from U.S. public R&D is fairly robust. The response to the nondefense R&D shock is somewhat higher in the two earliest sample windows than in later samples, and the response to the defense R&D shock is modestly lower for the first and last windows than for the intermediate windows. Some variation in the point estimates is to be expected: the effective sample is held fixed across rolling regressions, but the identifying variation on the right-hand side changes modestly with each window. There is no evidence from the top row of Figure 5 that international spillovers from U.S. public R&D strengthened further as globalization and information technology deepened over this period—trends that might be expected to progressively facilitate such spillovers.

Within the original OECD subsample, the $h = 10$ impulse response of TFP to the nondefense R&D appropriations shock is positive and statistically significant at the 95 percent level in the first three subsample windows, with positive but marginally insignificant estimates (only significant at the 90 percent level) in the later three windows; the point estimates are nearly monotonically decreasing across the six staggered windows. This is consistent with the baseline results in Figure 1: unlike the more persistent TFP responses for the full panel or non-OECD countries, the OECD response peaked at $h = 7$ and was already receding by $h = 10$, which naturally yields more marginally significant rolling estimates—particularly as the effective sample shrinks for this smaller subgroup. The response to the defense R&D shock is never statistically significant and the point estimates are always modestly negative, again consistent with the $h = 10$ impulse response for original OECD countries in Figure 1.

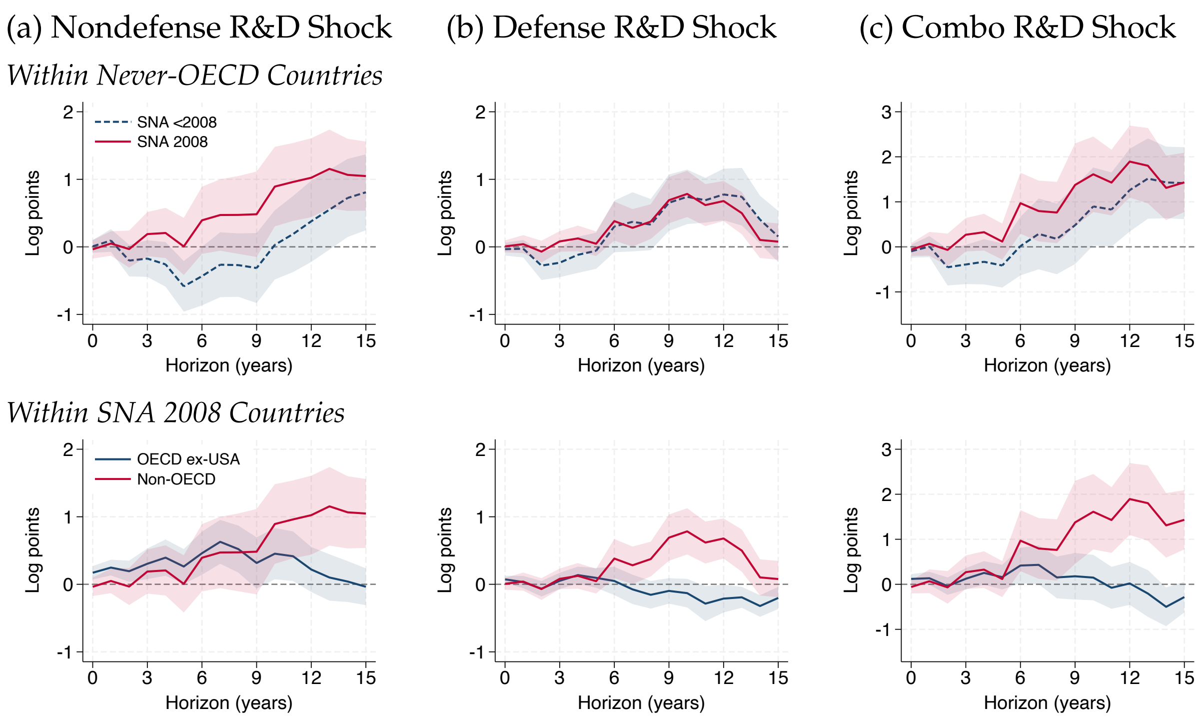
Conversely, the rolling regressions in the bottom row of Figure 5 show that the more persistent positive TFP responses for the non-OECD countries seen in Figure 1 are quite robust to the shorter staggered subsample windows. Every $h = 10$ estimate is significant at the 95 percent confidence level, save for the response to the nondefense R&D shock in the third rolling window (1984–2013). Deepening globalization and improved information technology would seem more important for facilitating productivity spillovers to non-OECD countries than to OECD countries, but as in the top row, there is no evidence that international spillovers from U.S. public R&D to non-OECD countries strengthened further over this period. The significant responses to the defense R&D and combined shocks in the full panel (top row) are also clearly driven by even larger but qualitatively similar responses in the non-OECD subsample. More broadly, Figure 5 robustly underscores the headline finding of larger and more persistent TFP spillovers from U.S. public R&D in the non-OECD countries than in the OECD countries—a finding we subject to further scrutiny in the next section, given that TFP measurement regimes vary meaningfully across these two subgroups.

4.3 TFP Measurement Robustness: SNA Vintages

We next test whether the stronger TFP response of non-OECD countries is driven by their disproportionate representation among SNA 1968 or 1993 vintage adopters, under which R&D investment is not capitalized into the aggregate capital stock and instead enters the PWT TFP residual directly—a pattern that could mechanically inflate non-OECD TFP responses through a measurement artifact rather than genuine technological diffusion.

The top row of Figure 6 splits non-OECD countries between those on SNA 2008 vintage (cranberry solid) and SNA 1968 or 1993 vintages (navy dashed). For both the nondefense and combined R&D appropriations shocks, the statistically significant TFP response is more immediate and larger for the SNA 2008 countries than for those using the earlier vintages—direct evidence against the measurement hypothesis that TFP mismeasurement is driving the headline non-OECD results. The TFP responses to the defense R&D shock are qualitatively similar across both non-OECD subgroups, though the response becomes statistically significant at shorter horizons for the SNA 2008 countries and is slightly less persistent.

Figure 6: Foreign TFP Response to an Increase in U.S. R&D Appropriations: Stability of Estimates Across and Within SNA Vintages



Notes: Estimates based on (1) using the narrative measure of changes in federal nondefense (left panel), defense (middle panel), and combined (right panel) U.S. R&D appropriations shocks. Shaded bands are 95 percent confidence intervals based on country-clustered standard errors. Impulses scaled to imply a 1 percent increase in U.S. R&D appropriations as a share of the respective federal R&D capital stock. The top row splits non-OECD countries between those on SNA 2008 vintage (cranberry solid) and SNA 1968 or 1993 vintages (navy dashed). The bottom row compares non-OECD SNA 2008 countries (cranberry solid) against original OECD excluding the U.S. (navy solid). *Sample*: 1980–2019.

The bottom row of Figure 6 provides a crosswalk between TFP responses—clean of differential measurement concerns—for original OECD economies excluding the U.S. (navy solid) and the SNA 2008 subset of non-OECD countries (cranberry solid). By construction, the cranberry lines are identical across both rows of Figure 6, and the navy solid lines replicate the headline impulse response functions for original OECD economies (excluding the U.S.) in the middle row of Figure 1. Figure 6 underscores a differential TFP response for non-OECD countries that is more delayed, larger, and more persistent than for original OECD economies—none of which is driven by R&D investment entering the

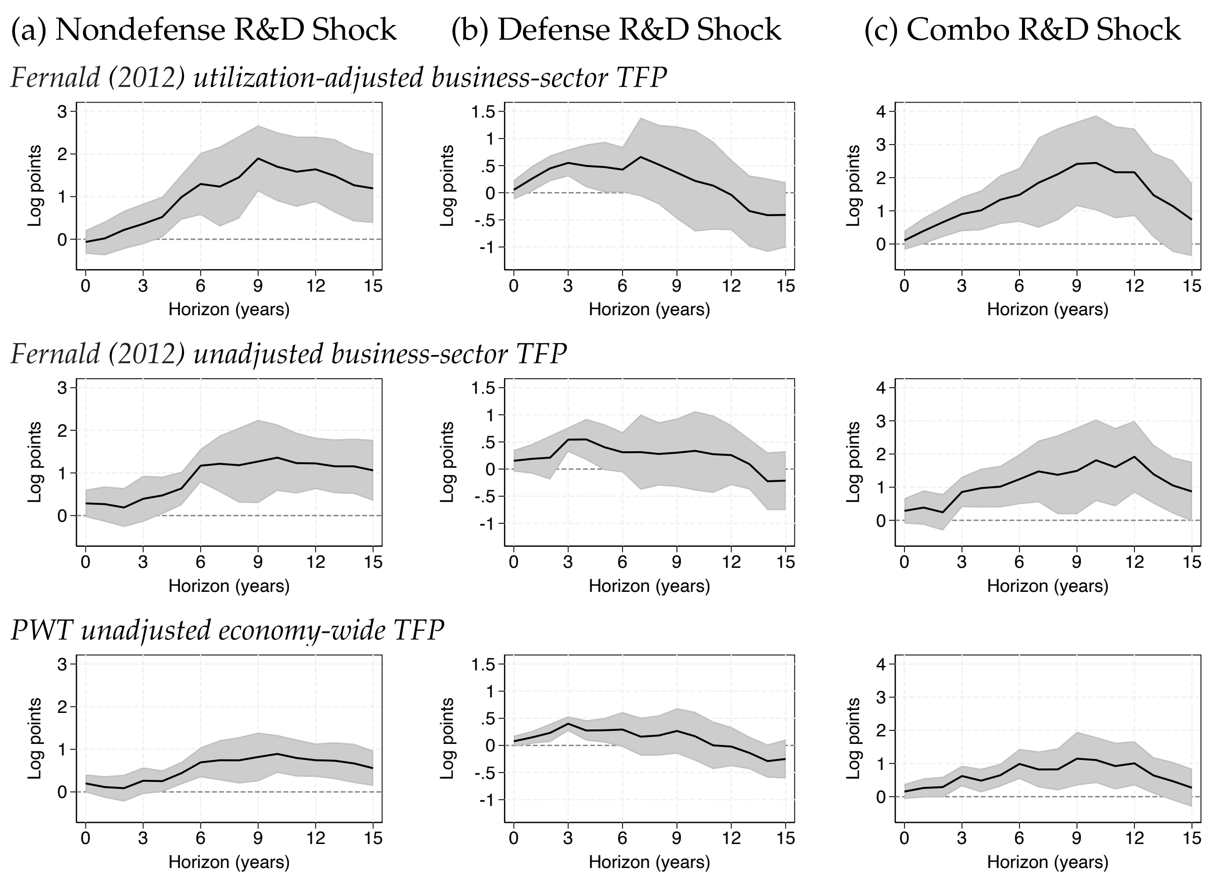
TFP residual under pre-2008 SNA conventions, as non-OECD countries operating under earlier SNA vintages are deliberately excluded from the cranberry series.

4.4 TFP Measurement Robustness: Factor Utilization Adjustments

It is harder to test how our headline results are influenced by a second concern about TFP measurement in the PWT: the lack of factor utilization adjustment. A large literature has shown that unadjusted TFP series can be misleading as measures of technology shocks and are noisier than utilization-adjusted alternatives (Basu et al., 2006; Fernald, 2012), although cyclical controls and country fixed effects may absorb some of this in a panel setting. Unfortunately, there is very little readily available utilization-adjusted TFP data for OECD countries beyond the United States and, to the best of our knowledge, none for non-OECD countries. Huo et al. (2023) construct a utilization-adjusted TFP database for 12 advanced economies, following Basu et al. (2006) in estimating utilization-adjusted TFP with unobserved variable factor utilization rates using external instruments to address endogeneity of factor inputs. However, their panel is only available through 2007, leaving just a 1980–2007 overlap with our baseline sample window and at most an 18-year effective estimation sample even when curtailing the impulse response horizon to $h = 10$—too short for reliable inference on the medium-run TFP responses of interest. Extending the HLP database over a longer sample and for additional countries represents a substantial research agenda in its own right—one constrained by the availability of granular industry-level capital data as well as suitable external instruments for identifying factor utilization rates—and is beyond the scope of this paper.

As an alternative illustration of how TFP utilization measurement influences our main results, Figure 7 reports impulse responses of U.S. TFP to the federal nondefense (left panel), defense (middle panel), and combined (right panel) U.S. R&D appropriations shocks across three types of measured TFP. The impulse responses are estimated from our baseline local projection specification (1), adapted to a time-series variant for just the United States: we drop the country fixed effect but retain two lags of the dependent (TFP) variable and U.S. real GDP per capita growth, as in our cross-country specification. We omit the Fernald (2012) utilization-adjusted TFP series that otherwise enters (1) as a lagged control, since it is the dependent variable itself in the top row and is highly correlated with the dependent TFP series in the remaining two rows (retaining it would introduce multicollinearity). The top row reports results for the Fernald (2012) utilization-adjusted TFP for the business sector, the middle row for the Fernald (2012) measure of business-sector TFP without adjustment for cyclical utilization rates of factor inputs, and the bottom row for the PWT TFP measure for the entire economy. The United States is the only country for which all three types of measured TFP are available over our full 1980–2019 estimation sample.

Figure 7: U.S. (Log) TFP Following an Increase in U.S. R&D Appropriations: Sensitivity to TFP Measurement



Notes: Estimates based on (1) using the narrative measure of changes in federal nondefense (left panel), defense (middle panel), and combined (right panel) U.S. R&D appropriations shocks. Shaded bands are 95 percent confidence intervals based on heteroskedasticity-robust standard errors. Impulses scaled to imply a 1 percent increase in U.S. R&D appropriations as a share of the respective federal R&D capital stock. *Sample*: 1980–2019.

Reassuringly, the positive and persistent U.S. TFP response to the nondefense R&D shock is qualitatively robust across all three measures, with point estimates largest for the utilization-adjusted Fernald business-sector series (top row) and attenuating toward the economy-wide PWT series (bottom row); the larger responses for the Fernald measures than for PWT TFP are to be expected, since publicly funded R&D is not capitalized with private business investment and any direct productivity effects of public capital are left in the TFP residuals. The combined shock also generates positive and broadly significant responses across all three measures, albeit less persistent than the nondefense R&D response at longer horizons. For both the nondefense and combined shocks, the utilization-adjusted Fernald series likewise yields larger point estimates than the unadjusted series, consistent with cyclical variation in factor utilization partially contaminating the unadjusted Solow residual at the time of R&D appropriations shocks and attenuating the measured productivity response. For the defense shock, all three measures yield an earlier positive response peaking near $h = 3$, but it is short-lived and the point estimates are not significantly different from zero at longer horizons, all suggesting that defense R&D appropriations do not generate lasting productivity gains—consistent with Fieldhouse and Mertens's (2026a) inconclusive findings for productivity effects of defense R&D over a longer postwar sample. Unsurprisingly, the responses to the nondefense R&D shocks seen here are less precisely estimated at longer horizons than those reported by Fieldhouse and Mertens (2026a); our baseline sample is roughly half as long as their 1948–2021 sample.

More broadly, the key takeaway from Figure 7 is that our headline results based on PWT TFP measures are, if anything, likely conservative relative to alternative estimates using utilization-adjusted or business-sector TFP measures—if those were broadly available for foreign countries.

5 Exploring Mechanisms

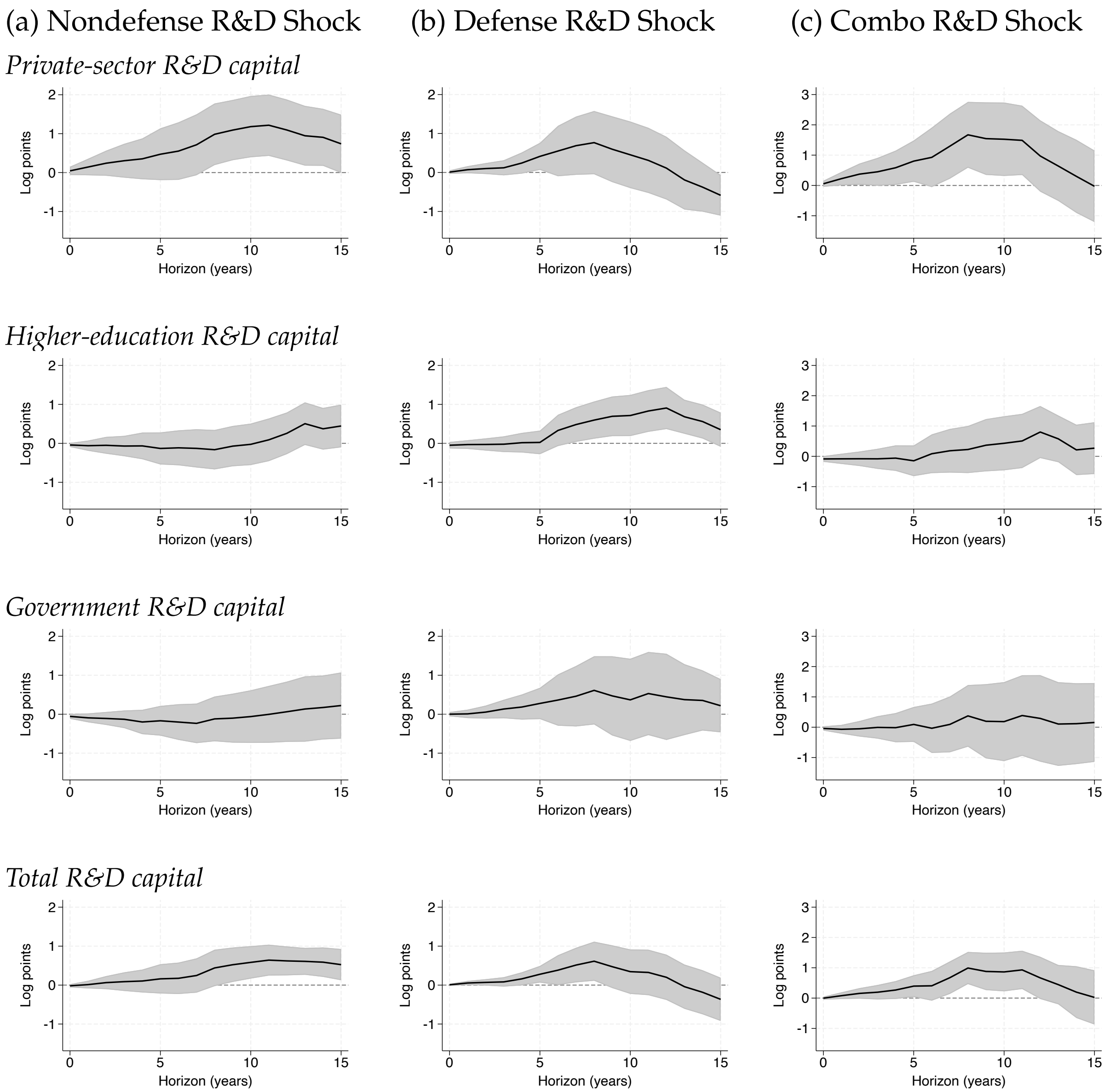
Section 4 contributes causal evidence of highly significant and persistent productivity responses abroad—particularly for non-OECD countries—in response to increases in U.S. public R&D. The heterogeneity is informative about mechanism: spillovers are larger where domestic absorptive capacity is lower in the aggregate, and concentrate in nondefense rather than defense R&D. Together, these patterns are difficult to reconcile with mechanisms that operate primarily through formal research absorption (e.g., a global extension of Cohen and Levinthal, 1990), and instead point toward channels that operate through (i) openly accessible content from frontier nondefense R&D, (ii) capital-embodied diffusion via trade (Eaton and Kortum, 2001; Caselli and Wilson, 2004; Ayerst, Ibrahim, MacKenzie, and Rachapalli, 2023), and (iii) leapfrogging by economies less anchored to currently available global technology (Brezis et al., 1993).

A further set of questions remains, and answering them requires evidence beyond the TFP responses themselves. One is first-order for innovation policy, particularly regarding international coordination: is increased U.S. public R&D simply being substituted for foreign R&D, or does foreign R&D investment also rise as a complementary input to the expanding global frontier? The second concerns measurement: in the advanced economies for which we can construct sectoral R&D capital, R&D spending is capitalized in the national accounts (SNA 2008), so any endogenous response of foreign R&D capital enters measured capital rather than the TFP residual. Any foreign R&D investment response would thus be a channel of international innovation spillovers from U.S. public R&D that the TFP estimates in Section 4 mechanically will not capture; studying this potential channel requires direct evidence on foreign R&D inputs. Section 5.1 takes up both questions using two measures of aggregate foreign R&D activity: the responses of R&D capital and R&D workers abroad to the U.S. R&D appropriations shocks.

5.1 Foreign R&D Crowd-in or Crowd-out

We estimate our baseline local projections in (1), separately rotating in the Coe et al. (2009)-style R&D capital stocks (in logs) for each performing sector on the left-hand side and into the vector of lagged controls. Figure 8 reports results for these foreign R&D capital stocks in response to the U.S. nondefense (left), defense (middle), and combined (right) R&D appropriations shocks for a panel of 23 modern OECD countries excluding the U.S. (all countries for which these stocks can be reliably constructed). In the top three rows, the capital stocks by performing sector are ordered based on their average shares of R&D capital across these sectors over 1980–2019, with private-sector R&D in the top row (67.3 percent), higher-education R&D in the second row (19.8 percent), and government R&D in the third row (12.9 percent). The bottom row reports results for the "total" R&D capital stock constructed by summing the private-sector, higher-education, and government R&D capital stocks for each country; we omit the private non-profit sector due to limited data availability, as explained in Section 2.3.

Figure 8: Foreign R&D Capital Responses to an Increase in U.S. R&D Appropriations



Notes: Estimates based on (1) using the narrative measure of changes in federal nondefense (left column), defense (middle column), and combined (right column) U.S. R&D appropriations shocks. Each row reports the response of a different measure of foreign R&D capital: the amounts performed by the private-sector (top row), higher-education (second row), and government (third row) sectors, and their total (bottom row). The total is the sum of the three sector-specific capital stocks, excluding the small private non-profit sector because of data limitations. Foreign R&D capital stocks are constructed from OECD MSTI R&D data, following Coe et al. (2009) and using a 16 percent depreciation rate; see Data Appendix for details. Shaded bands are 95 percent confidence intervals based on country-clustered standard errors. Impulses scaled to imply a 1 percent increase in U.S. R&D appropriations as a share of the respective federal R&D capital stock. *Sample*: 23 modern OECD economies (CHH09 R&D-capital panel, ex-USA), 1980–2019.

The top row of Figure 8 shows that the response of foreign private-sector R&D capital to the nondefense shock rises gradually, becoming statistically significant after roughly eight years and peaking at about 1.2 percent around eleven years before receding at longer horizons. The response to the defense shock is weaker: it is statistically significant only around year five, and although the point estimate peaks near 0.8 percent at eight years, this peak is not statistically distinguishable from zero, with the response attenuating and drifting marginally negative at longer horizons. The response of private-sector R&D capital to the combined U.S. R&D shock rises more strongly than either component alone, peaking at roughly 1.7 percent after eight years before reverting toward zero at longer horizons. These results provide evidence that increased U.S. public R&D, particularly nondefense R&D, crowds in foreign private-sector R&D investment, contrary to the hypothesis of cross-border substitution of research inputs.

The second row shows that the response of foreign higher-education R&D capital to the U.S. nondefense shock is fairly flat and imprecisely estimated for the first ten years, then gradually rises over the last five years, though it never quite reaches statistical significance. By contrast, the U.S. defense R&D shock crowds in foreign higher-education R&D capital: the response is statistically significant starting around year seven, peaking at roughly 0.9 percent around year twelve, and is qualitatively similar to the response of private-sector R&D capital to the nondefense shock, particularly in their gradual, delayed timing.

The third row of Figure 8 shows that there is no significant response of government (performed) R&D capital at any horizon for any of the three U.S. public R&D appropriations shocks. The point estimates are modestly negative at shorter horizons for the nondefense shock, modestly positive for the defense shock, and fairly flat for the combined shock. There is no evidence that increased U.S. public R&D spending crowds out R&D performed by foreign governments, at least among OECD peer countries.

The responses of the total R&D capital stock to the U.S. R&D appropriations shocks in the bottom row underscore that, in the aggregate, increased U.S. public R&D crowds in foreign R&D. The U.S. nondefense R&D shock causes a gradual rise in total foreign R&D capital, which becomes significant after about nine years and peaks at roughly 0.6 percent. Foreign R&D capital sees a less delayed and less persistent response to the U.S. defense R&D shock, also peaking at around 0.6 percent after eight years before tapering off and losing statistical significance. The combined shock has a gradual, hump-shaped positive effect that is statistically significant from around year five through year eleven. There is no evidence in Figure 8 of U.S. public R&D crowding out total foreign R&D across any horizon or shock. More broadly, with the sole exception of the $h = 15$ response of private-sector R&D to the defense shock, there is no evidence of U.S. public R&D crowding out foreign R&D performed by any of the three major performing sectors at any horizon or shock.

As a robustness check, Appendix Figure B.3 replicates the results from Figure 8 using alternative capital stock series constructed (otherwise identically) under a 5 percent depreciation rate instead of a 16 percent depreciation rate; we choose 5 percent because it is substantially lower than our baseline and matches the depreciation rate assumed in Coe et al. (2009), whose methodology we otherwise follow in constructing these foreign R&D capital stocks. This exercise underscores that the headline results reported here are robust to alternative assumptions about how quickly knowledge capital depreciates: under either depreciation rate, U.S. public R&D significantly crowds in foreign R&D capital, particularly for private-sector R&D in response to the nondefense shock and higher-education R&D in response to the defense shock.

Taken together, the responses in Figure 8 rule out the displacement channels that motivate our analysis: increases in U.S. public R&D crowd *in* foreign R&D rather than crowding it out, with no indication that peer foreign countries free-ride by scaling back their own R&D investment or that the U.S. bids away scarce research inputs. In particular, U.S. nondefense R&D gradually and persistently draws in foreign private-sector R&D (the performing sector most closely tied to measured productivity), consistent with our TFP findings of much stronger international diffusion from U.S. nondefense than defense R&D. Because nondefense R&D largely funds fundamental research, its benefits would be expected to diffuse to foreign innovation only slowly; both the foreign TFP responses in Figure 1 and the private-sector R&D responses in Figure 8 share this gradual, delayed timing. Fieldhouse and Mertens (2026a) also find that U.S. public and private R&D are complements, though they measure R&D by *funding* sector (as in the NIPAs) rather than *performing* sector (as we do here), so the two are not directly comparable: a dollar increase in defense R&D capital crowds in roughly 85 cents of U.S. private R&D capital, whereas a dollar increase in nondefense R&D capital crowds in 20 cents.[15] Despite these measurement differences, the two sets of results together suggest a stronger response of domestic than foreign private-sector R&D to U.S. defense R&D, consistent with less international diffusion of defense work owing to national-security prerogatives.

For this group of OECD countries (all of which use SNA 2008), the measured capital-service contribution of this endogenous crowd-in of foreign R&D performed by the higher education and private sectors is netted out of the TFP residual; as discussed in Section 2.1, however, broader productivity effects remain in the impulse responses reported in Section 4.1, which are therefore not purely direct technology spillovers. Indeed, the endogenous response of foreign R&D capital is evidence of broader indirect international innovation spillovers from U.S. public R&D to other OECD countries. Moreover, higher foreign R&D investment should raise innovation and output, even if not captured by measured TFP. This broader output channel is illustrated in Appendix Figure B.4, which overlays the TFP responses from Figure 1 (dashed navy lines) on the responses of real GDP per capita for the same country groups and shocks (solid black lines with shaded bands). For the OECD economies, the output response runs well above the more modest TFP response, as would be expected if induced R&D capital is capitalized out of the residual under SNA 2008. We interpret this, however, as corroboration rather than a clean test of the measurement channel, since the output-TFP gap reflects all factor accumulation and is sizable even for the catching-up non-OECD economies, where R&D mostly remains in the residual. The sharpest test of the measurement effect is instead the within-non-OECD comparison of countries using SNA 2008 versus earlier vintages in Section 4.3.

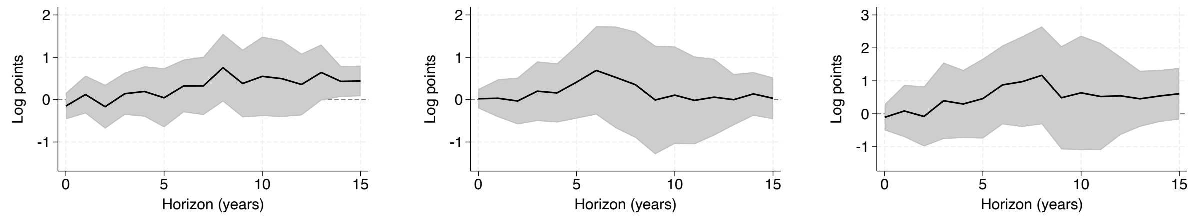
Lastly, Figure 9 reports estimated responses of R&D researchers (full-time equivalent employment) constructed from the OECD MSTI database. The responses are largely imprecise, but the point estimates are generally positive and gradually rise over longer horizons for the nondefense and combined shocks, following dynamics similar to the responses of private-sector R&D capital; the nondefense response reaches statistical significance at the longest horizons. While broadly imprecise, there is no evidence in Fig-

[15]When measuring R&D capital by funder, any federally funded R&D work performed by private firms (e.g., aerospace contracts from NASA) shows up as government R&D; the private R&D response to nondefense shocks in Fieldhouse and Mertens (2026a) thus understates the response of R&D performed by private firms, the measurement concept used in Figure 8. The distinction between measuring by funder or performer matters relatively more for government-funded than privately funded R&D; over our sample, more than 80 percent of total R&D performed by U.S. businesses was self-funded.

Figure 9: R&D Researcher Responses to an Increase in U.S. R&D Appropriations

(a) Nondefense R&D Shock (b) Defense R&D Shock (c) Combo R&D Shock

Subset of Original OECD, ex-USA (14 countries)



Notes: Estimates based on (1) using the narrative measure of changes in federal nondefense (left panel), defense (middle panel), and combined (right panel) U.S. R&D appropriations shocks. Shaded bands are 95 percent confidence intervals based on country-clustered standard errors. Impulses scaled to imply a 1 percent increase in U.S. R&D appropriations as a share of the respective federal R&D capital stock. R&D researcher data are from the OECD MSTI; see Data Appendix for details. *Sample*: 1980–2019.

ure 9 suggesting that the U.S. R&D appropriations shocks—which cause significant albeit transitory increases in U.S. employment of R&D researchers (Fieldhouse and Mertens, 2026a)—operate by reallocating foreign STEM talent to the U.S., as would be expected if U.S. public R&D were crowding out foreign R&D activity in the aggregate.

6 Global Social Returns to U.S. Public R&D

Economists have long believed that the social returns to R&D (meaning the aggregate economic benefits, including all spillover effects) are likely substantially higher than the private returns to R&D accruing to the performer (Griliches, 1979; Jones and Williams, 1998; Bloom et al., 2013; Jones and Summers, 2022). More recent empirical evidence suggests in turn that the social returns to public R&D are even higher than the social returns to private R&D or economy-wide R&D investments; see Fieldhouse and Mertens (2026b) for a survey. To the best of our knowledge, however, this literature has quantified aggregate social returns only within one country, itself already a considerable data and identification challenge. We contribute the first causal evidence quantifying the *global* social returns to public R&D, which are strictly higher than the *domestic* social returns, as illustrated by the significant foreign TFP responses to U.S. public R&D in Section 4.

Fieldhouse and Mertens (2026a) structurally estimate the aggregate *domestic* rates of return to U.S. government R&D using a Cobb-Douglas production function augmented for public capital, after first netting out public infrastructure capital from the TFP residual. The exogenous R&D appropriations shocks identify the dynamic causal responses of both utilization-adjusted business-sector TFP and the government R&D capital stock, and they use the System Projections on Instrumental Variables (SP-IV) methodology of Lewis and Mertens (2024) (a GMM estimator in the impulse-response space) to recover the structural relationship between these two endogenous variables over 15 years. Their estimates yield rates of return to government R&D ranging from 140 to 210 percent for the nondefense R&D appropriations shocks, depending on the specification and estimation method.[16] Their rate-of-return framework, however, cannot be extended to our international panel, due to both the lack of government R&D capital data for most countries in our sample and the variable treatment of R&D capital in the PWT TFP residual.

We therefore adopt a conservative back-of-the-envelope approach that is designed to yield a lower bound for the wedge between global and domestic returns to public R&D. Each return is a productivity gain per dollar of direct U.S. public R&D spending; because the global and domestic gains arise from the same shock, that common spending cancels in the ratio, and the wedge can be formed from the estimated TFP impulse responses alone.[17] Because the U.S. TFP response to the defense R&D shock is imprecisely estimated and roughly zero at longer horizons, both in Fieldhouse and Mertens's (2026a) 1948–2021 sample and in our 1980–2019 crosswalk (Figure 7), we focus exclusively on the wedge for public *nondefense* R&D, computed from the nondefense impulse responses in column (a) of Figure 1 for the 40 non-OECD countries, Figure 7 for the U.S., and Appendix Figure B.1 for the other 29 modern OECD countries.

The global-to-domestic spillover wedge is ψ, the ratio of the present value of global TFP gains from the U.S. nondefense R&D appropriations shock to the present value of the U.S. gains alone,

$$\psi = \frac{PV_{US} + PV_{RW}}{PV_{US}} = 1 + \frac{PV_{RW}}{PV_{US}}, \tag{2}$$

[16] They directly estimate rates of return and also calculate them from estimated production function elasticities, both of which yield a similar range of estimates.

[17] The cost that cancels is the direct federal R&D appropriations the shock measures; any endogenous responses of private or foreign public R&D enter the benefits rather than this common cost. These responses are captured in the TFP impulse responses for countries on earlier SNA vintages, where R&D is not capitalized or netted out of the residual. For SNA 2008 countries, where R&D is capitalized into the aggregate capital stock, the capital-service contribution of any induced R&D accumulation (as documented for OECD countries in Section 5) raises output beyond the measured TFP responses, reinforcing ψ as a lower bound; spillovers and excess returns are instead already embedded in those TFP responses.

where PV_{US} (PV_{RW}) are the gains to the U.S. (rest of the world excluding the U.S.). Due to data limitations, we restrict focus for the "rest of the world" to the 69 countries (excluding the U.S.) in our PWT panel; to the extent that U.S. nondefense R&D also raises productivity in countries outside our sample, this understates the global gain and makes ψ a conservative estimate. Given the striking divergence between the TFP responses for OECD and non-OECD countries, we also disaggregate PV_{RW} into PV_O (the 29 modern OECD members excluding the U.S.) and PV_N (the 40 never-OECD countries) to exploit our more granular estimates. Whereas the pooled panel regressions in (1) weight every country equally within the full panel or subgroups, here we weight each group's TFP response by that group's aggregate real GDP, w_g, so that PV_{RW} measures the dollar value of the foreign output gain. Because the non-OECD countries with the largest TFP responses account for a relatively small share of world output, this GDP weighting is conservative relative to alternatives that would place more weight on non-OECD countries, e.g., by population.

But given the degree of catch-up by non-OECD countries relative to OECD countries over our sample, our results are naturally sensitive to choices about when to define these weights. For transparency, we report results for three sets of real GDP weights calculated at roughly the start, middle, and end of our sample, taking three-year averages of years chosen to avoid major global recessions or economic crises: 1977–79 (start, preceding the Volcker disinflation and the 1980–82 recessions); 1994–96 (middle, avoiding the East Asian crisis, Russian default, and Argentine depression); and 2017–19 (end, predating the COVID-19 pandemic and global recession).

For each country group $g \in \{US, O, N\}$, define the present value of the cumulative output gain from a nondefense U.S. R&D appropriations shock as

$$PV_g \;=\; w_g \sum_{h=0}^{H} \frac{\hat{\beta}_{h,g}}{(1+r)^h}, \tag{3}$$

where $\hat{\beta}_{h,g}$ is the estimated TFP impulse response of group g at horizon h, r is the discount rate, and $H = 15$. The global-to-domestic spillover wedge of interest is thus

$$\psi = 1 + \frac{PV_O + PV_N}{PV_{US}}. \tag{4}$$

For our main results, we follow the Congressional Budget Office (CBO, 2025) in assuming a discount rate of $r = 7\%$, based on the rate of return on capital in the U.S. national income and product accounts, which it uses to discount the future benefits of federal R&D funding. A 7 percent discount rate lies toward the upper end of the rates commonly applied to public investment around the world, but the effect of r on ψ is a priori ambiguous: the non-OECD gains materialize later than the U.S. gains, which on its own would make a higher rate lower ψ, while the earlier-peaking OECD gains pull in the opposite direction.

Given the gradual build-up of the TFP responses, which remain small and often statistically insignificant at shorter horizons, we also evaluate (4) at a single longer horizon ($h = 10$ or $h = 15$), using the impulse responses at just that horizon in place of their discounted sum, in which case the discount rate drops out entirely:

$$\psi_h = 1 + \frac{w_O \hat{\beta}_{h,O} + w_N \hat{\beta}_{h,N}}{w_{US} \hat{\beta}_{h,US}}. \tag{5}$$

We evaluate (4) and (5) by weighting each group's estimated nondefense TFP response by that group's aggregate real GDP, separately for each of the three weighting windows.

Table 1: Wedge Between Global and Domestic Social Returns to U.S. Nondefense R&D

	Global-to-domestic spillover wedge, ψ			U.S. share
Weighting window	Discounted PV	$h = 10$	$h = 15$	of global gain
1977–79	1.77	1.64	1.49	56.6%
1994–96	1.91	1.81	1.84	52.3%
2017–19	2.22	2.25	2.93	45.0%

Notes: The spillover wedge is the ratio of the global to the domestic social return to a U.S. nondefense R&D appropriations shock, computed as a GDP-weighted ratio of TFP impulse responses. The U.S. response uses Fernald (2012) utilization-adjusted business-sector TFP and the foreign responses use PWT economy-wide TFP. Countries are partitioned into the United States, 29 modern OECD members excluding the United States, and 40 non-OECD economies. Weights are average annual real GDP at chained PPPs (PWT `rgdpo`) over the indicated three-year window. The discounted present value applies a 7% annual discount rate to the impulse-response path through $h = 15$; the snapshot columns evaluate the wedge at a single horizon. The U.S. share equals the inverse of the discounted-PV wedge.

Table 1 reports implied values of the wedge between global and domestic social returns to U.S. public R&D, ψ, ψ_{10}, and ψ_{15}, for each of the three weighting windows, along with the U.S. share of the discounted present value of global social returns in the final column. Unsurprisingly, the global social returns rise monotonically relative to the domestic returns when moving from earlier to later weighting windows, as the larger TFP response in the non-OECD countries receives greater relative weight. For our preferred discounted present value of the cumulative returns over 15 years, the global-to-domestic spillover wedge ranges from roughly 1.8 for the early 1977-79 GDP weighting to 2.2 for the later 2017-19 weighting, or roughly 2.0 on average. The snapshot wedges are fairly similar to the baseline results at $h = 10$, but range more widely at $h = 15$, from 1.5 to 2.9. Across all specifications and weightings, the wedges between global and domestic social returns indicate that the *global* social returns to public R&D considerably exceed the domestic returns the related literature has focused on to date. Our preferred specification discounting the cumulative TFP response over 15 years suggests that the U.S. captures only between 45 percent and 57 percent of the global benefit of its public nondefense R&D investments.

Appendix Table B.1 shows that the wedges between global and domestic social returns to U.S. public R&D reported in Table 1 are deliberately quite conservative relative to certain modeling choices and robust to others. As a baseline for comparison, the first column replicates the discounted present value wedges from Table 1. The second column of Appendix Table B.1 reports alternate results when measuring the U.S. TFP response with PWT economy-wide TFP (matching the foreign TFP measures used in the numerator of (4)) rather than the Fernald (2012) utilization-adjusted business-sector TFP measure. As expected, substituting in the more muted PWT TFP response (bottom row of Figure 7) for the larger Fernald TFP response (top row) shrinks the denominator of (4) and increases the global spillover wedges, which range from 2.4 to 3.3 instead of 1.8 to 2.2. This exercise is not meant to suggest that PWT TFP is the proper way to measure U.S. or global TFP spillovers, but rather to underscore that our headline international TFP responses would likely be larger, and the spillover wedges with them, if Fernald's (2012) utilization-adjusted business-sector TFP measure were available and used for all countries in our panel instead of the PWT TFP series.

We report results using population weights instead of real GDP weights in the third column of Appendix Table B.1. This specification again captures the potential for international TFP spillovers to raise living standards, weighting by number of residents rather than dollars of economic activity. As expected, weighting by population upweights the larger TFP response of the (more populous) non-OECD countries, thus increasing the wedge between global and domestic social returns to U.S. public R&D: ψ rises to between 4.6 and 5.1, or roughly 5 on average—about two-and-a-half times the baseline wedges in Table 1.[18] This underscores that weighting the TFP impulses by real GDP is conservative relative to alternative weighting schemes that would upweight non-OECD countries relative to OECD countries.

The last two columns of Appendix Table B.1 show that the baseline wedges are entirely robust to two other alternative modeling choices: either using the TFP response for the pooled full sample (excluding the U.S.) instead of the OECD and non-OECD breakouts in (4) or using an alternative discount rate (5 percent instead of 7 percent). The wedges range from 1.7 to 2.2 under either alternate specification, virtually unchanged from the baseline results. The pooled full sample (excluding the U.S.) TFP impulse response in the top row of Figure 1 already equal-weights all 69 foreign countries, blending the OECD and non-OECD responses, so it is unsurprising that the wedge is largely unaffected along this dimension. As noted earlier, the influence of r on ψ was unclear ex ante and turns out not to matter much ex post. The total global TFP spillovers from U.S. public R&D cannot be fully quantified, and broader country coverage would likely push our estimates of the global social returns higher still. But taken together, the results in Appendix Table B.1 demonstrate that our calculated wedges between the global and domestic social returns to U.S. public R&D are robust to certain modeling choices and deliberately conservative along other dimensions, and should thus be treated as a lower bound: our measurement and weighting choices each err on the conservative side.

7 Conclusion

This paper adapts the narrative IV approach developed by Fieldhouse and Mertens (2026a) to an international panel of countries, estimating the causal effects of U.S. public R&D on foreign productivity and global growth. We find that shocks to U.S. R&D appropriations cause large and persistent increases in foreign TFP. The effects of a 1 percent shock to total

[18]The population share of non-OECD countries within our panel ranges between 71 percent and 76 percent across the three weighting windows, whereas these countries account for only 18 percent to 46 percent of real GDP; both weights are monotonically increasing, which is why the wedges rise monotonically over time.

U.S. public R&D appropriations, relative to the federal R&D capital stock, build gradually in other countries and peak at about a 1 percent increase 12 years after the shock. The estimates imply a substantial wedge between international and domestic productivity spillovers. Across our three GDP-weighting windows, the global social returns to U.S. nondefense public R&D are at least 1.8 to 2.2 times the domestic returns, so the U.S. captures at most 45 to 57 percent of the productivity benefits of its own nondefense R&D. Fieldhouse and Mertens (2026a) estimate that the domestic social returns to U.S. nondefense government R&D appropriations shocks lie between 140 and 210 percent; our estimates suggest that the global social returns to these shocks are roughly 250 to 460 percent or higher.

We document substantial heterogeneity in these spillovers: nondefense R&D causes larger and more persistent increases in foreign TFP than defense R&D, and benefits accrue disproportionately to non-OECD countries and OECD countries with smaller domestic R&D capital stocks. These dynamic causal effects can help discipline and calibrate models of international growth. We interpret our heterogeneity results as consistent with diffusion occurring through openly accessible scientific content, capital-embodied trade, and leapfrogging or imitation by countries further from the technological frontier.

Our results have several important implications for policy. First, the large international spillovers from domestic innovation suggest a potentially important cost of deglobalization policies that may impede international technology flows. If a substantial proportion of global productivity growth relies on innovations that originate in other countries, then policies such as tariffs, export controls, and migration restrictions may impose dynamic costs on economic growth that far exceed their static costs. The greater spillovers from nondefense than defense R&D suggest that the open exchange of ideas, goods, and technologies plays a key role in the international growth benefits of innovation. Our results narrow the plausible channels of diffusion, but more fully characterizing how technology spreads across countries remains an important area for future research.

Second, the wedge between the global and domestic returns to R&D speaks directly to the optimal level of public investment in science and technology. The standard rationale for governments to subsidize research is that market failures cause the social returns to R&D to exceed the private returns, leading private firms to underinvest relative to the social optimum. We find that the same is true at the country level, with the U.S. capturing roughly half of the global benefits of its own nondefense R&D investments. Even if U.S. public R&D investment were set to maximize domestic welfare, it would still fall well below the globally efficient level—a gap that no single government can close unilaterally.

Our results suggest a possible benefit to international coordination among advanced economies to internalize these positive externalities across countries. Our crowd-in results speak directly to the potential payoff from such coordination: because U.S. public R&D and foreign R&D behave as complements rather than substitutes—the U.S. R&D appropriations shocks endogenously crowd in R&D investment in peer OECD countries—coordinated expansions of public research across advanced economies would reinforce one another rather than offset, raising the return to each participant's own investment. Existing international basic science consortia such as the European Organization for Nuclear Research (CERN) or the Human Genome Project could serve as models for such coordination. We caution, however, that the prospects for mutually beneficial international cooperation are complicated by the asymmetry implied by our results: the productivity benefits of U.S. R&D accrue disproportionately to middle- and lower-income countries that conduct relatively little frontier research themselves, potentially limiting their ability to reciprocate. In that sense, the magnitude of our estimated effects suggests that U.S. public R&D is plausibly among the largest unintentional development programs in the world. How these dynamics evolve as the next generation of frontier technologies emerges, and the global policy landscape governing the flow of ideas and goods shifts, will play a major role in shaping the trajectory of global productivity growth in the coming decades.

References

ABRAMOVITZ, M. (1986): "Catching Up, Forging Ahead, and Falling Behind," *Journal of Economic History*, 46, 385–406.

ACEMOGLU, D., P. AGHION, AND F. ZILIBOTTI (2006): "Distance to Frontier, Selection, and Economic Growth," *Journal of the European Economic Association*, 4, 37–74.

AGHION, P. AND P. HOWITT (1992): "A Model of Growth through Creative Destruction," *Econometrica*, 60, 323–351.

AKCIGIT, U., S. T. ATES, AND G. IMPULLITTI (2026): "Innovation and Trade Policy in a Globalized World," *Journal of Political Economy Macroeconomics*, 4, 211–255.

AKCIGIT, U., D. HANLEY, AND N. SERRANO-VELARDE (2020): "Back to Basics: Basic

Research Spillovers, Innovation Policy, and Growth," *The Review of Economic Studies*, 88, 1–43.

ALVAREZ, F. E., F. J. BUERA, AND R. E. LUCAS (2013): "Idea Flows, Economic Growth, and Trade," NBER Working Paper 19667, National Bureau of Economic Research.

AYERST, S., F. IBRAHIM, G. MACKENZIE, AND S. RACHAPALLI (2023): "Trade and diffusion of embodied technology: an empirical analysis," *Journal of Monetary Economics*, 137, 128–145.

BASU, S., J. G. FERNALD, AND M. S. KIMBALL (2006): "Are Technology Improvements Contractionary?" *American Economic Review*, 96, 1418–1448.

BILIR, L. K. AND E. MORALES (2020): "Innovation in the Global Firm," *Journal of Political Economy*, 128, 1566–1625.

BLOOM, N., M. SCHANKERMAN, AND J. VAN REENEN (2013): "Identifying Technology Spillovers and Product Market Rivalry," *Econometrica*, 81, 1347–1393.

BRANSTETTER, L. G. (2001): "Are Knowledge Spillovers International or Intranational in Scope? Microeconometric Evidence from the U.S. and Japan," *Journal of International Economics*, 53, 53–79.

BREZIS, E. S., P. R. KRUGMAN, AND D. TSIDDON (1993): "Leapfrogging in International Competition: A Theory of Cycles in National Technological Leadership," *American Economic Review*, 83, 1211–1219.

BUERA, F. J. AND E. OBERFIELD (2020): "The Global Diffusion of Ideas," *Econometrica*, 88, 83–114.

CAI, J., N. LI, AND A. M. SANTACREU (2022): "Knowledge Diffusion, Trade, and Innovation across Countries and Sectors," *American Economic Journal: Macroeconomics*, 14, 104–145.

CASELLI, F. AND D. J. WILSON (2004): "Importing Technology," *Journal of Monetary Economics*, 51, 1–32.

CBO (2025): "Preliminary Analysis of How Federal Investment in Nondefense Research and Development Affects the Economy and the Federal Budget," Tech. rep., Congressional Budget Office.

COE, D. T. AND E. HELPMAN (1995): "International R&D Spillovers," *European Economic Review*, 39, 859–887.

COE, D. T., E. HELPMAN, AND A. W. HOFFMAISTER (1997): "North-South R&D Spillovers," *Economic Journal*, 107, 134–149.

——— (2009): "International R&D spillovers and institutions," *European Economic Review*, 53, 723–741.

COHEN, W. M. AND D. A. LEVINTHAL (1990): “Absorptive Capacity: A New Perspective on Learning and Innovation,” *Administrative Science Quarterly*, 35, 128–152.

EATON, J. AND S. KORTUM (2001): “Trade in Capital Goods,” *European Economic Review*, 45, 1195–1235.

FARROKHI, F. AND H. S. PELLEGRINA (2023): “Trade, technology, and agricultural productivity,” *Journal of Political Economy*, 131, 2509–2555.

FEENSTRA, R. C., R. INKLAAR, AND M. P. TIMMER (2015): “The Next Generation of the Penn World Table,” *American Economic Review*, 105, 3150–3182.

FERNALD, J. G. (2012): “A Quarterly, Utilization-Adjusted Series on Total Factor Productivity,” Working Paper Series 2012-19, Federal Reserve Bank of San Francisco.

FIELDHOUSE, A. J. AND K. MERTENS (2023): “A Narrative Analysis of Federal Appropriations for Research and Development,” Federal Reserve Bank of Dallas Working Paper 2316.

——— (2026a): “The Returns to Government R&D: Evidence from U.S. Appropriations Shocks,” Federal Reserve Bank of Dallas Working Paper 2305.

——— (2026b): *The Social Returns to Public R&D*, University of Chicago Press, 9–39.

GERSCHENKRON, A. (1962): *Economic Backwardness in Historical Perspective: A Book of Essays*, Cambridge, MA: Belknap Press of Harvard University Press.

GRILICHES, Z. (1979): “Issues in Assessing the Contribution of Research and Development to Productivity Growth,” *The Bell Journal of Economics*, 10, 92–116.

HOWITT, P. (2000): “Endogenous Growth and Cross-Country Income Differences,” *American Economic Review*, 90, 829–846.

HSIEH, C.-T., P. J. KLENOW, AND I. B. NATH (2023): “A Global View of Creative Destruction,” *Journal of Political Economy Macroeconomics*, 1, 243–275.

HUO, Z., A. A. LEVCHENKO, AND N. PANDALAI-NAYAR (2023): “Utilization-adjusted TFP across countries: Measurement and implications for international comovement,” *Journal of International Economics*, 146, 103753.

IARIA, A., C. SCHWARZ, AND F. WALDINGER (2018): “Frontier Knowledge and Scientific Production: Evidence from the Collapse of International Science,” *Quarterly Journal of Economics*, 133, 927–991.

JONES, B. F. AND L. H. SUMMERS (2022): “A Calculation of the Social Returns to Innovation,” in *Innovation and Public Policy*, ed. by A. Goolsbee and B. Jones, University of Chicago Press, 13–59.

JONES, C. I. AND J. C. WILLIAMS (1998): “Measuring the Social Return to R&D,” *The Quarterly Journal of Economics*, 113, 1119–1135.

KLENOW, P. J. AND A. RODRÍGUEZ-CLARE (2005): "Externalities and Growth," in *Handbook of Economic Growth*, ed. by P. Aghion and S. N. Durlauf, Elsevier, vol. 1, chap. 11, 817–861.

LEWIS, D. J. AND K. MERTENS (2024): "Dynamic Identification Using System Projections and Instrumental Variables," Working Paper 2204, Federal Reserve Bank of Dallas.

LUCAS, R. E. (2009): "Trade and the Diffusion of the Industrial Revolution," *American Economic Journal: Macroeconomics*, 1, 1–25.

MORETTI, E., C. STEINWENDER, AND J. VAN REENEN (2025): "The Intellectual Spoils of War? Defense R&D, Productivity, and International Spillovers," *Review of Economics and Statistics*, 107, 14–27.

RAMEY, V. A. AND S. ZUBAIRY (2018): "Government Spending Multipliers in Good Times and in Bad: Evidence from US Historical Data," *Journal of Political Economy*, 126, 850–901.

ROMER, P. M. (1990): "Endogenous Technological Change," *Journal of Political Economy*, 98, S71–S102.

A DATA APPENDIX

A.1 Data Sources for Productivity Growth

Data sources for international TFP are as follows:

- **Penn World Table.** From the Penn World Table (PWT) v11.0 (Feenstra et al., 2015), published October 2025, we use economy-wide TFP at constant national prices ('rtfpna', 2021=1), real GDP at constant 2021 national prices ('rgdpna'), and population ('pop'), from which we construct real GDP per capita ('rgdppc').
https://www.rug.nl/ggdc/productivity/pwt/?lang=en

- **Penn World Table capital detail.** From the PWT v11.0 capital-detail breakout (Feenstra et al., 2015) we use the nominal current-cost net capital stock of other assets ('Nc_Other') and the capital stock price index for other assets ('Np_Other') to construct the real net capital stock of other assets ('rnn_other'). The PWT 'Other' capital category covers assets beyond residential and non-residential structures, machinery and (non-transport) equipment, and transport equipment; this residual category comprises software, other intellectual property (including R&D), and cultivated biological assets.
https://dataverse.nl/api/access/datafile/554029

- **SNA vintage classification.** We classify each country by its System of National Accounts (SNA) vintage, distinguishing those reporting on SNA 2008 from those on earlier vintages, based on the 'SNA System' variable in the UN Statistics Division's National Accounts Official Country Data Database, obtained via a custom data request (received May 19, 2026). Taiwan is not covered by the UN database; its classification instead uses data from Taiwan's Directorate-General of Budget, Accounting and Statistics, which retroactively applied SNA 2008 to its historical series in its 2014 revisions.

- **International utilization-adjusted TFP.** We use the 'HLP' utilization-adjusted TFP from the replication files of Huo et al. (2023):
https://www.dropbox.com/s/if6yv793jazqoro/utilization_adjusted_tfp.xlsx?dl=0

- **U.S. business-sector TFP.** We use the Fernald (2012) utilization-adjusted and unadjusted TFP series for the U.S. business sector, as maintained by the Federal Reserve Bank of San Francisco, available at:
www.frbsf.org/.../data-and-indicators/total-factor-productivity-tfp

A.2 Data Sources for R&D Activity

- **R&D capital stocks.** We construct R&D capital stocks by performing sector for 24 advanced economies (all current OECD members), following the perpetual-inventory methodology of Coe et al. (2009). The underlying R&D expenditures are from the OECD Main Science and Technology Indicators (MSTI) database by performing sector: business enterprise (which we refer to as private-sector R&D), government, and

higher education, available at https://www.oecd.org/en/data/datasets/research-and-development-statistics.html.
Nominal expenditures are converted to constant 2000 purchasing power parity (PPP) dollars by deflating to constant national prices with the World Bank GDP deflator ('NY.GDP.DEFL.ZS') and converting to a common currency with OECD purchasing-power-parity conversion factors for GDP ('PPP_B1GQ', OECD National Accounts), available at https://data.worldbank.org/indicator/NY.GDP.DEFL.ZS and https://data-explorer.oecd.org/.
For each country and sector, we set a benchmark stock in the year t_0 preceding the first reliable year of data, $S_{t_0} = R_{t_1}/(g+\delta)$, where t_1 is the first reliable year, R_{t_1} is expenditure in that year, g is the average log-growth rate of expenditure over an initial window, and δ is the depreciation rate. Capital stocks are then accumulated forward through 2023 as $S_t = (1-\delta)S_{t-1} + R_t$. Single-year gaps in expenditure are linearly interpolated; for Austria, where sectoral expenditure is unavailable before 2004, we impute it from each sector's share of gross domestic expenditure on R&D (GERD), following Coe et al. (2009). We omit the private non-profit sector because of sparse reporting (see Section 2.3), and construct a "total" series by summing the private-sector, government, and higher-education stocks.

- **R&D depreciation rates.** Our baseline assumes $\delta = 16\%$, the rate the Bureau of Economic Analysis applies to intellectual property for universities and colleges, scientific R&D services, and a number of government R&D series, available at https://apps.bea.gov/national/pdf/BEA_depreciation_rates.pdf.
For robustness we also construct all stocks using $\delta = 5\%$, following Coe et al. (2009).

- **R&D personnel.** Data are from the OECD MSTI 'R&D personnel' database for 1981–2023 (by sector of employment and function) and the OECD's discontinued 'Historical R&D expenditure series' for 1963–1980 (by sector of employment and occupation), both available at
https://www.oecd.org/en/data/datasets/research-and-development-statistics.html

Table A.1: Baseline panel countries and SNA system in use, 1978–2019

Country	ISO3	SNA 1968	SNA 1993	SNA 2008
Original OECD (24)				
Australia	AUS	–	–	1978–2019
Austria	AUT	–	1978–1994	1995–2019
Belgium	BEL	–	1978–1994	1995–2019
Canada	CAN	–	–	1978–2019
Denmark	DNK	–	1978–1994	1995–2019
Finland	FIN	–	1978–1994	1995–2019
France	FRA	–	1978–1994	1995–2019
Germany	DEU	–	1978–1994	1995–2019
Greece	GRC	–	1978–1994	1995–2019
Iceland	ISL	–	1978–1994	1995–2019
Ireland	IRL	–	1978–1994	1995–2019
Italy	ITA	–	1978–1994	1995–2019
Japan	JPN	–	1978–1994	1995–2019
Luxembourg	LUX	–	1978–1994	1995–2019
Netherlands	NLD	–	1978–1994	1995–2019
New Zealand	NZL	–	1978–1994	1995–2019
Norway	NOR	–	1978–1994	1995–2019
Portugal	PRT	–	1978–1994	1995–2019
Spain	ESP	–	1978–1994	1995–2019
Sweden	SWE	–	1978–1994	1995–2019
Switzerland	CHE	–	1978–1994	1995–2019
Türkiye	TUR	–	1978–1994	1995–2019
United Kingdom	GBR	–	1978–1994	1995–2019
United States	USA	–	–	1978–2019

(continued on next page)

(continued from previous page)

Country	ISO3	SNA 1968	SNA 1993	SNA 2008
Later OECD entrants (6)				
Chile	CHL	1978–1994	–	1995–2019
Colombia	COL	1978–1991	1992–2004	2005–2019
Costa Rica	CRI	1978–1990	–	1991–2019
Israel	ISR	–	–	1978–2019
Mexico	MEX	–	–	1978–2019
Republic of Korea	KOR	–	1978–1994	1995–2019
Never-OECD economies (40)				
Argentina	ARG	–	–	2004–2019
Barbados	BRB	1978–1989	1990–2019	–
Bolivia	BOL	1978–2019	–	–
Brazil	BRA	1980–1989	1990–1994	1995–2019
Burkina Faso	BFA	1978–1985, 1989–1993	1995–1998	1999–2019
Cameroon	CMR	1978–1989	1990	1993–2019
China	CHN	–	–	1978–2019
China, Hong Kong SAR	HKG	–	–	1978–2019
Cyprus	CYP	1978–1994	–	1995–2019
Côte d’Ivoire	CIV	1978–2000	2005–2014	2015–2019
Dominican Republic	DOM	1978–1990	–	1991–2019
Ecuador	ECU	–	–	1978–2019
Egypt	EGY	1978–1981, 1986–1991	1996–2019	–
Guatemala	GTM	1978–2000	2001–2012	2013–2019
Honduras	HND	1978–1999	2000–2019	–
India	IND	–	–	1978–2019
Indonesia	IDN	1978–1999	2000–2009	2010–2019

(continued on next page)

(continued from previous page)

Country	ISO3	SNA 1968	SNA 1993	SNA 2008
Jamaica	JAM	1978–1997	1998–2019	–
Jordan	JOR	1978–2007	2008–2019	–
Kenya	KEN	1978–1995	1996–2005	2006–2019
Lesotho	LSO	1978–1986	1987–2019	–
Malaysia	MYS	1978–1999	2000–2009	2010–2019
Malta	MLT	1978–1994	–	1995–2019
Mauritius	MUS	1978–1991	1992–2005	2006–2019
Morocco	MAR	1978–1997	1998–2006	2007–2019
Namibia	NAM	–	–	1980–2019
Niger	NER	1978–1997	1998–2004	2005–2019
Panama	PAN	1978–1995	1996–2017	2018–2019
Paraguay	PRY	1978–1993	1994–2006	2007–2019
Peru	PER	1978–1990	1991–1994	1995–2019
Philippines	PHL	1978–1997	–	1998–2019
Senegal	SEN	1978–1995	1996–1998	1999–2019
Singapore	SGP	–	–	1978–2019
South Africa	ZAF	–	1978–1992	1993–2019
Sri Lanka	LKA	1978–1997	1998–2009	2010–2019
Taiwan	TWN	–	–	1978–2019
Thailand	THA	1978–1989	1990–1992	1993–2019
Tunisia	TUN	1978–1996	1997–2014	2015–2019
U.R. of Tanzania: Mainland	TZA	1978–1993	1994–2011	2012–2019
Uruguay	URY	1978–1996	1997–2015	2016–2019

Notes: Year ranges show the System of National Accounts vintage in use for each country-year over 1978–2019 (our 1980–2019 baseline sample plus two annual lags), as classified by the UN Statistics Division's National Accounts Official Country Data Database; see Appendix A.1 for details. Where multiple SNA vintages are reported for a country-year, we take the most recent vintage. Year-range gaps within a country reflect missing years in the UN data rather than country exits from the panel.

Table A.2: Below-median indicators for heterogeneity splits of baseline panel subgroups

Country	ISO3	GDP per capita	R&D capital	Population
Original OECD, ex-USA (23)				
Australia	AUS			
Austria	AUT		•	•
Belgium	BEL	•		
Canada	CAN			
Denmark	DNK		•	•
Finland	FIN	•	•	•
France	FRA	•		
Germany	DEU			
Greece	GRC	•	•	•
Iceland	ISL	•	•	•
Ireland	IRL	•	•	•
Italy	ITA			
Japan	JPN	•		
Luxembourg	LUX		•	•
Netherlands	NLD			
New Zealand	NZL	•	•	•
Norway	NOR		•	•
Portugal	PRT	•	•	•
Spain	ESP	•	•	
Sweden	SWE			•
Switzerland	CHE			•
Türkiye	TUR	•	•	
United Kingdom	GBR	•		

(continued on next page)

(continued from previous page)

Country	ISO3	GDP per capita	R&D capital	Population
Never-OECD (40)				
Argentina	ARG			
Barbados	BRB		•	•
Bolivia	BOL	•		•
Brazil	BRA			
Burkina Faso	BFA	•	•	
Cameroon	CMR	•	•	
China	CHN	•		
China, Hong Kong SAR	HKG			•
Cyprus	CYP			•
Côte d'Ivoire	CIV	•		
Dominican Republic	DOM		•	•
Ecuador	ECU			
Egypt	EGY	•	•	
Guatemala	GTM		•	•
Honduras	HND	•	•	•
India	IND	•		
Indonesia	IDN	•		
Jamaica	JAM		•	•
Jordan	JOR		•	•
Kenya	KEN	•		
Lesotho	LSO	•	•	•
Malaysia	MYS		•	
Malta	MLT			•
Mauritius	MUS	•	•	•
Morocco	MAR	•		

(continued on next page)

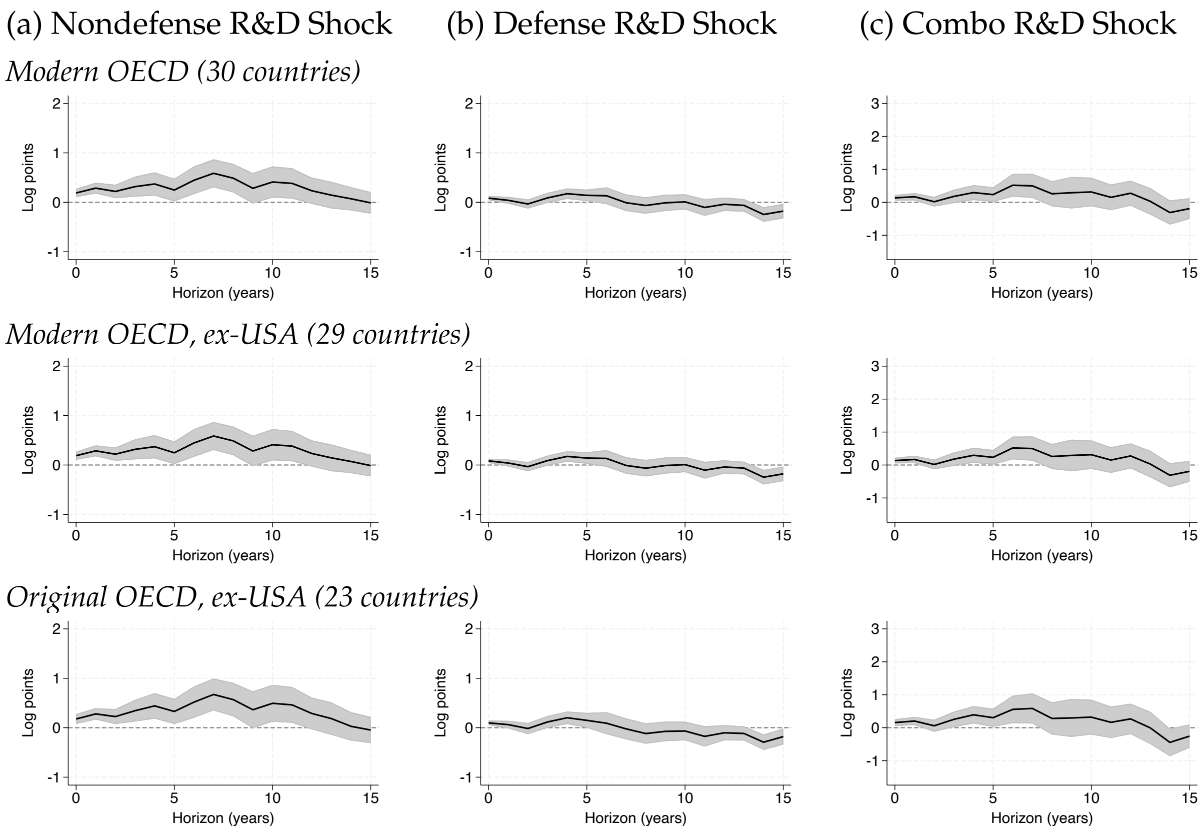
(continued from previous page)

Country	ISO3	GDP per capita	R&D capital	Population
Namibia	NAM		•	•
Niger	NER	•	•	•
Panama	PAN		•	•
Paraguay	PRY		•	•
Peru	PER			
Philippines	PHL	•		
Senegal	SEN	•	•	•
Singapore	SGP		•	•
South Africa	ZAF			
Sri Lanka	LKA	•	•	
Taiwan	TWN			
Thailand	THA	•		
Tunisia	TUN	•	•	•
U.R. of Tanzania: Mainland	TZA	•		
Uruguay	URY			•

Notes: Entries marked • indicate the country falls below the within-group median for that variable sort. Medians are computed separately within each country group (Original OECD ex-USA; never-OECD), based on 1977–79 pre-sample averages of the log of each variable relative to the corresponding log U.S. value. *GDP per capita*: real GDP per capita ('rgdppc'). *R&D capital*: real IPP and software capital ('rnn_other', constructed from the PWT capital detail), a proxy for absorptive capacity. *Population*: population ('pop'). Countries at exactly the within-group median are assigned to the below-median group. See Data Appendix A.1 for variable construction notes and sources.

B Additional Results

FIGURE B.1: Foreign TFP Response to an Increase in U.S. R&D Appropriations: Full OECD and OECD Subgroups

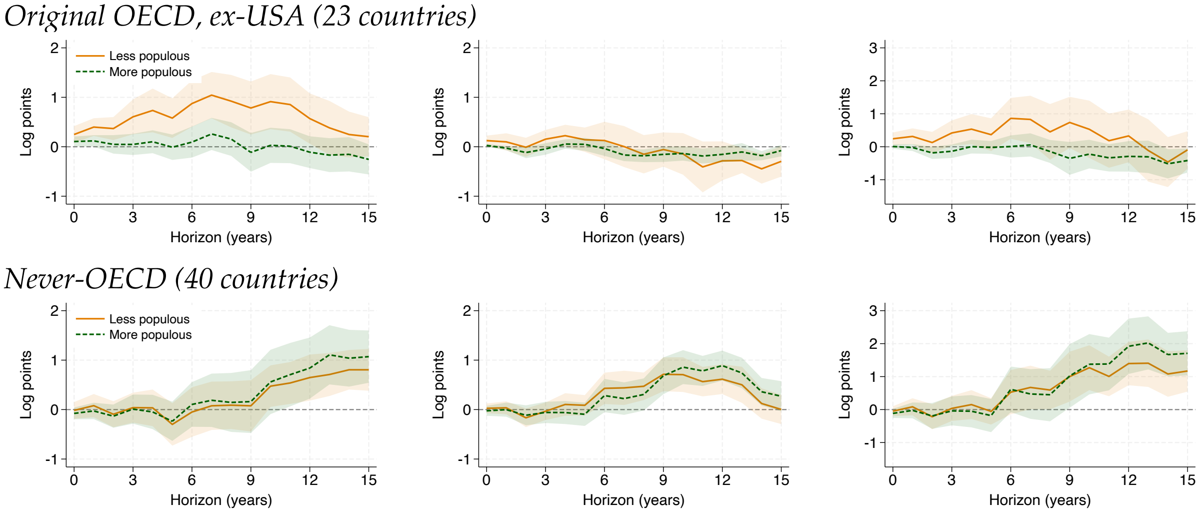


Notes: Estimates based on (1) using the narrative measure of changes in federal nondefense (left panel), defense (middle panel), and combined (right panel) U.S. R&D appropriations shocks. Shaded bands are 95 percent confidence intervals based on country-clustered standard errors. Impulses scaled to imply a 1 percent increase in U.S. R&D appropriations as a share of the respective federal R&D capital stock. *Sample*: 1980–2019.

The main paper focuses on the 'original' OECD countries excluding the U.S. when exploring heterogeneity across OECD versus non-OECD countries—in part for more consistent TFP measurement and in part because other key variables of interest are only available for a closely overlapping set of advanced economies, e.g., the Coe et al. (2009)-style R&D capital stocks. As a robustness check, Figure B.1 compares impulse response functions of foreign PWT TFP, estimated from (1), to the federal nondefense (left panel), defense (middle panel), and combined (right panel) U.S. R&D appropriations shocks across three OECD subgroups: the full modern OECD including the U.S. (top row), the modern OECD excluding the U.S. (middle row), and the 'original' OECD countries excluding the U.S. (bottom row, replicated from Figure 1). The key takeaway is that the TFP responses to each of the three U.S. R&D appropriations shocks are virtually identical across the three OECD cohorts, both quantitatively and qualitatively, when expanding the pool to include the U.S. or other modern entrants alongside the original OECD countries excluding the U.S.; the headline OECD results in the main paper are therefore robust to these alternative OECD cohort definitions.

FIGURE B.2: Foreign TFP Response to an Increase in U.S. R&D Appropriations: Heterogeneous Responses by Relative Population

(a) Nondefense R&D Shock (b) Defense R&D Shock (c) Combo R&D Shock

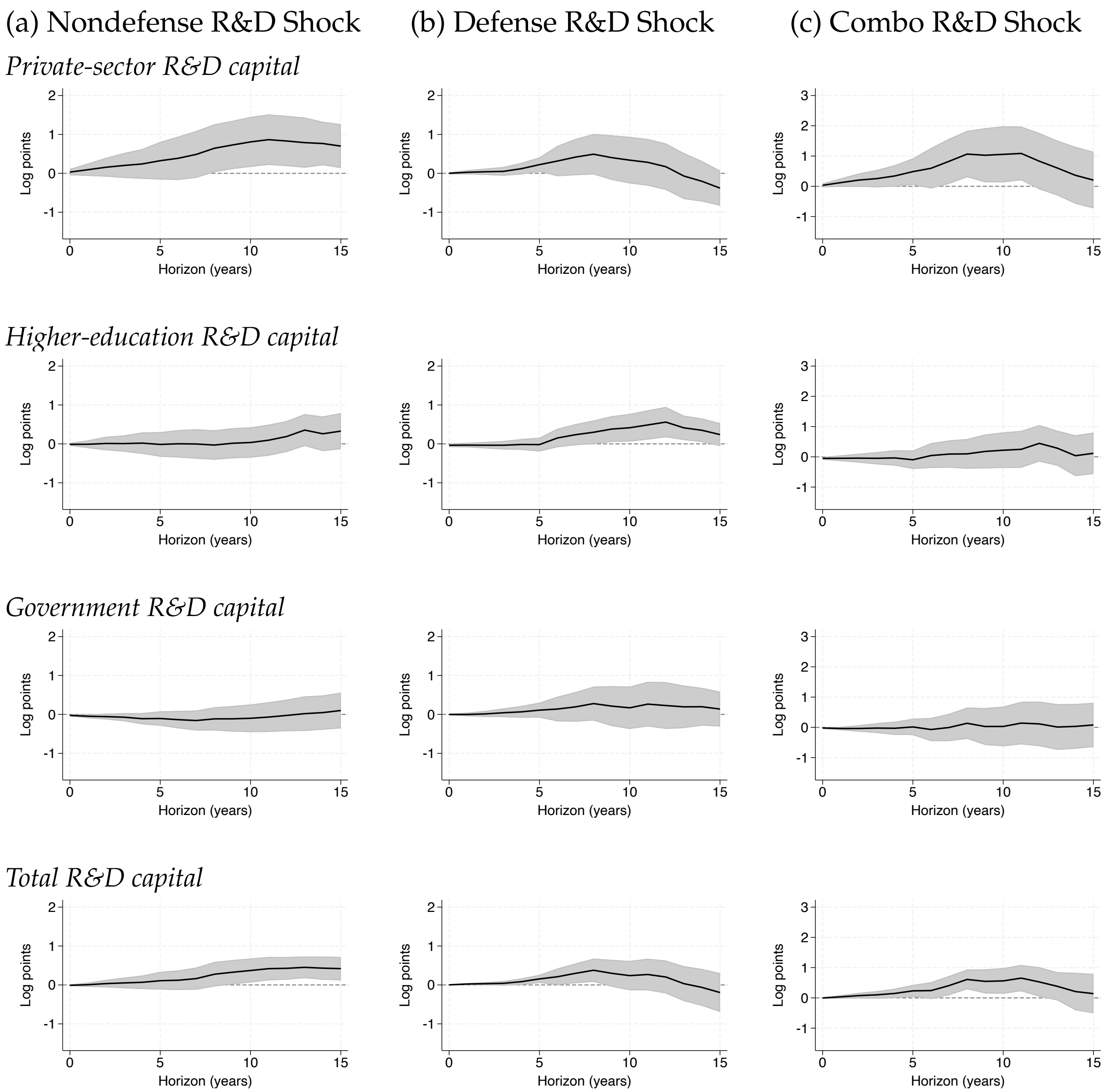


Notes: Estimates based on (1) using the narrative measure of changes in federal nondefense (left panel), defense (middle panel), and combined (right panel) U.S. R&D appropriations shocks. Shaded bands are 95 percent confidence intervals based on country-clustered standard errors. Impulses scaled to imply a 1 percent increase in U.S. R&D appropriations as a share of the respective federal R&D capital stock. Solid maroon (dashed teal) lines correspond to countries below (above) the sample median of each country's log population relative to the United States (PWT *pop*), averaged over 1977–1979 within that country group. *Sample*: 1980–2019.

The main paper reports heterogeneous responses within the OECD and non-OECD country groups when sorting countries by above/below median real GDP per capita (Figure 2) and IPP and R&D capital (Figure 3). Analogously, Figure B.2 reports TFP responses for each subsample split by population. For the OECD countries, a gradual hump-shaped TFP response to the nondefense R&D appropriations shock is statistically significant for the less populous countries but is never significantly different from zero for the more populous countries, while the responses to the defense R&D shock are relatively flat for both population bins. For the non-OECD countries, the TFP responses to all three shocks are qualitatively similar across the two population bins, but stronger and more persistent at longer horizons for the more populous countries. See Section 4.1 for discussion of how these patterns fit into the broader absorptive-capacity reading of the heterogeneity results.

Figure B.3 replicates Figure 8 from the main paper using foreign R&D capital stocks constructed under a 5 percent depreciation rate instead of a 16 percent depreciation rate;

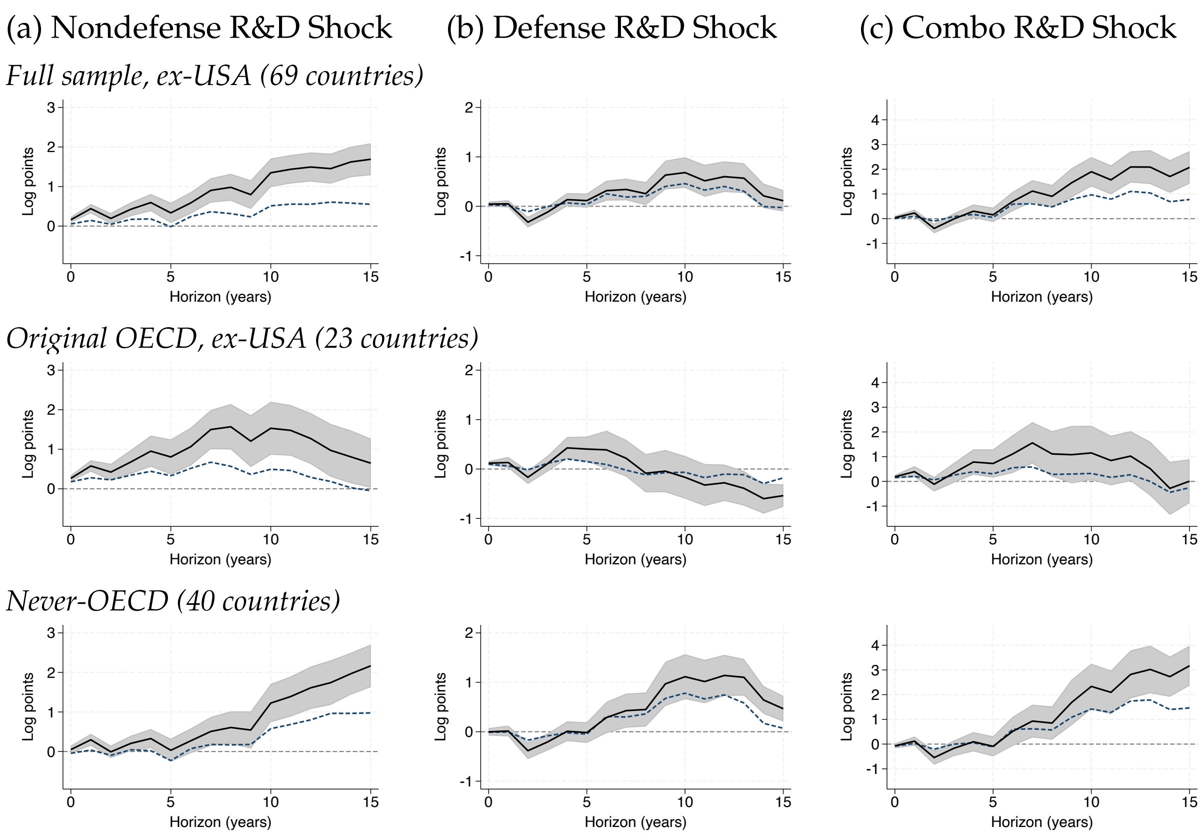
FIGURE B.3: Foreign R&D Capital Responses to an Increase in U.S. R&D Appropriations: Alternative Assumed Depreciation Rate ($\delta = 5\%$)



Notes: See notes to Figure 8 for specification, construction, and sample details. The only difference is that the foreign R&D capital stocks here are constructed using a 5 percent rather than 16 percent depreciation rate; see Data Appendix for details.

it underscores that the headline results regarding U.S. public R&D crowding in foreign R&D are robust to alternative assumptions about depreciation rates.

FIGURE B.4: Foreign GDP per Capita and TFP Responses to an Increase in U.S. R&D Appropriations



Notes: Estimates based on (1) using the narrative measure of changes in federal nondefense (left panel), defense (middle panel), and combined (right panel) U.S. R&D appropriations shocks. The solid lines and shaded bands report the real GDP per capita responses and their 95 percent confidence intervals (country-clustered standard errors); the dashed lines report the TFP point estimates from Figure 1 (confidence intervals omitted here for clarity). Impulses scaled to imply a 1 percent increase in U.S. R&D appropriations as a share of the respective federal R&D capital stock. *Sample*: 1980–2019.

Figure B.4 overlays responses of foreign real GDP per capita (solid black lines, with 95 percent confidence bands) on the corresponding TFP responses from Figure 1 (dashed navy lines), for the full ex-U.S. sample (top row), the original OECD excluding the U.S. (middle row), and the non-OECD countries (bottom row). Because real GDP per capita reflects both factor accumulation and total factor productivity, the gap between the two lines captures the part of the output response operating through inputs rather than the measured productivity residual. Across all three country groups the output response exceeds the TFP response at most horizons. For the original OECD countries, the nondefense output response builds to roughly 1.6 percent—about two-and-a-half times the corresponding TFP response of roughly 0.7 percent—with a similar wedge for the combined shock. This is consistent with the R&D-capital crowd-in documented in Figure 8: these economies report on the SNA 2008 vintage, under which R&D investment is capitalized into the aggregate capital stock and thus removed from the PWT TFP residual, so induced foreign R&D capital raises output without registering in TFP. The non-OECD countries show even larger output responses (exceeding 3 percent for the combined shock) and a sizable output-TFP gap of their own; for this group, however, the gap chiefly reflects the broad factor accumulation that accompanies catch-up growth rather than the accounting convention, since many of these countries remain on earlier SNA vintages under which R&D flows into the residual. We therefore interpret Figure B.4 as evidence that, for the OECD countries, U.S. public R&D spillovers raise foreign output through input accumulation only partially visible in measured TFP—though the output-TFP wedge cannot, on its own, attribute that accumulation specifically to the R&D-investment channel rather than other forms of factor deepening. The sharpest test of the SNA measurement effect itself is instead the within-non-OECD comparison of SNA 2008 versus earlier SNA vintage countries in Section 4.3.

Table B.1: Robustness of the Global-to-Domestic Spillover Wedge

Weighting window	Baseline	U.S. PWT	Population	Pooled	$r = 5\%$
1977–79	1.77	2.42	4.60	1.71	1.74
1994–96	1.91	2.68	4.98	1.87	1.89
2017–19	2.22	3.26	5.09	2.24	2.23

Notes: Each cell is the discounted present-value global-to-domestic spillover wedge ψ for U.S. nondefense R&D, as in Table 1, using a 7% discount rate unless noted. The first column reproduces the Table 1 baseline (Fernald (2012) utilization-adjusted U.S. response, GDP-weighted modern-OECD/non-OECD breakout, 7% rate). Each remaining column changes one modeling assumption: *U.S. PWT* measures the U.S. response with PWT economy-wide TFP rather than utilization-adjusted business-sector TFP, matching the foreign TFP measures; *Population* weights country groups by PWT population rather than real GDP; *Pooled* replaces the breakout with the pooled full-sample-ex-U.S. response applied to total foreign GDP (the top row of Figure 1); and $r = 5\%$ lowers the discount rate from 7% to 5%.

Table 1 in the main paper reports lower bounds on the wedge between the global and domestic social returns to U.S. nondefense public R&D. Table B.1 assesses the robustness of these global-to-domestic spillover wedges to alternative modeling choices, as discussed in Section 6. The first column reproduces the baseline discounted-present-value wedges from Table 1; each remaining column changes a single ingredient of the construction. The wedges are substantially larger when using an apples-to-apples U.S. TFP measure (economy-wide PWT TFP rather than Fernald's (2012) utilization-adjusted business-sector TFP) or when weighting the TFP impulse responses using population instead of real GDP. The wedges are essentially unchanged when pooling the full foreign sample or lowering the discount rate from 7 percent to 5 percent. These patterns confirm that the baseline wedges should be interpreted as conservative lower bounds.