CPAG

RESEARCH REPORT

R&D Composition and Absorptive Capacity in the European Union

Panel Evidence on Catch-Up Growth in
Romania, Czechia, Hungary and Poland

27 EU MEMBER STATES	1995–2024 PANEL COVERAGE	5 COMPOSITION DIMENSIONS	1 SURVIVES THE PROTOCOL

Ella Kállai
CONSILIUM POLICY ADVISORS GROUP

Executive Summary

27
EU member states in the panel, 1995–2024

5
dimensions of R&D composition tested

1
dimension survives the full robustness protocol

+10pp
capital-expenditure shift, within Romania's own historical range

WHAT THE EVIDENCE SHOWS

- **One lever, not five.** Of the five compositional dimensions the literature suggests should matter for absorptive capacity, only the division between capital and current expenditure withstands systematic testing. The remaining four are not shown to be irrelevant; they are shown to be unsupported by this panel, which is a weaker and more honest claim.

- **The lever is conditional on distance.** The association between a higher capital share and faster convergence is concentrated in economies well behind the frontier and disappears near it. Any recommendation drawn from this work therefore has to be addressed to a particular kind of economy rather than to the Union as a whole.

- **Composition is not a substitute for level.** The estimates hold total research intensity constant. They speak to how a given research budget is allocated, not to whether that budget is large enough, and they offer no support for reducing overall research effort.

- **Methodological choices carry substantive weight.** Replacing the conventional accession-based grouping of member states with groupings built from measured distance to the frontier removed most of the apparently significant results. Studies in this literature that rely on the historical grouping should be read with that fragility in mind.

- **Romania has both the largest projected gain and the least room for error.** It is the economy furthest from the frontier in the sample and the only one of the four with an implied effect distinguishable from zero, while also being the one whose research statistics have the thinnest historical coverage.

- **What would strengthen the evidence.** A credible source of exogenous variation in research composition would move these results from association to causation. The European Commission funding share was examined as a candidate instrument and rejected as too weak; firm- or project-level data on equipment purchases would be a more promising route.

WHAT IS AT STAKE

For a country at Romania's distance from the European productivity frontier, the composition of research spending is one of the few levers that can be moved inside a single budget cycle, without new money and without institutional reform. That makes it unusually cheap to try relative to most growth policy — and, on the evidence assembled here, the only compositional lever with a defensible claim to work.

— In brief: what this means for policy

Research policy is usually argued about as a single number: how much a country spends on research, as a share of its economy. This study asks a different question. If the amount stays the same, does it matter what the money is spent on?

FIVE WAYS TO SPLIT A RESEARCH BUDGET — ONLY ONE PREDICTS CATCH-UP

The split	What it compares	Predicts catch-up?
Who does the work	Firms vs. universities	✕ No
Who pays	Government vs. business	✕ No
Type of research	Basic vs. applied	✕ No
Field	Engineering vs. social sciences	✕ No
What the money buys	Equipment vs. salaries	✓ Yes

THE SAME SHIFT, TWO DIFFERENT COUNTRIES

A country still far behind

Catches up faster

Equipment brings in technology invented elsewhere. Buying it and learning to run it is the quickest way to import knowledge the country has not produced itself.

A country already at the top

No gain, or a loss

There is nothing left to import. The same money does more when it goes to the people who create new knowledge rather than to the machines that carry existing knowledge.

ROMANIA HAS ALREADY MADE THIS SHIFT ONCE

6% EQUIPMENT SHARE, 1995 → **24%** EQUIPMENT SHARE, 2024

Romania is the furthest behind of the four Central and Eastern European economies examined, and the only one where the effect is statistically clear. A further move is well inside familiar territory.

TWO THINGS THE FINDING DOES NOT SAY

- **It is not an argument for spending less on researchers.** The comparison holds the total research budget fixed. It is about the mix, not the size.

- **It is a pattern, not a proven cause.** The evidence comes from what 27 countries actually did between 1995 and 2024, not from an experiment. The projected gains are scenarios, not guarantees.

ABSTRACT

This paper investigates whether the composition of gross domestic R&D expenditure (GERD), not merely its intensity, shapes an economy's absorptive capacity and hence its speed of productivity catch-up to the EU frontier. Using an unbalanced panel of 27 EU countries spanning 1995-2024, we estimate a Nelson-Phelps-style catch-up specification in which productivity growth depends on a country's distance from the EU productivity frontier, interacted with the composition of its R&D spending across sector of performance, funding source, type of research, field of research, and the current-versus-capital expenditure split, holding overall R&D intensity (R&D as a share of gross value added) constant. We find that the capital-expenditure share of R&D spending is the only composition variable whose relationship with catch-up speed is robust to a systematic test battery: exclusion of two economies (Ireland, Luxembourg) with statistically distorted productivity series, replacement of historical accession-based country groupings with data-driven groupings based on each country's actual distance from the frontier, lag lengths of one to three years, and three separate treatments of the 2008-2009 financial crisis and 2020-2021 COVID-19 shock. The result additionally survives a wild-cluster bootstrap addressing the small number of country clusters, a randomization-inference test and leave-one-country-out re-estimation addressing the data-driven grouping's construction, and two Granger-style precedence checks addressing reverse causality; an exploratory instrumental-variables approach using European Commission R&D funding share as a candidate instrument proved too weak to be informative and is reported for transparency rather than pursued. A higher capital-expenditure share is associated with faster catch-up in economies far from the frontier and is neutral-to-negative in economies close to it. Several other candidate findings, including an initially significant result for business-performed R&D share and a partially significant result for government-funded R&D share, do not survive this protocol and are explicitly retracted. We discuss implications for Romania, Czechia, Hungary and Poland, the four Central and Eastern European economies at the empirical center of this study.

Keywords: absorptive capacity, R&D composition, technology catch-up, Central and Eastern Europe, national innovation systems

JEL classification: O30, O33, O47, O52

01 Introduction

A country's capacity to absorb, adapt, and exploit external knowledge, its absorptive capacity, is widely held to be as important to productivity growth as the volume of domestic R&D spending itself (Cohen and Levinthal, 1990). Yet empirical work on absorptive capacity has tended to treat R&D expenditure as a scalar, typically R&D as a share of GDP or gross value added, collapsing a great deal of policy-relevant structure into a single number. Two countries with identical R&D intensity can differ sharply in whether that spending is directed toward basic or applied research, performed by firms or universities, funded by government or business, embodied in laboratory equipment or in research salaries, and concentrated in engineering or in the social sciences. Whether these compositional choices matter for the speed at which an economy catches up to the technological frontier is the question this paper addresses.

We exploit a purpose-built panel dataset covering all 27 current EU member states from 1995 to 2024, combining labour productivity and R&D statistics at a level of disaggregation, sector of performance, funding source, type of research, field of research, and expenditure type, that is rarely brought together in a single cross-country panel. Our empirical strategy follows the catch-up tradition initiated by Nelson and Phelps (1966), in which the growth-enhancing effect of a country's capacity variables is allowed to depend on its distance from the technological frontier. We extend this framework by interacting frontier-distance not with human capital, as in the original formulation, but with each dimension of R&D composition in turn, while holding R&D intensity constant so that composition effects are identified net of the overall scale of R&D effort.

A central methodological contribution of this paper is as much cautionary as substantive. An initial pass at the data, using the conventional historical grouping of EU member states into the pre-2004 "EU-15" and the post-2004 accession countries, produced several apparently significant findings. Subjecting each of these findings to a systematic robustness protocol, sequentially excluding two small economies (Ireland and Luxembourg) whose productivity statistics are known to be distorted by multinational profit-shifting, and replacing the historical grouping with a data-driven grouping based on each country's actual measured distance from the frontier, caused four of five candidate composition effects to weaken to statistical insignificance or disappear entirely. Only the capital-versus-current expenditure split survives this protocol. We report this process in full, including the retracted findings, in the interest of transparency and because the pattern of fragility is itself informative about the risks of relying on historical political groupings as proxies for economic structure in panel work of this kind.

The remainder of the paper is organized as follows. Section 2 situates the analysis within the absorptive capacity and catch-up growth literatures. Section 3 describes the data. Section 4 sets out the empirical strategy, including the robustness protocol. Section 5 presents results. Section 6 discusses implications for Romania, Czechia, Hungary and Poland. Section 7 addresses limitations, and Section 8 concludes.

> **KEY FINDING**
>
> Two economies spending the same share of output on R&D can differ in how that spending converts into catch-up; the question this paper puts to the data is whether the composition of R&D expenditure, and not only its level, moves the speed of convergence.

02 Theoretical Framework and Literature Review

2.1 Absorptive capacity and the two faces of R&D

Cohen and Levinthal (1990) argue that R&D spending serves two distinct functions: it generates new knowledge directly, and it builds the capacity to recognize, assimilate, and exploit knowledge generated elsewhere. This "two faces" framing implies that the return to R&D spending is not fixed but depends on the stock of external knowledge available to be absorbed, and that R&D directed at basic, general-purpose understanding may contribute disproportionately to the second function relative to applied, narrowly targeted research. Subsequent work has extended this framework in two directions relevant here. Zahra and George (2002) reconceptualize absorptive capacity as a dynamic capability with potential and realized components, moving the construct beyond a single R&D-based proxy and toward the kind of multidimensional treatment this paper's five composition dimensions attempt empirically. At the country level, Griffith, Redding and Van Reenen (2004) provide the direct empirical precedent for treating R&D as serving both an innovation and an absorption role in a cross-country panel, finding that R&D-based absorptive capacity contributes to productivity catch-up across a panel of OECD industries independent of its role in generating innovation directly; more recently, Harris, Krenz and Moffat (2021) construct enterprise-level absorptive capacity estimates across European countries and find that the returns to absorptive capacity vary systematically across countries, including specifically documenting lower absorptive capacity among Central and Eastern European firms relative to Western Europe, a pattern consistent with this paper's own findings in Section 6.

2.2 Catch-up growth and the distance to frontier

The idea that backwardness itself creates growth opportunity dates to Gerschenkron (1962), who argued that latecomer economies can leapfrog stages of technological development by importing already-proven methods rather than repeating the frontier economies' path of discovery. Nelson and Phelps (1966) formalize a version of this intuition: a country's capacity variables, in their original formulation, human capital, accelerate the diffusion of existing frontier technology into domestic use rather than only contributing to indigenous invention, so a given capacity endowment is mechanically more valuable to a country further from the frontier, where more unexploited technology remains to be diffused. Abramovitz (1986) develops the parallel notion of "social capability", the institutional and human capital preconditions that determine whether a country can actually exploit its catch-up potential, cautioning that distance from the frontier is a necessary but not sufficient condition for convergence. Barro and Sala-i-Martin (1992) formalize conditional convergence empirically, showing that growth rates decline with initial distance from a steady state only once other structural characteristics are held constant, a result consistent with the Nelson-Phelps mechanism operating unevenly across countries with different capacities. Benhabib and Spiegel (1994) test the Nelson-Phelps specification directly against cross-country data and find that human capital's growth effect operates primarily through this diffusion channel, an econometric strategy, interacting a capacity variable with distance to frontier, that this paper adapts to R&D composition rather than human capital.

Most directly relevant to the present paper is Vandenbussche, Aghion and Meghir (2006), who extend the Nelson-Phelps logic to argue that not just the level but the composition of a capacity variable should depend on distance to frontier: they show that skilled human capital matters disproportionately for growth near the frontier, where innovation is the binding constraint, while unskilled human capital matters more for growth further away, where imitation and adoption dominate. Their empirical strategy, interacting a compositional split of a capacity variable with distance to frontier, is structurally identical to the one used here, with R&D composition substituted for human capital composition.

This paper follows Vandenbussche, Aghion and Meghir (2006) and the broader catch-up growth literature in estimating a static country-fixed-effects panel by OLS rather than a dynamic panel GMM estimator (Arellano and Bond, 1991; Blundell and Bond, 1998). This choice is not neutral: gap, the frontier-distance regressor, is highly persistent and mechanically related to lagged productivity, which is exactly the setting in which Nickell (1981) shows fixed-effects estimation of models with persistent regressors can be biased in finite samples. We do not implement a dynamic panel correction here, both because the composition variables of interest, not the persistence of gap itself, are this paper's object of interest, and because Arellano-Bond-type estimators are themselves poorly behaved in short, unbalanced panels with the coverage gaps documented in Section 3.2. We return to this as an open limitation in Section 7 rather than treating the static specification as bias-free.

2.3 National innovation systems and the structure of R&D

The national innovation systems literature (Freeman, 1987; Lundvall, 1992; Nelson, 1993) argues that innovative performance is shaped by the institutional structure through which knowledge is produced, funded, and diffused, not simply by the aggregate volume of resources devoted to R&D. This tradition motivates treating sector of performance and funding source as substantively distinct dimensions rather than collapsing them into a single R&D-intensity figure: an innovation system in which business performs most R&D funded by business capital differs structurally, in its incentives, time horizons, and diffusion channels, from one in which government or university-performed research funded by public budgets predominates, even at identical aggregate R&D intensity.

2.4 Sector of performance and funding source: public and private R&D

Griffith, Redding and Van Reenen (2004) provide firm- and industry-level evidence that R&D affects productivity growth through both an innovation channel and a technology-transfer channel, with the relative importance of each varying systematically with a sector's distance from the technological frontier, directly motivating this paper's sector-of-performance variable. A companion paper, Griffith, Redding and Van Reenen (2003), develops the underlying absorptive-capacity mechanism more formally and shows that R&D's effect on the speed of technology transfer from the frontier is distinct from, and in their data larger than, its direct effect on domestic innovation. A separate empirical literature addresses whether publicly funded R&D complements or substitutes for privately funded R&D. David, Hall and Toole (2000) review the econometric evidence and find it decidedly mixed, with public R&D crowding in private R&D in some contexts and crowding it out in others, depending on the specific mechanism (direct funding, tax incentives, public research infrastructure) and the technological maturity of the sector involved. Guellec and van Pottelsberghe de la Potterie (2003), using OECD panel data, find that public R&D expenditure has a positive but time-lagged effect on business R&D, with the relationship weakening as an economy's business R&D base matures, a pattern broadly consistent with the frontier-distance-dependent effect for government funding examined in this paper, albeit one that this paper finds only partial statistical support for once outlier economies are excluded (Section 5.3).

2.5 Type and field of research: basic science and general-purpose capacity

A separate strand of the literature asks whether basic, applied, or experimental research contributes differentially to absorptive capacity. Mansfield (1991) provides early evidence that academic, largely basic, research contributes directly to industrial innovation with a measurable, if long, time lag, while Salter and Martin (2001) review the wider literature on the economic returns to publicly funded basic research and conclude that its benefits operate substantially through the general-purpose human capital and problem-solving capability it builds in trained researchers, a channel conceptually close to Cohen and Levinthal's absorptive-capacity mechanism. This literature motivates the type-of-research and field-of-research variables tested in this paper, though, as discussed in Section 5.3, we find no robust evidence that either dimension predicts catch-up speed in the EU panel once outlier economies and data-driven country groupings are accounted for, an empirical finding this paper reports as a genuine null result rather than a specification failure.

2.6 Capital versus current expenditure: embodied technology diffusion

A distinct literature links the composition of investment, rather than R&D specifically, to growth through the channel of embodied technological change. De Long and Summers (1991) show that equipment investment in particular has a growth effect substantially larger than other forms of investment, especially in developing and catching-up economies, and argue that new capital goods are a primary vehicle through which frontier technology is physically transferred into an adopting economy. Eaton and Kortum (2001) formalize this mechanism at the international level, modeling trade in capital goods as a direct channel of cross-country technology diffusion and showing that a substantial share of measured productivity growth in non-frontier economies can be attributed to imported, equipment-embodied technology rather than domestic invention. This literature provides the theoretical basis for treating the capital-versus-current split of R&D expenditure as a distinct compositional dimension, and for the expectation, borne out in Section 5.2, that a higher capital-expenditure share should matter more for catch-up the further a country sits from the technological frontier.

2.7 Innovation systems in transition and catch-up economies

A body of work specific to Central and Eastern Europe examines how the innovation systems of former centrally planned economies were restructured after 1989 and what this implies for their catch-up trajectories. Radosevic (1999) documents the transformation of the region's inherited science and technology systems, oriented under central planning primarily toward large-scale, government-performed, capital-intensive research, into more market-oriented innovation systems, and argues that this transition was uneven across the region and left many CEE economies with a residual R&D structure poorly matched to the needs of catch-up growth under market conditions. Kravtsova and Radosevic (2011) formalize this concern econometrically, finding that Eastern European countries generate less productivity than their R&D, innovation and production capabilities alone would predict, and attributing part of this shortfall to inefficiencies in how national innovation systems convert R&D inputs into output, a finding consistent with this paper's own emphasis on composition rather than intensity as the more informative margin. Radosevic and Yoruk (2018) develop this further into a technology-upgrading framework in which production capability and capital-embodied technology absorption, rather than domestic knowledge creation narrowly defined, are the binding constraint for middle-income European economies, directly motivating this paper's focus on the capital-versus-current expenditure margin as opposed to R&D intensity alone. A closely related body of firm-level evidence traces this capital-embodied channel to foreign direct investment: Damijan, Knell, Majcen and Rojec (2003) find that FDI and imported technology, rather than domestic R&D accumulation, is the primary channel of productivity-relevant technology transfer in eight transition economies, while Majcen, Radosevic and Rojec (2009) and Narula and Guimon (2010) show that foreign-owned subsidiaries in CEE new member states have historically driven a disproportionate share of measured productivity growth and R&D activity relative to domestically owned firms, a pattern this paper's Section 6 discussion of Romania's FDI-linked manufacturing integration draws on directly. This literature motivates treating CEE economies as a group whose R&D-composition relationships with growth may plausibly differ structurally from those of long-standing market economies, the heterogeneity tested formally in Section 4.2, while also cautioning, consistent with this paper's methodological findings in Section 5.3, against assuming that any such difference detected empirically reflects a genuine structural distinction rather than an artifact of the particular country grouping used.

2.8 Positioning of the present study

Taken together, this literature motivates five distinct compositional dimensions along which R&D spending might plausibly affect absorptive capacity and catch-up speed: sector of performance and funding source (Sections 2.3-2.4), type and field of research (Section 2.5), and the capital-versus-current expenditure split (Section 2.6), with the CEE-specific innovation systems literature (Section 2.7) motivating an explicit test for group heterogeneity in each relationship. To our knowledge, no existing study has jointly tested all five dimensions against a common Nelson-Phelps-style catch-up specification in a single, current, EU-wide panel. This is the gap the present paper addresses; equally, by systematically testing each resulting finding against outlier exclusion and data-driven country regrouping, we aim to demonstrate that a compositional finding motivated by this literature is not, on its own, sufficient evidence of a genuine effect without such scrutiny.

> **KEY FINDING**
>
> The literature motivates five compositional dimensions — sector of performance, funding source, type of research, field of research, and the capital-versus-current split — each with its own plausible channel to absorptive capacity.

03 Data

3.1 Sample and variables

The dataset covers all 27 current EU member states over 1995-2024 (unbalanced panel; coverage varies by variable, see Section 3.2). All series are sourced from Eurostat. The dependent variable, growth, is the year-over-year log change in labour productivity (constant 2019 EUR). The frontier-distance variable, gap, is defined as one minus labour productivity expressed as a percentage of the EU frontier level, so that gap equals zero at the frontier and rises with distance behind it. R&D intensity, rd_gva, is gross domestic R&D expenditure as a percentage of gross value added.

R&D composition is captured across five dimensions, each expressed as a share of total R&D expenditure: (i) sector of performance (business enterprise, government, higher education); (ii) funding source (business enterprise, government, higher education, private non-profit, rest of world); (iii) type of research (basic, applied, experimental development); (iv) field of research (natural sciences, engineering and technology, medical and health sciences, agricultural sciences, social sciences, humanities); and (v) expenditure type (current versus capital expenditure). All composition variables and the R&D intensity variable enter the regressions lagged by one year to reduce simultaneity concerns.

Variable	Definition	Source
growth	Year-over-year log change in labour productivity (2019 EUR)	Eurostat
gap	1 - (labour productivity as % of EU frontier)	Eurostat, authors' calculation
rd_gva	R&D expenditure as % of gross value added	Eurostat
sector_*	R&D performed by business / government / higher education, % of total	Eurostat
fund_*	R&D funded by business / government / higher education / private non-profit / rest of world, % of total	Eurostat
type_*	R&D by basic / applied / experimental development, % of total	Eurostat
field_*	R&D by field of science, % of total	Eurostat
exp_current, exp_capital	Current vs. capital R&D expenditure, % of total	Eurostat

3.2 Data quality and coverage

Romania, Czechia, Hungary and Poland have near-complete coverage (24-30 of 30 sample years) across every variable block. Coverage is markedly weaker for the type-of-research and field-of-research blocks in several EU-15 economies: Germany, Finland, France, Belgium, Austria, Ireland and Sweden report 0-2 usable years each for these two blocks, meaning results for those dimensions are estimated predominantly from Central, Eastern and Southern European observations. Eurostat's missing-value placeholder in the R&D/GVA series required explicit handling; left uncorrected it silently converts the affected column to a categorical type, which materially corrupts ordinary least squares estimation. All results reported below reflect the corrected series. We checked systematically for the same placeholder across every variable block used in this paper, sector of performance, funding source, type of research, field of research, and current/capital expenditure, in addition to R&D/GVA and the frontier-distance series; the placeholder occurred only in the R&D/GVA series (35 country-year cells), and no other block required correction.

Ireland and Luxembourg have negative average gap values over the sample period, meaning their measured labour productivity exceeds the EU frontier benchmark used here. This is a well-documented statistical artifact arising from multinational profit-shifting and its effect on headline GDP and productivity statistics in both economies, rather than genuine productivity leadership. Their treatment is addressed directly in the robustness protocol (Section 4.3).

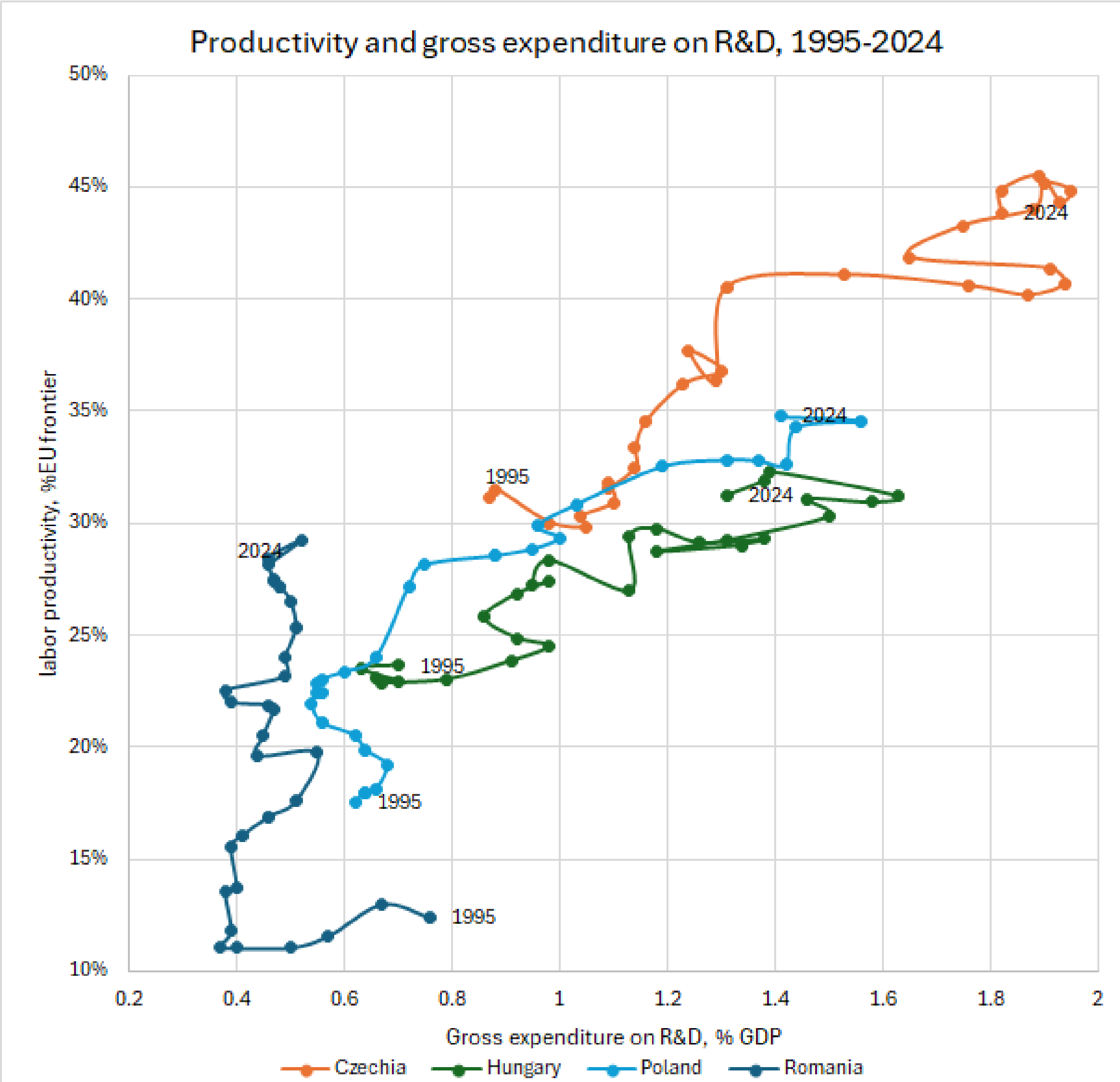
3.3 CEE in depth

Before turning to the econometric results, this section examines the raw R&D intensity and compositional trends for the four countries at the empirical centre of the paper. The purpose is threefold: to show that R&D intensity alone traces a strikingly loose relationship with relative productivity across the four countries, foreshadowing the weak intensity effect reported in Section 5.1; to document substantial cross-country structural differences across all five composition dimensions that motivate treating R&D composition as more than a single scalar; and to identify features of the underlying time series, in particular a shared hump around the 2007-2013 EU multiannual financial framework, that are relevant to interpreting the regression results in Section 5.

3.3.1 R&D intensity and productivity: joint trajectories, 1995-2024

The evolution of R&D spending together with the labour productivity over three decades in CEE countries (Romania, Czechia, Hungary and Poland) reveals different patterns. All four countries moved up and to the right over the period, meaning they simultaneously increased R&D spending as a share of GDP and closed (part of) the productivity gap with the most productive EU economies. But the starting points, paths, and end points differ substantially.

FIGURE 1



Source: own computations based on Eurostat data

Plotting each country's R&D intensity against its relative productivity level, year by year, makes visible a relationship that a simple correlation of levels would obscure: the path each country actually took was neither smooth nor monotonic in either dimension, and the four countries occupy strikingly different regions of this space by 2024 despite all four starting under 1% of GDP in 1995.

Czechia begins 1995 already ahead of the other three on both dimensions (0.87% of GDP, 31% of frontier) and ends 2024 far ahead of them on both (1.8% of GDP, 44% of frontier), but the path between the two is far from a straight line: the trajectory shows repeated hairpin reversals, including a pronounced pullback in R&D intensity around 2015-2019 that coincides with little corresponding loss in relative productivity, consistent with the post-2015 decline in capital-expenditure share documented in Section 3.3.2. Poland's path is comparatively smoother, rising from approximately 0.6% of GDP and 18% of frontier in 1995 to 1.4% of GDP and 35% of frontier by 2024, interrupted by a well-documented dip in R&D intensity around 2001-2007 to roughly 0.55-0.6% of GDP. Hungary's path contains the most pronounced reversal of the four: relative productivity rises to approximately 24% of frontier by around 0.7% of GDP, then falls back to approximately 27% even as R&D intensity continues rising toward 1.1-1.2% of GDP, a stall consistent with the 2008-2009 financial crisis and subsequent fiscal consolidation in Hungary, before resuming its climb to end at approximately 1.3% of GDP and 31% of frontier.

Romania is the most consequential case for this paper's argument. Its 1995 starting point, approximately 0.75% of GDP, is comparable to or higher than Poland's or Hungary's starting intensity, but R&D intensity then collapses through the late 1990s and 2000s to lows near 0.35-0.4% of GDP, consistent with the dismantling of inherited state research institutes documented in Section 3.3.3, and remains below 0.55% of GDP for essentially the entire sample thereafter. By 2024, Romania's R&D intensity, approximately 0.5% of GDP, is lower than its own 1995 starting value, the only one of the four countries to end the sample less R&D-intensive than it began. Yet Romania's relative productivity rises from approximately 12% of frontier in 1995 to approximately 28% by 2024, more than doubling and the largest proportional gain of the four countries. This combination, the largest relative productivity gain achieved with a falling rather than rising R&D intensity, is difficult to reconcile with a story in which R&D intensity itself drives catch-up, and is consistent with Section 5.1's finding that R&D intensity carries only a weak, marginally significant relationship with catch-up speed once composition is controlled for. It suggests Romania's convergence has leaned more heavily on channels external to domestic R&D effort, plausibly capital-embodied technology transfer, FDI-linked manufacturing integration, and EU accession itself, than on deepening its own research base, a possibility the composition-specific evidence in the remainder of this section and in Section 5 bears out.

3.3.2 Features of R&D expenditures

- ***R&D expenditure split in current expenses and capital expenses***

All four countries show a pronounced hump-shaped path in the capital-current expenditure ratio, rising through the late 2000s to a peak around 2011-2013 and declining thereafter. Poland's capital share is consistently the highest of the four throughout the sample, rising from roughly 24-30% in the mid-1990s to a peak near 42% in 2011, before falling back to the low twenties by the end of the series. Hungary follows a similar but lower-amplitude path, peaking near 34% in 2001 and again near 22-26% in the early 2010s. Czechia's capital-current expenditure ratio was comparable to Hungary's for most of the 1995-2010 period before rising sharply to a peak near 30% around 2011-2015 and then falling to single digits (roughly 9%) by 2023, the steepest post-peak decline of the four. Romania is both the lowest and the most volatile series: starting near 6-7% in the mid-1990s, it shows two distinct spikes, one around 2007 (approximately 24%) and a larger one around 2011 (approximately 22%), before falling back to single digits for much of 2017-2021 and recovering modestly to around 13% by 2023.

The shared timing of the post-2011 peak across all four countries is consistent with the absorption profile of 2007-2013 EU cohesion and structural funds, a substantial share of which was directed toward research infrastructure and equipment in new member states, followed by a slower-disbursing 2014-2020 framework and the disruption of the COVID-19 period. This shared timing is one reason Section 5.4 tests the paper's central finding explicitly against crisis-period and EU-funding-cycle-sensitive specifications (full year fixed effects, crisis-year exclusion) rather than relying on the baseline specification alone.

FIGURE

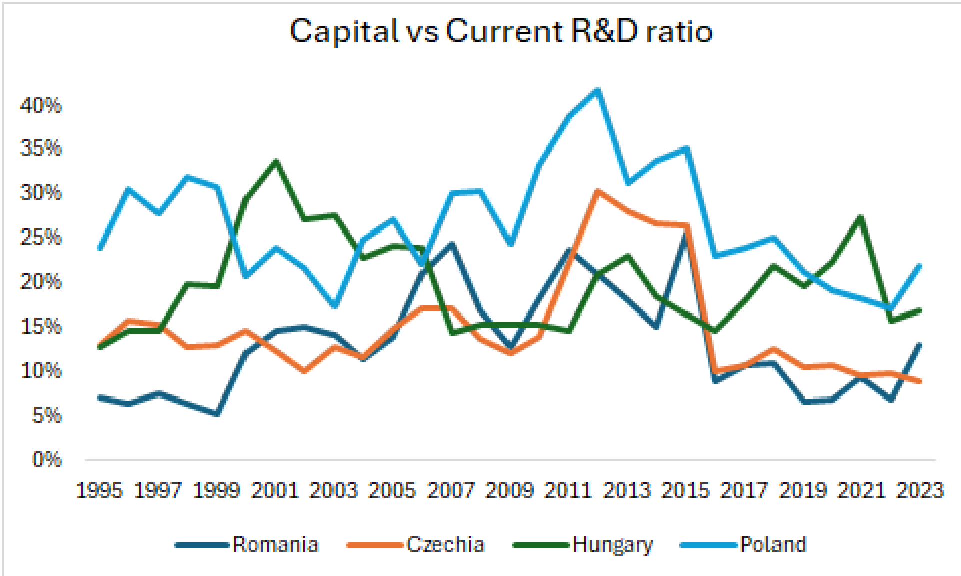


Figure 2. R&D capital-current expenditure ratio, Romania, Czechia, Hungary and Poland, 1995-2023.

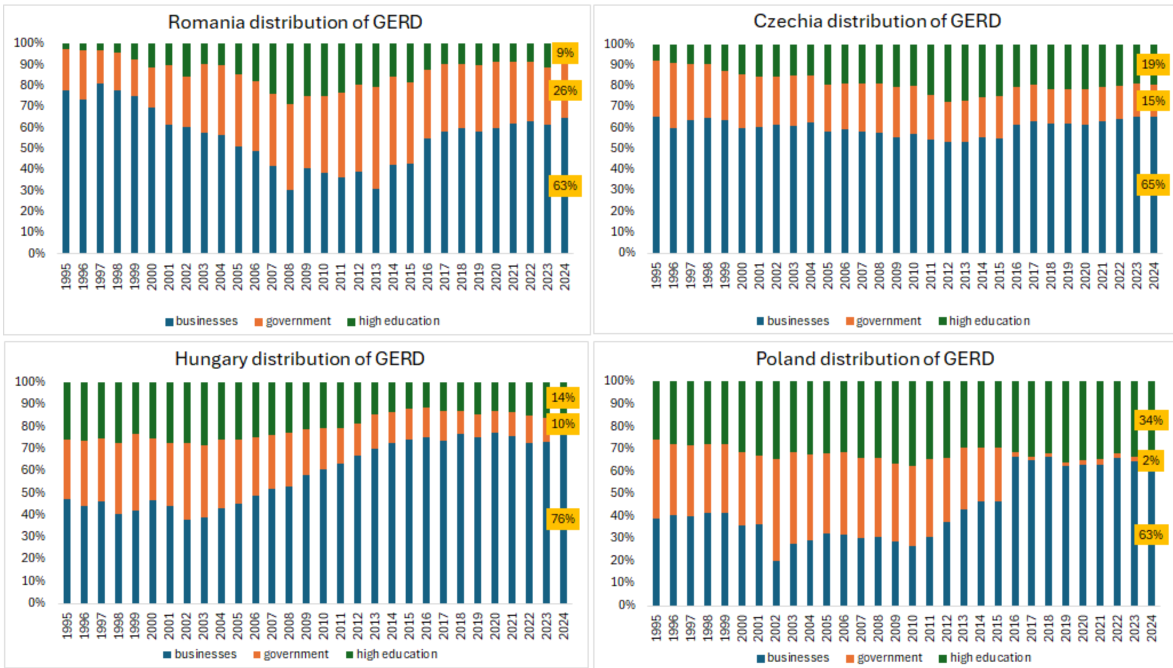
Source: Eurostat, Gross domestic expenditure on R&D split in current and capital expenses

R&D expenses by sectors of performance

The four countries display markedly different trajectories in who performs R&D. Czechia is the only one of the four with a consistently business-dominated R&D system throughout the sample, with businesses performing 60-75% of GERD in every year and ending at roughly 65% business, 15% government, 19% higher education in 2024. Romania shows a pronounced U-shape: the business share collapsed from around 80% in the mid-1990s, a legacy of centrally planned industrial research institutes, to a low of roughly 35-40% around 2008-2013 as that inherited structure was dismantled, before recovering to approximately 61% by 2024, with government absorbing much of the remaining share (26%) and higher education a smaller residual (9%).

Hungary and Poland both show a steady, near-monotonic rise in the business-performed share over the full sample, converging toward, and in Hungary's case exceeding, Czechia's level: Hungary rises from approximately 44% in 1995 to roughly 76% by 2024, while Poland rises from a low of around 18% in the early 2000s to approximately 63% by 2024. Poland's endpoint composition is structurally distinct from the other three, however: its higher-education-performed share (approximately 34%) is far larger than Romania's (9%), Czechia's (19%), or Hungary's (14%), while its government-performed share has fallen to a residual 2%, implying that Poland's R&D system has effectively reallocated the role played by government research institutes elsewhere in the region to its university sector rather than to business.

FIGURE 3



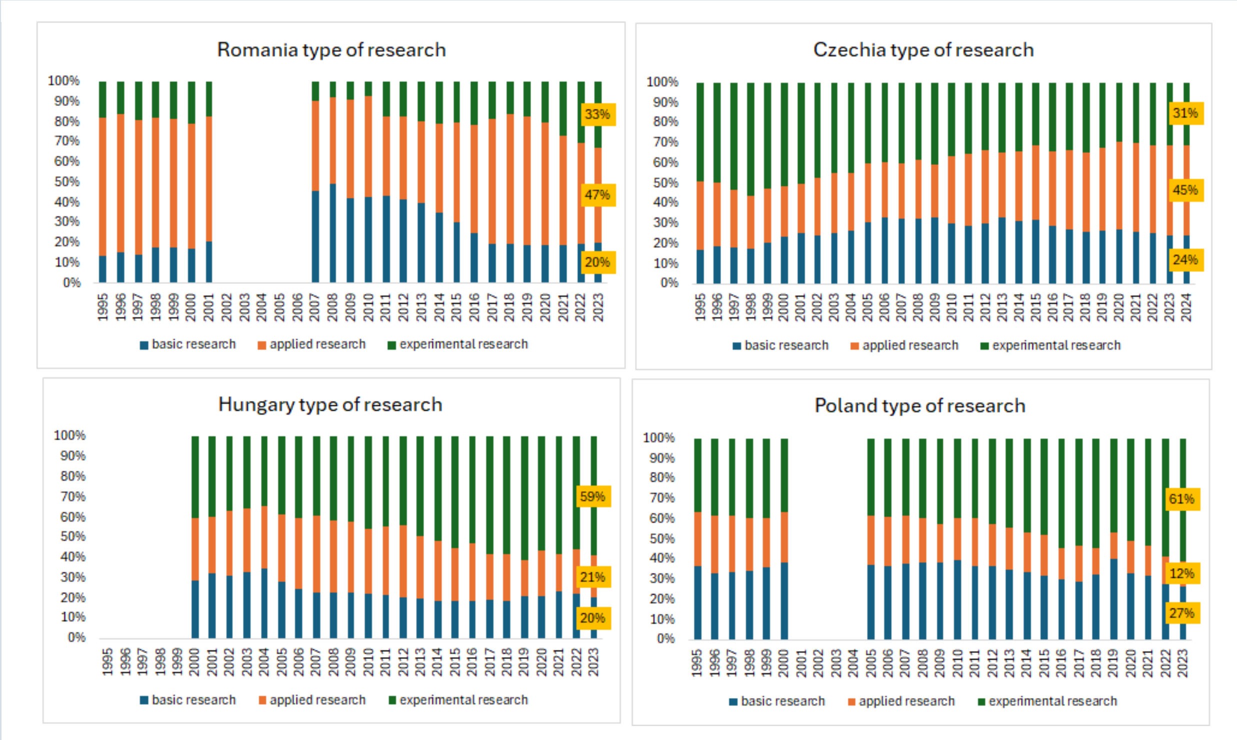
Source: Eurostat, Gross domestic expenditure on R&D split by sector of performance

▸ *R&D by type of research*

Romania's series has a visible gap for 2002-2006, consistent with the coverage limitation noted in Section 3.2. Where data exist, two distinct structural patterns emerge. Czechia and Romania both show a relatively balanced split weighted toward applied research: Czechia ends 2024 at approximately 24% basic, 45% applied, 31% experimental, while Romania ends 2023 at approximately 20% basic, 47% applied, 33% experimental, the two most similar profiles of the four on this dimension. Hungary and Poland, by contrast, have shifted heavily toward experimental development: Hungary ends at approximately 20% basic, 21% applied, 59% experimental, and Poland at approximately 27% basic, 12% applied, 61% experimental, giving Poland both the lowest applied-research share and the highest experimental-development share among the four countries.

This applied-versus-experimental divide broadly tracks the sector-of-performance divide in Section 3.3.3: Hungary and Poland, the two countries with the steepest rise in business-performed R&D, are also the two with the heaviest tilt toward experimental development, the type of research most closely associated with near-market, firm-level product and process development, while Czechia's more stable, longer-standing business-R&D base retains a comparatively larger applied-research component.

FIGURE 4



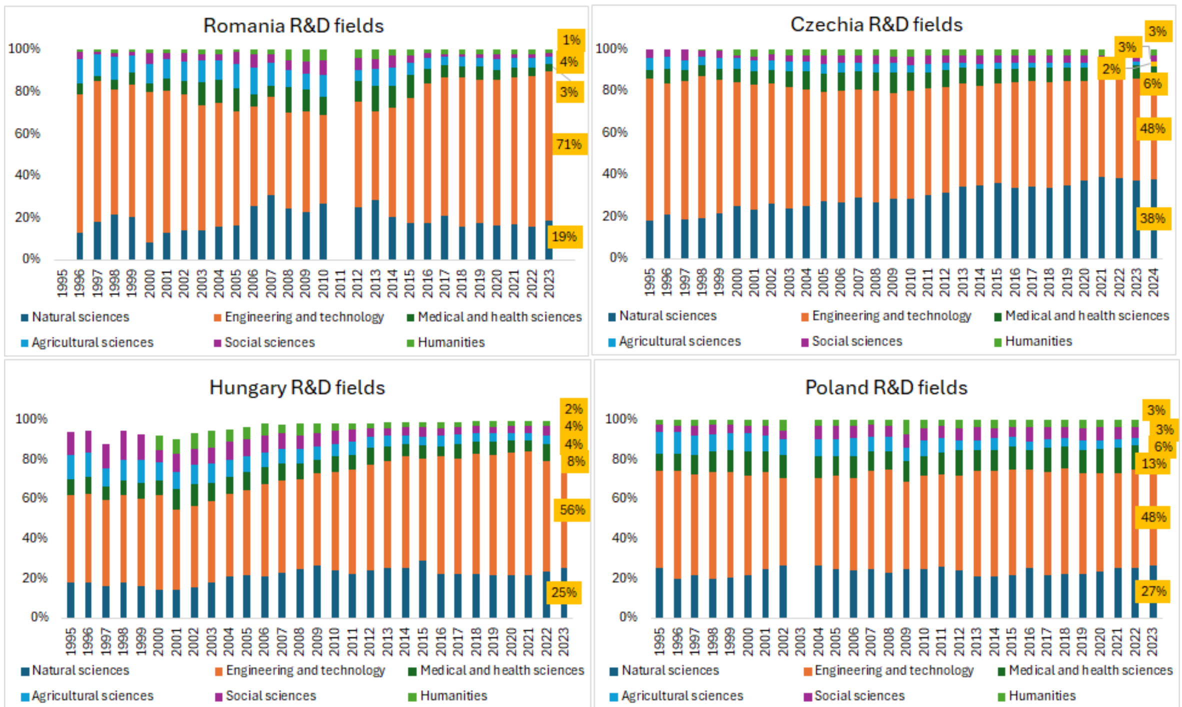
Source: Eurostat, Gross domestic expenditure on R&D split by type of research

R&D expenses by research field

All four countries are engineering-and-technology dominated, consistent with the region's manufacturing-oriented industrial base, but the degree of concentration varies substantially. Romania is the most concentrated by far, with engineering and technology accounting for approximately 71% of R&D expenditure by 2023 and natural sciences a distant second at 19%, leaving little room for medical, agricultural, social science, or humanities research (each in low single digits). Hungary (56% engineering and technology, 25% natural sciences, 8% medical and health sciences) and Poland (approximately 48% engineering and technology, 27% natural sciences) are less concentrated, with Hungary in particular retaining a visibly larger medical and health sciences allocation than the other three. Czechia is the least concentrated of the four: 49% engineering and technology and 37% natural sciences, the smallest gap between the two largest fields among the four countries, alongside small but non-trivial agricultural, social science and humanities shares.

Hungary's earliest observations (mid-to-late 1990s) show a markedly larger social sciences share than any other country or period in the sample, a residual feature of its research infrastructure not shared by Romania, Czechia or Poland, which had largely disappeared by the mid-2000s.

FIGURE 5



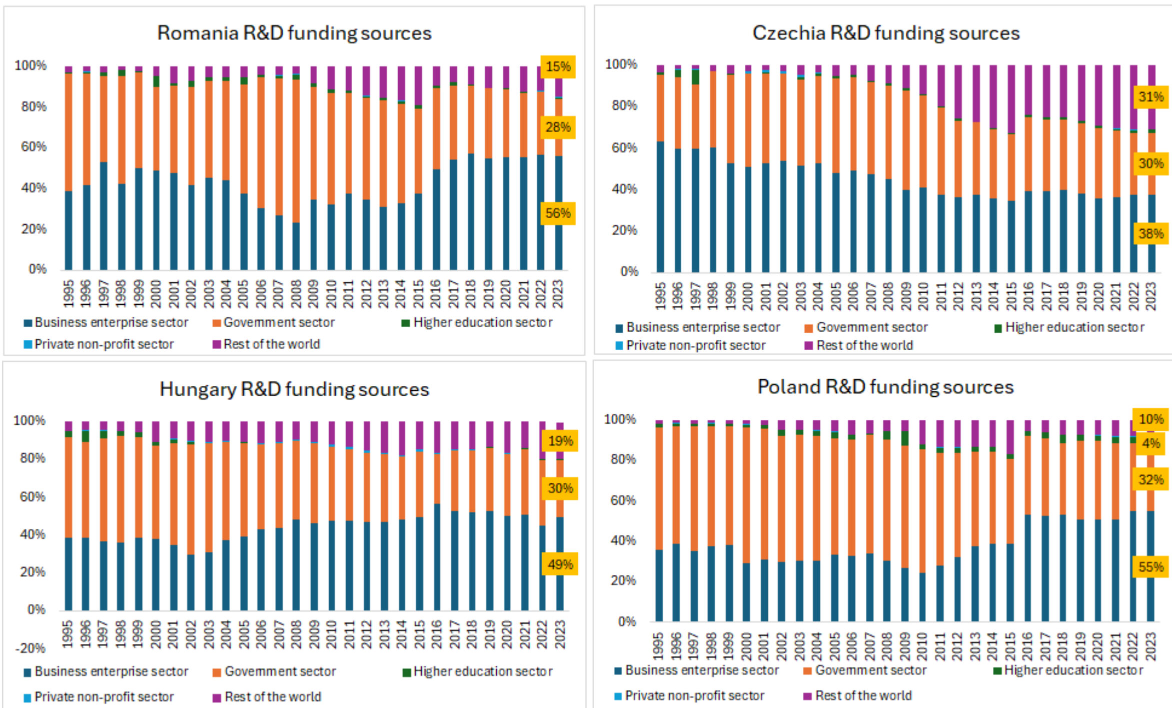
Source: Eurostat, Gross domestic expenditure on R&D split by fields of research

R&D expenses by funding sources

Funding-source composition reveals a pattern not visible in the sector-of-performance data alone. Czechia's business-enterprise funding share fell substantially over the sample, from roughly 60% in 1995 to approximately 38% by 2023, even as its business-performed share (Figure 3) stayed broadly stable, implying that a growing share of R&D actually performed by Czech businesses is funded from elsewhere. The counterpart is a rest-of-world funding share that rose to approximately 31% by 2023, the highest of the four countries, consistent with substantial multinational-affiliated R&D investment and EU fund inflows being channelled through business performers without originating in domestic business capital.

Romania's rest-of-world funding share also rose over the sample, reaching approximately 15% by 2023, alongside a business-funded share that recovered to roughly 56% and a government-funded share of approximately 28%. Hungary's funding structure at the endpoint (approximately 49% business, 30% government, 19% rest of world) sits between Czechia's and Romania's. Poland is the clear outlier on this dimension: its rest-of-world funding share remains low throughout, around 10% by the endpoint, the smallest of the four, while business funding (approximately 55%) and government funding (approximately 32%) account for nearly all the remainder. Combined with Poland's unusually large higher-education-performed share noted above, this suggests Poland's R&D system is comparatively more domestically self-contained, funded and performed with less reliance on foreign capital or multinational R&D centres, than Czechia's or Romania's.

FIGURE 6



Source: Eurostat, Gross domestic expenditure on R&D split by funding sources

3.3.7 Synthesis

Table 1 summarizes the endpoint value of each dimension for the four countries. The R&D intensity-productivity trajectories in Section 3.3.1 already suggest that raw R&D effort is a poor guide to catch-up outcomes on its own, most starkly in Romania's case, and the five composition dimensions in Sections 3.3.2-3.3.6 point toward the same conclusion from a different angle: two broad groupings emerge, though neither aligns perfectly with the far-tercile grouping used in the econometric analysis (Section 4.3), a reminder that structural R&D composition and frontier distance are related but distinct axes of variation. Hungary and Poland resemble one another most closely: both show a steep, near-monotonic rise in business-performed R&D, a heavy tilt toward experimental development, and comparatively lower applied-research shares. Czechia stands apart as the most structurally stable of the four, with a business-performed R&D base already well established at the start of the sample and a more balanced type-of-research and field-of-research profile than the other three. Romania is the outlier on multiple dimensions simultaneously: the most engineering-concentrated field profile, the only country with a clear U-shaped sector-of-performance trajectory, a falling rather than rising R&D intensity over the full sample, and, alongside Czechia, comparatively higher capital-expenditure volatility around the 2007-2013 EU funding cycle. Poland is distinguished less by trajectory than by its endpoint institutional structure, an unusually large higher-education-performed share and an unusually low reliance on foreign R&D funding, that sets it apart from the other three even where its type-of-research and capital-expenditure trends resemble Hungary's.

Indicator (latest available year)	Romania	Czechia	Hungary	Poland
Gross R&D expenditure, % of GDP	~0.5%	~1.8%	~1.3%	~1.4%
Labour productivity, % of EU frontier	~28%	~44%	~31%	~35%
Capital/current expenditure ratio	13%	9%	17%	22%
Business-performed share of GERD	63%	65%	76%	63%
Higher-education-performed share of GERD	9%	19%	14%	34%

Business-funded share of R&D	56%	36%	49%	55%
Rest-of-world-funded share of R&D	15%	31%	19%	10%
Applied research share	47%	45%	21%	12%
Experimental development share	33%	31%	59%	61%
Engineering and technology field share	71%	49%	56%	48%

Note: figures are approximate values read from the underlying Eurostat-sourced series for the most recent available year per country (2023 for Romania, 2024 for Czechia, Hungary and Poland); the R&D intensity and productivity rows are read from Figure 1. All figures are reported here for descriptive purposes. The econometric analysis in Section 5 uses the full annual time series, one-year-lagged, rather than these single-year endpoints; in particular, the sample-average frontier gap used in Section 6 (Czechia 0.62, Hungary 0.72, Poland 0.74, Romania 0.80) differs from the 2024-endpoint productivity figures shown here because it is averaged over the full 1995-2024 period rather than measured at a single point in time.

KEY FINDING

Romania, Czechia, Hungary and Poland have near-complete coverage across every variable, while Ireland and Luxembourg carry productivity series distorted enough by multinational profit-shifting to require exclusion.

04 Empirical Strategy

4.1 Baseline specification

The baseline estimating equation, for country i and year t, is:

```
growth_it = α + β1 gap_i,t-1 + β2 rdgva_i,t-1 + β3 (gap_i,t-1 × rdgva_i,t-1) + β4
X_i,t-1 + β5 (gap_i,t-1 × X_i,t-1) + μ_i + ε_it
```

where X is an R&D-composition share (one dimension at a time, to limit multicollinearity among mechanically related shares within a block), μ_i is a country fixed effect, and standard errors are clustered by country. The coefficient of interest is β5: a positive estimate indicates that a higher share of X is associated with faster catch-up the further a country sits from the frontier, consistent with a Nelson-Phelps-style diffusion-acceleration channel specific to that composition variable.

4.2 Testing for group heterogeneity

A specification pooling all 27 countries under a common β5 implicitly assumes the catch-up mechanism operates identically in, for instance, Germany and Romania. We test this assumption by adding a group indicator G interacted with X and its gap-interaction:

```
growth_it = α + β1 gap_i,t-1 + β2 rdgva_i,t-1 + β3 (gap_i,t-1 × rdgva_i,t-1) + β4
X_i,t-1 + β5 (gap_i,t-1 × X_i,t-1) + β6 (G_i × X_i,t-1) + β7 (G_i × gap_i,t-1 ×
X_i,t-1) + μ_i + ε_it
```

A significant β7 indicates that the catch-up slope on X differs systematically between group G and the reference group.

4.3 Robustness protocol

Each candidate composition variable that reached significance in the baseline specification was subjected to four tests: (i) the heterogeneity specification of Section 4.2 with G defined as the historical post-2004 EU-accession grouping ("CEE": Romania, Czechia, Hungary, Poland, Bulgaria, Estonia, Latvia, Lithuania, Croatia, Slovakia and Slovenia. Malta and Cyprus, though also 2004 entrants, are Mediterranean rather than Central or Eastern European economies and are excluded from the CEE indicator, remaining in the reference group throughout); (ii) the same specification with Ireland and Luxembourg excluded from the sample; (iii) the same specification with G redefined as a data-driven indicator equal to one for the top ⌈25/3⌉ = 9 of the 25 remaining countries with the highest average gap over the sample period, and zero otherwise; and (iv) a fully pooled specification with no group interaction at all, on the outlier-excluded sample. A finding is treated as robust only if the group-heterogeneity result (or, in the absence of heterogeneity, the pooled result) remains significant across specifications (ii) through (iv). Because a substantially larger set of composition sub-variables was screened at the baseline stage than the small number carried into this protocol, Section 5.3 reports this multiple-testing exposure explicitly rather than treating baseline significance alone as evidence.

> **KEY FINDING**
>
> Every candidate result must clear the same four-part protocol: exclusion of the distorted outlier economies, data-driven rather than historical country groupings, lag lengths of one to three years, and three separate treatments of the crisis periods.

05 Results

5.1 R&D intensity

R&D intensity itself (rd_gva) shows a marginal, weakly significant interaction with frontier distance (coefficient -0.028, p = .046) in the baseline 27-country specification, with the level term insignificant. This motivates its inclusion as a control throughout rather than as a headline result: composition effects reported below are estimated net of this weak intensity effect.

5.2 Capital expenditure share

Table 2 reports the full progression of estimates for capital-expenditure share, from the pooled 27-country baseline through each stage of the robustness protocol.

Specification	N	Level	× gap	Group×X	Group×gap×X	R^2
(1) Pooled, 27 countries, no group	642	0.045** (.016)	-0.027 (.045)	-	-	0.184
(2) Historical CEE grouping, 27 countries	642	0.053*** (.015)	-0.048 (.044)	-1.174*** (.200)	1.704*** (.291)	0.238

(3) Historical CEE grouping, IE/LU excluded	611	-0.005 (.085)	0.105 (.196)	-0.832*** (.205)	1.136*** (.315)	0.269
(4) Data-driven far-tercile grouping, IE/LU excluded	611	-0.023 (.086)	0.073 (.193)	-0.694*** (.195)	1.003*** (.298)	0.266

Notes: cluster-robust standard errors in parentheses, clustered by country. All specifications include R&D intensity, its interaction with gap, and country fixed effects (not shown). R^2 is the overall (between- and within-country) coefficient of determination; the within-country R^2 for specification (4), net of the explanatory power of the country fixed effects themselves, is 0.151. *** p<.01, ** p<.05, * p<.10.

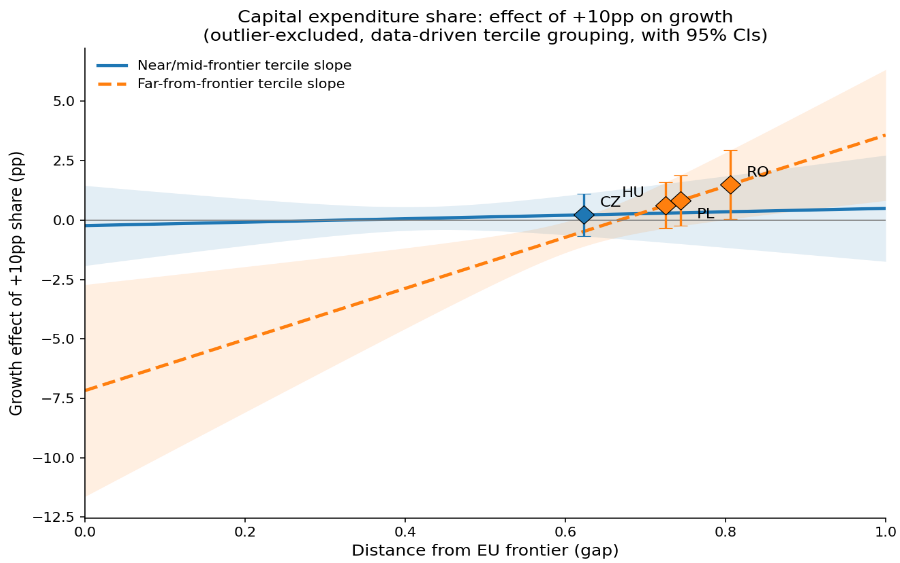
The group-interaction terms are significant at the 1% level in specifications (2) through (4), regardless of whether the historical or data-driven grouping is used and regardless of whether Ireland and Luxembourg are included. This is, to our knowledge, the most robust single result in the paper: economies in the far-from-frontier group derive a significantly larger growth benefit from a capital-heavy R&D composition, relative to the frontier group, than the pooled model alone would suggest.

With only 25 clusters in specification (4) and 27 in specification (2), asymptotic cluster-robust inference sits below the conventional comfort threshold discussed in Section 7. We address this directly for the headline result with a restricted-residual wild-cluster bootstrap (Cameron, Gelbach and Miller, 2008), using Rademacher weights and 999 replications, applied to the Group×gap×X coefficient. The bootstrap p-value is 0.004 for specification (4) and below 0.0001 for specification (2), against asymptotic cluster-robust p-values of 0.00077 and below 0.0001 respectively. The headline capital-expenditure result is therefore robust to small-cluster correction, not merely to the asymptotic standard errors reported in Table 2. The data-driven grouping in specification (4) assigns the top $\lceil 25/3 \rceil = 9$ of the 25 remaining countries, ranked by 1995-2024 average frontier gap, to the far group; the remaining 16 are pooled as the reference group rather than split into three separate terciles. This threshold is a genuine researcher degree of freedom, so we test sensitivity to coarser splits: replacing the top-9 tercile cut with a top-6 (quartile) or top-5 (quintile) cut, keeping the same formula and sample. The Group×gap×X coefficient falls from significant at the 1% level under the tercile cut (p<.001) to marginal under the quartile cut (p=.094-.098 depending on whether 6 or 7 countries are used to approximate a quarter of 25) and insignificant under the quintile cut (p=.24-.28). This attenuation is consistent with reduced statistical power from a smaller treated group (5-7 countries rather than 9) rather than evidence against the underlying relationship, but it means the headline result's statistical significance is specific to a three-way split of the sample and should not be read as robust to arbitrarily coarse groupings of the far-from-frontier economies.

Because the far-tercile grouping is constructed from the same productivity data used elsewhere in the specification, we additionally implement a randomization-inference check (Fisher, 1935) directly on the grouping itself. We randomly reassign 9 of the 25 non-outlier countries to the far group 2,000 times, holding the rest of specification (4) fixed, and re-estimate the Group×gap×X coefficient under each random grouping. The actual data-driven grouping's coefficient (1.003) exceeds all but 8 of the 2,000 random groupings in absolute value, a randomization-inference p-value of .004, and sits at the 99.7th percentile of the permutation distribution, whose mean is close to zero (0.002) as expected under a genuinely uninformative grouping. This is reassuring on the specific concern that the grouping's construction from the dependent variable's own history could mechanically generate a large coefficient: an arbitrary same-sized grouping essentially never does.

We also re-estimate specification (4) 25 times, each time excluding one country, to confirm the result is not driven by any single economy. The Group×gap×X coefficient ranges from 0.859 (excluding Bulgaria) to 1.197 (excluding Poland) across all 25 leave-one-out re-estimates, remaining significant at the 1% level in every case (maximum p-value across all 25 re-estimates: .005). No single country, including Romania, whose implied effect is the largest in Table 4, is responsible for the headline result.

FIGURE



5.3 Other composition variables

Table 3 summarizes the robustness protocol outcome for the remaining four composition variables that reached significance in some specification. These four were drawn from a substantially larger set of R&D-composition sub-variables, eighteen mutually distinct shares across the five composition dimensions described in Section 3.1, screened at the baseline stage; we did not apply a formal multiple-testing correction at that screening step. Re-estimating the group-heterogeneity specification separately for all eighteen sub-variables shows that considerably more than the roughly one hit expected by chance alone under a global null at the 5% level reach nominal significance, and that several of these do not survive a Benjamini-Hochberg correction at a 5% false discovery rate. This is precisely why the four-stage robustness protocol below, rather than baseline significance itself, is this paper's standard of evidence, and why Table 3's verdicts, not the raw baseline p-values, should be read as the paper's findings.

Variable	(i) Historical	(ii) IE/LU excl.	(iii) Tercile, IE/LU excl.	(iv) Pooled	Verdict
Business-performed R&D share	p = .004	p = .27	p = .22-.44	p = .63-.87	Not robust
Government-funded R&D share	p < .001	p = .054/.005	p = .10/.02	p = .38	Not robust (see Section 5.4)
Higher-education-funded R&D share	p = .002	p = .89	p = .84	p = .89	Not robust
Social sciences field share	p = .44	p = .38	p = .70	p = .18	Never significant

Business-performed R&D share exhibited a significant heterogeneity result under the historical grouping (p = .004) that was not present under either outlier-exclusion or data-driven regrouping. The most plausible interpretation is that Ireland's and Luxembourg's distorted productivity statistics, both classified under the historical grouping as "EU-15" alongside genuine frontier economies, generated a spurious group contrast. Higher-education-funded R&D share shows an even sharper pattern: not merely the group contrast but the entire pooled relationship (specification iv) loses significance once the two outlier economies are removed, indicating the original result was driven almost entirely by those two observations rather than by any broader EU-15/CEE structural difference. Social sciences field share was not significant under any specification, including the original, and should not have been treated as a candidate finding in the first analytical pass; we retain it in Table 3 for completeness and as a methodological illustration.

Government-funded R&D share initially appeared to occupy an intermediate position: the slope-difference term (group × gap × X) remained significant at the 5% level under the strictest data-driven specification ($p = .02$), while the level-difference term did not ($p = .10$) and the fully pooled coefficient with no grouping was insignificant ($p = .38$). Section 5.4 subjects this surviving term to two further checks, lag-length variation and crisis-period controls, that it does not pass: it loses significance under two of three crisis-period treatments. We therefore retract this result as well, rather than retaining it with a low-confidence caveat as an earlier version of this analysis did.

Type of research (basic, applied, experimental development) and the remaining fields of research (engineering and technology, natural sciences, medical and health sciences, agricultural sciences) did not reach significance in the baseline specification for any variant tested and were not carried into the robustness protocol.

5.4 Additional robustness checks: lag length and crisis-period controls

The two variables surviving to this stage, capital-expenditure share (Section 5.2) and the single surviving government-funded share coefficient (Section 5.3), were subjected to two further checks not part of the original four-specification protocol: sensitivity to the choice of a one-year lag, and sensitivity to the 2008-2009 global financial crisis and 2020-2021 COVID-19 shock, both of which coincide with unusual swings in productivity growth and, potentially, R&D composition across the sample.

Lag-length sensitivity re-estimates the data-driven far-tercile specification (Table 2, column 4, and its government-funded-share equivalent) with the composition variable and frontier-gap lagged one, two, and three years.

Variable (slope-difference term)	Lag 1	Lag 2	Lag 3
Capital expenditure share	$p < .001$	$p < .001$	$p = .014$
Government-funded share	$p = .022$	$p = .007$	$p = .003$

Both results are stable across lag length; neither finding is an artifact of the specific one-year lag used elsewhere in the paper.

Crisis-period sensitivity applies three separate treatments to the data-driven far-tercile specification: adding year dummies for the crisis years, replacing country fixed effects with full country-and-year fixed effects (a stronger control that also absorbs EU structural-fund disbursement cycles), and re-estimating on a sample with crisis years dropped entirely.

Variable (slope-difference term)	+ Crisis dummies	+ Full year FE	Crisis years excluded
Capital expenditure share	p = .003	p < .001	p < .001
Government-funded share	p = .28 (not sig.)	p = .009	p = .33 (not sig.)

Capital-expenditure share is significant under all three crisis-period treatments and is, at this point, the only composition variable in the paper to survive every robustness dimension tested: historical and data-driven grouping, outlier exclusion, three lag lengths, and three treatments of the two major crisis episodes in the sample period. Government-funded share fails two of the three crisis-period checks, surviving only under the most aggressive control (full year fixed effects). Combined with its already-marginal status in Section 5.3, we interpret this as evidence that the one term which appeared to survive the original robustness protocol was likely an artifact of unusual co-movement between government R&D funding shares and productivity growth during the financial crisis in several sample countries, rather than a genuine structural relationship. We retract this result.

5.5 Granger-style precedence check

One-year lagging of the composition regressor reduces mechanical simultaneity but does not rule out reverse causality, in which growth shocks themselves drive subsequent R&D-composition decisions, or a common factor driving both series together. We do not have a source of exogenous variation in capital-expenditure share suitable for instrumental-variables estimation, so we instead implement two precedence checks in the spirit of Granger (1969), which test the timing of the relationship rather than establishing causal identification.

The first checks whether past growth predicts future capital-expenditure share, which would indicate the reverse-causality channel the referee process flagged as a concern: fast-growing economies mechanically reallocating R&D toward or away from capital spending. We regress capital-expenditure share on its own lag, lagged frontier gap, and one-year-lagged growth, with country fixed effects and cluster-robust standard errors: the coefficient on lagged growth is 0.041 (se = 0.069, $p = .55$, N = 525), not distinguishable from zero. Past growth does not appear to drive subsequent capital-expenditure allocation.

The second adds the one-year-ahead (lead) value of capital-expenditure share to the full specification (4) alongside the existing one-year lag, testing whether growth is already anticipating a composition shift that has not yet occurred, a pattern that would be difficult to reconcile with the lagged term alone driving the relationship. With both lag and lead included, the lead term is not significant (coefficient = 0.036, se = 0.024, $p = .14$, N = 556, four country-years lost per country relative to specification 4 owing to the lead), while the headline Group×gap×X interaction remains significant at the 1% level (coefficient = 0.837, se = 0.313, $p = .008$), close to its magnitude in specification (4) itself (1.003). Neither check rules out reverse causality with certainty, since a Granger-style precedence test can never fully distinguish a genuine causal channel from an unobserved factor moving both series on a compatible timeline, but both point in the same reassuring direction, and the headline result is not sensitive to including a forward-looking control that would be expected to absorb much of any anticipation-driven contamination.

We also explored whether the funding-source breakdown available in the underlying data could support an instrumental-variables approach. The share of R&D funded directly by the European Commission (competitive Framework Programme grants, distinct from the government-channelled Cohesion and

Structural Funds) is a candidate: it correlates with capital-expenditure share ($r = 0.39$) and has a statistically significant first-stage coefficient ($p = .004$) in a country-fixed-effects regression of capital-expenditure share on this instrument, gap, and R&D intensity. However, the cluster-robust first-stage F-statistic is approximately 8, below the conventional threshold of 10 associated with reliable asymptotic inference (Staiger and Stock, 1997), and the instrument correlates non-trivially with the frontier-gap regressor itself ($r = 0.36$), which is a plausible violation of the exclusion restriction: European Commission funding formulas are partly a function of a country's relative income level, the same underlying force gap is designed to capture, and Framework Programme grants are competitively awarded on the basis of existing research capacity rather than mechanically assigned. Instrumenting the level term of capital-expenditure share with this variable in a simplified two-stage least squares specification (omitting the group interactions, for which additional excluded instruments would be required) yields a coefficient of -0.100 (se = 0.100, $p = .32$), reversing sign and losing significance relative to the OLS estimate in the same restricted sample, the pattern expected when a weak instrument amplifies sampling noise rather than sharpening identification (Staiger and Stock, 1997). We do not pursue this instrument further and report the exploration here in the same spirit as our treatment of retracted composition variables elsewhere in this paper: a negative methodological result is still worth recording.

5.6 Summary of the robustness battery

Table 5 collects every check applied to the headline capital-expenditure result across Sections 4, 5.2, 5.4 and 5.5, alongside its outcome, for a reader who wants the full battery at a glance rather than assembled from six subsections.

Table 5. Summary of all robustness checks applied to the capital-expenditure share result.

Check	Result	Verdict
Pooled baseline, 27 countries (Table 2, spec 1)	Level term significant, $p < .001$	Confirmed
Historical CEE grouping, full sample (Table 2, spec 2)	Group×gap×X significant, $p < .001$	Confirmed
Historical CEE grouping, IE/LU excluded (Table 2, spec 3)	Group×gap×X significant, $p < .001$	Confirmed
Data-driven far-tercile grouping (Table 2, spec 4, headline)	Group×gap×X significant, $p < .001$	Confirmed
Lag length 1/2/3 years (Section 5.4)	Significant across all three, $p < .001$ to $p = .014$	Confirmed
Crisis-period dummies / full year FE / crisis years excluded (Section 5.4)	Significant under all three treatments, $p = .003$ to $p < .001$	Confirmed
Wild-cluster bootstrap, Rademacher, B=999 (Section 5.2)	$p = .004$ (spec 4), $p < .0001$ (spec 2)	Confirmed
Quartile / quintile grouping sensitivity (Section 5.2)	Significant at tercile only; marginal at quartile ($p \approx .09$), insignificant at quintile ($p \approx .24$-$.28$)	Caveat: not robust to coarser splits
Randomization inference on grouping, 2,000 permutations (Section 5.2)	Observed coefficient at 99.7th percentile of permutation distribution, $p = .004$	Confirmed
Leave-one-country-out jackknife, 25 re-estimates (Section 5.2)	Significant in all 25 (max $p = .005$); no single country drives the result	Confirmed
Granger-style precedence check (Section 5.5)	Past growth does not predict future X ($p = .55$); lead X does not predict growth controlling for lag ($p = .14$)	Reassuring; not causal proof

Instrumental-variables exploration, EC funding share (Section 5.5)	First-stage F ≈ 8, below conventional weak-instrument threshold; not pursued	Not usable

KEY FINDING

Only the capital-expenditure share survives the full battery. The business-performed and government-funded results weaken under the data-driven grouping and are retracted rather than omitted.

06 Implications for Romania, Czechia, Hungary and Poland

Applying the coefficients from the most robust specification (Table 2, column 4) at each country's sample-average distance from the frontier (Czechia 0.62, Hungary 0.72, Poland 0.74, Romania 0.80) yields the projected association of a 10 percentage point increase in capital-expenditure share with productivity growth shown in Table 4; because the underlying estimates are drawn from an observational panel rather than a design with an explicit source of exogenous variation in R&D composition (Section 7), these figures are best read as scenario projections conditional on the estimated relationship rather than as causal policy effects. A 10pp change is not a hypothetical extreme: Romania's own capital-expenditure share moved from roughly 6% to roughly 24% of R&D spending over the sample period (Section 3.3.2), so a 10pp shift sits well within the range of variation the country has already experienced. Under the data-driven grouping, Czechia falls in the middle tercile rather than the far-from-frontier tercile that its historical EU-accession-date grouping would imply, which materially changes its projected coefficient relative to Hungary, Poland and Romania, all three of which remain in the far-from-frontier tercile.

Country	Frontier gap (sample avg.)	Tercile group	Implied effect of +10pp capital-expenditure share	95% CI
Czechia	0.62	Middle	+0.22 pp	[−0.68, 1.12]
Hungary	0.72	Far-from-frontier	+0.61 pp	[−0.34, 1.59]
Poland	0.74	Far-from-frontier	+0.79 pp	[−0.25, 1.89]
Romania	0.80	Far-from-frontier	+1.46 pp	[0.04, 2.94]

The pattern is consistent with a staged view of technological catch-up: Romania, Poland and Hungary, furthest from the EU productivity frontier, appear to derive the largest growth benefit from directing R&D spending toward capital equipment and infrastructure, consistent with a channel in which imported, equipment-embodied technology drives early-stage diffusion. This is consistent with firm-level evidence from the region: Damijan, Knell, Majcen and Rojec (2003) find that imported technology and FDI, rather than domestic R&D accumulation, were the dominant channel of productivity-relevant technology transfer in CEE transition economies, and Majcen, Radosevic and Rojec (2009) find that foreign-owned subsidiaries, which are disproportionately capital-intensive relative to domestically owned firms in the same economies, accounted for a substantial share of measured productivity growth in the region, a pattern that plausibly extends to the capital-versus-current R&D composition margin examined here. Czechia, closer to the frontier and reclassified into the middle tercile under the data-driven grouping, shows a much smaller, though still positive, implied effect, consistent with the broader catch-up literature's expectation that the marginal value of capital-embodied technology transfer declines as an economy approaches the frontier and the constraint shifts toward domestic innovation capacity (Radosevic and Yoruk, 2018).

Given the retraction of the government-funded R&D share result (Section 5.4), we do not extend country-specific projections to that variable. No projections are offered for business-performed share, higher-education-funded share, or social sciences field share, none of which survived the robustness protocol. The 95% confidence intervals reported alongside each point estimate in Table 4, obtained via the delta method from the covariance matrix of specification (4), show that only Romania's implied effect is distinguishable from zero at conventional levels; the intervals for Czechia, Hungary and Poland all include zero, even though their point estimates are positive and consistent with the group-level Group×gap×X coefficient being significant. This is expected: each individual country's implied effect is a linear combination of several estimated coefficients evaluated at a single gap value, so it carries more sampling uncertainty than the underlying group-level interaction term itself, and readers should not treat the Czechia, Hungary and Poland point estimates in Table 4 as separately statistically distinguishable from no effect.

KEY FINDING

Romania, Poland and Hungary, the economies furthest from the frontier, show the largest projected gain from shifting R&D spending toward capital equipment; of the four, only Romania’s implied effect is distinguishable from zero.

07 Limitations

Several limitations qualify these results. First, with 25-27 country clusters, cluster-robust inference sits below the 30-40 cluster threshold conventionally regarded as comfortable for asymptotic validity; we address this directly for the headline capital-expenditure result via a wild-cluster bootstrap (Section 5.2), which confirms significance at conventional levels under Rademacher-weight resampling, but the same correction has not been applied to every secondary specification reported in Sections 5.1, 5.3 and 5.4, so those results should still be read as suggestive rather than definitive. Second, this remains an observational panel analysis. Country fixed effects absorb time-invariant heterogeneity, and Section 5.4 shows the capital-expenditure result is stable to explicit crisis-period controls and full year fixed effects, but other time-varying confounders beyond the 2008-2009 financial crisis and the COVID-19 shock are not separately addressed and could in principle drive both R&D composition and productivity growth. Third, reverse causality cannot be fully excluded: fast-growing economies may adjust their R&D composition in response to growth opportunities rather than the reverse. We do not have a source of exogenous variation suitable for instrumental-variables estimation, so we cannot rule this out with certainty, but the Granger-style precedence checks in Section 5.5 are reassuring on this specific point: past growth does not significantly predict subsequent capital-expenditure share ($p = .55$), and the one-year-ahead value of capital-expenditure share does not significantly predict current growth once the existing one-year lag and the full interaction structure are controlled for ($p = .14$), while the headline result itself remains significant under that more demanding specification. These checks test timing rather than establishing causal identification, so the projections in Section 6 are best read as scenario projections conditional on the estimated relationship rather than as causal policy effects, consistent with the framing used there. Fourth, the type-of-research and field-of-research results rest on a sample skewed toward Central, Eastern and Southern European countries due to poor EU-15 reporting coverage in those two data blocks, limiting the generalizability of even the null results obtained there. Fifth, this paper follows the catch-up growth literature in using a static fixed-effects panel rather than a dynamic panel GMM estimator, despite gap being a highly persistent regressor mechanically related to lagged productivity, precisely the setting in which Nickell (1981) bias is a concern; we did not implement an Arellano-Bond or Blundell-Bond correction, both because composition rather than the persistence of gap itself is this paper's object of interest and because such estimators tend to perform poorly in short, unbalanced panels of the kind documented in Section 3.2, but this remains an open validity question rather than one this paper resolves. Sixth, the data-driven grouping's statistical significance is itself sensitive to how coarsely the sample is split: the headline Group×gap×X coefficient is significant at the 1% level under a tercile split of the 25 non-outlier countries but only marginal under a quartile split and insignificant under a quintile split (Section 5.2), consistent with reduced power from smaller treated groups rather than a different underlying relationship, but a reader should not treat the tercile-based result as robust to arbitrarily coarse groupings. Finally, the robustness protocol, while more systematic than typical practice, is not exhaustive; alternative outlier treatments or grouping rules could in principle yield different conclusions, though the capital-expenditure result's consistency across grouping, lag-length, and crisis-period variants (Sections 4 and 5.4) is reassuring on this point.

> **KEY FINDING**
>
> With 25–27 country clusters and no source of exogenous variation in R&D composition, these figures are scenario projections conditional on the estimated relationship, not causal policy effects.

08 Conclusion

This paper set out to test whether the composition of R&D expenditure, beyond its intensity, predicts the speed of productivity catch-up in the European Union. Of five composition dimensions and multiple candidate variables examined, only the capital-versus-current expenditure split survives a robustness protocol designed to guard against four specific threats: statistically distorted outlier economies, reliance on a historical political grouping as a proxy for economic structure, sensitivity to the choice of lag length, and confounding from the two major crisis episodes of the sample period. Economies furthest from the EU productivity frontier, prominently including Romania, Poland and Hungary in our sample, appear to benefit disproportionately from directing R&D spending toward capital equipment and infrastructure rather than current expenditure, while this compositional lever appears largely neutral for economies close to the frontier. Equally important is a methodological finding: five of six other candidate results that appeared significant under a conventional historical country grouping or an initial robustness pass did not survive full scrutiny, including a government-funded R&D share result that passed the initial grouping-based protocol but failed on crisis-period controls, underscoring the value of testing panel results against data-driven groupings, alternative lag structures, and crisis-period controls, rather than a single specification, before treating them as policy-relevant.

KEY FINDING

For economies far from the technological frontier, the capital content of R&D spending is a policy lever distinct from the level of R&D effort — and the only compositional lever this study can defend.

References

Abramovitz, M. (1986). Catching up, forging ahead, and falling behind. Journal of Economic History, 46(2), 385-406.

Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. Review of Economic Studies, 58(2), 277-297.

Barro, R. J., & Sala-i-Martin, X. (1992). Convergence. Journal of Political Economy, 100(2), 223-251.

Benhabib, J., & Spiegel, M. M. (1994). The role of human capital in economic development: Evidence from aggregate cross-country data. Journal of Monetary Economics, 34(2), 143-173.

Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. Journal of Econometrics, 87(1), 115-143.

Cohen, W. M., & Levinthal, D. A. (1990). Absorptive capacity: A new perspective on learning and innovation. Administrative Science Quarterly, 35(1), 128-152.

Damijan, J. P., Knell, M., Majcen, B., & Rojec, M. (2003). The role of FDI, R&D accumulation and trade in transferring technology to transition countries: Evidence from firm panel data for eight transition countries. Economic Systems, 27(2), 189-204.

David, P. A., Hall, B. H., & Toole, A. A. (2000). Is public R&D a complement or substitute for private R&D? A review of the econometric evidence. Research Policy, 29(4-5), 497-529.

De Long, J. B., & Summers, L. H. (1991). Equipment investment and economic growth. Quarterly Journal of Economics, 106(2), 445-502.

Eaton, J., & Kortum, S. (2001). Trade in capital goods. European Economic Review, 45(7), 1195-1235.

Eurostat. R&D expenditure and labour productivity statistics [data set]. European Commission.

Fisher, R. A. (1935). The design of experiments. Oliver and Boyd.

Freeman, C. (1987). Technology policy and economic performance: Lessons from Japan. Pinter Publishers.

Gerschenkron, A. (1962). Economic backwardness in historical perspective. Harvard University Press.

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. Econometrica, 37(3), 424-438.

Griffith, R., Redding, S., & Van Reenen, J. (2003). R&D and absorptive capacity: Theory and empirical evidence. Scandinavian Journal of Economics, 105(1), 99-118.

Griffith, R., Redding, S., & Van Reenen, J. (2004). Mapping the two faces of R&D: Productivity growth in a panel of OECD industries. Review of Economics and Statistics, 86(4), 883-895.

Guellec, D., & van Pottelsberghe de la Potterie, B. (2003). The impact of public R&D expenditure on business R&D. Economics of Innovation and New Technology, 12(3), 225-243.

Harris, R., Krenz, A., & Moffat, J. (2021). The effects of absorptive capacity on innovation performance: A cross-country perspective. Journal of Common Market Studies, 59(3), 589-607.

Kravtsova, V., & Radosevic, S. (2011). Are systems of innovation in Eastern Europe efficient? Economic Systems, 36(1), 109-126.

Lundvall, B.-Å. (Ed.). (1992). National systems of innovation: Towards a theory of innovation and interactive learning. Pinter Publishers.

Majcen, B., Radosevic, S., & Rojec, M. (2009). Nature and determinants of productivity growth of foreign subsidiaries in Central and East European countries. Economic Systems, 33(2), 168-184.

Mansfield, E. (1991). Academic research and industrial innovation. Research Policy, 20(1), 1-12.

Narula, R., & Guimon, J. (2010). The R&D activity of multinational enterprises in peripheral economies: Evidence from the EU new member states. UNU-MERIT Working Paper 2010-048.

Nelson, R. R. (Ed.). (1993). National innovation systems: A comparative analysis. Oxford University Press.

Nelson, R. R., & Phelps, E. S. (1966). Investment in humans, technological diffusion, and economic growth. American Economic Review, 56(1/2), 69-75.

Nickell, S. (1981). Biases in dynamic models with fixed effects. Econometrica, 49(6), 1417-1426.

Radosevic, S. (1999). Transformation of science and technology systems into systems of innovation in central and eastern Europe: The emerging patterns and determinants. Structural Change and Economic Dynamics, 10(3-4), 277-320.

Radosevic, S., & Yoruk, E. (2018). Technology upgrading of middle-income economies: A new approach and results. Technological Forecasting and Social Change, 129, 56-75.

Salter, A. J., & Martin, B. R. (2001). The economic benefits of publicly funded basic research: A critical review. Research Policy, 30(3), 509-532.

Staiger, D., & Stock, J. H. (1997). Instrumental variables regression with weak instruments. Econometrica, 65(3), 557-586.

Vandenbussche, J., Aghion, P., & Meghir, C. (2006). Growth, distance to frontier and composition of human capital. Journal of Economic Growth, 11(2), 97-127.

Zahra, S. A., & George, G. (2002). Absorptive capacity: A review, reconceptualization, and extension. Academy of Management Review, 27(2), 185-203.

Disclaimer

The material presented in this report is intended to encourage better understanding of economic policy and financial markets. Neither the information nor the opinions expressed herein constitute, or are to be construed as, an offer or solicitation of an offer to buy or sell any commodities, securities or investments.

Although the information compiled herein is considered reliable, its accuracy is not guaranteed and is not to be taken in substitution for the exercise of judgment by any recipient. This document includes forward-looking statements, such as prospects or expectations, that are based on current views and assumptions and subject to known and unknown risks and uncertainties.

Actual results, performance figures or events may differ significantly from those expressed or implied in such forward-looking statements. Any person using this report's material does so solely at their own risk, and the author shall be under no liability whatsoever in respect thereof.