Cardiff Economics Working Papers

Working Paper No. E2023/16

Specialisation precedes diversification: R&D productivity effects

James Foreman-Peck and Peng Zhou

May 2023

ISSN 1749-6010

Cardiff Business School
Cardiff University
Colum Drive
Cardiff CF10 3EU
United Kingdom
t: +44 (0)29 2087 4000
f: +44 (0)29 2087 4419
business.cardiff.ac.uk

Specialisation precedes diversification: R&D productivity effects[1]

James Foreman-Peck[1], Peng Zhou[2,*]

[1] Cardiff Business School, Cardiff University, UK; Email: foreman-peckj@cardiff.ac.uk
[2] Cardiff Business School, Cardiff University, UK; Email: zhoup1@cardiff.ac.uk
[*] Corresponding author. Address: D47, Aberconway Building, Colum Drive, Cardiff, CF10 3EU. Tel: +44 (0)2920 688778. Fax: +44 (0)292087 4000.

Abstract: We model how R&D enters the innovation system in four ways (intramural, extramural, cooperative, and spillover). Despite measuring three different spillovers together, for a very large sample of European enterprises we conclude that the productivity effects of spillovers were at best smaller than intramural R&D productivity effects. We also find that building on the greater skills and experience of enterprises already undertaking R&D (intensity) raised labour productivity more than providing support for those beginning R&D (extensity). Optimal extramural R&D intensity was higher than the actual level; sample firms could boost productivity either by abandoning extramural R&D or by doing much more. There were substantial differences in our sample between enterprises and countries in terms of R&D spillovers. Greater multinational corporation incidence in new EU members accounted for these countries' high direct R&D intensity productivity, regardless of their generally low overall labour productivity. Absorptive capacity made little difference to the utilisation of spillovers.

Keywords: R&D; innovation; knowledge spillover.

JEL: L53, L21, H71, H25

Highlights:

- R&D direct productivity effects are greater than spillovers in our sample countries.
- Productivity differences between new and old EU members dampened by multinationals.
- Promoting R&D intensity tends to be more effective than boosting R&D extensity.
- The optimal extramural R&D intensity is higher than the actual level.
- Extramural R&D tends to be more effective than cooperative R&D.

[1] We would like to sincerely thank the Editor, Adam Jaffe, and the anonymous reviewers in *Research Policy* for their insightful and constructive comments, which significantly improved the paper. We also benefited from discussions with Anna Grosman, Huw Dixon, Patrick Minford, Andy Henley, Xiaodong Chen, and Wei Yin at Loughborough University London, Cardiff University, Xi'an Jiaotong University, and Southeast University. The opinions and conclusions contained in this document are the sole responsibility of the authors. The anonymised data used in the analysis of this paper were obtained on CDROM from Eurostat as part of the research proposal 'European National Innovation Systems'. The results and the conclusions are those of the authors. They represent their opinion and not necessarily those of Eurostat, the European Commission or any of the statistical authorities whose data have been used.

1 Introduction

Policy makers attempt to encourage innovative strategies by supporting private sector R&D in both developed (Freitas & Von Tunzelmann, 2008) and transition countries (Szczygielski et al., 2017). Support is often justified by appeal to spillovers, benefits to other firms than the one undertaking the R&D investment (Hashi & Stojčić, 2013b). In the resource-based view of the firm, an enterprise will achieve and sustain a position of competitive advantage only if their resources and capabilities are difficult to imitate (Barney, 1991). R&D can be such a strategic resource. But if a firm's R&D is not distinctive or is substitutable then the R&D of another business may make it redundant, creating a negative spillover. Estimates of R&D spillovers have frequently been made at the industry or country level, but these process effects are more accurately measured at the firm level. Aggregation to the country level involves averaging out firm-level spillovers, so understating the firm-level spillovers that may be the object of support policy. Moreover, focusing only on one type of spillover and neglecting other types may underestimate their importance. We therefore aim to separately measure various effects of R&D, in terms of extensity (the whether-or-not decision), of intensity (the how-much decision), and including the effects of different spillover pools with different spans.

A spillover pool is defined as all *potential* sources of knowledge. Our method is essentially the correlation of one enterprise's labour productivity with its own R&D and the R&D of other firms. Although we cannot find point estimates of spillover effects, we can establish upper and lower bounds. The lower bound is determined by statistical significance of the coefficient that requires there to be at least one recipient of a spillover. The upper bound is set by convergence of our expression as the pool increases. With our comprehensive approach, we can offer a more convincing answer to the research question: *what are the maximum productivity effects of various levels of spillovers, controlling for other possible effects of R&D*?

The countries in our data set of more than 600,000 enterprises are very heterogenous, ensuring the coverage of potentially diverse behaviour patterns and therefore greater generalizability of the results. At the extremes the countries include Bulgaria in 2006 (GDP per capita $5600) and Germany in 2018 ($43,000)[2], some mature market economies and others still transitioning. We show that for a firm-level representative sample of these European countries between 2006 and 2018, state support boosted R&D and R&D increased labour productivity. This R&D support increased productivity mainly in two ways, the first via intramural R&D and the second via other firms' R&D in the spillover pools. The R&D spillover effects are smaller than the substantial positive intramural R&D effects. Our spillovers are net effects of (negative) "business stealing" and of (positive) R&D impacts.

[2] World Bank data base, constant 2015 US $.

Promoting R&D intensity tends to be more effective than extensity[3]. Moreover, the optimal extramural R&D intensity is higher than the actual level, so firms should either do nothing or do much more extramural R&D. Separating the sample into three "old" EU members (Germany, Spain, and Portugal) and "new" EU members (former centrally planned economies) we find the "old" are more efficient in intramural R&D extensity, because the R&D intensity is already optimised. The "new" are more efficient in intramural R&D intensity because the average R&D intensity is lower than optimal. The "old" have more positive spillover effects, while the firms in "new" EU members have insignificant or even negative spillover effects. We are unable to establish that absorptive capacity of enterprises had much effect on their utilisation of spillovers.

Our theoretical contributions in this paper include identifying three types of net R&D spillover at the level of the enterprise, allowing for nonlinearities, and estimating these effects in a modified structural CDM model (Crepon et al., 1998). Our paper differs from spillover literature utilising trade flows, foreign personnel, FDI and/or input-output industry linkages such as Hashi & Stojčić (2013b), Vujanović, et al. (2021), and Vujanović, et al. (2022). By contrast, we provide a way of defining spillovers *on the firm* of industry-level, country-level, and EU-level spillover pools.

Having set the research question in introduction, we briefly review some of the relevant literature and formulate three testable hypotheses in section 2. Section 3 lays out our model before the data are described in section 4. The results are presented in section 5, followed by the conclusion in section 6.

2 Literature Review

A resource-based view of competitive advantage maintains that R&D capabilities should be internalized in the firm. But according to developments of the view, through collaboration the firm gains access to external resources such as equipment, expertise, and information (Knudsen & Nielsen, 2010; Iammarino et al., 2012; Fuller, 2018). A qualification is that the success of collaborative R&D relationships depends on the type and the quality of partners involved and proximity between them (Stojčić, 2021). There are also unintended consequences of R&D capabilities; the knowledge generated by R&D is not invariably entirely private to the R&D investing firm. Rather it may spread, or spill over, to other firms through various channels. If spillovers are important, then failing to model them in innovation studies could severely incorrectly estimate the impact of government support. For example, in an earlier study (Hashi & Stojčić, 2013a), in some respects similar to the present exercise, support was found to create additional spending on innovation by firms directly. However, the effects of some of this

[3] Tevdovski et al. (2017) refer to extensity as engagement.

additional innovation spending may have spilled over to other firms, especially since Hashi & Stojčić (2013b), with many similar countries, found an important role played by knowledge spillovers generated through international trade, horizontal and vertical interactions in the domestic market, and within-group exchange of knowledge. So, integrating the two types of study might produce rather different policy results.

As Wieser (2005) notes, researchers usually implicitly assume that all knowledge is embodied R&D or that the usage of knowledge between industries mirrors the usage of commodities, foreign personnel or FDI between industries. Typically, R&D spillovers have been modelled by adding to a knowledge production or cost function an indicator of R&D, trade or FDI external to the firm (Hall et al., 2010; Ugur et al., 2020). This spillover variable might be measured as a weighted sum of the R&D stocks from sources outside the firm. The weights would be proportional to some flows or proximity measures between the receiver of R&D spillover (firm, industry, or country i) and the source of R&D spillover (firm, industry, or country j). Flow related weights have included intermediate input transactions, investments in capital goods, hiring of R&D personnel, attendance at workshops, seminars or trade fairs, collaborations, adoption of new technologies, flows of patents (Luintel & Khan, 2017) or innovation.

Alternatively, instead of using R&D to calculate spillovers, foreign inflows have been used to measure them in a productivity or innovation output equation. Hashi & Stojčić (2013b) use trade indicators of spillovers supplemented by group membership and university cooperation. Vujanović et al. (2021) and Vujanović et al. (2022) calculate more complex spillovers variables based on FDI inflows. They distinguish a horizontal spillover variable that represents within industry knowledge transmission from foreign to domestic firms. They approximate this variable by the presence of foreign firms in an industry measured by employment or sales. Vertical spillovers are the knowledge transmitted between firms in upstream and downstream industries. Backward spillovers are the knowledge home suppliers obtain through supplying foreign customers. The knowledge domestic customers acquire buying inputs from foreign suppliers are forward spillovers. To calculate vertical spillovers Vujanović et al. (2022) use input-output coefficients multiplied by the corresponding horizontal spillovers. The spillovers measured by this approach do not include those between home country enterprises.

The foregoing studies are unusual in measuring several spillovers. Most studies focus on only one type; only one of Ugur et al. (2020)'s sample of 60 studies addresses two types of spillover and none of these do this with firm level data. Neglect of other types may underestimate the importance of spillovers. Moreover, it is arguable that the spillover effect is most accurately measured at the firm level because there may be spillover effects both within and across industries and countries.

Positive (technology) spillovers have been distinguished from negative business stealing effects of product market rivalry, and the net impact computed (Bloom et al., 2013). As more firms undertake R&D, business stealing may increase and dominate the technology spillover impact of low-quality repetitive innovations. However, the sign of the net effect is indeterminate. For example, a firm at or near the technological frontier, thanks to high intramural R&D might have less possibility of finding and therefore absorbing useful spillovers. It is more likely that product market rivalry will dominate; the resulting net spillover coefficient would be negative. Negative spillovers are common for some countries. Hashi & Stojčić (2013b) using data from a broadly similar collection of countries to ours find negative spillovers from imports and from intra-group knowledge. One Korean firm's innovation negatively affects sales of other rival firms even after identifying the positive and negative roles separately (Lee & Kim, 2019). In Spanish firms, intra-industry externalities have a negative coefficient in the sample as a whole and inter-industry externalities play an ambiguous role (Goya et al., 2016). On the other hand, with fast-moving technology a research-intensive firm may be in a strong position to absorb beneficial innovations adopted elsewhere.

If some capacity affects the ability of a firm to utilise net spillovers, we expect to find measures of this absorptive capacity such as R&D intensity to be positively associated with spillover marginal effects (Cohen & Levinthal, 1990; Cassiman & Veugelers, 2002). R&D extensity as a correlate of spillover marginal effects, by contrast, may measure the pressure of competition on absorptive capacity. Substantial absorptive capacity itself has been measured as occurring with firms in which more than 25% of personnel possess a tertiary degree of education or by knowledge flows from related firms indicated by a firm's membership of an enterprise group (Stojčić et al., 2020). Firms participating in Research Joint Ventures may benefit more from technology spillovers because of greater absorptive capacity, or they may be more resilient to the effects of product market rivalry (Banal-Estañol et al., 2022).

Turning to the magnitudes calculated for spillover effects, the survey by Hall et al. (2010) reports that spillover and own-R&D effects are similar, although there is a risk of selection bias (Griliches, 1992). R&D spillover estimates for the US economy are typically large. For a panel of US firms Bloom et al. (2013) estimate that the gross social returns to R&D were at least twice as high as the private returns. They find diverse effects of spillovers in different industries (especially computers, pharmaceuticals, and telecommunications) and size classes of firms, identifying wide variation in social and private returns to R&D. Technology spillovers were found in all sectors, but smaller firms had significantly lower social returns because they tended to operate in technological "niches". Updating this panel, Lucking et al. (2019) find that the ratio of the social return to the private return to R&D was about four to one. A study of R&D spillovers created by grants to small firms from the US Department of Energy established that for every patent produced by grant recipients, three more were produced by spillover beneficiaries (Myers & Lanahan, 2021).

However, using the Science Policy Research Unit's database of innovative UK manufacturing firms, Wakelin (2001) established that there were no significant spillover effects of R&D from related sectors. The results from Geroski (1991) confirm these conclusions. Cincera (1998)'s study of large manufacturing firms between 1980 and 1994 determined that European firms did not particularly benefit from national and international sources of spillovers. But later, EU countries have been found to generate significant correlations between TFP growth and external R&D knowledge stock growth (Corrado et al., 2017; Goodrich et al., 2017). For Norway R&D tax credits also were estimated to generate large spillovers. But R&D capital stocks increased only slowly, delaying knowledge flows to other industries and spillover effects both from abroad and from domestic sources (von Brasch et al., 2021). By contrast despite recent substantial increases in R&D of the emerging world, spillovers from there are virtually non-existent (Luintel & Khan, 2017). Even spillovers from the industrialised to the emerging world are modest.

A meta regression survey by Neves & Sequeira (2018) utilises a database for their baseline analysis comprising 170 estimates of the effect-size—the spillover effect -, from 15 different primary studies. They concluded that the average spillover effect was less than but close to one and was highly significant. The larger survey by Ugur et al. (2020) was more pessimistic, finding that the productivity effect of R&D spillovers was usually smaller and estimated with lower precision than the effect of own R&D. This study involved 983 productivity estimates for spillovers and 501 estimates for own-R&D from 60 empirical studies and concluded that the spillover effect is too small to be practically significant. Controlling for observable sources of heterogeneity and best-practice research, the meta-effect is insignificant in the full sample but significant and large among OECD firms/industries/countries.

As Hall et al. (2010) remind us, we should not necessarily expect similar impacts for spillovers with different technologies and areas. Moreover, the literature uses very different definitions of "spillover". Findings of large spillovers tend to combine unpaid-for knowledge with intentionally purchased/franchised knowledge or patents. In our study and many other studies which find small spillovers (Jaffe, 1986, 1989; Pessoa, 2005; Luintel & Khan, 2009), spillover is defined as purely unintentional on the part of the provider and an unpaid-for externality for the recipient. Following the literature on spillovers reviewed, we propose our first hypothesis:

> (H1) **The Smaller Spillover Hypothesis**. The unintentional spillover effects of R&D on productivity *at different levels* are smaller than the effects of intentional R&D activities.

To accurately estimate unintentional spillovers, mechanisms by which intentional R&D activities affect productivity must be controlled. For example, if a firm does not have adequate intramural R&D capacity, a common choice is to outsource the job to specialised R&D firms or directly acquire existing innovative products (extramural R&D). Sometimes, it may be more cost effective to purchase an existing solution, so then firms face a choice between "knowledge use" and "knowledge creation" (Cirrera & Maloney, 2017). Furthermore, acquisition can target hardware (e.g., machinery and equipment) or software (e.g., know-how and patents). Innovation in emerging economies takes place mostly using existing knowledge and embedded technology rather than by the generation of new products and processes (Vujanović et al., 2022). But this tendency can be ameliorated by public procurement for innovation (Stojčić et al., 2020).

As a special type of firm, multinational corporations (MNC) can facilitate knowledge flows between subsidiaries and the headquarter (Gupta & Govindarajan, 2000), so MNC subsidiaries in different countries are expected to have similar intramural R&D effects. This dampens the productivity discrepancies across countries (Vujanović et al., 2021; Vujanović et al., 2022).

> (H2) **The Dampening Hypothesis**. Multinational firms have similar *intramural* R&D effects in different countries, resulting in converging productivity.

Specifically, heterogeneity in spillover effect can result from the absorptive capability for external knowledge (Escribano et al., 2009). For example, Castellacci & Natera (2013) find a coevolution (cointegration) relationship between innovative and absorptive capacity at the country level. Crescenzi & Gagliardi (2018) discovered that only firms that actively combine internal and external knowledge can take advantage of a favourable external environment. Potential absorptive capacity is composed of external knowledge acquisition and assimilation (Zahra & George, 2002). Based on the literature on absorptive capacity, we form the following hypothesis.

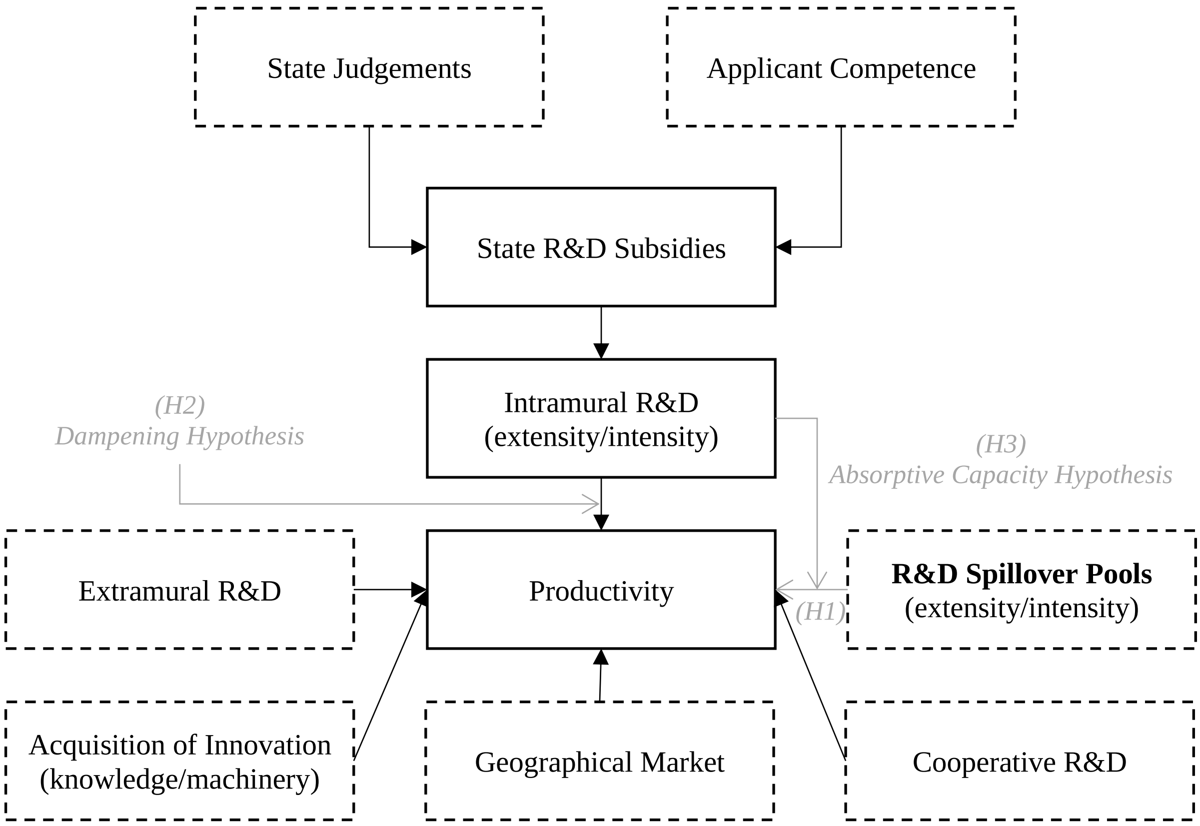
> (H3) **The Absorptive Capacity Hypothesis**. Industries with greater absorptive capability tend to have greater incoming spillover effects.

3 Model

To disentangle the contributions of R&D to productivity, we distinguish four types of effects by whether the effect is intentional and where the R&D is carried out: (i) the effect of intramural R&D, which is planned and undertaken internally by the firm, (ii) the effect of cooperative R&D, which is intentional and undertaken partially outside the firm, (iii) the effect of extramural R&D, which is intentional and performed entirely externally, and (iv) the spillover effect of R&D, which is unplanned by the donor and absorbed from outside the recipient enterprise.

These effects can also be defined for extensity and for intensity. Extensity is measured by whether a firm engages in R&D, while intensity is measured by how much R&D a firm undertakes. The former is a zero-one dummy, and the latter here is a percentage of turnover. A 1% rise in extensity means more firms are doing R&D activities (not necessarily but probably with more resources in the total allocated to R&D), and a 1% rise in the intensity means more resources are devoted to R&D activities (but not necessarily by more firms). Firms already conducting R&D may be more productive than those undertaking R&D for the first time, so R&D state support should be more effective for the first category than the second.

Figure 1 Representation of the firm-level model with endogenous R&D subsidies



Notes: Solid edged boxes indicate endogenous variables, and dash edged boxes indicate exogenous variables.

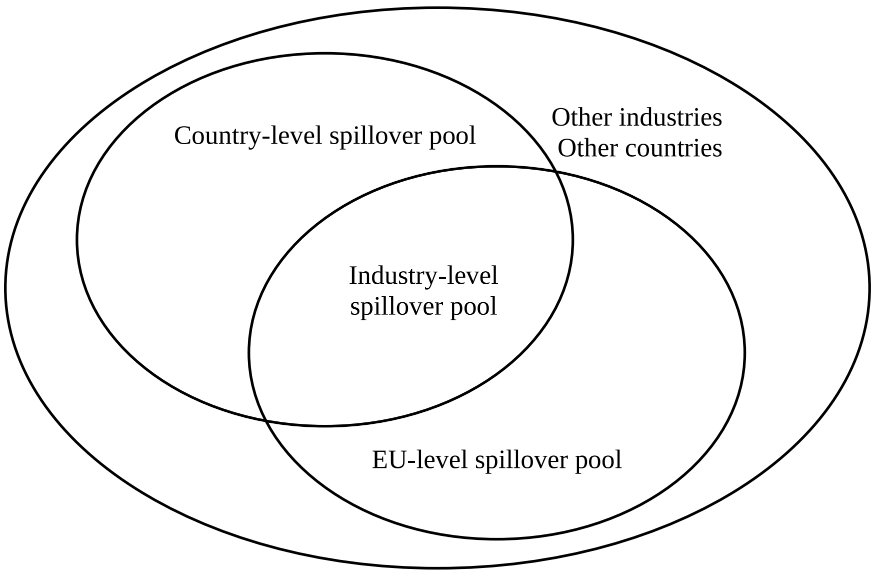
To model the effects of R&D on productivity, we use a modified CDM model (Crepon et al., 1998)[4]. This features an additional state support equation (Foreman-Peck & Zhou, 2022) to model the R&D funding process endogenously, includes inverse Mill's ratios (IMR) to correct for self-selection or endogeneity bias (Maddala, 1983; Greene, 2000), and measures four types of R&D effects. The modified CDM model is represented in Figure 1. The three hypotheses developed in the literature review can be tested by estimating the structural model. The small spillover hypothesis (H1) is directly embodied in the productivity equation. The dampening

[4] A brief review on the CDM model can be found in Appendix 2.

hypothesis (H2) can be tested by re-estimating the model with the MNC subsample (subsection 5.1), while the absorptive capacity hypothesis (H3) can be tested using the pseudo-panel data constructed by the CDM model (subsection 5.2).

To measure spillover effects, we distinguish three levels of spillover pools within which the unintentional influence takes effect: (i) the industry-level spillover pool that includes R&D activities in the same industry of a country, (ii) the country-level spillover pool that includes R&D activities across industries of a country, and (iii) the EU-level spillover pool that includes R&D activities in the same industry across countries. A spillover pool is the knowledge bank of a set of firms that between them are potential sources of the knowledge. But every enterprise in the pool faces slightly different pools because the firm in question is not included in the pools it faces (it cannot receive knowledge from itself as a spillover); the spillover pool is not symmetric between firms[5]. The asymmetry also applies to higher level spillover pools.

Figure 2 The Venn diagram of different levels of spillover pools



By definition the country-level spillover pool includes any spillovers from firms of all industries of a given country, so it contains the industry-level spillover pool[6]. We only include firms of the same industry from available countries in the EU-level spillover pool. If we include firms of all industries of all sample countries in the EU-level spillover pool, the "grand" pool would be almost the same for all firms, because the pool is so large relative to an individual firm.

[5] And for that reason, the weights of the firms in the pool must be different for each firm as well as the membership of the pool.
[6] The primary purpose of distinguishing different pools is to compare spillover effects at different levels with an increasing pool size, i.e., the industry level, the country level, and the EU level. Overlapping is not an issue because lower-level pool is part of higher-level pools by definition. Alternatively, if we define mutually exclusive spillover pools, it would be easier to add up the effects, but it would be less straightforward to compare the spillover effects as the pool size gets bigger.

Hence, collinearity or lack of variation would ensure that this spillover effect could not be identified[7]. The topology of the three spillover pools is illustrated by Figure 2 which shows the total spillover effect is a sum of the effects of the country- and EU-level pools minus the effect of the industry-level spillover pool.

To estimate spillover effects, we need a measure of knowledge stock for each spillover pool. Close to steady state growth the R&D flow is related to the stock by the depreciation and growth of R&D parameters of the perpetual inventory method. This means the knowledge stock is some multiple of the flows from which it is calculated. In this paper we use the annual R&D flow as a proxy for the stock[8]. Following Hall et al. (2010) a knowledge stock $\tilde{X}_i$ for a unit (a firm, an industry, or a country) is defined as:

$$\tilde{X}_i = \sum_{i \neq j} w_{ij} X_j, \text{ where } \sum_{i \neq j} w_{ij} = 1. \tag{1}$$

One simplification is to set the knowledge flow weight w_{ij} equal to $\frac{1}{N-1}$ where N is the total number of units in the spillover pool. In general, the weight w_{ij} of each unit X_j in the spillover pool differs for different i, because j indicates all the other units except for the unit i in the pool. To account for the heterogeneities of units in a pool, we define variable weights w_{ij} based on R&D shares in the spillover pool:

$$w_{ij} = \frac{1}{N-1} \text{ or } w_{ij} = \frac{Turnover_j}{\sum_{i \neq j} Turnover_{ij}} \text{ or } w_{ij} = \frac{R\&D_j}{\sum_{i \neq j} R\&D_{ij}}. \tag{2}$$

The magnitude of the spillover effect depends on the size of the potential pool, which is also known as the scale effect. Intuitively, a larger pool means more firms might benefit from the same spillover effect. Nevertheless, we can show that the scale effect eventually converges to a maximum level as the size of the pool gets bigger. Assume there are N firms in the spillover pool. If one additional firm becomes R&D active, the intramural effect raises its own productivity by γ, and the spillover might affect $N-1$ firms in the pool. Note that the estimated spillover effect in the model is defined as "how much another firm's productivity changes (δ) if the average ratio of R&D-active firms (over N) rises by 1 percentage point." Thus, the starting of one firm's R&D activity raises the average ratio by $\frac{1}{N}$. The actual spillover effect of this change on another firm in the pool is therefore $\frac{1}{N} \times \delta$. This spillover effect is received by $N-1$ firms

[7] No attempt is made to measure spillovers from, e.g., EU members excluded from the sample or the US.
[8] This will not bias the estimates for the following reason. The true relationship between productivity and knowledge stock is ln(productivity) = a + b * ln(stock) + error. The law of motion for the stock is stock(t) = stock(t-1) * (1-d) + flow(t). Assume there is a growth rate of g for the stock, then stock(t) = stock(t-1) * (1+g). Combine the two to create the perpetual inventory formula: stock(t) = flow(t)/(g+d). If we write in logs, we have ln(stock) = ln(flow) - ln(g+d). Substitute this into the first equation: ln(productivity) = a + b * ln(stock) - ln(g+d) + error. Since the term - ln(g+d) will be combined with the constant term, the estimated beta is unbiased.

in the pool, so the total spillover effect is equal to $\frac{\delta(N-1)}{N}$. In other words, as N rises, the spillover effect rises (the scale effect), but the magnitude converges to δ. Theoretically, the intended intramural effect can be greater, equal to, or smaller than the spillover effect, but in the present sample, we will show that $\gamma > \delta$.

The sign of spillover effect is ambiguous. On the one hand, there is positive spillover of R&D due to knowledge sharing. Some R&D leads to process innovations in logistics and distribution, which can promote productivity of the entire supply chain. On the other hand, we do not separately identify "business stealing" under monopolistic competition. This means that other firms lose customers and profits from the entry of a new competitor, which imposes a negative externality on existing firms. R&D duplication has a similar effect. The firm that first successfully concludes the R&D process renders obsolete or redundant competitor firms' R&D directed to the same end—the Schumpeterian "creative destruction". These competitors suffer a negative externality. Therefore, the overall spillover effect depends on which spillover dominates.

There are three endogenous variables (productivity, R&D activity, and government support) in the structural equation system (Figure 1), where R&D enters the system in four ways (intramural, extramural, cooperative, and spillover). Intramural R&D effects are endogenously influenced by state R&D subsidies, which in turn are endogenously determined by state judgements and applicant competence. The self-selection of R&D extensity is captured by an inverse Mills ratio (IMR) when the "treatment" (D) is a dummy, i.e., whether R&D is done in our case.

$$IMR = \begin{cases} \frac{\phi(X\beta)}{\Phi(X\beta)} \text{ if } D = 1 \\ -\frac{\phi(X\beta)}{1-\Phi(X\beta)} \text{ if } D = 0 \end{cases}$$

For the continuous treatment case (R&D intensity), we simply employ IV regressions. The excluded instruments in our model of R&D intensity are state support and the IMR calculated in the funding equation. To capture nonlinearity, we include both linear and quadratic terms for R&D intensity measures (ratios of turnover). We also control for other influences on labour productivity namely geographical market, cooperative R&D agreements and the acquisition of knowledge and innovative machinery, which are treated as exogenous because Banal-Estañol et al. (2022) find no direct effect of collaboration (research joint ventures) on R&D.

4 Data

Our data consist of national Community Innovation Surveys (CIS), with the majority of questions standardised across Europe[9]. There are, however, some differences in variable definition and value categorisation. To achieve consistent data over all surveys, we utilise a smaller

[9] In the majority of countries, the survey is a combination of census and sample survey. Bulgaria conducts only a census. The survey is mandatory in most countries but voluntary in Germany.

number of variables compared with many CIS studies and re-categorise some variables[10]. A detailed mapping of variables across countries and years is available in Appendix 1. European national statistical offices generally take a sample from all establishments, stratifying the sample by sector, establishment size and possibly region. For the size classes, a portion of all establishments below a certain size threshold is selected, but in most countries all large establishments receive a questionnaire. The survey is conducted at the enterprise level every two years, with certain questions covering a period of three years, such as "*During the three years from xxx to xxx, did your enterprise receive any public financial support….?*"

Firms that organise their business activities into separate legally defined units can be sampled several times. According to the CIS2018 questionnaire, R&D is broadly defined to cover the creation of new knowledge and the solving of scientific or technical problems (including software development that meets this requirement). Nonetheless, most firms do not engage in any R&D activities (only 18% of our sample do such intramural R&D). Furthermore, for those firms that do engage (RRDIN=1), there is another decision to make on how much R&D to undertake, R&D intensity (RRDINX). On average, the turnover ratio of intramural R&D is 7.8% among R&D active firms in our sample. Enterprises in Bulgaria in 2006 had a ratio of 4.9% and those in Germany in 2018 achieved a similar ratio, indicating far higher R&D expenditures because of the discrepancy in turnovers and extensity[11]. The ratio measure has the merit of approximately controlling for price changes. 13.31% of firms in our sample opt for an eclectic approach of cooperative R&D to share the responsibilities and costs of R&D. In our data, we can distinguish collaborators from enterprises of the same group, suppliers, clients, competitors, consultants, universities, and governments. Our key variables (Table 1) are:

- FUN: **R&D support** dummy includes EU, central government, and local government support. Romania has the smallest proportion of sample funded firms.
- RRDIN: **Extensive margin**. Germany has the largest proportion of enterprises undertaking intramural R&D in our sample.
- RRDINX: **Intensive margin**. The intensity varies significantly across countries and across industries. Hungary and Spain report greatest intensity and Slovenia the least.
- LPROD: (ln) **Productivity** is calculated by turnover divided by estimated employment size. Bulgaria has the lowest labour productivity.

[10] For example, the sample of the German CIS2018 has been stratified by eight "person employed" size groups: 5-9, 10-19, 20-49, 50-99, 100-249, 250-499, 500-999, and 1000. The Bulgarian CIS2018 includes in the target population three size groups 10-49, 50-249, and more than 250. Hence, this necessitates restricting to three employment size categories our entire sample.

[11] R&D measures by other researchers such as Griffith et al. (2006) and Tevdoski et al. (2017) have employment rather than turnover in the denominator and so are not readily compared with ours. The restriction of the analysis to manufacturing business tends to give higher average R&D spends compared to data such as ours that includes services as well in the analysis.

Table 1 Descriptive statistics of key variables

Country	Code	No. of Obs.	FUN	RRDIN	RRDINX	LPROD
Bulgaria	BG	104,171	0.039	0.035	0.095	9.815
Czechia	CZ	40,778	0.122	0.258	0.042	10.959
Germany	DE	37,070	0.167	0.420	0.036	12.017
Estonia	EE	13,703	0.104	0.245	0.050	10.722
Greece	EL	10,642	0.160	0.258	0.061	11.146
Spain	ES	233,694	0.158	0.219	0.110	11.310
Croatia	HR	15,797	0.102	0.163	0.035	10.616
Hungary	HU	41,360	0.096	0.126	0.136	10.786
Lithuania	LT	15,992	0.110	0.160	0.080	10.285
Latvia	LV	9,418	0.049	0.101	0.039	10.136
Portugal	PT	46,315	0.138	0.237	0.033	10.962
Romania	RO	58,127	0.027	0.056	0.081	10.309
Slovenia	SI	9,254	0.110	0.303	0.029	11.207
Slovakia	SK	19,537	0.038	0.117	0.032	10.909
Overall		**655,858**	**0.111**	**0.180**	**0.078**	**10.863**

INN_LAST (firm undertook some innovative activities in the previous period) is generated from the variables INABA (innovation abandoned or suspended before completion) and INONG (innovation still ongoing at the end of the survey year). CIS2018 has different definitions from the earlier surveys, but similar variables can be found. Further details of the data and its treatment are in Appendix 1.

5 Results

The estimated coefficients of the modified CDM model (Figure 1) are shown in Table 2. To allow for heterogeneities, both intramural R&D (RRDIN) effects and spillover effects can vary by industries, countries, years, and firm types. Equations (1) and (2) are estimated with a two-step endogenous treatment procedure (Lee, 1978). Equation (3) uses an inverse Mill's ratio (IMR) generated from (2) to correct for endogeneity of RRDIN. By contrast, RRDINX in equations (2)' and (3)' is a continuous variable, so the 2SLS approach is applied with inverse Mill's ratios used as excluded instruments.

Not surprisingly, government support for innovation (FUN) was selective for enterprises which undertook some innovative activities in previous period (INN_LAST) in Table 2 column (1). The largest negative impact on support is if the enterprise was headquartered in another EU country. Both support and engagement with in-house R&D are characteristics of larger firms. The effects of support on R&D extensity and intensity are positive as shown in column (2) and (2)'. The nonlinearity of (1), (2), and (2)' ensures that the marginal effects are different from the estimated coefficients. The two productivity equations (3) and (3)' show that the most

Table 2 Estimated coefficients of the structural model

	(1) FUN	(2) RRDIN	(2)' RRDINX	(3) LPROD	(3)' LPROD
FUN		5.435***	2.443***		
IMR1		-2.319***	-0.956***		IV
IMR2				-0.028***	IV
INN_LAST	1.163***				
Part of a group	0.059***	0.199***	0.030***	0.609***	0.608***
Headquarter in EU	-0.317***	0.164***	0.144***	0.262***	0.28***
Headquarter in other	-0.069***	0.017	0.191***	0.175***	0.167***
Size 50-249	0.171***	0.062***	-0.302***	-0.318***	-0.292***
Size 250+	0.256***	0.228***	-0.511***	0.697***	0.774***
Local market				-0.031***	-0.035***
National market				0.289***	0.277***
EU market				0.325***	0.326***
Other market				0.282***	0.264***
RRDIN (extensity)				12.959***	
RRDINX (intensity)					8.65***
$RRDINX^2$ (intensity)					-1.448*
RRDEX (extensity)				0.197***	
RRDEXX (intensity)					-3.578***
$RRDEXX^2$ (intensity)					3.629***
ROEK (extensity)				0.136***	
ROEKX (intensity)					-3.054***
$ROEKX^2$ (intensity)					2.205***
RMAC (extensity)				0.135***	
RMACX (intensity)					-0.501***
$RMACX^2$ (intensity)					-0.273**
Coop: Group				0.109***	0.14***
Coop: Supplier				0.12***	0.156***
Coop: Client				-0.133***	-0.147***
Coop: Competitor				-0.015	0.009
Coop: Consultant				0.093***	0.145***
Coop: University				-0.026**	-0.021
Coop: Government				-0.015	-0.005
Constant	Yes	Yes	Yes	Yes	Yes
Spillover pools	Yes	Yes	Yes	Yes	Yes
Country fixed effects	Yes	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Interactive terms				Yes	Yes
No. of Obs.	644,574	644,574	644,574	578,530	612,834
Model	probit	probit	fractional reg	fractional reg	IV reg

Notes: Significance * 5%, ** 1%, *** 0.1%. INN_LAST = the firm had innovative activities in the last period. IMR1 = IMR based on FUN equation. IMR2 = IMR based on RRDIN equation. The interactive terms are slope dummies (year, country, industry) for RRDIN, RRDINX, and spillover pools. RRDEX variables: extramural R&D. ROEK variables: engagement in acquisition of knowledge. RMAC variables: acquisition of machinery.

productive enterprises were group members, employers of more than 250 persons, and exporters to another EU market. The first labour productivity equation (3) estimates the effect of R&D engagement or extensity, (3)' estimates the impact of inhouse R&D spending or intensity. The negative contribution of cooperation with universities (as Hashi & Stojčić (2013b) find with innovation output) and governments presumably reflects the pursuit of other objectives than productivity. The spillover coefficients are not reported in Table 2 because there are too many interactive terms. Instead, their marginal effects are reported in Table 3.

Table 3 Estimated marginal effects of R&D extensity and intensity on productivity

		Extensity	Intensity
Intramural	**Overall**	**12.959*** **	**21.283*** **
Cooperative	Group	10.902***	
	Supplier	12.038***	
	Client	-13.334***	
	Competitor	-1.494	
	Consultant	9.341***	
	University	-2.58**	
	Government	-1.544	
	Overall	**1.9041*** **	
Extramural	Outsourced R&D	19.731***	-3.525***
	Acquired knowledge	13.601***	-3.016***
	Acquired machinery	13.451***	-0.512***
	Overall	**15.5943*** **	**-2.351*** **
Spillover	Industry-level	-0.138**	-0.283***
	Country-level	-0.274	-0.341***
	EU-level	0.573***	0.64***
	Overall	**0.436** **	**0.582*** **

Notes: Significance, * 5%, ** 1%, *** 0.1%. For intramural, cooperative, and extramural rows, effects of R&D extensity are based on a discrete change from 0 to 1, and effects of R&D intensity are based on a 1% increase. For spillovers, effects of R&D extensity are based on a 1% increase of R&D-active firm ratio, and effects of R&D intensity are based on a 1% increase of R&D expenditure-turnover ratio. The overall cooperative marginal effect is an average rather than a sum because no firms engage in all cooperative R&D activities at the same time. Similar logic applies to extramural R&D. In contrast, the overall spillover effects are a sum because every firm faces three spillover pools. Only extensity measures of cooperative R&D are available, so the intensity equation also uses these extensity measures to control for cooperative R&D, but the estimates (very close to the extensity equation estimates, see Table 2) are omitted from Table 3.

The estimated coefficients of the structural equation models in Table 2 can be used to infer the signs of the effects but not their magnitudes due to nonlinear model specifications and incommensurable variable definitions. To provide an intuitive interpretation and comparison, we calculate marginal effects of R&D activities on productivity in Table 3. For example, if an average firm engages in intramural R&D, its productivity is expected to rise by 13% compared to if the firm does not. And if it increases its intramural R&D expenditure-turnover ratio by 1%, then its productivity is expected to rise by 21%. The survey by Hall et al. (2010) found research elasticities ranging from 0.01 to 0.25 but centred on 0.08 or so. For comparison our results must be multiplied by the average ratio of R&D to turnover, so with a 5% ratio our response of 21% is at the bottom of the Hall et al. (2010) range.

Our first finding is that the smaller spillover hypothesis (H1) is confirmed. To see this, we calculate productivity effects of a typical support[12] via intramural R&D and spillovers respectively to make a fair comparison. For intramural R&D effects, we need to multiply the effects of support on R&D extensity/intensity[13] (Table 2) with the effects of R&D extensity/intensity on productivity (Table 3). The results are that a typical support boosts productivity by 7.26% via extensity and 163.9% via intensity over a two-year period[14]. In contrast, spillover effects of a typical support are smaller than the intended effects.

We provide a numerical calculation to demonstrate this. To compare like with like, we evaluate the two effects based on a typical government support. This way, we can see how much effect of the support is due to the intended intramural mechanism and how much is due to the unintended spillover effect. Assume a spillover pool has N firms with the same productivity (normalised as 1) initially. According to Column (2) of Table 2, the R&D intensity of a firm rises by 2.443% after receiving a typical support.

The **intended intramural effect** for the supported firm is equal to 12.595% (extensity) plus 2.443*21.283% (intensity) = 64.58%. The **unintended spillover effect** for the other firms is calculated as follows.

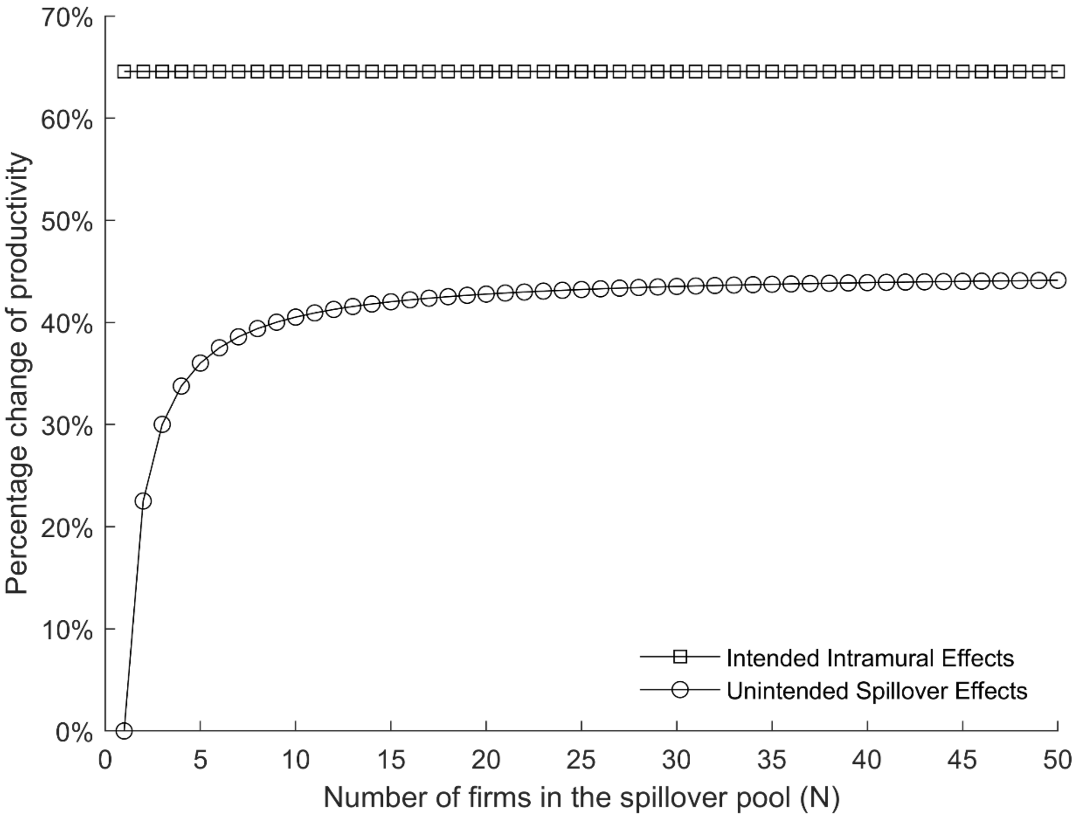
- Extensity: If $1\% * N$ more firms engage in support-induced R&D, then the spillover effect of one R&D firm on another firm in the pool is equal to $\frac{0.436\%}{1\% * N}$. The spillover is received by all-but-one firms, so the aggregate spillover effect is equal to $\frac{0.436\% * (N-1)}{1\% * N}$.
- Intensity. A typical support raises the R&D expenditure-turnover ratio by 2.443 percentage points (Table 2), and each percentage in turn raises productivity by 0.582%. But this effect is diluted by all firms in the pool, so the effect on the average expenditure-turnover ratio is $\frac{2.443 * 0.582\%}{N}$. Again, all-but-one firms receive the spillover, so the aggregate spillover effect (intensity) is $2.443 * 0.582\% * \frac{N-1}{N}$.

To summarise, the total unintended spillover effect is equal to $0.436\% * \frac{N-1}{1\% * N} + 2.443 * 0.582\% * \frac{N-1}{N}$. As N rises, the aggregate spillover effect (extensity) rises. However, *the scale*

[12] Unfortunately, the data only disclose whether each firm receives support, not how much support it receives. Therefore, the effect of support is calculated over a discrete change from 0 to 1.
[13] The coefficients of columns (2)-(2)' of Table 2 imply that the marginal effect of support on R&D extensity is $\Pr(RRDIN = 1|FUN = 1) - \Pr(RRDIN = 1|FUN = 0) = 56\%$, and the marginal effect of support on R&D intensity is $RRDINX(FUN = 1) - RRDINX(FUN = 0) = 7.8\%$.
[14] $56\% \times 12.959 = 7.26\%$ and $7.8\% \times 21.283 = 163.6\%$.

effect converges to a fixed level because $\frac{N-1}{N}$ *converges to 1*. The intended and unintended effects have the following empirical relationship—the intramural effect of one typical support on one firm's productivity is greater than the spillover effects on all other firms combined. The gap is smaller as N gets bigger, but the scale effect diminishes and converges to around 45% which is still substantially lower than the intended effect 64.58% as shown in Figure *3*. The spillover effect lies between 22% to 45% (boundaries), smaller than the intramural effect (H1). This finding is in line with the survey of Ugur et al. (2020)[15].

Figure 3 Comparison between intended intramural effects and unintended spillover effects



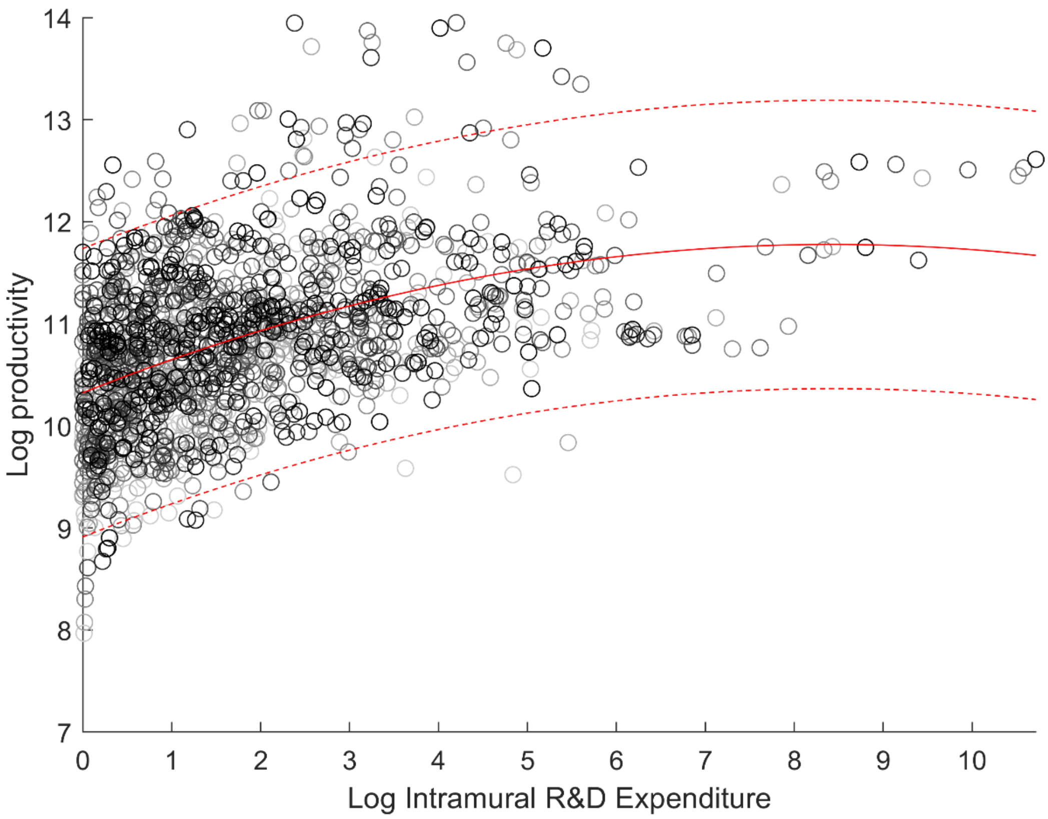
The second result concerns where R&D support should most effectively be given. Finding that R&D intensity can promote productivity more than R&D extensity indicates one direction. This result was most likely because enterprises already undertaking R&D had greater relevant skills and experience than those beginning R&D. The finding suggests that, compared to trying to make everyone innovative, it would be more effective to support those with greater competence and richer experience in R&D. In this regard, efficient state capitalism may exacerbate inequality in productivity. Another direction for effective support is that collaborating in R&D with suppliers benefits the firm but doing so with a client does the reverse. This may suggest that in our sample R&D reduces costs but does not raise revenue, so the R&D effect on the

[15] If new research opportunities arise exogenously in a given area, then all firms in that area will do more R&D and may improve their productivity, an effect that may upwardly bias the spillover measure. Correcting the effects of such shocks would merely shrink further the small spillover estimates we find.

client is positive while that on the supplier is negative. For R&D extensity extramural R&D tends to be more effective than cooperative R&D[16]. We conjecture that the cooperative link requires more complex, hard-to-manage relationships.

Our third result is that R&D effects are nonlinear with respect to the intensity of R&D expenditure, as shown in Table 4 and in the scatter plot of Figure 4. Specifically, the productivity return to intramural R&D expenditure declines only after the sample average (at turning point 2.987% compared with 1.8% sample average). This finding suggests that most sample firms underinvest in intramural R&D and have yet to reach the optimal level.

Figure 4 Nonlinearity in the productivity effect of intramural R&D



Notes: Each circle stands for a country-industry pair in a particular year. A darker shade indicates a more recent observation (2006-2018).

One seemingly odd result is the negative productivity effects of extramural R&D intensity in Table 3. Possibly the thinness of experience may be a reason; there is very little extramural R&D. Note that in the model of R&D intensity, marginal effects are estimated at the observed level of R&D. In the light of the nonlinearity finding above, in Table 4 we separately estimate the R&D effect before and after the turning points. The results suggest that the estimated marginal effects are negative because firms invest too little in extramural R&D in the form of

[16] Cooperative R&D intensity cannot be measured for lack of appropriate data.

outsourced R&D and acquisition of knowledge. Should the extramural R&D intensity go beyond the turning points, the effects become positive (the U trajectory). The only exception is the acquisition of machinery, which always has a negative effect on productivity (the inverted U trajectory), indicating that on average firms do not have the capacity to fully utilise the advanced machines they aquire. Therefore, it would be more productive if R&D expenditures were limited to purchasing innovative software (knowledge) rather than hardware (machinery).

Table 4 Nonlinearity in marginal effects of R&D intensity

	Intramural R&D	**Extramural**		
		Outsourced R&D	Acquired knowledge	Acquired machinery
Linear coeff	8.65***	-3.578***	-3.054***	-0.501***
Quadratic coeff	-1.448**	3.629***	2.205***	-0.273***
Shape	Inverted U	U	U	Inverted U
Turning point	2.987**	0.493***	0.693***	-0.918**
Effect (<Turning)	20.601***	-3.532***	-3.021***	NA
Effect (>Turning)	14.55***	2.674***	1.049***	-0.512***
Effect (Overall)	20.241***	-3.525***	-3.016***	-0.512***

Notes: Significance, * 5%, ** 1%, *** 0.1%.

5.1 Heterogeneities and Robustness

Separating the sample into “old” EU members (Germany, Spain, and Portugal) and “new” EU members (former centrally planned economies) in Table 5, we can see some differences: there are cross-sectional heterogeneities in both R&D intensity and R&D quality. The “old” are more effective in intramural R&D extensity, presumably because from longer market experience they know better which firms to pull into subsidy regimes. The two groups are about equally effective in intramural R&D intensity. For the average enterprise in both groups a 1% increase in the intramural R&D ratio raises productivity by about 16%. The “old” have more positive spillover effects; their R&D is complementary with other firms’ R&D. The “new” have insignificant or even negative spillover effects, most likely because their R&D is substitutable for other firms’ R&D.

A multinational corporation (MNC) plant is here identified as a member of a larger group of enterprises with a headquarters in another country. There is strong intra-group financing of R&D due to a large presence of multinationals based in the “new” members of the EU (European Investment Bank, 2018 p13). We find no difference between “old” and “new” in intramural R&D effects on productivity, consistent with Vujanović et al. (2022). There are more MNCs in the new (13% of enterprises in the “new” members of the EU are MNCs and 8% in the “old”) and MNCs are more intensity productive than non-MNC enterprises. These findings confirm the dampening hypothesis (H2).

For the “old”, a wider distribution of firms doing R&D raises productivity by more than for “new” EU member countries; the old are more extensity productive. “New” economies are much less extensity productive than “old” which is also explained by a greater contribution of MNCs because MNCs are less extensity productive. Because the “new” are generally less productive across their economies, extensity productivity is much greater for the “old”. As relative GDP per capita suggests, Germany (“old”) is more directly R&D productive than the average of the rest (mainly “new”).

Table 5 Heterogeneity and Robustness of Marginal Effects

	Intramural R&D		**Spillover**	
Subsample	Extensity	Intensity	Extensity	Intensity
Old	23.803***	16.098***	2.7881***	1.8142***
New	10.398***	16.506***	-0.728***	0.1220
MNC	9.1111***	19.997***	1.0749***	1.4867***
non-MNC	14.322***	15.825***	0.0152	0.4034***
2018	23.952***	16.693***	-0.729*	1.4408***
non-2018	12.334***	16.358***	0.2269	0.4039**
Germany	41.417***	19.041***	5.2773***	-2.613***
non-Germany	11.955***	16.249***	-0.191	0.7315***

Notes: These marginal effects are estimated on the entire sample with heterogeneity across four dimensions: MNCs, country, industry, and year. By contrast, Table 3 is estimated with heterogeneity only across the last three dimensions.

Spillover sensitivity continues to be small (at least relative to direct effect when significantly positive). It is biggest for German extensity R&D productivity but small compared with huge direct German extensive productivity. Germany also has the largest negative spillover. Spillovers are larger for MNCs than non-MNCs, though still proportionately smaller compared to direct R&D effects.

To test the robustness of our estimates we compare the effects of dropping the data for the year 2018 (Table 5). R&D intensity productivity is similar for both the data of year 2018 and the sample without this data. In both estimates the spillovers are small though the intensity spillover is larger for 2018 and for the sample excluding 2018. In 2018 full recovery from the financial and debt crises may have pushed up extensity to account for the big difference in extensity while maintaining similar intensity.

5.2 Absorptive Capacity vs. Absorptive Possibility

Enterprises able to derive a competitive advantage from knowledge of their environment have a strong absorptive capacity (Cohen & Levinthal, 1990). Such businesses may be in a better position to utilise spillovers. (Estrada et al., 2010; Harris et al., 2021). We measure two types of enterprise absorptive capacity: the influence of a firm's R&D intensity on spillover effects of other firms' R&D intensity and the influence of R&D extensity on spillover effects of other firms' R&D extensity. With few firms to copy from close to the frontier of technology the possibility of absorbing spillovers is low with limited benefits and the spillovers are likely expensive to adapt. Hence, the costs of absorption can exceed the benefits; the estimated R&D coefficient can be negative.

To quantify the magnitudes of absorptive capacity effect, we regress the estimated spillover effects on the average intramural R&D activities (extensity or intensity). Thanks to the heterogeneity along country, industry, and year in the structural model (interactive terms or slope dummies), we can construct a spillover effect for each country-industry pair in each year. To explain the variations of these spillover effects, RRDIN (extensity) and RRDINX (intensity) are averaged over country-industry pairs and years to construct a pseudo-panel (Guillerm, 2017)[17]. Compared to the original firm-level data which do not have observations of past R&D activities, the pseudo-panel data can capture lag effects of the same country-industry unit.

Table 6 R&D spillovers in a pseudo-panel

	Extensity			Intensity		
	Industry	Country	EU	Industry	Country	EU
RRDIN(t)	0.2386*	-5.8607***	-0.3188			
RRDIN(t-1)	0.3559***	-3.6475*	-2.2470***			
RRDINX(t)				0.4412	-1.1439	-0.8225
RRDINX(t-1)				1.9581***	2.8613*	-4.2682***
Constant	-0.2808***	1.4937***	0.9880***	-0.3037***	-0.1120***	0.6803***
No. of Obs.	1073	1073	1073	1115	1115	1115

Notes: Any enterprise observation with the same country, industry, year, and MNC attributes should share marginal effects. That is why we can collapse the data into a pseudo panel with one or more combinations of these dimensions. Here we collapse the data into a pseudo panel in terms of country-industry pair and year. Then the marginal effects are averaged over the subsamples old/new, MNC/non-MNC, 2018/non-2018, Germany/non-Germany.

The pseudo-panel regression with the country-industry pair fixed effects with an allowance for a lag is reported in Table 6. A negative coefficient means the absorptive possibility dominates the capacity effect. The mixed results suggest that the absorptive capacity hypothesis (H3) holds with conditions. If a firm is already at the technological frontier, it is less likely to copy or absorb from others, i.e., the negative absorptive possibility effect dominates the positive absorptive capacity effect. Table 6 shows that the impact of R&D intensity and extensity on the spillover marginal effects are small compared to direct effects and the largest numbers are

[17] A pseudo-panel here is observations over time of cohorts, stable groups of enterprises, rather than enterprises themselves. Individual variables are replaced by their intra-cohort means.

negative. The intensity coefficients mean intramural R&D more than two years ago[18] helps absorb spillovers in the same country and in the same industry but not EU spillovers. Generally absorptive capacity does not matter much if at all for spillover, compared with direct effects.

6 Conclusion

We modelled how R&D enters the innovation system in four ways (intramural, extramural, cooperative, and spillover) and in two dimensions (extensity and intensity). We found four key results and corresponding policy implications of the R&D effects on productivity based on a very large sample of European enterprises.

First, we simultaneously estimated three levels of R&D spillover effects on labour productivity to avoid underestimation, finding that the spillovers were smaller than direct R&D productivity effects (H1). We prove that spillover effects are bounded because the scale effect of a larger pool eventually converges to a fixed level. It is a better use for support to focus on promoting intramurally planned R&D activities, at least if policymakers know more than does the market. Second, extramural R&D tended to be more efficient than cooperative R&D, probably because cooperation required more complex, and costly to manage, relationships. Thus, it is advisable that businesses should outsource R&D rather than trying to cooperate in fields where they do not have expertise. Third, the R&D extensity effect on labour productivity is considerably less than the R&D intensity effect. Promoting R&D intensity would have been more effective than supporting extensity, because enterprises already undertaking R&D usually have greater relevant skills and experience than those beginning R&D. Therefore, support should focus on boosting intensity rather than extensity promotion. Fourth, the nonlinearity of the R&D intensity effect suggests that the optimal extramural R&D intensity was higher than the actual level. Consequently, sample firms could boost productivity more effectively either by abandoning extramural R&D or by doing much more. In sum, all key results support one general principle—*specialisation precedes diversification*. Instead of widely spreading the limited resources, firms and governments should identify where the use of resources is more effective, either to R&D activities with greater effects or to R&D firms with specialised expertise.

There were substantial differences in our sample between enterprises and countries in terms of R&D spillovers. However, surprisingly the intramural R&D intensity productivity was similar between "new" and "old" EU members. We attribute this to the greater incidence of multinational corporations in the new members spreading the expertise of the old. Multinational investments appear to compensate for the absence of spillovers in Eastern Europe by dampening cross-country differences in R&D productivity effectiveness and therefore should be welcomed

[18] The lag is one CIS cross-section which is on average two years.

by policymakers (H2). The absorptive capacity effect in utilising spillovers, measured by R&D, was found to be conditional on whether the firm is at the technological frontier (H3).

A qualification to these results is that our approach does not take into account technological proximity between firms which may determine their spillover effects. Nor does the approach necessarily allow a long enough period for the results to be fully evaluated; the survey data specification gives a maximum of three years which may not always be sufficient for commercial R&D or its spillovers to bear fruit, as our absorption estimates pseudo-panel suggest. There are some possible incoming spillovers that we have not attempted to measure, namely those from outside our EU sample. R&D spillovers have been found especially substantial in the US. However, most economies in the present sample are much smaller and are relative newcomers to the market economy, which may help explain the different spillover findings. Another possible reason for the small size of our spillovers is that they include potentially offsetting business stealing or competition effects.

In the literature, spillovers in knowledge flows are proxied by a variety of physical flows. It would be helpful for future research to establish which physical flows are most accurate measure. Research based on firm-level panel data would be able to address the dynamics of R&D spillovers, which is beyond the scope of this paper. Including all those joining the EU between 2004 and 2007 and all the older members, thus, widening the coverage of the data could further test the robustness or our results. There is a widespread interest in quantifying the return to innovation subsidies, so including adequate information about support in the Community Innovation Surveys could facilitate future research.

References

Banal-Estañol, A., Duso, T., Szücs, F. (2022) R&D spillovers through RJV cooperation, Research Policy 51 (4)

Barney, J. (1991). Firm Resources and Sustained Competitive Advantage. Journal of Management, 17(1), 99-120. https://doi.org/10.1177/014920639101700108

Bloom, N., Schankerman, M., Van Reenen, J. (2013). Identifying technology spillovers and product market rivalry. Econometrica 81 (4), 1347–1393.

Freitas, I., & Von Tunzelmann, N. (2008). Mapping public support for innovation: A comparison of policy alignment in the UK and France. Research Policy, 37(9), 1446-1464. https://doi.org/https://doi.org/10.1016/j.respol.2008.05.005

Cassiman, B., Veugelers, R. (2002) R&D Cooperation and Spillovers: Some Empirical Evidence from Belgium. Am. Econ. Rev. 92 4 1169-1184

Castellacci, F., Natera, J. (2013). The dynamics of national innovation systems: A panel cointegration analysis of the coevolution between innovative capability and absorptive capacity. Research Policy, 42(3), 579-594. https://doi.org/10.1016/j.respol.2012.10.006

Cincera, M. (1998) Economic and technological performances of international firms, ULB Institutional Repository 2013/212081, Universite Libre de Bruxelles.

Cirera, X., Maloney, W. (2017). The innovation paradox. Developing-country capabilities and the unrealized promise of technological catch-up. Washington D.C.: World Bank.

Cohen, W., Levinthal, D. (1990) Absorptive Capacity: A New Perspective on Learning and Innovation. Administrative Science Quarterly, 35, 1, 128-52.

Conte, A., Vivarelli, M. (2014). Succeeding in innovation: key insights on the role of R&D and technological acquisition drawn from company data. Empirical Economics, 47(4), 1317-1340. https://doi.org/10.1007/s00181-013-0779-1

Corrado C., Haskel, J., Jona-Lasinio, C. (2017) Knowledge Spillovers, ICT and Productivity Growth. Oxford Bulletin of Economics and Statistics, 79, 4

Crepon, B., Duguet, E., Mairesse, J. (1998). Research, innovation and productivity: an econometric analysis at the firm level. Econ. Innov. New Technol. 7 (2), 115–158.

Crescenzi, R., Gagliardi, L. (2018). The innovative performance of firms in heterogeneous environments: The interplay between external knowledge and internal absorptive capacities. Research Policy, 47(4), 782-795. https://doi.org/10.1016/j.respol.2018.02.006

Escribano, A., Fosfuri, A., Tribó, J. (2009). Managing external knowledge flows: The moderating role of absorptive capacity. Research Policy, 38(1), 96-105. https://doi.org/10.1016/j.respol.2008.10.022

Estrada, I., de la Fuente, G., Martín-Cruz, N. (2010) Technological joint venture formation under the real options approach. Research Policy, 39, 9, 1185-97.

European Investment Bank (2018) Innovation investment in Central, Eastern and South-Eastern Europe, December EIB Regional study.

Foreman-Peck, J. and Zhou, P. (2022) R&D subsidies and productivity in eastern European countries, Economic Systems, 46, 2, 100978

Fuller , A. F. (2018) Toward a Perspective on R&D Outsourcing: RBV and Firm Performance, International Journal of Innovation and Technology Management,15, 5, 1850040

Geroski, P., (1991). Innovation and the sectoral sources of UK productivity growth. Economic Journal 101, 1438–1451

Goodridge, P., Haskel, J., Wallis, G. (2017) Spillovers from R&D and other intangible investment: evidence from UK industries. Review of Income and Wealth Series 63, Supplement 1, February

Goya, E., Vayá, E., and Suriñach, J. (2016). Innovation spillovers and firm performance: Micro evidence from Spain (2004–2009). Journal of Productivity Analysis, *45*(1), 1–22

Greene, W. (2000). *Econometric Analysis* (4th ed.). London: Macmillan.

Grliches, Z. (1992) The Search for R&D Spillovers. Scandinavian Journal of Economics 94, supplement: S29-S47.

Guillerm, M. (2017) Pseudo-panel methods and an example of application to Household Wealth data, Economie et Statistique / Economics and Statistics No 491-492

Gupta, A., Govindarajan, V. (2000). Knowledge flows within multinational corporations. Strategic Management Journal, 21(4), 473-496.

Hall, B., Mairesse, J., Mohnen, P. (2010) Measuring the Returns to R&D, in Handbook of the Economics of Innovation, Vol. 2, 1033-1082 Chapter 24.

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). pp. 589-607.

Hashi, I., Stojčić, N. (2013a) The impact of innovation activities on firm performance using a multi-stage model: Evidence from the Community Innovation Survey 4, Research Policy 42 353– 366

Hashi, I., Stojčić, N. (2013b) Knowledge spillovers, innovation activities and competitiveness of industries in EU member and candidate countries. Economic Annals, 58(198)

Iammarino, S., Piva, M., Vivarelli, M., Von Tunzelmann, N. (2012). Technological Capabilities and Patterns of Innovative Cooperation of Firms in the UK Regions. Regional Studies, 46(10), 1283-1301. https://doi.org/10.1080/00343404.2012.679259

Jaffe, A. (1986) Technological opportunity and spillovers of R&D: evidence from firms’ patents, profits and market value. Am. Econ. Rev. 76 (5), 984–1001.

Jaffe, A. (1989) Real effects of academic research. Am. Econ. Rev. 79 (5), 957–970.

Knudsen, L., Nielsen, B. (2010). Collaborative capability in R&D alliances: exploring the link between organisational- and individual-level factors. International Journal of Knowledge Management Studies, 4(2), 152-175. https://doi.org/10.1504/ijkms.2010.03079

Kogut, B., Zander, U. (1992). Knowledge of the firm, combinative capabilities, and the replication of technology. Organization Science, 3, 383-397. https://doi.org/10.1287/orsc.3.3.383

Lavie, D. (2006). The Competitive Advantage of Interconnected Firms: An Extension of the Resource-Based View. The Academy of Management Review, 31(3), 638-658. http://www.jstor.org/stable/20159233

Lee J. H. and Kim, (2019) Competition and Innovation Spillover, Seoul Journal of Economics, 32, 4, 468-491

Lee, L.-F. (1978). Unionism and Wage Rates: A Simultaneous Equations Model with Qualitative and Limited Dependent Variables. International Economic Review, 19(2), 415-433.

Lööf, H., Mairesse, J. and Mohnen, P. (2017) CDM 20 years after, Economics of Innovation and New Technology, 26:1-2, 1-5

Lucking, B., Bloom, N., Van Reenen, J. (2019) Have R&D Spillovers Declined in the 21st Century? Fiscal Studies, 40, 4, pp. 561–590 0143-56712

Luintel, K., Khan, M. (2017) Ideas production and international knowledge spillovers: Digging deeper into emerging countries. Research Policy. 46, 10, 1738-1754

Luintel, K., Khan, M. (2009). Heterogeneous ideas production and endogenous growth: an empirical investigation. Can. J. Econ. 42 (3), 1176–1205

Maddala, G. (1983). *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge: Cambridge University Press.

Mathews, J. (2003). Competitive dynamics and economic learning: An extended resource-based view. Industrial and Corporate Change, 12(1), 115-145.

Myers, K. Lanahan, L. (2022) Estimating spillovers from publicly funded R&D: Evidence from the US Department of Energy – Am. Econ. Rev. 112 7 2393-2423

Neves, P., Sequira, T. (2018) Spillovers in the production of knowledge: A meta-regression analysis, Research Policy 47, 4, 750-767

Pessoa, A., (2005). Ideas driven growth: the OECD evidence. Portuguese Econ. J. 4, 46–67.

Stojčić, N. (2021). Collaborative innovation in emerging innovation systems: Evidence from Central and Eastern Europe. The Journal of Technology Transfer, 46(2), 531-562. https://doi.org/10.1007/s10961-020-09792-8

Stojčić, N., Srhoj, S., Coad, A. (2020). Innovation procurement as capability-building: Evaluating innovation policies in eight Central and Eastern European countries. European Economic Review, 121, 103330. https://doi.org/10.1016/j.euroecorev.2019.103330

Szczygielski, K., Grabowski, W., Pamukcu, M., Tandogan, V. (2017). Does government support for private innovation matter? Firm-level evidence from two catching-up countries. Research Policy, 46(1), 219-237. https://doi.org/10.1016/j.respol.2016.10.009

Tevdovski, D., Kosevska-Trpcevska, K., Disoska, E. (2017) What is the role of innovation in productivity growth in Central and Eastern European countries? Economics of Transition and Institutional Change, 25(3), 527–551

Ugur M, Churchill SA and Luong HM (2020) What do we know about R&D spillovers and productivity? Meta-analysis evidence on heterogeneity and statistical power, Research Policy, 49 1 103866

Von Brasch, T., Cappelen, A., Hungnes, H., Skjerpen, T. (2021) Modeling R&D spillovers to productivity: the effects of tax credits. Economic Modelling 101 105545

Vujanović, N., Radošević, S., Stojčić, N., Hisarciklilar, M., Hashi, I. (2022). FDI spillover effects on innovation activities of knowledge using and knowledge creating firms: Evidence from an emerging economy. Technovation, *118*, 102512.

Vujanović, N., Stojčić, N., Hashi, I. (2021). FDI spillovers and firm productivity during crisis: Empirical evidence from transition economies. Economic Systems, *45*(2), 100865.

Wakelin, K. (2001) Productivity growth and R&D expenditure in UK manufacturing firms. Research Policy 30: 1079–1090

Wieser, R. (2005) Research and development, productivity and spillovers: empirical evidence at the firm level. Journal of Economic Surveys 19, 4, 587-621

Zahra, S. George, G. (2002). Absorptive Capacity: A Review, Reconceptualization, and Extension. The Academy of Management Review, 27(2), 185-203. https://doi.org/10.2307/4134351

Zahra, S., George, G. (2002) Absorptive capacity: a review, reconceptualization, and extension. Academy of Management Review, 27, 185–203.

Appendix 1: Notes on CIS data

Turnover. The data on current turnover (TURNOVER) and previous turnover (TURNOVER_LAST) are recorded separately in CIS2006 and CIS2008. Later years only have current turnover, but previous turnover can be derived based on the variable TURN_GROWTH. Note that the Germany 2010 dataset uses a different unit (in thousands), so the position of the decimal point in the data for Germany was corrected.

R&D Expenditure Ratios. There are four intensity ratios, RRDINX_RAT, RRDEXX_RAT, RMACX_RAT, and ROEKX_RAT. They are directly available from 2010-2016. The datasets in 2006-2008 are derived by expenditures (RRDINXM, RRDEXXM, RMACXM, and ROEKXM) divided by TURNOVER. The dataset in 2018 is coded differently with more details. A mapping between CIS2018 and previous datasets are listed in Table A1. Of particular interest is the ratio of R&D expenditure over turnover, defined as RND = the sum of RRDINX_RAT and RRDEXX_RAT. All ratios are capped at 100%.

R&D Activities. Before CIS2018, RRDIN and RRDEX are intramural and extramural R&D activities respectively. The counterparts in the CIS2018 datasets are INNA_IH_RND and INNA_RND_CONTR. The variable RDENG refines the two cases (continuous or occasional engagement) when RRDIN is equal to 1. This refinement is included in CIS2018 as two separate variables INNA_IH_RND_CONT and INNA_IH_RND_OCC. Other measurements of R&D activities include RMAC (acquisition of machinery or "hardware") and ROEK (acquisition of knowledge or "software"), the counterparts of which are summarised in Table A2. Unless completely missing in the entire year for a country, we do the following to fill the missing values.

- RRDIN. If RDENG is equal to 1 or 2, then missing RRDIN observations are filled with 1. If RRDINX_RAT is positive, then RRDIIN observations are filled with 1. Remaining missing RRDIN observations are filled with 0.
- RRDEX, RMAC, ROEK. If ratio counterparts of these variables (RRDEXX_RAT, RMACX_RAT, ROEKX_RAT) are positive, then the observations are filled with 1. The remaining missing values are filled with 0.

External Funding. There are three broad categories of public funding sources for innovation activities: local government (FUNLOC), central government (FUNGMT), and the EU (FUNRTD and FUNEU). We combine the EU's Research and Technological Development framework programme (FUNRTD) with other EU fundings (FUNEU) into one category. In CIS2018, FUNRTD is renamed as Horizon Programme 2020 (FUND_EU_HP2020). We also do not distinguish R&D funding from others since this refinement is only available for CIS2018. Details of the mapping are shown in Table A3.

Industry Code. In CIS2006-CIS2008 datasets, industry is classified using NACE pro, which is a combination of letters and numbers. From CIS2010 onwards, two-digit NACE codes are recorded. To keep consistency, we convert all two-digit NACE codes to NACE pro categories as shown in Table A4.

Employment Size. The measure of employment size evolves over time. In CIS2006-CIS2008, there are three categories in the measure of current and past employment sizes (EMP04, EMP06, EMP08). Throughout CIS2010-CIS2018, three measures are available containing different refinements of (current) employment size categories. SIZE_2 has two categories, SIZE_3 has three, and SIZE_4 has four. To keep consistency, we use SIZE_3 as the baseline and convert other size measures to three categories in a new variable (SIZE). To calculate productivity later, we follow Tevdovski et al. (2017) to assign a numerical value to employment size (EMP) by the middle value of the employment size category to which the firm belongs to. For example, if SIZE of a firm is “<50”, then we set EMP to 25 for that observation. The mapping among the categorical and numerical measures is shown in Table A5.

Cooperation. Throughout CIS2006-CIS2016, the cooperation variables roughly correspond to source-of-information variables: enterprise or group (CENTG), suppliers (CO2* or SSUP), clients/customers (CO3* or SCLI), competitors (CO4* or CCOM), consultants (CO5* or CINS), universities (CO6* or CUNI), and governments (CO7* or CGMT). In CIS2018, more detailed information is provided but new variables can be mapped into old ones.

Table A1: Mapping between CIS datasets on R&D expenditure

CIS2006-CIS2016 Definitions	**Variable Names**	**Variable Names**	**CIS2018 Definitions**
Expenditures on intramural R&D	RRDINX_RAT = RRDINXM/TURNOVER	EXP_INNO_RND_IH_RAT	Expenditure in R&D performed in-house - share of real turnover
Expenditures in extramural R&D	RRDEXX_RAT = RRDEXXM/TURNOVER	EXP_INNO_RND_CONTR_OUT_RAT	Expenditure in R&D contracted out to others (including enterprises in own group)
Expenditures in acquisition of machinery	RMACX_RAT = RMACXM/TURNOVER	EXP_TOT_ACQ_MEBTA_RAT	Expenditure on acquisition of machinery, equipment, buildings and other tangible assets
Expenditures in acquisition of external knowledge	ROEKX_RAT = ROEKXM/TURNOVER	EXP_TOT_MKT_RAT	Expenditure on marketing, brand building, advertising (include in-house costs and purchased services)
		EXP_TOT_TNG_RAT	Expenditure on training own staff (include all in-house costs including wages and salaries of staff while being trained, and costs of purchased services from others)
		EXP_TOT_PRD_DESG_RAT	Expenditure on product design (include in-house costs and purchased services)
		EXP_TOT_SOFT_DBA_RAT	Expenditure on software development, database work and data analysis (include in-house costs and purchased services)
		EXP_TOT_IPR_RAT	Expenditure on registering, filing and monitoring own Intellectual Property Rights (IPRs) and purchasing or licensing IPRs from others

Table A2: Mapping between CIS datasets on R&D activities

CIS2006-CIS2016 Definitions	Variable Names	Variable Names	CIS2018 Definitions
Engagement in intramural R&D	RRDIN	INNA_IH_RND	In-house R&D activities
Type of engagement in intramural R&D: 1 continuously, 2 occasionally	RDENG	INNA_IH_RND_CONT	Continuous in-house R&D activities
		INNA_IH_RND_OCC	Occasional in-house R&D activities
Engagement in extramural R&D	RRDEX	INNA_RND_CONTR_OUT	R&D contracted out to other enterprises (include enterprises in own group) or to public or private research organisations)
Engagement in acquisition of machinery	RMAC	PUR_MES_SAME	Purchase of machinery, equipment or software based on the same or improved technology used before in the enterprise
		PUR_MES_NEW	Purchase of machinery, equipment or software based on new technology not used before in the enterprise
Engagement in acquisition of external knowledge	ROEK	CKNO_CONF_TRDF_EXHIB	Acquisition of knowledge by: Conferences, trade fairs or exhibitions
		CKNO_JRNST_TRDP	Acquisition of knowledge by: Scientific/technical journals or trade publications
		CKNO_ASS_PROF_IND	Acquisition of knowledge by: Information from professional or industry associations
		CKNO_PAT_PUBL	Acquisition of knowledge by: Information from published patents
		CKNO_DOC_STD_COM	Acquisition of knowledge by: Information from standardisation documents or committees
		CKNO_WEB_NET_CDS	Acquisition of knowledge by: Social web-based networks or crowd-sourcing
		CKNO_B2B_OS	Acquisition of knowledge by: Open business-to-business platforms or open-source software
		CKNO_RE	Acquisition of knowledge by: Extracting knowledge or design information from goods or services (reverse engineering)

Table A3: Mapping between CIS datasets on funding

CIS2006-CIS2016 Definitions	Variable Names	Variable Names	CIS2018 Definitions
Public funding from local or regional authorities	FUNLOC	FUND_AUT_LOC_REG	Financial support received from local/regional authorities
		FUND_AUT_LOC_REG_RNDINN	Financial support from local/regional authorities used partly or fully for R&D or other innovation activity
Public funding from central government	FUNGMT	FUND_GOV_CTL	Financial support received from national governments
		FUND_GOV_CTL_RNDINN	Financial support from national government used partly or fully for R&D or other innovation activity
Funding from EU's Framework Programme	FUNRTD	FUND_EU_HP2020	Financial support received from EU 2020 Horizon Programme
		FUND_EU_HP2020_RNDINN	Financial support from EU 2020 Horizon Programme used partly or fully for R&D or other innovation activity
Public funding from the EU	FUNEU	FUND_EU_OTH	Financial support received from other EU institutions
		FUND_EU_OTH_RNDINN	Financial support from EU institutions used partly or fully for R&D or other innovation activity

Table A4: Mapping between NACE 2 digit (CIS2010-CIS2018) and NACE pro (CIS2006-CIS2008)

NACE	NACE Pro			
CIS2010+	CIS2006	CIS2008	Definitions	Category
1		A	Agriculture, forestry, and fishing	1
2				
3				
5	C	B	Mining and quarrying	2
6				
7				
8				
9				
10	DA	C10-C12	Manufacture of food, beverages, and tobacco	3
11				
12				
13	DB	C13-C15	Manufacture of textiles, apparel, and leather	4
14				
15	DC			
16	20-21	C16-C18	Manufacture of wood, paper, and media	5
17				
18	22			
19	DF-DG	C19-C23	Manufacture of fuel, chemical, pharmaceutical, and plastic	6
20				
21				
22	DH			
23	DI			
24	27	C24-C25	Manufacture of metals	7
25	28			
26	DL	C26-C30	Manufacture of electronic, electric, machinery, vehicles	8
27				
28	DK			
29	DM			
30				
31	DN	C31-C33	Manufacture of furniture and others	9
32				
33				
35	E	D	Electricity, gas, steam, and AC	10
36		E	Water, sewerage, waste	11
37				
38				
39				
40				
41	F	F	Construction	12
42				
43				
45	50	G	Wholesale and retail	13
46	51			

47	52			
49	60	H49-H51	Transport	14
50	61			
51	62			
52	63	H52-H53	Warehousing and courier	15
53				
55		I	Accommodation	16
56				
58		J58-J60	Publishing, motion picture, and TV	17
59				
60				
61	64	J61-J63	Telecom and programming	18
62				
63				
64	J	K	Financial and insurance	19
65				
66				
68	70	L	Real estate	20
69		M69-M70	Legal, accounting, and consulting	21
70				
71	73-74	M71-M73	Research	22
72				
73				
74		M74-M75	Design, photographic, translation, and veterinary	23
75				
77	71, 72	N	Administrative	24
78				
79				
80				
81				
82				

Table A5: Mapping between categorical and numerical employment size measures

CIS2006-CIS2008	**CIS2010-CIS2018**	**CIS2010-CIS2018 excluding CIS2014**	**CIS2014**	**CIS2010-CIS2018**		**CIS2006-CIS2018**	
EMP04/EMP06/EMP08	SIZE_1	SIZE_2	SIZE_2	SIZE_3	SIZE_4	SIZE	EMP
<50	No information	<50		<50	<50	<50	25
50-249		>=50	10-249	50-249	50-249	50-249	150
>=250			>=250	>=250	250-499 >=500	>=250	275

Appendix 2: Notes on CDM model

The original CDM model (the initials of the three authors Crépon, Duguet, and Mairesse) was a static model estimated on cross-sectional French data. Crépon et al. (1998) estimated their model by asymptotic least squares (also known as minimum distance estimator). Their framework introduced a four-equation structural model that related productivity to innovation output, innovation output to research, and research to its determinants. The paper used new information similar to that provided by the European Community Innovation Surveys, in particular the share of sales of innovative products. It also developed an explicit modelling framework, to apply appropriate estimation methods in the presence of sample selectivity (due to the firm's choice of whether or not to undertake R&D), potential endogeneity of some of the dependent variables, and the qualitative nature of some of the dependent variables (binary or categorical).

Taken together, R&D, innovation, and productivity equations formed a recursive nonlinear system, for which there were two versions, one with patent counts and the other with the share of innovative sales. The first two equations were the same in the two versions. The two dependent variables were k^*, the latent research capital per employee and g', where a firm was observed to invest in research if g' was positive or larger than some threshold. The next equation was either: n^*, a patent equation, as a heterogeneous count data process with an expectation conditional on research and other variables or an equation for t, innovative sales. The final equation was for productivity as a function of one or other of patents or innovative sales.

Only a small proportion of firms engage in research activities and/or apply for patents; productivity, innovation, and R&D are endogenously determined; research investment and capital are truncated variables, patents are count data and innovative sales are interval data.

CDM found the probability of engaging in research (R&D) for a firm increased with its size (number of employees), its market share and diversification, and with the demand pull and technology push indicators. The research effort (R&D capital intensity) of a firm engaged in research increased with the same variables, except for size (its research capital being strictly proportional to size).

Later research in the field modifies the CDM model in various ways such as adding more endogenous variables and/or using alternative endogenous variables. Variants of the CDM model are estimated mainly using CIS data, such as Lööf & Heshmati (2002) for Sweden, Griffith et al. (2006) for Germany, Spain, France, and the UK, and Conte & Vivarelli (2014) for Italy. Specifically, the closest model to our paper is the one by Foreman-Peck & Zhou (2022), who extend the CDM model with an additional funding equation.