

Blind, Knut; Schubert, Torben

Article — Published Version

Estimating the GDP effect of Open Source Software and its complementarities with R&D and patents: evidence and policy implications

The Journal of Technology Transfer

Provided in Cooperation with:

Springer Nature

Suggested Citation: Blind, Knut; Schubert, Torben (2023) : Estimating the GDP effect of Open Source Software and its complementarities with R&D and patents: evidence and policy implications, The Journal of Technology Transfer, ISSN 1573-7047, Springer US, New York, NY, Vol. 49, Iss. 2, pp. 466-491,
<https://doi.org/10.1007/s10961-023-09993-x>

This Version is available at:

<https://hdl.handle.net/10419/312272>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>



Estimating the GDP effect of Open Source Software and its complementarities with R&D and patents: evidence and policy implications

Knut Blind^{1,2} · Torben Schubert^{1,3}

Accepted: 30 December 2022 / Published online: 28 February 2023
© The Author(s) 2023

Abstract

Open Source Software (OSS) has become an increasingly important knowledge asset in modern economies. However, the economic impact of OSS on countries' GDP is ambivalent due to its public good character. Using a cross-country panel from 2000 to 2018, including 25 of the largest EU countries plus the USA, Japan, Korea, Canada, China, Norway, and Switzerland, matching OSS commits to GitHub to macroeconomic data provided by the OECD, our results confirm the dual nature of OSS. On the one hand, the open-access character creates great learning potential by providing a commonly accessible productive resource for all countries. On the other hand, it creates outward-directed spillovers associated with own OSS contributions. Accordingly, on average, we find that countries experience an increase in GDP when the world stock of OSS grows. However, smaller countries experience a decline in GDP resulting from their own contributions due to knowledge spillovers. The net effect is nonetheless positive. If no country contributed to OSS development, GDP for the average country would be 2.2% lower in the long run. Moreover, the losses associated with unintended spillovers are lower for countries with a higher R&D and patenting intensity. Based on our findings, we derive implications for policies and regulations concerning OSS.

Keywords Open Source Software · GitHub · R&D · Patents · GDP

✉ Knut Blind
knut.blind@isi.fraunhofer.de

Torben Schubert
torben.schubert@isi.fraunhofer.de

¹ Fraunhofer Institute for Systems and Innovation Research ISI, Karlsruhe, Germany

² TU Berlin, Berlin, Germany

³ Department of Design Sciences, CIRCLE - Centre for Innovation Research, Lund University, Lund, Sweden

1 Introduction

Departing from older growth models treating technological progress as exogenously determined (Solow 1956), new growth theory explained growth as the result of endogenously determined investments into research and development (R&D) (Grossman & Helpman, 1994; Romer, 1990; Aghion & Howitt, 1992). These models of economic growth have in common that appropriation of the returns to R&D through patenting the resulting knowledge is essential. Indeed, several studies have shown a positive impact of patents on economic growth (Akçomak & Ter Weel, 2009; Crosby 2000; Gould & Gruben, 1997; Hasan & Tucci, 2010; Lach, 1995). However, studies have increasingly revealed overprotection of knowledge through too strong or too long patent protection can slow down growth because it limits access to and diffusion of new valuable knowledge (Philippe & Nyssen 2000, Iwai-sako and Futagami 2003). As a result, scholars have begun emphasizing the crucial role of public knowledge pools, including scientific publications (Arora et al., 2021) or technical standards for growth (in't Veld, 2019; Swann, 2000; Blind & Jungmittag, 2008; Blind et al., 2021b; ISO, 2021).

Complementary to the emphasis on public knowledge pools, firms' investment into intangibles, like research and development, has been identified as being an increasingly relevant source of economic growth and productivity (Corrado et al., 2009). Here, software is a major and increasingly relevant pillar among the different types of intangibles. However, the literature about intangibles does so far not explicitly consider Open Source (Nagle, 2019b).

While the creation and provision of Open Source Software (OSS), defined as software released under a license complying with the Open Source Initiative's definition¹, as a specific form of public knowledge pool, is not a new phenomenon. It has gained increasing traction in the last two decades. More than 65% of firms now use or contribute to OSS, according to the Black Duck Software Survey (2016). Indeed, for many important commercial software packages, there exist open access solutions, including free operating systems (Linux, Android), statistical software (R), or image editing (GIMP). In some cases, the OSS solutions are even cutting-edge, with no equivalent commercial competitors. The Python packages TensorFlow and Keras for deep learning are good examples. While this suggests that OSS is becoming a decisive production factor, its macroeconomic outcomes have, besides very few exceptions (Ghosh, 2006), hardly been addressed empirically, resulting in a lack of knowledge on both the overall size of the effects as well as the specific mechanisms channeling them.

Therefore, we derive the following research questions. (1) How do the domestic and global contributions to OSS influence a country's Gross Domestic Product (GDP)? (2) Which complementary activities related to research and development, but also patents, influence the impact of OSS on GDP?

In this paper, we contribute to answering the above-derived research questions by exploiting rich empirical data, which can capture not only the overall macroeconomic effects of OSS but which also allow identifying spillovers and externalities as well as the relation between OSS and protectable knowledge investments and assets, R&D and patents, in particular. We build this framework as an adapted extension of the panel-cointegration model developed by Bottazzi and Peri (2007). To operationalize the model, we rely on the

¹ See details here <https://opensource.org/osd>.

recent availability of long-time cross-country series of OSS contributions through GitHub as the largest repository of OSS, which we match to country-level macroeconomic statistics. Relying on data from 2000 to 2018,² we estimate the long-run relationship between OSS commits to GitHub and GDP. Our results show strong evidence that OSS has a dual role to play. On the one hand, the home country commits reduce GDP in the originating country, consistent with the view that knowledge spills over without compensation. However, the larger countries, the less they suffer from these knowledge spillovers. On the other hand, we find that the global pool of commits increases GDP in all countries. Yet, the second effect strongly dominates the first, implying that OSS commits are associated with considerable net economic returns. Moreover, we show that higher R&D and patenting intensities can help countries to reduce the GDP losses generated by their own OSS contributions. The contribution of this paper is threefold: on a theoretical level, we systematize the thinking about the dual role of public knowledge pools for growth. On an empirical level, we present the first framework to estimate the economic returns of OSS as the probably most rapidly growing public pool of knowledge. Specifically, we document the overall value of OSS and corroborate the duality of high social returns and large uncompensated spillovers. Building on this insight, our third contribution relates to developing policy recommendations. Because economic policy has a national focus, but knowledge spillovers occur globally, designing effective policies related to OSS is far from trivial. Specifically, we propose measures that we can classify as strengthening incentives to contribute and increase the skill base.

The remainder of the paper is structured as follows: Sect. 2 provides a literature review and the derivation of the hypotheses. Section 3 presents the empirical strategy and data, followed by Sect. 4, which describes the results being discussed in Sect. 5. Finally, Sect. 6 concludes with policy implications and the limitations of our approach.

2 Literature and hypotheses

In general, we base the derivation of our hypotheses on the nascent economic but growing literature about OSS. The literature focuses broadly on two separate topics. The first relates to the incentives to contribute to the provision of OSS despite its public good nature and thereby addresses the supply side. The second one analyzes the economic effects/value of OSS for firms or the economy, i.e., the demand side. For our empirical research, we will eventually rely only on the latter strand of the literature because we do not consider the generation of OSS. Beyond the economic works, there are, of course, substantial bodies of literature dealing with the management and organization of OSS projects. Although relevant to understanding OSS in their own respect, these works will not be covered here (see a recent overview in Blind et al., 2021a) because they are more relevant for the supply side.

While using or generating OSS also implies private returns, early on, authors have highlighted the public good nature of OSS, which is important for our analysis of its impact on GDP. For example, Johnson (2002) characterized the development of OSS as the dynamic provision of a public good, referring to the example of the GNU-Linux operating system. In his approach, individual user-programmers invest their effort to contribute to software

² Technically, GitHub was founded only in 2008. However, backward migration implies that there is data also before its foundation year.

that will become a public good. However, he also shows that free-riding may prevent the development of valuable OSS code. Hawkins (2004) modifies the perception of Johnson (2002) and defines OSS as a quasi-public good in contrast to a true public good, in which the cost of production is small compared to the social benefit but large compared to the private benefit. Taking the open source HTTP-server project Apache as an example, Hawkins (2004) argues that it is more profitable for IBM to invest in the OSS code than to keep and maintain its proprietary solution because IBM is not bearing the entire cost of providing this quasi-public good (Bonaccorsi & Rossi, 2003; Bitzer et al., 2007) perceive OSS also as a public good. Still, they understand OSS development as the private provision of a public good, which is driven by play value, homo ludens payoff, user programmers, and gift culture benefits. Kubiszewski et al. (2010) understand OSS as an information good being a subcategory of a public good, which is upgraded through use. For them, the status-driven incentive structures based on individuals' reputations derived from their contributions are the main promoters for developing OSS.

While the literature about the public good character of OSS has not made an effort yet to assess the macroeconomic value of OSS, it still provides important elements for such analyses. Specifically, highlighting the public good nature implies a dual role of OSS characterized by simultaneously high value and high externalities. This renders OSS a complicated resource whose macroeconomic impacts are a priori unclear. Therefore, one way to assess its relevance is to integrate it into the macroeconomic production function framework. Lerner and Schankerman (2010) provided a previous contribution to this direction by putting software in general and OSS in particular in the context of the new growth theory. In principle, they argue that OSS could provide the best available software at essentially zero cost, exploiting its non-rivalry property's full advantage. Thus, OSS could have a large impact on economic development. However, they do not empirically quantify the economic impact of OSS derived from their theoretical considerations.

Contributing to closing this gap in the quantitative impact assessment of OSS, Ghosh (2006) integrated OSS into a simulation model to explain labor productivity. In this model, a hypothesized duplication of OSS investment leads to a 0.1% increase in GDP. However, this finding appears to depend strongly on the modeling assumptions, which remained untested. Moreover, the finding has never been empirically validated. Again, the main argument by Ghosh (2006) explores the significant savings related to the development of software, which is beneficial for economic development. For example, Mockus (2007) finds that 50% of popular OSS code is often reused in several projects. The economic rationale for this cost-saving effect is elaborated by Riehle (2007) in a microeconomic model. In addition, Ghosh (2006) argues that OSS potentially saves industry investments in software development, resulting in increased profits or more usefully spent on further innovation activities. Robbins et al. (2018) quantify the resource cost and, therefore, also the savings for some popular OSS packages.

Summarizing the theoretical arguments about the economic characteristics of OSS, we can conclude that not only single firms' (Nagle, 2019b) but also countries' macroeconomic production is expected to benefit from an increasing pool of OSS, which can be considered as public good similar to open standards (e.g., Blind et al., 2021b).

Therefore, we integrate, in addition to capital, labor, and other investment in knowledge or technological progress, explicitly contributions to OSS in macroeconomic product functions to answer our two research questions. In particular, we are interested in the implications of OSS being a public good for countries' GDP. Due to digitalization, OSS can be considered a global public good, which can benefit all countries in general. However, complementary knowledge assets have also to be considered. Furthermore, the public good

character of OSS also has implications for the countries contributing to OSS because they generate knowledge spillovers that cannot be appropriated domestically.

Since the globally available pool of OSS can be considered a public good, we expect that all countries' GDPs benefit from it. Therefore, we derive our first hypothesis:

Hypothesis 1 Increases in the global pool of OSS contributed by the rest of the world affect the countries' GDP positively.

Nagle (2018) argues that in addition to traditional learning by doing via applying a specific technology, which has significant implications for growth also at the economy level (Arrow, 1962 or Romer 1990), contributing may itself be a learning activity. It is, for example, possible that the advice received by senior experts (e.g., Lakhani & von Hippel 2003) generates continuous feedback and thus facilitates learning and, therefore, provides additional economic benefits. However, we have to consider the knowledge spillovers to all the other countries from the rest of the world, which are going to benefit by increasing their productivity. Consequently, their enhanced productivity and competitiveness hamper domestic competitiveness and, therefore, reduce, as shown by Nagle (2018) for individual companies' value added, the domestic GDP. Nevertheless, we assume that the former learning by doing and learning by contributing impact is a prerequisite to using OSS effectively and is, therefore, stronger than the possible, but not necessarily realized negative spillover created by a country. Consequently, we derive the following second hypothesis:

Hypothesis 2 Countries contributing more to OSS development have a higher GDP than countries contributing less.

A further aspect that drives own OSS contributions relates to complementarities within the regular innovation processes. First, employees active in R&D can use the available OSS for their own work, e.g., in improving the increasingly software-based research processes, eventually increasing the productivity of R&D expenditures for the country as a whole. In a similar vein, Reisinger et al. (2014) argue that firms, although not benefitting directly from contributing to OSS, use their investments into OSS to upgrade complementary goods or services and thereby sell more of them at a higher price. A particularly important driver of this complementarity is that resulting from hardware-software interrelation (Di Gaetano, 2015). Thus, there is probably a strong complementarity between OSS and R&D investments which results from the fact that OSS is often embedded in complementary products or services (Amiri-Kordestani & Bourdouce, 2017; Krogh & Spaeth, 2007), implying a higher customer utility. Interestingly, because the complementary goods may indeed be patentable (although OSS itself is by definition not), there may also exist a complementarity with patent stocks. Consistent with this argumentation, Aksoy-Yurdagül (2015) shows that for firms with higher patent stocks, there is a stronger effect of OSS on firm value. Another source of complementarity may be that between OSS and proprietary software (Lerner & Schankermann, 2010; O'Reilly 1999), which may be protected by software patents. As Bessen and Hunt (2007), these patents, in particular in the US, are of considerable importance, having accounted for 15% in the 2000s. Indeed, they were found to be important drivers of firm value (Hall & MacGarvie, 2010). Finally, patents represent the proprietary knowledge base of a country, whereas OSS is—as elaborated above—a public good. The complementarity between proprietary and publicly available knowledge can be beneficial for the productivity of companies (David et al., 2000), for example, because

maintaining proprietary knowledge pools may reflect overall higher technological capabilities or higher absorptive capacities (Cohen & Levinthal, 1990). Likewise, contributions to OSS code may be a form of selective revealing of knowledge (Alexy et al., 2013) or generative appropriability (Ahuja et al., 2013) and may help to extract value from external knowledge provided by other companies or organizations. Consequently, we derive the following third hypothesis.

Hypothesis 3 (a) For countries with a higher R&D intensity, the associated GDP losses of own contributions are lower. (b) For countries with a higher patent intensity, the associated GDP losses of own contributions are lower.

3 Data and methodology

In this paper, we aim to estimate the macroeconomic effects of OSS on GDP. To do that, we will exploit variation in contributing to OSS across countries and time. Accordingly, we build a country-level panel dataset linking OSS-related activities to economic outcomes in terms of GDP.

3.1 Data

As our measure of OSS, we will use commits to GitHub aggregated at the country level. GitHub is an internet-based system for hosting software and maintaining accurate version control, which can also be accessed through a web interface. The related OSS data obtained from the GitHub developer platform is collected by TU Delft in the context of the GHTorrent project (<https://ghtorrent.org/>). GitHub was launched in early 2008, but code from various previously existing repositories has been transferred to GitHub, implying that data is available from 2000 onwards. After its launch, GitHub quickly became the primary repository for OSS projects, with more than 1.3 billion OSS lines of code or commits in 2018. Meanwhile, these commits were contributed by more than 32 million users in 2018 compared to 15 million in 2016 (Ojanperä et al., 2019), originating from more than 680,000 organizations. Earlier empirical studies rely on SourceForge (Von Engelhardt & Freytag, 2013; Von Engelhardt et al., 2013; Lakka et al., 2015). Meanwhile, platforms such as SourceForge, with 3.7 million users (SourceForge, 2016), and Launchpad, with 3.1 million users (Launchpad, 2016), have far fewer users than GitHub (Ojanperä et al., 2019) and are therefore of relatively minor importance. Finally, the archive data provided by SourceForge is not up to date anymore, which does not allow an adequate assessment of the current impact of OSS. Finally, our analyses cover only data until 2018. Consequently, the implications of the takeover of GitHub by Microsoft in the same year are marginal for our analyses. While GitHub commits are not necessarily used by others, they still should reflect reasonable proxies for uptake. As outlined in the literature review, OSS can be considered as user innovation or a form of co-creation between developers and users. In the second step, it also promotes the learning of the contributors as they receive feedback from the crowd of more experienced users and are, therefore, able to better capture value from using the goods (Nagle, 2018). Indeed, in other contexts, several authors have used code contributions to GitHub or SourceForge as measures of OSS (Nagle, 2019a, Wright et al., 2020; Engelhardt et al., 2013; Lakka et al., 2015).

The largest part of economic data in the model was comparably easy to collect as processed and cleaned country-level data is provided by the OECD and the World Bank for License Payments. As a measure for the output Y , the total value-added of a country, GDP, is used. An exception was a measure of physical capital stocks, which was taken from the compilation of Berlemann and Wesselhöft (2017). As a result, the authors provide the only available capital stock indicator being consistent over countries and uniformly covering the long-term panel dimension of almost 30 years.

Panel data is used, which is available from 2000 to 2018, and all EU members except Croatia, Cyprus, and Malta. In addition, several other countries, which are either located in Europe or contributing massively to GitHub, are also included, specifically the USA, Japan, Korea, Canada, China, Norway, and Switzerland.

To get more intuition for the resulting data, we present the basic descriptives of the variables used in this paper in Table 1. All monetary variables are in constant US-\$. All other variables are in natural units. However, since all regressions are in log-log format, the regression coefficients can be interpreted as elasticities so that scaling does not affect the results. Observe that our sample covers a set of 576 time-year observations (32 countries observed over 18 years). The effective regression observations are reduced because dynamic OLS constructs dynamic control variables of leads and lags that create a further drop-out in the effective regression sample to 480 observations.

Moreover, we present some results on the development of GitHub commits over time in the left panel of Fig. 1 and a scatterplot of GDP per employee and GitHub commits per employee on the right panel. Overall, we see that GitHub commits have been strongly increasing over time, in particular between 2010 and 2016. A slight slowdown was observed in 2017–2018. If we look at the bivariate correlations, we also see that the GDP per employee and GitHub commits per employee are positively correlated. Obviously, the positive correlation should not be understood to imply causation. A more sophisticated approach to measuring the effects of GitHub commits on GDP follows in Sect. 4.

Looking at a breakdown by countries year in Table 2, we see that the US strongly dominates GitHub. In 2001, 65.2% of all commits came from the US. Interesting to see is that this share, while still being very large, dropped to 41.0% in 2018. While many countries, such as Germany (8.5% in 2018), Italy (1.3% in 2018), or the Netherlands (2.3% in 2018), kept relatively stable shares, the countries increasing their contributions relatively included in particularly the UK (from 1.5 to 8.0%) and China (from 0.0 to 8.2%) (see also Wachs et al., 2022).

3.2 Empirical approach

In this section, we outline the econometric approach used to identify the economic effects of OSS on GDP. For this, we create a macro-econometric regression framework built on by Bottazzi and Peri (2007) and Jungmittag et al. (1999). Then, applying the approach used by Jungmittag et al. (1999) to calculate the impact of standardization or by Nagle (2018) to analyze the influence of OSS on the micro-level of US companies, the baseline model relies on a simple Cobb–Douglas production function as follows:

$$Y_{it} = A_{it-1} K_{it}^{\alpha} L_{it}^{\beta} F(.) \quad (1)$$

where Y denotes GDP, K denotes capital, and L denotes labor in country i at time t , where the coefficients α and β refer to measuring their respective production elasticities. $F(.)$ contains further log-linearised input factors or control variables. Most importantly, A_{it-1}

Table 1 Descriptive statistics

Variable	Source	Obs	Mean	Std. Dev.	Min	Max
GDP (millions constant \$)	The World Bank WDI	576	1489463.74	2944491.83	14888.74	17856476.00
Capital Stock (millions constant \$)	Berlemann and Wesselhöft (2017)	576	4251825.82	7971769.53	23158.76	44959844.00
Employment (millions)	OECD STAN Rev. 3	576	40.10	134.45	0.19	787.40
Payments for the use of IP (millions constant \$)	World Bank World Development Indicators	576	6801.60	11898.28	10.18	85119.03
R&D expenditures (millions constant \$)	OECD Science, Technology, and R&D Statistics	576	34915.32	79031.74	70.64	500268.84
R&D expenditures by ROW (millions constant \$)	OECD Science, Technology, and R&D Statistics	576	1083028.40	199662.23	507514.78	1448891.63
Transnational patent applications	PASSAT	576	6561.57	13417.88	2.00	69662.00
GitHub commits	GitHub/ GHTorrent project (https://ghtorrent.org/)	576	617938.12	2722514.42	1.00	31797050.00
GitHub commits by ROW	GitHub/ GHTorrent project (https://ghtorrent.org/)	576	19200811.20	27768428.61	37721.00	77325880.00

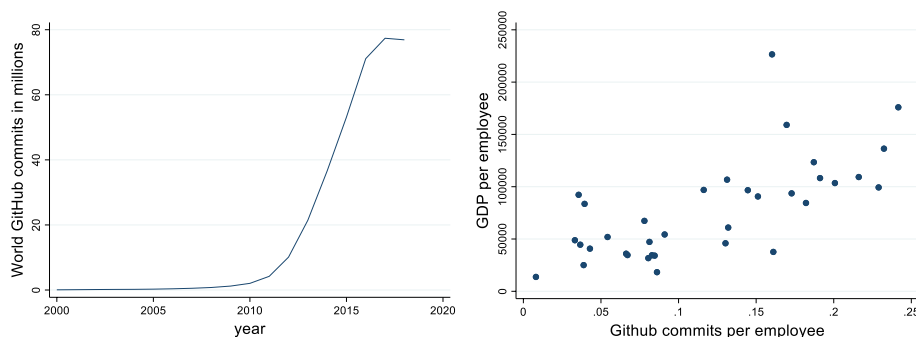


Fig. 1 World GitHub commits (left) & GDP-GitHub scatterplot in 2018 (right)

denotes the knowledge stock, which is modeled based on a structural approach proposed by Bottazzi and Peri (2007). In this approach, the evolution of the knowledge stock is modeled as a function of R&D and the existing knowledge stock. If it is further allowed that there may be differential effects from foreign and domestic R&D expenditures, it is assumed the following log-linear function:

$$\log(A_{it})' = \varepsilon_1 \log RD_{it-1} + \varepsilon_1 \log RD_{it-1}^{ROW} + \varepsilon_1 \log A_{it-1} \quad (2)$$

where $(A_{it})'$ refers to the change in the knowledge stock, RD_{it} is the R&D expenditures, and the superscripts ROW refers to the rest of the world. Including ROW variables are particularly important because this can separate spillover from privatized effects. When taking logs of Eq. (1), approximating the change of the knowledge stock by the number of annual patents, our central equation of interest can be rewritten as follows:

$$\log Y_{it} = \gamma_1 \log RD_{it-1} + \gamma_2 \log RD_{it-1}^{ROW} + \gamma_3 \log PAT_{it} + \alpha \log K_{it} + \beta \log L_{it} + \log F(.) \quad (3)$$

Now assuming that among the additional factors that constitute the other input factors and control variables also contain OSS, our estimation model can be rewritten as follows:

$$\begin{aligned} \log Y_{it} = & \gamma_1 \log RD_{it-1} + \gamma_2 \log RD_{it-1}^{ROW} + \gamma_3 \log PAT_{it} + \alpha \log K_{it} + \beta \log L_{it} \\ & + \gamma_4 \log OSS_{it-1} + \gamma_5 \log OSS_{it-1}^{ROW} + \log x_{it\mu} \end{aligned} \quad (4)$$

where $\log x_{it}$ are logged versions of generic control variables (see below), and OSS_{it-1} refers to measures approximating national and global OSS contributions. H1 and H2 would then suggest that both γ_4 and that γ_5 are positive.

To test H3a/b, we extend the model in Eq. (4) to follow up on how OSS differs by the country's level of own contributions to OSS and by its R&D as well as patent intensity. To do this, we allow that the coefficient γ_4 in Eq. (4) differs for countries with below median R&D intensity (R&D expenditures divided by workforce) and patent intensity (patents divided by workforce). We implement this methodology by creating a dummy variable, d , indicating whether the intensity variable is above the median.

Table 2 Country shares of GitHub commits over time

Country	2001	2018
Austria	1.18	0.69
Belgium	0.01	0.86
Bulgaria	0.01	0.37
Canada	2.13	4.57
China	0.05	8.27
Croatia	0.00	0.16
Cyprus	0.04	0.03
Czech Republic	1.54	0.92
Denmark	0.69	0.73
Estonia	0.35	0.15
Finland	1.12	0.81
France	3.25	5.69
Germany	8.28	8.54
Greece	0.01	0.34
Hungary	0.01	0.41
Ireland	0.24	0.52
Italy	1.18	1.31
Japan	5.05	3.11
Korea, Rep.	0.06	1.22
Latvia	0.00	0.10
Lithuania	0.31	0.16
Luxembourg	0.00	0.06
Malta	0.02	0.04
Netherlands	2.77	2.39
Norway	0.45	0.87
Poland	0.10	1.98
Portugal	0.00	0.55
Romania	0.01	0.45
Slovak Republic	0.40	0.15
Slovenia	0.00	0.12
Spain	0.20	2.32
Sweden	2.20	1.52
Switzerland	1.46	1.49
United Kingdom	1.59	8.08
United States	65.28	41.01

Then we include the following two interactions $\log\text{OSS}_{it-1} \cdot d$ and $\log\text{OSS}_{it-1} \cdot (1 - d)$, and test for the equality of the coefficients for both variables.

3.3 Identification strategy

The choice of estimation procedure for Eq. (4) and the extended models depend on the assumptions of the variables. Most importantly, because our dataset has a relatively large T compared to N dimension, regular panel data methods may fail. An important issue

relates to non-stationary time series. Typically, when time series are non-stationary, regular OLS-type regressions lead to inconsistency because usual asymptotic theorems (such as the law of large numbers or central limit theorems) no longer apply. Many time series, such as GDP, are known to be non-stationary. Likewise, the results in Bottazzi and Peri (2007) show that the relationship expressed in Eq. (2) contains non-stationary variables. Moreover, given the vastly increased volume of OSS, the OSS is very unlikely to follow a stationary trend. A common approach to dealing with non-stationary data is integrating (differentiating) them until they become stationary. However, this approach can be problematic if the non-stationary time series are co-integrated, i.e., there is a linear combination of them such that this combination is stationary. Economically speaking, such a stationary combination often results from an economic law binding two non-stationary time series together in the long run. If, for example, OSS and GDP followed such a law - as, for example, the production law in Eq. (4)—the special co-integration estimators are preferable to differentiation.

However, choosing co-integration estimators always requires showing that the specific conditions necessary for co-integration analysis are met. Specifically, co-integration techniques require that the relevant time series are non-stationary and that they control indeed for a long-term stationary relationship. The equations above are the long-term growth equations and, in sum, reflect the requirement to combine technological and economic indicators (Castellacci, 2007; Bottazzi & Peri, 2007) devise a model in which it can be expected that patent stocks, international patent stocks, and R&D are co-integrated. We followed a regular two-step procedure to show the validity of a co-integration panel estimator. First, we tested the hypothesis that all time series are non-stationary using regular panel-unit root tests. Second, we tested whether the non-stationary time series are co-integrated using panel-co-integration tests. In particular, we relied on the panel/group t-tests, which are known to outperform alternative tests in terms of power and size in finite samples. Finally, the co-integrating relationships are estimated based on the extensions of the Bottazzi and Peri (2007) model using alternative panel co-integration estimators, in particular DOLS (Dynamic OLS).

4 Empirical results

4.1 Panel co-integration tests

As concerns the question of co-integration, we test all variables in Eq. (4) for whether they contain unit roots. For none of the variables in the tests presented in Table 3, we were able to reject non-stationarity. We then continued with a panel co-integration test. This test presented at the bottom of Table 3 rejects the null hypothesis of no co-integration. Thus, overall, all conditions necessary to apply panel co-integration are met.

4.2 The baseline relationship

The results of the basic macroeconomic production function are presented in Table 4. In Cobb–Douglas production functions, the coefficients represent elasticities and the production shares that the production factors receive as compensation in terms of wages our capital payments. The coefficients lie in reasonable ranges between 0.57 and 0.65 for capital

Table 3 Stationarity and Co-integration Tests

	Inverse normal stat.	p-val.
Stationarity tests (H0: all panels contain unit roots)		
Log GDP	2.68	0.99
Log R&D expenditures	2.03	0.97
Log R&D expenditures by ROW	9.84	1.00
Log transnational patent applications	1.66	0.95
Log GitHub commits	5.00	1.00
Log GitHub commits by ROW	2.23	0.98
Co-integration test (H0: No co-integration)		
Panel-t Stat for joint co-integration of all non-stationary variables	-3.16***	0.00

and 0.27 and 0.39 for labor. Similar findings have, for example, been found by Schubert and Neuhäusler (2018).

With respect to our H2, the impact of national investments into OSS measured by the commits of the users, which can be attributed to a country, is significantly negative for national GDP in both Columns 1 and 3. This is likely to reflect that OSS investments produce costs (see also Nagle, 2019b), which must be privatized while the rest of the world socializes the benefits. Specifically, the development costs are not immediately compensated by increased productivity or international competitiveness because every other country has free access to this OSS code. At the same time, the public good character of the OSS code is confirmed by the significantly positive impact of the contributions to OSS by the rest of the world in Columns 2 and 3. Therefore, the national GDP is significantly benefitting from the global investment in OSS. Thus, we corroborate H1 but not H2.

Moreover, we note that from the coefficients in Column 3 of Table 4, the net effect of OSS is indeed high and positive. If no country contributed (i.e., both stocks would be reduced by 100%), this would imply a GDP-gain (evaded loss) of 2.7% ($-100\% \cdot -0.027$), but the country would also experience a loss (foregone gain) of -4.9% ($-100\% \cdot 0.049$). Thus, without any OSS contributions, the GDP of the average country would lose 2.2% of its GDP.

Several other variables are interesting in this model. Notably, the role of technological progress is represented by a set of variables. First, the import of foreign technologies, measured by the payments for the use of intellectual property covering licensing payments for patents and copyrights, including software, is a significant driver of GDP. Likewise, the domestic R&D expenditures are positive, albeit weakly related to GDP. The R&D expenditures of the rest of the world are negative for domestic growth, most likely because they push the competitiveness of the other countries harming the domestic balance of trade, i.e., both negative for exports and positive for imports. Finally, like R&D, also national patent applications are positive. Their relationship appears to be much stronger for GDP, potentially indicative that R&D is still uncertain in its outcomes. At the same time, for patent

Table 4 Impact of OSS commits on GDP (Dynamic OLS, all countries)

	(1)	(2)	(3)
	Log GDP	Log GDP	Log GDP
Log capital stock	0.61073*** (0.06211)	0.57971*** (0.06259)	0.65714*** (0.05823)
Log employment	0.39544*** (0.12445)	0.27717** (0.12626)	0.27397** (0.11671)
Log payments for use of IP	0.03102*** (0.01201)	0.03915*** (0.01233)	0.04485*** (0.01145)
L.Log R&D expenditures	0.06579* (0.03455)	0.06217* (0.03489)	0.03966 (0.03221)
L.Log R&D expenditures by ROW	0.09855 (0.09070)	− 0.33582*** (0.12191)	− 0.27525** (0.11276)
Log transnational patent applications	0.03572*** (0.00378)	0.04011*** (0.00384)	0.04467*** (0.00355)
L.Log GitHub commits	− 0.01571*** (0.00418)		− 0.02746*** (0.00462)
L.Log GitHub commits by ROW		0.02646*** (0.00704)	0.04907*** (0.00770)
Country dummies	Yes	Yes	Yes
Observations	480	480	480
R^2	0.882	0.900	0.917
N_g	32	32	32

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

applications, at least the technological risk, i.e., the risk of technological (rather than market) failure, should be minimal.

It has to be pointed out that this effect is different from the impact of global investment into R&D, which is hampering the national GDP, because here the results are not public and freely available due to secrecy measures or they are protected by intellectual property rights, like patents and other rights. In this respect, our findings underline that OSS measured by the code contributed to GitHub represents a pool of knowledge, which is accessible and usable by all companies and individuals worldwide, and, therefore, represents a public good, which can be of considerable economic value for the economies and societies as such.

4.3 The role of learning by contributing, R&D, and patenting

In addition to the presented basic models, we also investigated whether R&D or patenting (H3a/b) play a role. The results are displayed in Table 5.

We indeed find the effect of home contributions is substantially more negative in the group of countries with below median R&D intensities (− 0.029) as compared to the countries with above median R&D intensities (− 0.017). The same pattern can be observed for patent intensities, where for the group of countries with below median intensities, the

coefficient is -0.028 , while the coefficient in the group of countries with above median patenting intensity is -0.021 . The difference in the R&D coefficients is significant at the 1%-level, which indicates that, indeed, the negative spillover effect of contributing is lower in countries that rely more on R&D. The difference is also significant for the patenting intensity, albeit only at the 10% level. Overall, we find strong support for H3a and somewhat weaker support for H3b.

4.4 Robustness checks

The robustness of the different models has been tested in several respects. Firstly, we have checked whether our results depend on the choice of the co-integration framework. In an alternative set-up, we have ignored the issue of non-stationarity and have applied regular fixed effects regressions. The results are fully corroborated and are even more significant. We have also confirmed that calculating heteroscedasticity robust standard errors did not lead to different conclusions in the fixed effects model.

Second, special attention is needed because our results are suggestive of the existence of spillover effects that lead to negative GDP effects for home country contributions. While the argument is similarly made and tested in the framework by Bottazzi and Peri (2007), it is, of course, true that spillovers occur not only across the boundaries of countries but also inside them, which is also true for OSS (see Wachs et al., 2022). With our country-level data, we are unable to identify intra-country spillover. However, it stands to reason that the degree of spillover-internalization should be increasing in the size of the country, whereas the cross-border spillovers should decrease in the size of the country. To test this, we estimate a fixed-effects model with the size of the country (measured in terms of workforce) as an additional moderator for OSS contributions from within the country and by the rest of the world. The results are found in Table 6. A visualization is in Fig. 2. Indeed, we see the predicted patterns corroborated, giving additional support to the argument that spillovers may play an important role in the provision of OSS. Furthermore, because for large countries, the effects of home contributions turn positive, we conclude that home contributions are not necessarily growth-reducing. Specifically, our results suggest that home contributions can become a driver of growth if a country is large enough to internalize a sufficient share of the spillovers.

Third, we have rerun all regressions using the number of contributors to GitHub as an alternative measure of OSS. On the one hand, the number of contributors may be seen as a more direct measure of labor input and may facilitate the interpretation in this respect. On the other hand, the individual contributors' allocations to GitHub differ widely, implying that this measure is also a noisy one. In either case, the results did not differ significantly, irrespective of the OSS measure we used.

Moreover, we have tested whether certain countries are responsible for the significant effects. We have, for example, excluded China or the USA without, however, observing different results. We have also run an analysis only for the sample of EU countries, found in Tables 7 and 8. In Table 7, we see that the positive role of OSS contributed by the rest of the world and the negative impact of the country-specific contributions on the growth of the EU Member States are almost identical to the results of the panel, including all countries, with some differences. Most notably, while the duality of OSS is confirmed, there are important differences in size. While the negative effect of own contributions is with -0.027 , roughly identical to the full sample, the positive spillover effect from contributions by the rest of the world is with 0.035 , much smaller than in

Table 5 Impact of OSS commits on GDP with group splits (Dynamic OLS, all countries)

	(1) Log GDP	(2) Log GDP
Log capital stock	0.66575*** (0.05609)	0.68381*** (0.05861)
Log employment	0.07083 (0.12502)	0.13823 (0.12622)
Log payments for use of IP	0.04065*** (0.01097)	0.04301*** (0.01127)
L.Log R&D expenditures	0.06438** (0.03128)	0.05082 (0.03185)
L.Log R&D expenditures by ROW	− 0.29538*** (0.10799)	− 0.30128*** (0.11165)
Log transnational patent applications	0.05382*** (0.00339)	0.04792*** (0.00350)
L.Log GitHub commits: high own commits		
L.Log GitHub commits: low own commits		
L.Log GitHub commits by ROW	0.04340*** (0.00756)	0.04603*** (0.00773)
L.Log GitHub commits: high R&D intensity	− 0.01736*** (0.00613)	
L.Log GitHub commits: low R&D intensity	− 0.02918*** (0.00440)	
L.Log GitHub commits: high patent intensity		− 0.02096*** (0.00623)
L.Log GitHub commits: low patent intensity		− 0.02893*** (0.00453)
Country dummies	Yes	Yes
Observations	480	480
R^2	0.926	0.923
N_g	32	32
Eq. coeffs. chi-sq(1)	6.97***	3.05*

Standard errors in parentheses

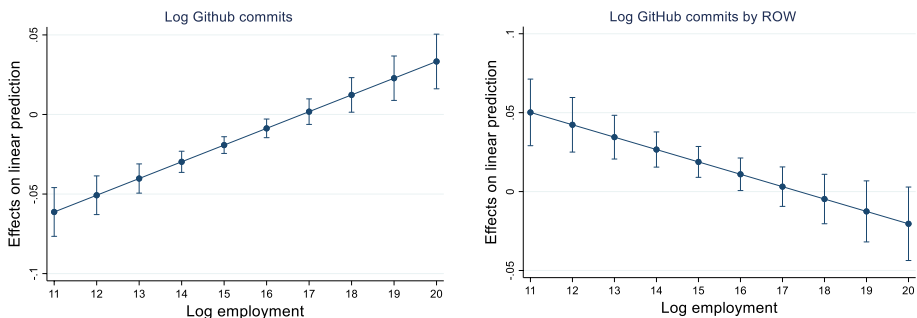
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

the overall sample (0.049). Thus, although the net gain of OSS is still positive for the EU countries, it is with 0.8% ($= [0.035 - 0.027]\%$) substantially smaller. We note that although Table 8 shows that the effects of GitHub contributions still differ depending on the countries' R&D intensities, the difference is smaller and only weakly significant at the 10% level. In the full sample, the difference was strongly significant at the 1% level. For the patenting intensity, the difference disappears altogether. Thus, for EU countries, higher R&D or patenting intensities seem to be less effective in controlling unintended knowledge spillovers.

Seventh, we have tested whether the inclusion of data before the official start of GitHub resulting from backward classification and transfers of older projects has any important

Table 6 Impact of OSS Commits on GDP (Fixed Effects with employment moderation, all countries)

	(1) Log GDP	(2) Log GDP	(3) Log GDP
Log capital stock	0.49601*** (0.02825)	0.47601*** (0.02889)	0.49012*** (0.02810)
Log payments for use of IP	0.06846*** (0.00654)	0.06573*** (0.00686)	0.07259*** (0.00647)
L.Log R&D expenditures	0.08174*** (0.01779)	0.09982*** (0.01842)	0.06408*** (0.01794)
L.Log R&D expenditures by ROW	− 0.04869 (0.04365)	− 0.27879*** (0.06877)	− 0.17755*** (0.06558)
Log transnational patent applications	− 0.01741*** (0.00331)	− 0.01824*** (0.00346)	− 0.01686*** (0.00324)
L.Log GitHub commits	− 0.08656*** (0.00748)		− 0.17686*** (0.02667)
Log employment	0.32798*** (0.05929)	0.27056*** (0.06283)	0.42021*** (0.06320)
L.Log GitHub commits # Log employment	0.00488*** (0.00047)		0.01051*** (0.00174)
L.Log GitHub commits by ROW		− 0.08299*** (0.01117)	0.13641*** (0.03477)
L.Log GitHub commits by ROW # Log employment		0.00561*** (0.00064)	− 0.00784*** (0.00226)
Constant	6.05614*** (1.44323)	13.38270*** (1.95068)	8.45385*** (1.91334)
Country dummies	Yes	Yes	Yes
Observations	576	576	576
R^2	0.869	0.856	0.874
N_{itg}	32	32	32

**Fig. 2** Visualization of the employment moderation

influence on the results. One can see in Table 9 shows the results for identical estimations using only the years since the GitHub start that all conclusions hold.

Table 7 Impact of OSS commits on GDP (Dynamic OLS, only EU)

	(1) Log GDP	(2) Log GDP	(3) Log GDP
Log capital stock	0.52866*** (0.06249)	0.51037*** (0.06639)	0.60373*** (0.05925)
Log employment	0.18393 (0.11531)	0.14181 (0.12207)	0.14061 (0.10897)
Log payments for use of IP	0.05044*** (0.01118)	0.05154*** (0.01187)	0.05676*** (0.01065)
L.Log R&D expenditures	0.02133 (0.03257)	0.03615 (0.03394)	0.01031 (0.03029)
L.Log R&D expenditures by ROW	0.31467*** (0.09923)	− 0.08159 (0.14089)	0.00383 (0.12691)
Log transnational patent applications	− 0.00001 (0.00381)	0.01309*** (0.00402)	0.01609*** (0.00361)
L.Log GitHub commits	− 0.02073*** (0.00405)		− 0.02714*** (0.00423)
L.Log GitHub commits by ROW		0.01458* (0.00752)	0.03533*** (0.00752)
Country dummies	Yes	Yes	Yes
Observations	375	375	375
R^2	0.847	0.863	0.891
N_g	25	25	25

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Finally, we probed our results with instrumental variable approaches where, following Wright et al. (2020), we use the human capital supply shocks as instruments and institutional differences as instruments, where we use the unemployment rate, the share of tertiary educated people, the interaction and varying institutional characteristics as IVs. Also, here the baseline results hold, and additional overidentification tests do not provide evidence of failures of instrument exogeneity (Table 10).

5 Discussion

Based on our regression results, we find that the contribution to OSS measured by the commits to GitHub has a significant impact on the development of countries' GDP. In detail, single economies benefit from the OSS contributions from the rest of the world, whereas the contributions of national developers or domestic companies have a negative impact, which disappears the larger countries. However, netting out the effects still indicates for all countries large positive gains of 2.2% in terms of GDP. Nonetheless, it has to be pointed out that the own contributions to a public good, like OSS, are creating costs [see the conceptual arguments by Nagle (2019b), and Blind et al., 2021a, 2021b for an assessment

Table 8 Impact of OSS Commits on GDP with group splits (Dynamic OLS, only EU)

	(1) Log GDP	(2) Log GDP	(3) Log GDP
Log capital stock	0.59472*** (10.21)	0.60727*** (10.36)	0.61178*** (10.18)
Log employment	0.17375 (1.58)	0.04318 (0.37)	0.06396 (0.53)
Log payments for use of IP	0.05865*** (5.66)	0.05112*** (4.84)	0.05445*** (5.10)
L.Log R&D expenditures	0.00408 (0.14)	0.03101 (1.02)	0.02647 (0.85)
L.Log R&D expenditures by ROW	0.03038 (0.24)	− 0.01592 (− 0.13)	− 0.02693 (− 0.21)
Log transnational patent applications	0.01137*** (3.23)	0.02742*** (7.73)	0.02462*** (6.81)
L.Log GitHub commits: high own commits	− 0.02837*** (− 5.66)		
L.Log GitHub commits: low own commits	− 0.02509*** (− 5.94)		
L.Log GitHub commits: high R&D intensity		− 0.02065*** (− 3.48)	
L.Log GitHub commits: low R&D intensity		− 0.02821*** (− 6.80)	
L.Log GitHub commits: high patent intensity			− 0.02242*** (− 3.66)
L.Log GitHub commits: low patent intensity			− 0.02782*** (− 6.60)
L.Log GitHub commits by ROW	0.03374*** (4.60)	0.03230*** (4.30)	0.03350*** (4.39)
Country dummies	Yes	Yes	Yes
Observations	375	375	375
R^2	0.893	0.896	0.894
N_g	25	25	25
Eq. coeffs chi-sq(1)	0.74	2.92*	1.38

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

of the investments by the EU), which are not immediately benefiting the domestic GDP of small countries in a measurable way. This effect is indicative of a non-trivial positive knowledge externality associated with the provision of OSS as a public good, which in turn may create an incentive, in particular for small countries, to engage in a free-rider strategy.

In addition, we observed strong complementarities with the innovation process in general. Notably, investing more in R&D reduced the costs associated with contributing. Although R&D expenditures, in particular for the developers' wages, might be needed to write software code, which is eventually uploaded at GitHub, we know from Nagle et al.

Table 9 Impact of OSS Commits on GDP (Dynamic OLS, only from the official start of GitHub)

	(1) Log GDP	(2) Log GDP	(3) Log GDP
Log capital stock	0.84568*** (0.07355)	0.84990*** (0.07158)	0.84447*** (0.07164)
Log employment	0.12477 (0.16142)	0.09545 (0.16395)	0.10516 (0.15611)
Log payments for use of IP	0.04243** (0.01726)	0.04163** (0.01759)	0.04456*** (0.01678)
L.Log R&D expenditures	0.01267 (0.03316)	– 0.00660 (0.03387)	– 0.00040 (0.03238)
L.Log R&D expenditures by ROW	0.52957*** (0.15231)	0.24905 (0.23102)	– 0.07781 (0.21897)
Log transnational patent applications	0.00477 (0.00405)	0.00801* (0.00481)	0.01221*** (0.00456)
L.Log GitHub commits	– 0.02681*** (0.00766)		– 0.02494*** (0.00926)
L.Log GitHub commits by ROW		– 0.00610 (0.01260)	0.03500** (0.01500)
Country dummies	Yes	Yes	Yes
Observations	288	288	288
R^2	0.861	0.865	0.873
N_g	32	32	32

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

(2020) that still many contributions are made in software programmers' free time driven by intrinsic motivation (Von Krogh et al., 2012), but also incentives to signal their own capacity to the demand side of the labor market (Lerner & Tirole, 2002, 2005). Furthermore, Blind et al. (2021a) reveal that small and even micro companies located in the EU without R&D departments are contributing a major share of the commits to GitHub. R&D expenditures can be interpreted as a country's absorptive capacity (Cohen & Levinthal, 1990), which can also be relevant for the absorption of domestic and global contributions to OSS. In addition, the evidence revealed by Nagle (2019b) that companies need specialized capabilities for the use of OSS to have a positive productivity impact can also be extrapolated to the country level.

Conceptually, R&D expenditures generate knowledge, which is in general not disclosed to the public but also not explicitly protected, therefore, contributing to the common knowledge pool driving endogenously economic growth (Romer, 1990). In contrast, OSS contributions are both publicly disclosed, e.g., at GitHub, and in principle, accessible to potential users, which might be restricted depending on the type of Open Source license. The complementarity between undisclosed knowledge generated by efforts in R&D and disclosed OSS has eventually positive impacts on countries' GDP, e.g., by combining proprietary software with OSS (e.g., West 2003 or Gambardella & Hall 2006).

Table 10 Impact of OSS Commits on GDP (Fixed effects with instrumental variables)

	(1) Log GDP	(2) Log GDP
L.Log GitHub commits	− 0.08137*** (0.01346)	− 0.07485*** (0.01230)
L.Log GitHub commits by ROW	0.09015*** (0.02838)	0.08981*** (0.02970)
Log capital stock	0.53299*** (0.04217)	0.48681*** (0.04818)
Log employment	0.44353*** (0.03915)	0.51607*** (0.10081)
Log payments for use of IP	0.06059*** (0.00978)	0.05570*** (0.00961)
L.Log R&D expenditures	0.05808** (0.02668)	0.02214 (0.02623)
L.Log R&D expenditures by ROW	− 0.15743 (0.31830)	− 0.17693 (0.33158)
Log transnational patent applications	− 0.00953* (0.00493)	− 0.00887* (0.00460)
Constant	6.31987 (8.54236)	7.86121 (7.96967)
Observations	504	504
N_g	28	28
Sargent Overid.	4.28 (pval: 11%)	4.52 (pval: 10%)

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In summary, there are several mechanisms at work justifying positive synergies at the country level, starting with R&D generating absorptive capacity for integrating OSS and cost efficiencies in producing, network effects in marketing, and competitive advantages between proprietary software and OSS. In addition, relying on Wright et al. (2020) finding a positive impact of commits to GitHub on the number of IT start-ups based on a large panel of countries, it can also be argued that OSS pushes the R&D expenditures in these R&D-intensive subpopulations of all companies and therefore economic growth.

Going beyond R&D expenditure, we discuss the relationship between patents and OSS. This relationship is indeed a tricky one. On the one hand, an inconsistency arises from the fact that OSS is generally disclosed and accessible to the use by interested stakeholders, i.e., companies and individuals, only restricted by the type of OSS license. In contrast, other companies can implement the knowledge codified in patents only after an explicit agreement with the patent owner. Therefore, OSS and patents may follow different and potentially conflicting regimes of intellectual property rights (e.g., Blind & Böhm 2019). This is supported by the observed decrease in software patents following a change in the French public procurement law favoring Open Source solutions (Nagle, 2019a) or the decline in OSS contributions following court

cases enforcing intellectual property rights, like patents (Wen et al., 2013). This inconsistency in intellectual property rights, however, relates only to the narrow protection of the focal software assets. We argued, therefore, that, on the other hand, there is a complementarity between OSS and patents that arises from the fact that OSS is often embedded in goods that may be patentable. An important driver may be the complementarity between hardware and software. Thus, the sources of rent extraction from OSS may be an indirect channel by selling protected goods with an OSS component.

In our analysis, patents and OSS indeed appeared to be complementary, which points towards the mutually reinforcing relationship between patentable goods, such as hardware, and OSS. It may also indicate a complementary relationship between OSS and proprietary software (Lerner & Schankermann 2010), which may be protected through software patents.

Finally, our results showed that the discussed complementarities between national R&D expenditures or patent applications, on the one hand, and the national contributions to OSS, on the other hand, lose large parts of their significance when we restrict the country samples to the EU Member States. First, this might be caused by the reduced number of observations. Second, it could be argued that the companies within the EU Member States are not able to exploit the commercial potential of the possible synergies because the OSS contributing companies are very small or even micro companies (Blind et al., 2021a), lacking complementary resources, market power or effective business models and strategies. A further tentative explanation could be that within the European Union, patent applications addressing software face higher hurdles to being granted following an intensive debate at the beginning of the century (Blind et al., 2005).

6 Policy conclusions

Based on theoretical considerations about the public good character of OSS, we derived several hypotheses related to OSS on countries' GDP. The hypotheses have been tested with a panel of 32 countries covering the period between 2000 and 2018. Contributions to OSS have been operationalized as the countries' commits contributed to GitHub. Eventually, we can reveal that countries' GDP benefits from OSS provided by the rest of the world but not from domestic contributions, which is in line with our theoretical considerations. However, countries' GDP would be more than 2% lower without OSS, showing a positive net effect. In addition, there is complementarity between OSS on the one hand and R&D expenditures and patents on the other hand concerning GDP. Our findings allow us to derive several policy implications, which we can roughly classify as strengthening incentives to contribute, increasing the skill base, and improving governance and regulation frameworks.

6.1 Strengthening incentives to contribute

The companies' investments in OSS are often limited by fears of spillovers. However, our results showed that to reap the benefits of the worldwide available repository of OSS, in particular, to save own development, the necessary absorptive capacity to benefit from this global knowledge pool needs to be strengthened by own contributions. To internalize these

positive externalities and reduce free-rider incentives (Nagle, 2021), supra-national coordination, e.g., at the level of the European Union (EU), should be considered. The following recommendations can be derived with a specific focus on the EU (see also Blind et al., 2021a for more details). On the one hand, the already existing framework programs to support research and innovation, like Horizon 2020 or currently Horizon Europe, could be further opened towards OSS projects. On the other hand, following the creation of OSS supported with public money, further measures could support its broad diffusion to exploit its public good characteristics. Since we face different Open Source licenses (e.g., Blind & Böhm 2019), OSS created with public funding should be explicitly in the public domain, i.e. anyone can modify and use the software without any restrictions. On the other hand, it is also possible to introduce tax breaks for individual and professional contributors, as Ghosh (2006) suggested.

6.2 Increasing the skill base

Since contribution to OSS is resource-intensive, in particular as it concerns human resources, already existing shortages in skilled labor may prevent companies from using and contributing to OSS (BITKOM 2020, 2021; Nagle et al., 2020). However, the development of software skills is an important factor both in absorbing OSS from all over the world and contributing to OSS, which is necessary to exploit the synergies with R&D and patenting. Therefore, the inclusion of Open Source (development, business models, and licensing) in the programs of Higher Education Institutions should be promoted. Moreover, since start-ups are also contributing to GDP, we can refer to the insights of the study by Wright et al. (2020), who reveals a significant impact of OSS on the founding of start-ups. Despite the massive involvement of individuals and micro companies in OSS by the Member States of the EU, there is a lack of successful entrepreneurship by EU actors. Therefore, relevant education should be provided, and a culture to foster Open Source based start-ups should be established.

Our analyses face several limitations. First, starting at the variables explaining the countries' GDP, it has to be admitted that all the software in the GitHub repository has not necessarily been accessible via a license complying with the above-cited OSS definition and does not reflect the complete OSS stock. In particular, GitLab has become more relevant in the last years. Even the variables used, i.e., the commits, are not completely covering all the contributions of the various countries to GitHub because only around half of the accounts have a link to a specific country. Overall, this limitation might underestimate the involvement and investment in OSS. In addition, not all relevant variables for explaining the various economic dimensions might have been considered, i.e., the omitted variable bias might lead to the attribution of effects to the included variables. Consequently, some effects of omitted variables, like scientific publications and standards, might have been attributed to the variables representing OSS, i.e., an overestimation is possible. However, the inclusion of R&D expenditure, which is a much wider concept than OSS, but also of patents should reasonably limit the size of the omitted variable bias. So far, only the contribution to OSS is used, but whether these investments have been successful is not known, i.e., whether the developed OSS code is eventually used in practice. Finally, the general problem of lacking observable and verifiable information about the current value of intangibles being not exchanged via market transactions (Corrado et al., 2009) is also valid for OSS, which is often produced within the firms that use OSS.

Despite these limitations, the paper provides the first systematic approach to estimating the macroeconomic effects of OSS, considering its public good character. The exploitation of gradually better and more complete data sources for OSS could help to provide further evidence on the role of OSS for macroeconomic outcomes.

Funding Open Access funding enabled and organized by Projekt DEAL. Knut Blind and Torben Schubert received funding from Directorate-General for Communications Networks, Content and Technology in the context of the study “The impact of Open Source Software and Hardware on technological independence, competitiveness and innovation in the EU economy”.

Declarations

Conflict of interest Knut Blind and Torben Schubert certify that they have NO affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers’ bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Aghion, P., & Howitt, P. (1992). A model of growth through creative destruction. *Econometrica*, 60(2), 323–351. <https://doi.org/10.2307/2951599>
- Ahuja, G., Lambert, M., & Novelli, C. E. (2013). The second face of appropriability: generative appropriability and its determinants. *Academy of Management Review*, 38(2), 248–269.
- Akçomak, İ. S., Ter Weel, B. (2009). Social capital, innovation and growth: Evidence from Europe. *European Economic Review*, 53(5), 544–567. <https://doi.org/10.1016/j.eurocorev.2008.10.001>
- Aksoy-Yurdagul, D. (2015). The impact of open source software commercialization on firm value. *Industry and Innovation*, 22(1), 1–17.
- Alexy, O., George, G., & Salter, A. (2013). Cui Bono? The selective revealing of knowledge and its implications for innovative activity. *Academy of Management Review*, 38(2), 270–291.
- Amiri-Kordestani, M., & Bourdouce, H. (2017). A survey on embedded open source system software for the internet of things. In *Free and Open Source Software Conference* (Vol. 2017).
- Arora, A., Belenzon, S., & Sheer, L. (2021). Knowledge spillovers and corporate investment in scientific research. *American Economic Review*, 111(3), 871–898.
- Arrow, K. J. (1962). The economic implications of learning by doing. *Review of Economic Studies*, 29(3), 155–173.
- Berlemann, M., & Wesselhöft, J. E. (2017). Aggregate capital stock estimations for 122 countries: An update. *Review of Economics*, 68(2), 75–92.
- Bessen, J., & Hunt, R. M. (2007). An empirical look at software patents. *Journal of Economics & Management Strategy*, 16(1), 157–189.
- BITKOM (2020). Open Source Monitor 2019, <https://www.bitkom.org/opensourcemonitor2019>
- BITKOM (2021). Open Source Monitor 2021, <https://www.bitkom.org/opensourcemonitor2019>
- Bitzer, J., Schrettl, W., & Schröder, P. J. (2007). Intrinsic motivation in open source software development. *Journal of Comparative Economics*, 35(1), 160–169.

- Black Duck Software (2016). The Tenth Annual Future of Open Source Survey. Accessed February 6, 2022, <https://www.slideshare.net/blackducksoftware/2016-future-of-open-source-survey-results>.
- Blind, K., Edler, J., & Friedewald, M. (2005). *Software Patents: Economic Impacts and Policy Implications*. New Horizons in Intellectual Property Series; Edward Elgar.
- Blind, K., & Jungmittag, A. (2008). The impact of patents and standards on macroeconomic growth: a panel approach covering four countries and 12 sectors. *Journal of Productivity Analysis*, 29(1), 51–60. <https://doi.org/10.1007/s11223-007-0060-8>.
- Blind, K., & Böhm, M. (2019). The Relationship Between Open Source Software and Standard Setting: <https://ec.europa.eu/jrc/en/science-update/relationship-between-open-source>.
- Blind, K., Böhm, M., Grzegorzewska, P., Katz, A., Muto, S., Pätsch, S., & Schubert, T. (2021a). The impact of Open Source Software and Hardware on technological independence, competitiveness and innovation in the EU economy, Final Study Report. Brussels. <https://doi.org/10.2759/430161>
- Blind, K., Ramel, F., & Rochell, C. (2021b). The influence of standards and patents on long-term economic growth. *Journal of Technology Transfer*. <https://doi.org/10.1007/s10961-021-09864-3>.
- Bonaccorsi, A., & Rossi, C. (2003). Why open source software can succeed. *Research Policy*, 32(7), 1243–1258.
- Bottazzi, L., & Peri, G. (2007). The international dynamics of R&D and innovation in the long run and in the short run. *The Economic Journal*, 117(518), 486–511. <https://doi.org/10.1111/j.1468-0297.2007.02027.x>.
- Castellacci, F. (2007). Evolutionary and new growth theories. Are they converging? *Journal of Economic Surveys*, 21(3), 585–627.
- Cohen, W. M., & Levinthal, D. A. (1990). Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly*, 35(1), 128–152. <https://doi.org/10.2307/2393553>
- Corrado, C., Hulten, C., & Sichel, D. (2009). Intangible capital and US economic growth. *Review of Income and Wealth*, 55, 661–685.
- Crosby, M. (2000). Patents, innovation and growth. *Economic Record*, 76(234), 255–262.
- 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. [https://doi.org/10.1016/S0048-7333\(99\)00087-6](https://doi.org/10.1016/S0048-7333(99)00087-6)
- Di Gaetano, L. (2015). A model of corporate donations to open source under hardware–software complementarity. *Industrial and Corporate Change*, 24(1), 163–190.
- Gambardella, A., & Hall, B. H. (2006). Proprietary versus public domain licensing of software and research products. *Research Policy*, 35(6), 875–892.
- Ghosh, R. A. (2006). *Economic impact of open source software on innovation and the competitiveness of the information and Communication Technologies (ICT) sector in the EU*. Maastricht: UNU-MERIT.
- Gould, D. M., & Gruben, W. C. (1997). *The role of intellectual property rights in economic growth. Dynamics of globalization and development* (pp. 209–241). Boston, MA: Springer.
- Grossman, G. M., & Helpman, E. (1994). Endogenous innovation in the theory of growth. *Journal of Economic Perspectives*, 8(1), 23–44.
- Hall, B. H., & MacGarvie, M. (2010). The private value of software patents. *Research Policy*, 39(7), 994–1009.
- Hasan, I., & Tucci, C. L. (2010). The innovation–economic growth nexus: Global evidence. *Research Policy*, 39(10), 1264–1276. <https://doi.org/10.1016/j.respol.2010.07.005>
- Hawkins, R. E. (2004). The economics of open source software for a competitive firm - Why give it away for free? *Netnomics*, 6, 103–117.
- in't Veld, J. (2019). The economic benefits of the EU single market in goods and services. *Journal of Policy Modelling*, 41(5), 803–818.
- International Organization for Standardization (2021). Standards & economic growth: ISO members' research on the impact of standards on their national economies, Geneva.
- Iwaisako, T., & Futagami, K. (2003). Patent policy in an endogenous growth model. *Journal of Economics*, 78(3), 239–258. <https://www.jstor.org/stable/41795028>.
- Johnson, J. (2002). Open Source Software: private provision of a public good. *Journal of Economics & Management Strategy*, 11, 637–662. <https://doi.org/10.1162/105864002320757280>.
- Jungmittag, A., Blind, K., & Grupp, H. (1999). Innovation, standardisation and the long-term production function: A cointegration analysis for Germany 1960–1996. *Zeitschrift für Wirtschafts- und Sozialwissenschaften*, 119, 205–222.
- Kubiszewski, I., Farley, J., & Costanza, R. (2010). The production and allocation of information as a good that is enhanced with increased use. *Ecological Economics*, 69(6), 1344–1354.

- Lach, S. (1995). Patents and productivity growth at the industry level: a first look. *Economics Letters*, 49(1), 101–108.
- Lakhani, K. R., & von Hippel, E. (2003). How open source software works: “Free” user-to-user assistance. *Research Policy*, 32(6), 923–943.
- Lakka, S., Stamati, T., Michalakelis, C., & Anagnostopoulos, D. (2015). Crossnational analysis of the relation of eGovernment maturity and OSS growth. *Technological Forecasting and Social Change*, 99, 132–147. <https://doi.org/10.1016/j.techfore.2015.06.024>.
- Launchpad (2016). People and teams in Launchpad. Retrieved August 27, 2016, from <https://launchpad.net/people>
- Lerner, J., & Schankerman, M. (2010). *The comingled code: Open source and economic development*. MIT Press Books.
- Lerner, J., & Tirole, J. (2002). Some simple economics of Open-Source. *Journal of Industrial Economics*, 50, 197–234.
- Lerner, J., & Tirole, J. (2005). The Economics of Technology sharing: open source and Beyond. *Journal of Economic Perspectives*, 19(2), 99–120. <https://doi.org/10.1257/0895330054048678>
- Michel, P., & Nyssen, J. (2000). On knowledge diffusion, patents lifetime and innovation based endogenous growth. In D. Encaoua, B. H. Hall, F. Laisney, & J. Mairesse (Eds.), *The economics and econometrics*, pp. 55–81. Springer US, Boston, MA.
- Mockus, A. (2007). Large-scale code reuse in open source software. In First International Workshop on Emerging Trends in FLOSS Research and Development (FLOSS’07: ICSE Workshops 2007), 7–7. IEEE.
- Nagle, F. (2018). Learning by contributing: gaining competitive advantage through contribution to crowdsourced public goods. *Organization Science*, 29(4), 569–587. <https://doi.org/10.1287/orsc.2018.1202>
- Nagle, F. (2019a). Government technology policy, social value, and national competitiveness (March 3, 2019). Harvard Business School Strategy Unit Working Paper No. 19–103. <https://doi.org/10.2139/ssrn.3355486>
- Nagle, F. (2019b). Open source software and firm productivity. *Management Science*, 65(3), 1191–1215. <https://doi.org/10.1287/mnsc.2017.2977>.
- Nagle, F. (2021). *The Problem of Social Benefit*. Stanford Social Innovation Review.
- Nagle, F., Wheeler, D. A., Lifshitz-Assaf, H., Ham, H., & Hoffman, J. L. (2020). *Report on the 2020 FOSS Contributor Survey*. The Linux Foundation & The Laboratory for Innovation Science at Harvard.
- Ojanperä, S., Graham, M., & Zook, M. (2019). The digital knowledge economy index: Mapping content production. *The Journal of Development Studies*, 55(12), 2626–2643. <https://doi.org/10.1080/00220388.2018.1554208>
- O’Reilly, T. (1999). Lessons from open-source software development. *Communications of the ACM*, 42(4), 32–37.
- Reisinger, M., Ressner, L., Schmidtknecht, R., & Thomes, T. P. (2014). Crowding-in of complementary contributions to public goods: Firm investment into open source software. *Journal of Economic Behavior & Organization*, 106, 78–94. <https://doi.org/10.1016/j.jebo.2014.06.005>.
- Riehle, D. (2007). The economic motivation of open source software: Stakeholder perspectives. *Computer*, 40(4), 25–32.
- Robbins, C. A., Korkmaz, G., Bayoán Santiago Calderón, J., Chen, D., Kelling, C., Shipp, S., & Keller, S. (2018). : Open Source Software as Intangible Capital: Measuring the Cost and Impact of Free Digital Tools. In a paper from 6th IMF Statistical Forum on Measuring Economic Welfare in the Digital Age: What and How, 19–20.
- Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economy*, 98(5), 71–102.
- Schubert, T., & Neuhäusler, P. (2018). *Can depleting technological opportunities explain the stagnation of productivity?* Panel data evidence for 11 OECD countries, Studien zum deutschen Innovations system, No. 11-2018, Expertenkommission Forschung und Innovation (EFI), Berlin
- Solow, R. M. (1956). A contribution to the theory of economic growth. *The Quarterly Journal of Economics*, 70(1), 65–94.
- SourceForge (2016). About [Page describing the platform]. Retrieved August 27, 2016, from <https://sourceforge.net/about>
- Swann, G. P. (2000). *The economics of standardization*. University of Manchester, Manchester, UK.: Final report for standards and technical regulations directorate department of trade and industry. Manchester, UK: University of Manchester.
- Von Engelhardt, S., & Freytag, A. (2013). Institutions, culture, and open source. *Journal of Economic Behavior & Organization*, 95, 90–110.

- Von Engelhardt, S., Freytag, A., & Schulz, C. (2013). On the geographic allocation of open source software activities. *International Journal of Innovation in the Digital Economy*, 4(2), 25–39.
- Von Krogh, G., Haefliger, S., Spaeth, S., & Wallin, M. W. (2012). Carrots and rainbows: Motivation and social practice in open source software development. *MIS Quarterly*, 36(2), 649–676.
- Von Krogh, G., & Spaeth, S. (2007). The open source software phenomenon: characteristics that promote research. *The Journal of Strategic Information Systems*, 16(3), 236–253.
- Wachs, J., Nitecki, M., Schueller, W., & Polleres, A. (2022). The geography of open source software: Evidence from GitHub. *Technological Forecasting and Social Change*, 176, 121478. <https://doi.org/10.1016/j.techfore.2022.121478>
- Wen, W., Forman, C., & Graham, S. (2013). The impact of intellectual property rights enforcement on open source software project success. *Information Systems Research*, 24(4), 1131–1146.
- West, J. (2003). How open is open enough?: Melding proprietary and open source platform strategies. *Research Policy*, 32(7), 1259–1285. [https://doi.org/10.1016/S0048-7333\(03\)00052-0](https://doi.org/10.1016/S0048-7333(03)00052-0)
- Wright, N., Nagle, F., & Greenstein, S. M. (2020). Open source software and global entrepreneurship. Harvard Business School Technology & Operations Mgt. Unit Working Paper No. 20–139.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.