ISSN 1936-5330

kcFED Research Working Papers

R&D Capital and the Idea Production Function

Jakub Growiec, Peter McAdam, and Jakub Mućk

May 2023
RWP 23-05
http://doi.org/10.18651/RWP2023-05

Federal Reserve Bank of Kansas City

R&D Capital and the Idea Production Function*

Jakub Growiec
SGH Warsaw School of Economics

Peter McAdam
Federal Reserve Bank of Kansas City

Jakub Mućk
SGH Warsaw School of Economics

Abstract

We supplement the 'Idea Production Function' (IPF) with measures of R&D capital. We construct the R&D capital stock in the US, and estimate the IPF with patent applications as R&D output, allowing for a flexible treatment of R&D productivity (over 1968-2019). The estimated substitution elasticity between R&D inputs is $0.7 - 0.8$. Hence R&D capital is an essential factor in producing ideas, complementary to R&D labor. There is a positive trend in R&D labor productivity ($\sim 1\%$) and cyclical variation of R&D capital productivity. Instead of 'ideas getting harder to find', there is a scarcity of R&D capital needed to find them.

Keywords: R&D, Patents, Long-Run Growth, Technical Change, Estimation, CES, Quality-adjustment, TFP.

JEL Codes: O30, O40, O47

*We thank Nick Bloom, John H. Cochrane, Alice von Ende-Becker, Cooper Howes and Pete Klenow as well as many seminars audiences for comments and discussions. This research has been financed by the Polish National Science Center (Narodowe Centrum Nauki) under grant OPUS 14 No. 2017/27/B/HS4/00189. The views expressed are those of the authors alone and not necessarily those of the Federal Reserve Bank of Kansas City or the Federal Reserve system. McAdam thanks the economics department at UC Berkeley for hospitality.

Corresponding Author: Peter McAdam. Email: peter.mcadam@kc.frb.org

1 Introduction

Technological change, due to purposeful R&D activities, is widely acknowledged as a fundamental driver of economic growth. Technological ideas, due to their non-rivalrous nature, essentially act as a source of increasing returns to scale allowing output to grow even when input usage is constant (Romer, 1990).

The question of how ideas and technical change impact long-run productivity growth is profoundly important. Unsurprisingly, there are many viewpoints. Bloom et al. (2020), Gordon (2016), and others have argued that the US economy may be running out of ideas.[1] If so, the consequence would be stagnant productivity and a lower growth rate in the coming decades even as, perhaps, the number of scientific researchers continues to grow.[2] In line with this, since the early 2000s the US indeed appears to have entered a low-productivity growth regime (Fernald, 2015; Fernald and Li, 2022). By contrast, Brynjolfsson and McAfee (2014) and Brynjolfsson, Rock and Syverson (2019) argue that the economy is on the cusp of major technological breakthroughs, and the current productivity slowdown reflects a period of sharpening those technologies to improve their applicability and diffusion across the production processes.[3]

However, whilst these precepts are well-established in the literature, the exact specification of the R&D process (the 'Idea Production Function', hereafter IPF) is subject to considerable dispute.[4] In particular, and perhaps surprisingly, the majority of the existing R&D-based growth literature assumes that researchers' labor is the only input into the R&D process (Romer, 1990; Jones, 1995, 1999; Ha and Howitt, 2007). Alternatively, some studies embrace the "lab equipment" specification of the R&D process, conditioning R&D output on the flow of R&D spending (Rivera-Batiz and Romer, 1991; Kruse-Andersen, 2017; Bloom et al., 2020).

In reality, though, both approaches may be limiting since it is likely that productivity in the R&D sector depends not just on the labor of researchers, but also on the services

[1] See also Ramey (2020) and Kim and Qureshi (2020).

[2] See Jones (2022*a*) for a discussion of future demographic patterns for growth and ideas.

[3] Lags and cycles between technological adaption and growth is one explanation for the US technology slowdown of the 1970s. See the excellent survey of growth theories and outcomes of Helpman (2010). See also Growiec, McAdam and Mućk (2018).

[4] Jones (2022*b*) demonstrates that the IPF can be retrieved from either the Romer or 'quality-ladder' endogenous growth approaches.

of *R&D capital*. R&D capital should be understood as a stock, accumulated over the years through targeted R&D investment. Indeed, anecdotal evidence suggests that R&D is an increasingly capital-intensive activity: new scientific ideas and technological blueprints increasingly rely on the effects of experimentation in sophisticated laboratories as well as on advanced numerical computation, rather than abstract philosophical reflection or pen-&-paper calculations. Modern R&D capital ranges from researchers' computing facilities to such extraordinary machinery as the Very Large Telescope (VLT) and the Large Hadron Collider.

Moreover, the practicality and complexity of research equipment has also undergone systematic, cumulative changes and productivity improvements over the decades and centuries. The difference for instance in usefulness of Ptolemy's astrolabe or Galileo's telescope, set against the VLT is breathtaking. Likewise, consider how early statisticians computed correlations and ran regressions without relying on computer processing capabilities. Modern R&D activity also increasingly uses AI algorithms, ranging from general-purpose tools like web search engines; and (recently) large language models such as GPT-4, integrated into ChatGPT and Microsoft Bing (see Korinek, 2023), to specific applications in genome sequencing or analysis of astronomical dataries; and sometimes even in solving long-standing problems, as in the case of DeepMind's AlphaFold which produced a major breakthrough in the protein folding problem.

Given this, how would the introduction of R&D capital affect our estimates and understanding of the economy's IPF? And what will be the implications of such an extension for major questions such as whether ideas are getting harder to find, or whether the recent slowdown in total factor productivity (TFP) growth constitutes secular stagnation, or rather a temporary downswing? (e.g., Cowen, 2011; Ramey, 2020).

In their influential paper Bloom et al. (2020) focus their attention around the following IPF:

$$\text{Research Output} = \underbrace{\text{Research Productivity}}_{\alpha_t} \times \text{Researchers}. \tag{1}$$

In other words, they postulate that research output, proxied by the rate of TFP growth in the economy, is proportional to the number of researchers. Since the latter rose dramatically over the post-war period whilst the former stayed fairly constant, this concentrates attention on how the α middle term, capturing the (potentially time varying) level of research productivity, has behaved. To achieve balance, the authors argue, α must have declined, indicating that research ideas have been getting harder to find. (See also our "A First Look at the Data" section below).

But would "idea TFP" still be strongly falling over time if the IPF also included R&D capital (in addition to R&D labor)? Consider as an illustration the following log-linear (Cobb Douglas) specification where $\dot{A}_t/A_t$, TFP growth (theoretically representing the flow of new ideas), is a function of R&D labor (i.e., number of researchers, $\mathcal{R}$) but now also R&D capital ($\mathcal{K}$):

$$\frac{\dot{A}_t}{A_t} = \Gamma_t \mathcal{K}_t^{\beta} \mathcal{R}_t^{1-\beta}, \tag{2}$$

where $\beta \in [0,1]$ captures the share of R&D capital in the production of ideas, and Γ (like α above) captures unit research productivity. Predictions for idea TFP based on (1) will differ from those based on (2) if the rate of change in R&D capital systematically differs from that of R&D labor.[5] If both variables grow at a common rate, then the dynamics of both "idea TFP" concepts (α_t and Γ_t) will be the same. Otherwise, they will differ and idea TFP, as in (1), may in fact be systematically mismeasured, if not misleading.

However, although more general than (1), IPF (2) is itself *still* quite restrictive: it implicitly imposes a unit elasticity of substitution between both R&D input factors and assumes factor productivity improvements behave in a neutral manner. In other words, if, say, R&D labor became relatively more expensive, on this basis firms could simply substitute 1-for-1 into R&D capital. Moreover, if R&D productivity changes over time (as it surely does) then the specification assumes that it impacts both R&D factors in the same manner. By contrast our empirical findings, on a less restrictive specification of the IPF (see equation (3) below), demonstrate that R&D capital is in fact an *essential, complementary, and relatively scarce* factor in the R&D sector. In such a scenario, the relative scarcity of R&D capital will constrain R&D output even when (as is the case) R&D labor is abundant and fast growing. Notwithstanding, if IPF specification (2) was correct, then using (1) instead would mechanically misattribute the observed discrepancy in growth rates between R&D labor and R&D output to falling idea TFP. Moreover, following standard omitted-variable

[5] Symbols $\mathcal{R}$ and $\mathcal{K}$ represent some fraction of, respectively, the aggregate labor force and aggregate capital stock – namely the amounts that are used in the R&D sector.

bias reasoning, doing so would also attribute an incorrect weight to R&D labor (the β), depending on the true correlation between R&D factors.

Another fundamental question is whether the growth rate of TFP is an appropriate measure of research output. TFP growth may in fact reflect many other phenomena than just research output. For example, Baqaee and Farhi (2020) find that improvement in allocative efficiency, due to the reallocation over time of market shares to high-markup firms, accounted for about half of aggregate US TFP growth over 1997–2015 (see also Oberfield and Raval, 2021). TFP measures may also conflate the cyclical volatility of capacity, quality improvements, and factor utilization rates, which are independent of technical progress.[6] This was understood as far back as Solow (1957) and underpins the seminal work of Fernald in correcting raw TFP (Fernald, 2018). In light of this, we opt instead for *patent applications*, a more direct measure of R&D output. An important implication of that choice is that patent applications, even relative to patents in force, have been growing over time in the US over the last decades while the TFP growth rate has declined. That in itself impacts our estimates of idea TFP and makes the conclusion that "ideas are getting harder to find" less likely.[7]

A First Look at the Data An initial glance at the US data (Table 1) suggests the following.[8] First, the number of new patent applications per researcher ($\Delta A/\mathcal{R}$) was gradually increasing over time. Maintaining exponential growth in idea production has indeed become more difficult, though: the ratio of new patent applications to patents in force ($\Delta A/A$) grew slower than R&D employment ($\mathcal{R}$). A similar conclusion is reached when replacing raw R&D employment with data on "effective R&D employment" (Ω/w), defined by Bloom et al. (2020) as the ratio of total R&D expenditure to the average R&D wage, which grew slightly faster than raw R&D labor ($\mathcal{R}$).

[6] See also Figure A.21 in Appendix A for the indexed profile of TFP for several advanced countries. Whilst some countries (e.g., US, France) have experienced a strong upward trajectory in TFP levels (albeit punctuated by low-growth episodes), other countries (e.g., Canada, Italy) have experienced trend breaks and decades-long stagnation of TFP. Taken at face this would suggest that those economies are in technical regress. Additional issues with TFP as a proxy of ideas are measurement issues, for example the provision of zero-price technologies.

[7] The systematic discrepancy between the rate of new patent applications and aggregate TFP growth can be understood as an upward trend in firms' propensity to patent their innovations. This may mean, among other interpretations, that either average patent quality is declining, or that there is a growing pool of ideas which have not been successfully commercialized yet but may be commercialized in the future. To address this concern, we consider alternative measures of R&D output – quality-adjusted TFP growth and breakthrough patents – as a robustness exercise.

[8] For more general discussion of recent US growth and productivity performance see Fernald and Wang (2016); Fernald et al. (2017); Fernald and Li (2022), and for greater historical scope see Gordon (2016). See also Grossman et al. (2017) for links to income-share developments.

Table 1: Summary Statistics of R&D Variables:
Average Growth Rates (1968-2019)

Variables	Symbol	Growth Rate
Patent Applications	ΔA	3.211
Patents-in-Force	A	2.410
Patent Applications Relative to Patents-in-Force	$\Delta A/A$	0.782
R&D Capital	$\mathcal{K}$	3.394
R&D Labor	$\mathcal{R}$	2.099
R&D Wage	w	0.848
R&D Expenditure (Real)	Ω	3.319
R&D Expenditure Relative to R&D Wage	Ω/w	2.450
R&D Capital Relative to Patents-in-Force	$\mathcal{K}/A$	0.961
R&D Labor Relative to Patents-in-Force	$\mathcal{R}/A$	-0.304
Patent Applications Relative to R&D Labor	$\Delta A/\mathcal{R}$	1.090
Patent Applications Relative to Ω/w	$\Delta A/(\Omega/w)$	0.743
Patent Growth Relative to R&D Labor	$(\Delta A/A)/\mathcal{R}$	-1.289
Patent Growth Relative to Ω/w	$(\Delta A/A)/(\Omega/w)$	-1.628

Source: Derived from *WIPO*, *IPUMS CPS*.

Second, “idea TFP” as defined in (1) depends crucially on the definition of research output. With patent applications as the output variable, the resultant measure of idea TFP ($\Delta A/(\Omega/w)$) is increasing over time. Declining idea TFP is only obtained once one identifies research output with patent applications relative to patents-in-force, for example as in Bloom et al. (2020) ($(\Delta A/A)/(\Omega/w)$, cf. last line of Table 1).

Third, R&D capital grew almost exactly in line with growth in patent applications (3.4% vs. 3.2% per year) – and noticeably faster than growth in R&D labor and the number of patents in force (3.4% per versus 2.1% and 2.4%, respectively). This indicates that “idea TFP” growth measures which disregard the accumulation of R&D capital, such as (1), are most likely misleading.

Contribution Whilst Bloom et al. (2020) provide one of the most systematic analyses of this topic with a detailed look at both the macro and micro data, in this paper we attempt to apply a more general theory of the IPF. By introducing R&D capital alongside R&D labor into the IPF, and then estimating it allowing for a non-unitary elasticity of substitution and non-neutral unit productivity, our study fills an important gap in the empirical literature on R&D-based economic growth. We find that the elasticity of substitution between R&D capital and R&D labor in the IPF is about $0.7 - 0.8$ and, importantly, significantly below unity. This implies that R&D capital should be considered an essential factor in producing ideas, and complementary to R&D labor. We also identify a systematic positive trend in R&D labor productivity at about 1% per year on average and a cyclical dynamic in R&D capital productivity. On average, *effective* supply of R&D capital was lagging behind that of R&D labor, constraining R&D output. Idea TFP, the Hicks-neutral component backed out from the IPF, has not been falling but rather oscillating around a constant mean.

Accordingly, our results imply that ideas, instead of getting harder to find per se, in fact *require more sophisticated lab equipment* to be found. This is a scarcity which can only be bridged by increased accumulation and development of R&D capital, and not necessarily by employing more R&D staff. Because investments in R&D equipment are an endogenous variable that can be influenced by policy and institutions, our results contribute to potentially lowering the assessment of the likelihood and inevitability of a future secular stagnation. One might therefore view our contribution as *complementary* to Bloom et al. (2020)'s albeit with a twist: namely, that if indeed ideas are harder to obtain, it reflects the fact that researchers need ever better and more productive capital to find them.

Organization Section 2 documents the construction of the time series of R&D capital as well as measurement of R&D labor and R&D output. We construct the stock of R&D capital in the post-war US economy, using the perpetual inventory method applied to BEA chain-type quantity indexes for R&D assets. Section 3 discusses the IPF and its estimation over 1968-2019, using a nonlinear system estimation technique with a flexible treatment of the unit productivity of R&D factors. Section 4 presents our results. We present several IPF forms, where R&D capital is included alongside R&D labor and where unit productivity in both R&D factors is modeled in an increasingly flexible manner. Section 5 takes our results and show how R&D can be decomposed over time into its constituent determinants; this illuminates which variables have or have not constrained the production of ideas. Thereafter, in Section 6 we derive idea TFP as the residual of the IPF and comment on its properties. To close our treatment, Section 7 provides some robustness by repeating our analysis but this time using alternative measures of ideas, namely quality-adjusted TFP (Fernald, 2018) and breakthrough patents (Kelly et al., 2021). These largely confirm our baseline results and inference. Section 8 concludes. Additional material is in the appendices.

2 Data and Measurement

We shall now discuss our empirical strategy of measuring capital, labor, and output in the R&D sector. A fundamental challenge here is to collect sufficiently long time series of acceptable proxies of the variables of interest. Since the available classification systems are not able to uniquely identify total R&D activity in the economy, we use a variety of auxiliary data sources that should provide conceptually close proxy variables for the concepts at hand. See Appendix A for more discussion of the construction of the series and some sensitivity analysis.

2.1 R&D Output

The choice of the output variable in the IPF is challenging. As one possible approach, since R&D encompasses activities that are aimed at reducing unit costs of production or increasing the variety of goods offered, one could measure the aggregate stock of knowledge/technology as the level of TFP in the economy. In turn, the flow of R&D output would be represented as increases in aggregate TFP over time.

However, although popular, the strategy of using TFP growth as a proxy of R&D output is problematic.[9] Changes in TFP might be driven by changes in technology but they may also result from other processes. For instance, reduction in misallocation could increase measured TFP (Baqaee and Farhi, 2020; Oberfield and Raval, 2021). Other potential causes include, e.g., production function misspecification or changes in the internal composition of production factors. Worryingly, all these additional sources of variation cause measured TFP to sometimes fall over time – while the functional form of the IPF expects consistently positive values of R&D output.[10] Factoring in these caveats, we include the TFP growth case as a robustness exercise.

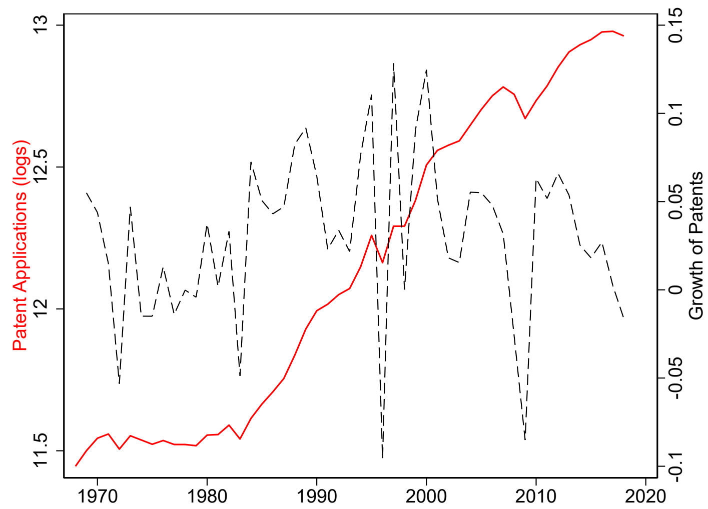
[9] There is also an open discussion as to whether one should use absolute or relative increases in TFP ($\dot{A}$ or $\dot{A}/A$, respectively) as the flow concept of R&D output (see Bloom et al., 2020).

[10] The problem of negative TFP growth in estimating IPFs can be overcome by considering an approximation of the IPF (Ha and Howitt, 2007) or by taking 5-year averages (Ang and Madsen, 2011).

For our baseline results, however, we measure the aggregate outcome of the R&D sector with patent data. Following a common practice in the literature, we use *patent applications* as the R&D output variable (Madsen, 2008; Ang and Madsen, 2011; Venturini, 2012). Since we are interested in long historical patent data, our principal measures are taken from Marco et al. (2015) which are updated with the recent WIPO (World Intellectual Property Organization) series.

Figure 1 plots the series in log levels and in growth rates. Consistent with other studies we witness a burst of patent activity from the mid-1980s onward which has been variously ascribed to computing and communication sectors, Kelly et al. (2021). The slowdown of growth rates around the mid-2000s, moreover, onward is consistent with Fernald and Wang (2016)'s assessment of the decline of US productive potential around that time.

FIGURE 1: New patent applications in the US, 1968-2019



Notes: In this figure we plot on the lhs axis the log of patent applications and on the rhs the growth rate of patents. Data derived from Marco et al. (2015).

2.2 R&D Capital

To estimate R&D capital we use Bureau of Economic Analysis (BEA) data. Unfortunately, the BEA does not measure the aggregate R&D capital stock directly, nor does it publish long-run series on fixed-weights aggregates of R&D investment or R&D stock.[11] The reason for that is there are long-run trends in relative prices of inputs, such as the secular decline in prices of equipment relative to structures (Greenwood, Hercowitz and Krusell, 1997).

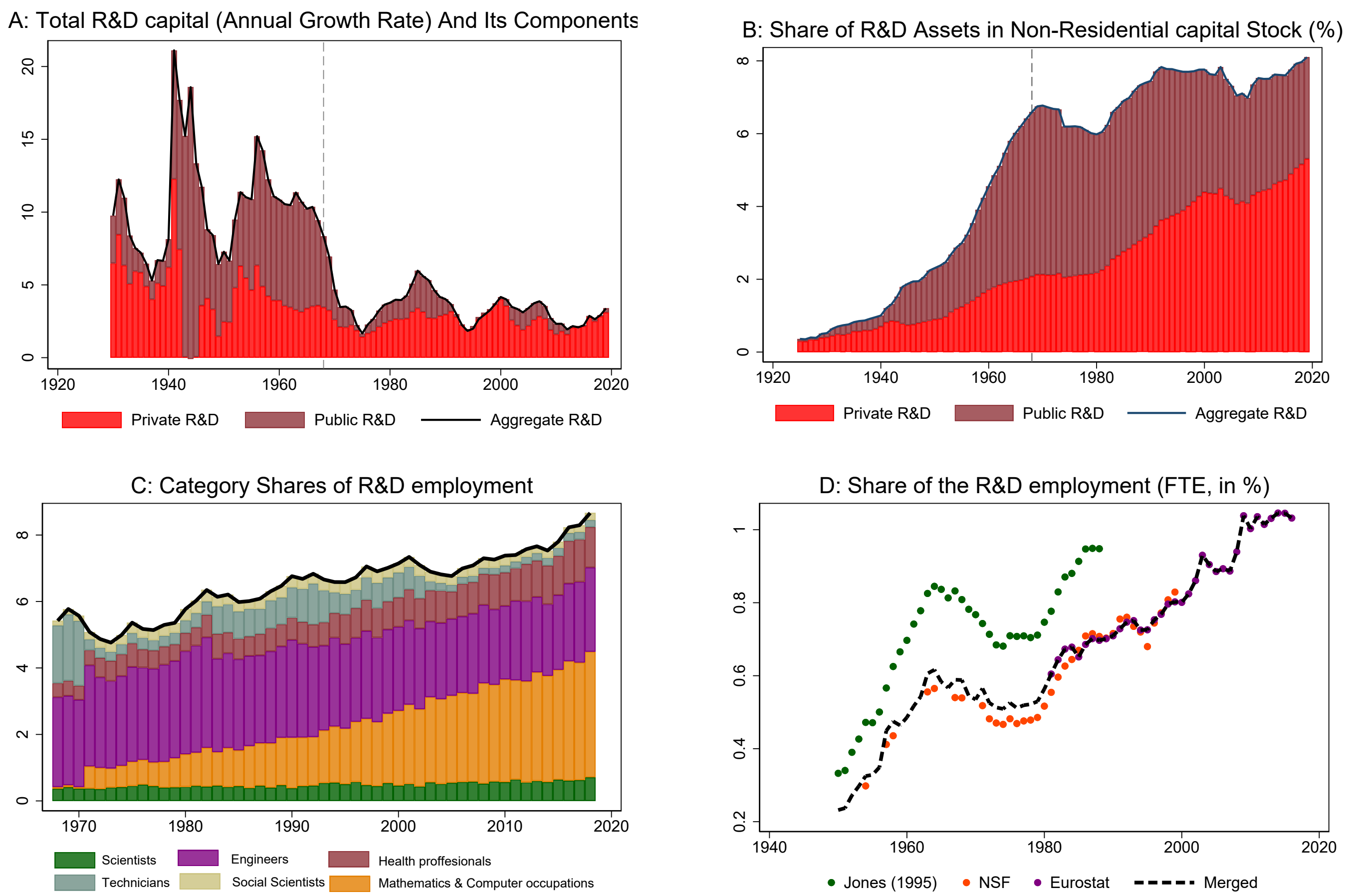
We construct the R&D capital stock using the perpetual inventory method. The capital stock is calculated as the sum of investment in period t and previously depreciated capital stock, $\mathcal{K}_t = (1-\delta)\,\mathcal{K}_{t-1} + I_t^{rd}$ where $\delta \in (0,1)$ is the depreciation rate of R&D capital and I_t^{rd} is real investment in R&D. This relationship is initialized in the standard manner: $\mathcal{K}_0 = I_0^{rd}/(g+\delta)$, where g is the long-run geometric growth rate of R&D investment. While the latter can be easily calculated from historical data, there is considerable uncertainty about the depreciation rate of R&D capital. We calibrate this rate at 15% per year. This number, which we understand as something of a consensus in the literature (Venturini, 2012), is higher than that pertaining to the aggregate capital stock because of a relatively larger share of fast-depreciating equipment in R&D, and an accordingly lower share of structures.[12]

To obtain a long-dated series of the total R&D capital and private R&D capital we proceed as follows. Since there are no measures of real R&D investment expressed in chained dollars we estimate it based on available series, i.e, nominal R&D investment data as well as price indexes. For the private sector, we divide nominal R&D investment (BEA code: `Y006RC`) by the price index of this asset (`Y006RG`). The same strategy is applied for the public sector (`Y057RC` and `Y057RG`, respectively). In addition, we also consider the following components of public investment: Federal Non-Defense (`Y069RC` and `Y069RG`), Defense (`Y076RC` and `Y076RG`) and state and local (`Y073RC` and `Y073RG`). The growth rate of the R&D capital stock in the US since 1929 is plotted in Figure 2, **Panel A**. In turn, the R&D capital share in the total nonresidential capital stock is shown in **Panel B**.

[11] The data (in constant dollars) starts in 1999. This time span is however too short to analyze long-run patterns in R&D productivity.

[12] The depreciation rate of R&D capital may be even higher than 15%. Bernstein and Mamuneas (2006) estimate it above 15%; Li and Hall (2019) place it even above 30%. In KLEMS (2019) the depreciation rate for R&D assets is fixed at 20%, see https://euklems.eu/wp-content/uploads/2019/10/Methodology.pdf In addition, the BEA publish historical series on depreciation of R&D assets in the US and, according to the BEA estimates, the implied depreciation rate is slightly above the consensual value of 15%. By contrast Griliches (1998) has argued that private and social depreciation rates of R&D may be quite different. Private R&D may depreciate rapidly as firms copy one another to an extent less prevalent in public ventures. See Appendix A for sensitivity analyses.

Figure 2: R&D Capital and R&D Labor



Notes: The gray vertical dashed line in Panels A and B indicates the start of the estimation sample.

2.3 R&D Labor

The second factor in the IPF is R&D labor. At the conceptual level – and in line with the definition from the Frascati Manual (OECD, 2015) – this category refers to all employees who undertake creative work aimed at general increases in the existing stock of knowledge. In practice, however, application of this definition requires very detailed information about tasks that are related to R&D activities. To the best of our knowledge such data are not available, making it effectively impossible to measure R&D labor directly.

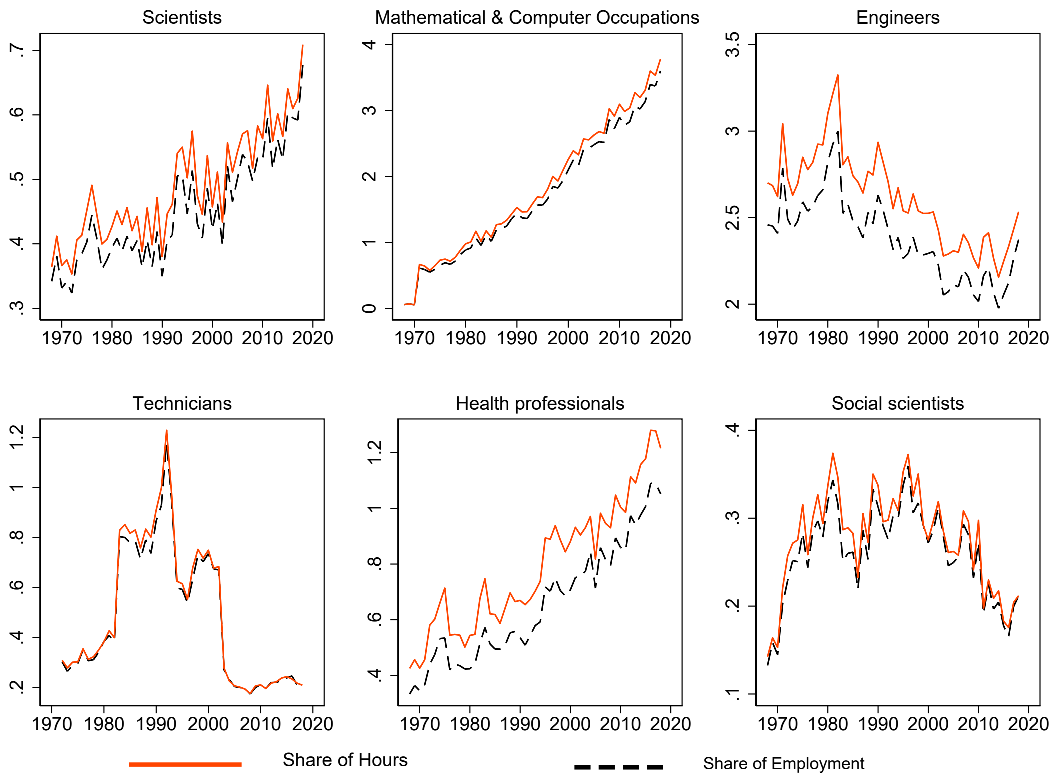
To circumvent this, we estimate the labor input in R&D activity using micro-data which contains information about the structure of occupations. To this end we use IPUMS Current Population Survey (CPS) data (Ruggles et al., 2019).[13] This database offers harmonized micro data from a monthly US labor force survey. Based on the conceptual definition of R&D personnel and scientists and engineers (S&E) we identify the following occupational groups as those whose work embodies R&D activity:[14] **`Scientists`**; **`Engineers`**; **`Health Professionals`**; **`Technicians`**; **`Social Scientists`**; and **`Mathematical & Computer Occupations`**. The relative shares of these groups over time can be seen in Figure 2, **Panel C**. For the same groups Figure 3 calculates aggregate hours worked.

As a robustness check we also narrow the set of occupations to Scientists, Mathematical & Computer occupations and Engineers. As a further robustness check we also take advantage of publicly available data on R&D employment and merge historical series. We begin with official estimates of R&D activity published by statistical offices, Eurostat and the OECD. The Eurostat/OECD series only begins in 1981. To overcome this, we use older data vintages to extrapolate the existing series. For the time period 1968-1980 we use data collected within the IRIS (Industrial Research and Development Information System) program conducted by the NSF (National Science Foundation). Moreover, based on historical data from Jones (1995) it is possible to further extrapolate the observations backwards, i.e., into the 1950s-60s (see Figure 2, **Panel D**). Robustness analysis is presented in Appendices A and D.

[13] See Appendix A for more discussion. See also https://cps.ipums.org/cps.

[14] The practical idea is to try to match the Eurostat definition of human resources in science and technology. According to this definition, scientists and engineers (S&E) are workers who conduct research, improve or develop concepts, theories and operational methods and/or apply scientific knowledge relating to their fields. This definition can be covered by following groups of `ISCO-08` occupations: Science and engineering professionals (21), Health professionals (22) and Information and communications technology professionals (25). See Appendix A.3 for a detailed list of included categories.

Figure 3: Share of R&D related occupation groups in total US employment and hours worked



2.4 R&D Rental Prices

Finally, identification of the elasticity of substitution between R&D factors and the nature of the unit productivities requires data on relative rental prices. We calculate the capital rental rate as the sum of the real interest rate and the R&D capital depreciation rate. Specifically, we use the interest rate on 10-year government bonds (FRED© code: `GS10`) deflated by the GDP deflator (`GDPDEF`).[15]

For the rental price of labor, we calculate the real hourly wage. The CPS data set contains sufficient information on wages, allowing us to build a long-dated series on real wages for our baseline measure of R&D labor. It also enables us to construct the series of real hourly wages for the alternative measure that uses a narrower set of occupations. In the case of the merged historical series on R&D labor, we use the same real wages as in our baseline since

[15] We have also experimented with CPI as the price proxy; results remain materially unchanged.

there is no publicly available long series on wages in the R&D sector.

3 The Idea Production Function

3.1 Constant Elasticity Specification

Following our earlier discussion, we estimate the following IPF:

$$\Delta \widetilde{A}_t = \left[\eta \left(\Gamma_t^{\mathcal{K}} \widetilde{\mathcal{K}}_t \right)^{\frac{\xi-1}{\xi}} + (1-\eta) \left(\Gamma_t^{\mathcal{R}} \widetilde{\mathcal{R}}_t \right)^{\frac{\xi-1}{\xi}} \right]^{\frac{\xi}{\xi-1}} \tag{3}$$

where ΔA_t is the flow of new ideas (as represented by new patent applications). The IPF, note, is written in 'normalized' (or indexed) form. Thus, $\widetilde{X}_t = X_t / X_z$ where $X_z > 0$ denotes the value of X at the point of normalization.[16]

Distribution parameter $\eta \in [0, 1]$ measures the steady-state level of the R&D capital share in total R&D income. Parameter $\xi \geq 0$ is the elasticity of substitution between R&D capital and R&D labor, with the special cases of Leontief, log-linear and linear forms, respectively, given by $\xi \to 0, 1, \infty$.[17] IPF form (3) relaxes the assumption of a unit elasticity of substitution and accommodates the possibility of factor-specific productivity improvements over time, whose paths are captured by $\Gamma_t^{\mathcal{K}}$ and $\Gamma_t^{\mathcal{R}}$ for R&D capital and labor, respectively.

It is well-known that estimation of production relationships is improved by joint estimation with the first order conditions (FOC), León-Ledesma, McAdam and Willman (2010). This is because such an approach combines information from different sides of the production framework (costs and volumes) and exploits cross-equation restrictions. In the case of the considered IPF and, after taking logs and combing the FOCs, this implies the proportionality:

$$\ln \left(\frac{r_t^{\mathcal{K}} \mathcal{K}_t}{w_t^{\mathcal{R}} \mathcal{R}_t} \right) = \left(\frac{\xi - 1}{\xi} \right) \ln \left(\frac{\Gamma_t^{\mathcal{K}} \widetilde{\mathcal{K}}_t}{\Gamma_t^{\mathcal{R}} \widetilde{\mathcal{R}}_t} \right) \tag{4}$$

where $r_t^{\mathcal{K}}$ is the real rental price of R&D capital, $w_t^{\mathcal{R}}$ denotes real wages in the R&D sector.[18]

---

[16] Without explicit normalization, parameter estimates in constant-elasticity functions can be shown to be scale dependent, arbitrary and non-robust. Normalization points are averages. For linear series such as a time trend, they are given by the arithmetic mean; otherwise, geometric averages are used. See León-Ledesma, McAdam and Willman (2010) for Monte-Carlo analyses, and La Grandville (1989) and Klump and de La Grandville (2000) for the seminal theoretical contributions.

[17] Thus, (2) emerges as the special (and testable) case: $\xi = 1$; and $\Gamma_t^{\mathcal{K}} = \Gamma_t^{\mathcal{R}}$.

[18] Estimating the capital and researcher R&D FOCs separately can be problematic. Accordingly, among these three equations (i.e., two first order conditions plus their ratio), we use the relative factor share equation (4), for the

Thus, estimation consists in the joint system estimation of parameters in the first-order condition (4) with IPF (3) (in the latter case, we also transform the specification into logs).

3.2 Specification of Unit Productivity Terms

Another important decision to make relates to the assumption about the trajectory of productivity improvements to the R&D factor inputs over time. The latent nature of both processes (and recalling the Diamond-McFadden impossibility theorem in standard production theory) requires assumptions need to be made about each.

We consider three increasingly more sophisticated assumptions about the growth in unit productivity of R&D factors in the IPF. The first two are nested in the Box–Cox form $\log \Gamma_t^j = \mathbf{B}(\gamma_j, \lambda_j; t)$. The log-level of productivity to each j R&D factor is thus increasing around a normalized or average growth rate γ_j, where parameter $\lambda_j \in \mathbb{R}$ determines shape. If $\lambda_j = 1$ then the level of unit productivity increases exponentially over time at a constant growth rate γ_j. Otherwise, growth accelerates ($\lambda_j > 1$) or decelerates ($\lambda_j < 1$) relative to the mean γ_j (see Appendix B for more details on the Box–Cox form).

The third case is where we consider a functional form that allows us to account for the possibility of unknown structural breaks (or long swings). This is a good robustness check in itself, but is also motivated by some evidence of structural instability in the patent growth process (see Appendix C). In that case, we use a Fourier expansion: $\log \Gamma_t^j = \mathbf{F}\left(\gamma_j, \kappa_j^{sin}, \kappa_j^{cos}; t\right)$.[19] Any possible structural breaks or cycles around its trend growth rate of γ_j will thus be captured by the κ parameters.

Due to substantial variation and a possible appearance of structural breaks, the normalization point for R&D factor shares seems far from obvious. We start by setting the distribution parameter η at $1/3$ (typical of the long-run average of the total capital income share). Making this assumption reflects the fact that there are no reliable data that allow

following reasons. First, the share equation contains information on both forms of factor productivities over time, rather than just one. Second, it does not require any information about the dynamics of markups. The individual FOCS are based on the assumption of perfect competition. The share equation remains useful if markups are positive but stable over time. However, recent empirical literature has documented a secular upward trend in markups in the US (De Loecker, Eeckhout and Unger, 2020), albeit aggregate (rather than R&D-specific) ones. This could potentially lead to a systematic bias in the estimation of the individual FOCs. At the same time, in the first order condition using the relative factor share (4) markups are eliminated. Third, condition (4) does not require any information about the prices of new ideas.

[19] See Christopoulos and León-Ledesma (2010) for a discussion of Fourier forms in economics. We follow Ludlow and Enders (2000) who showed that a single frequency is invariably sufficient to approximate the Fourier expansion in the bulk of empirical applications.

us to estimate factor shares in the R&D sector. However, we also include cases where η is estimated.

4 Results

4.1 Baseline Results

The first section of Table 2 presents the various parameter estimates and the expression of the R&D productivity terms. Abbreviations **B**, **F**, and "Exp." denote the Box–Cox, Fourier, and exponential ($\lambda = 1$) forms, respectively. The middle section presents tests of relevant parameter restrictions, and the final section shows estimation diagnostics. The first two rows in that final section refer to the bootstrapped ADF test of the unit root null associated with the errors in equations (4) and the log form of (3). Finally ll, bic and $rmse$ denote, respectively, the Log Likelihood, the Bayesian Information Criterion, and the Root Mean Square Error.

Case 1 estimates an IPF with only R&D labor (akin to equation (1)), followed in cases 2 and 3 by the IPF augmented with R&D capital (equation (2)), without the unit-elasticity constraint. All forms produce superficially not entirely unreasonable results: The first yields a power coefficient of 1.43, the second and third imply a growth rate of (Hicks neutral) R&D unit productivity of around $1.2 - 1.3\%$ per year (close to the R&D labor rates in subsequent specifications). However, the diagnostics suggest a poor fit to the data; or at least that these cases are statistically dominated by the subsequent cases.

Cases 4 and 6 introduce the Box-Cox unit productivity forms, first for R&D labor then for both R&D input factors, whilst case 5 imposes simple exponential productivity growth for both factors. The case for a unitary substitution across these cases is mixed: case 6 illustrates the severe and well-known issue of identifying productivity terms when $\xi \approx 1$ (Sato, 1970); column 4 produces an unusually high elasticity value. All three cases unsurprisingly suffer diagnostic issues, for instance, the residuals exhibit extreme persistence and nonstationarity.

Table 2: Baseline Results

Parameter/Case	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
β	1.428^{***}							
	(0.069)							
ξ		1.000	0.844^{***}	2.531^{***}	0.737^{***}	0.986^{***}	0.793^{***}	0.760^{***}
		$(-)$	(0.169)	(0.920)	(0.139)	(0.169)	(0.019)	(0.062)
$\gamma_{\mathcal{R}}$		0.013^{***}	0.012^{***}	0.012^{***}	0.011^{***}	-0.013	0.011^{***}	0.011^{***}
		(0.002)	(0.002)	(0.003)	(0.001)	(0.308)	(0.001)	(0.001)
$\lambda_{\mathcal{R}}$				2.890^{***}		6.453		
				(0.766)		(19.078)		
$\gamma_{\mathcal{K}}$					0.004	0.060	-0.016^{***}	-0.013^{***}
					(0.004)	(0.615)	(0.003)	(0.004)
$\lambda_{\mathcal{K}}$						5.208		
						(3.885)		
$\kappa_{\mathcal{K}}^{sin}$							0.556^{***}	0.438^{***}
							(0.045)	(0.137)
$\kappa_{\mathcal{K}}^{cos}$							-0.427^{***}	-0.337^{***}
							(0.028)	(0.109)
η								0.418^{***}
								(0.121)
R&D Labor Productivity	no	Exp.	Exp.	**B**	Exp.	**B**	Exp.	Exp.
R&D Capital Productivity	no	no	no	no	Exp.	**B**	**F**	**F**
$\xi = 1$			[0.357]	[0.096]	[0.058]	[0.934]	[0.000]	[0.000]
$\lambda_{\mathcal{R}} = 1$				[0.014]		[0.775]		
$\lambda_{\mathcal{K}} = 1$						[0.279]		
$\gamma_{\mathcal{R}} = \gamma_{\mathcal{K}}$					[0.086]	[0.937]	[0.000]	[0.000]
$\kappa_{cos}^{\mathcal{K}} = \kappa_{sin}^{\mathcal{K}} = 0$							[0.000]	[0.006]
res_4		[0.262]	[0.086]	[0.066]	[0.101]	[0.020]	[0.006]	[0.008]
res_3	[0.413]	[0.531]	[0.095]	[0.237]	[0.085]	[0.051]	[0.001]	[0.000]
ll	16.5	81.1	78.7	96.9	76.5	100.7	133.2	134.2
bic	-29.0	-150.4	-141.7	-174.2	-133.3	-173.9	-239.0	-237.0
$rmse\_4$		0.140	0.140	0.124	0.137	0.129	0.089	0.089
$rmse\_3$	0.177	0.139	0.138	0.116	0.138	0.097	0.049	0.049

Notes: This table estimates equations (3) and (4) as a system; all parameters are as defined there. In addition, we use in columns (7) and (8) the Fourier form to capture factor-augmenting technical progress. Symbols **B** and **F** denote the Box-Cox and Fourier forms respectively, and "Exp." denotes exponential (i.e., $\lambda = 1$). Cases 2 and 3 (the Cobb Douglas cases) assume Hicks neutrality. In the second section of the table, we present Wald tests of various parameter restrictions. In the (final) diagnostic section of the table, the first two rows refer to ADF test of the unit root null associated to the errors in equations (4) and the logged form of (3) and the p-values are obtained by bootstrapping distribution. Finally, terms ll, bic and $rmse$ denote, respectively, the Log Likelihood, and the Bayesian Information Criterion, and the Root Mean Square Error.

The final two cases are the most data congruent (witness the dramatic improvement in diagnostic measures). We impose constant growth in R&D labor productivity (consistent with the results of column 6) and allow R&D capital productivity to follow the Fourier form (the difference between cases 7 and 8 is that the latter freely estimates distribution parameter η). Both Fourier parameters are statistically significant, and of opposite signs implying a somewhat cyclical trajectory for R&D capital productivity (the point estimates of the normalized productivity growth of R&D capital are negative, but this is precisely an average over a cyclical trajectory).

Indeed, the role of structural breaks and swings is actually predominant over the sample (see also the next section) such that there is no visible downward trend in R&D capital augmentation. R&D labor productivity is increasing by 1.1% per year. In contrast to previous estimates, non-stationarity in residuals can be decisively rejected. Finally, the substitution elasticity is significantly below unity (around $0.7 - 0.8$). Thus, to summarize, R&D capital and R&D labor are gross complements in the IPF. Unit labor productivity is increasing in the R&D sector, while unit capital productivity exhibits strong non-linear variability.

5 Decomposition of Ideas Growth

An instructive exercise is to use our preferred estimates (from Section 4) to decompose the sources of ideas production into its constituent elements: R&D factors and R&D productivity. This can illuminate which elements constrain or encourage the production of ideas over time.

Specifically, using the IPF (3) we can decompose growth in new patent applications as follows:

$$g_{\Delta \widetilde{A}_t} = \Pi_{K,t}(g_{\Gamma_t^{\mathcal{K}}} + g_{\widetilde{\mathcal{K}}_t}) + \Pi_{R,t}(g_{\Gamma_t^{\mathcal{R}}} + g_{\widetilde{\mathcal{R}}_t}) \tag{5}$$

where

$$\Pi_{K,t} = \eta \left(\frac{\Gamma_t^{\mathcal{K}} \widetilde{\mathcal{K}}_t}{\Delta \widetilde{A}_t} \right)^{\frac{\xi-1}{\xi}}, \tag{6}$$

$$\Pi_{R,t} = (1-\eta) \left(\frac{\Gamma_t^{\mathcal{R}} \widetilde{\mathcal{R}}_t}{\Delta \widetilde{A}_t} \right)^{\frac{\xi-1}{\xi}} \tag{7}$$

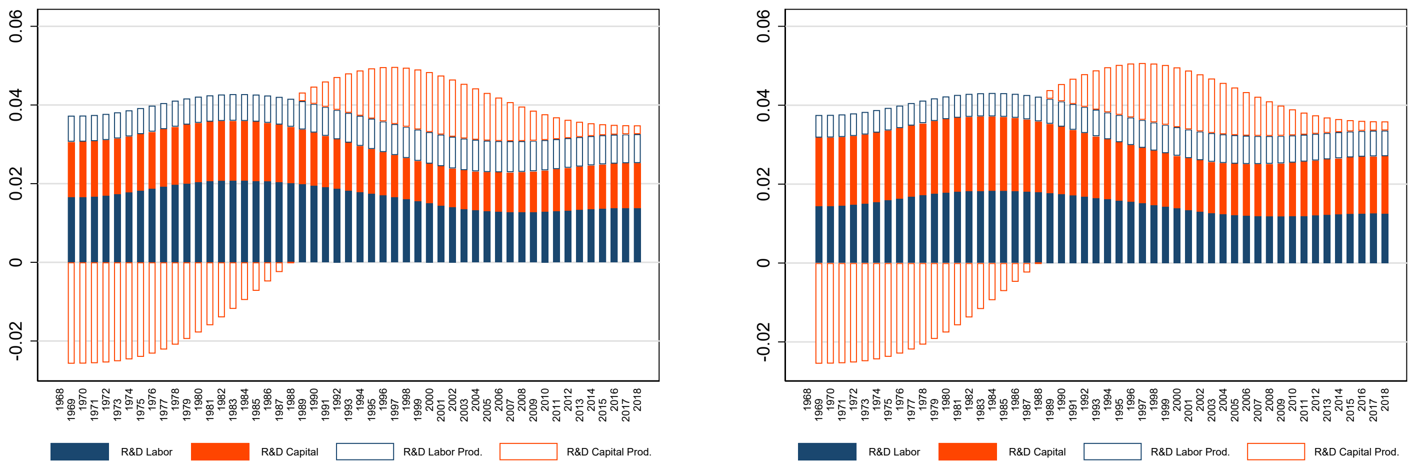
are the respective factor shares in R&D. We use the theoretical values for new patent applications, that is, the values explained by the IPF, excluding regression residuals. Furthermore, to capture secular trends in ideas production rather than high frequency fluctuations, data on R&D inputs and output have been filtered with HP prior to decomposition (using the annual smoothing parameter of 6.25).

Figure 4, **Panel A** decomposes the growth of ideas into factor and productivity terms, using the last two specifications in Table 2. From this we can perhaps identify three main phases of ideas growth: (i) sluggish growth in ideas (up to the early 1980s), (ii) sharp acceleration in ideas growth (1980s-2000s), and (iii) slowdown in ideas growth (since the 2000s). The relative contribution of R&D capital vs. R&D labor depends on the model specification – either they are roughly equal or R&D labor is somewhat more important – but in any case the time trends of both contributions are largely parallel, namely both are relatively steady over time, with just a minor increase around the early 1980s and a minor decrease in late 2000s. Furthermore, as labor productivity is growing uniformly at $\sim 1\%$ per year, its contribution to ideas production is also steady and quantitatively somewhat less important than the contribution of input growth. The three phases of ideas growth are accounted for exclusively by the strong cyclical dynamic of R&D capital productivity. The contribution of that factor to ideas growth was strongly negative in phase (i), then sharply increased into the positive domain in phase (ii), and then gradually fell back to about zero in phase (iii).

Figure 4: Characterization of the Estimated IPF

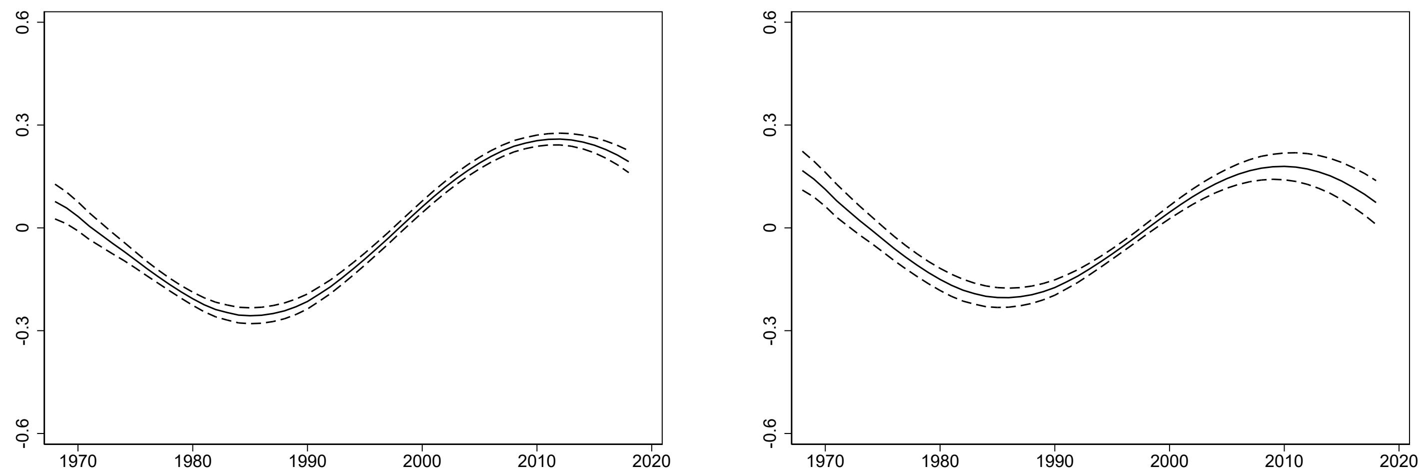
A. The idea growth decomposition (Annual change on HP-filtered contributions)

Specification (7) Specification (8)



B. Hicks-neutral idea TFP backed out from the IPF (in logs, with 95% CI)

Specification (7) Specification (8)



Interestingly, the timing of the three phases coincides with the adoption of ICTs as major general-purpose technologies in the 1980s-2000s (Bresnahan and Trajtenberg, 1995; Jorgenson, 2005; Aum, Lee and Shin, 2018). Universities and research laboratories in the United States were among the first adopters of both technologies. In that same period, R&D capital productivity increased markedly. In turn, according to our results, the episode of R&D capital productivity growth ended around the time of the global financial crisis. It is conceivable that it will resume one day in the future, perhaps after a breakthrough in artificial intelligence (Brynjolfsson, Rock and Syverson, 2019; Growiec, 2022*a*,*b*).

6 Is Idea TFP Falling Over Time?

With a constant elasticity IPF specification, there is no unique idea TFP. Instead, the unit productivity of each factor $(\Gamma_t^{\mathcal{K}}, \Gamma_t^{\mathcal{R}})$ is identified separately. With this in mind, however, one can nevertheless calculate a joint "idea TFP" factor capturing Hicks-neutral technical change in R&D.[20] Specifically, we calculate the log of the idea TFP from IPF (3) as follows:

$$\log(\widetilde{TFP}) = \frac{\xi}{\xi-1}\log\left[\frac{\eta\left(\Gamma_t^{\mathcal{K}}\widetilde{\mathcal{K}}_t\right)^{\frac{\xi-1}{\xi}}+(1-\eta)\left(\Gamma_t^{\mathcal{R}}\widetilde{\mathcal{R}}_t\right)^{\frac{\xi-1}{\xi}}}{\eta\left(\widetilde{\mathcal{K}}_t\right)^{\frac{\xi-1}{\xi}}+(1-\eta)\left(\widetilde{\mathcal{R}}_t\right)^{\frac{\xi-1}{\xi}}}\right]. \tag{8}$$

The results are plotted in Figure 4, **Panel B**. In contrast to Bloom et al. (2020) we do not find a sharp drop in idea TFP, rather a wave oscillating around a constant mean. Along with the three phases in ideas growth, discussed in Section 5, idea TFP first falls (until 1980s), then grows (from 1980s up to about 2010), and then begins to fall again.

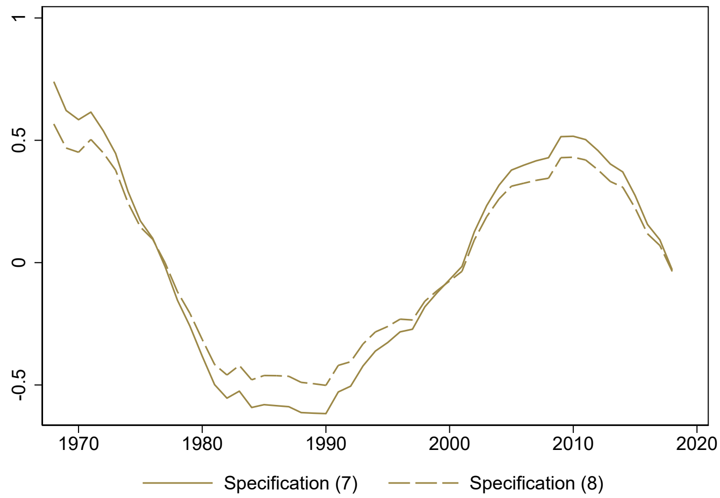
We interpret our results as an indication that R&D capital is an essential complementary factor in R&D activity. In R&D, as in the aggregate economy, capital accumulation markedly outweighs long-term growth in labor supply. However, in effective terms, factoring in the systematic increases in R&D labor productivity and the more erratic behavior of R&D capital productivity over 1968-2019, average growth in R&D labor outran that of R&D capital.

On top of this trend, the effective R&D capital-to-labor ratio also exhibited a clear cyclical

[20] The mapping from the pair $(\Gamma_t^{\mathcal{K}}, \Gamma_t^{\mathcal{R}})$ to Hicks-neutral idea TFP is not invertible. There is a second dimension of technical change, absent from the concept of the idea TFP: *factor bias* in technical change (see Klump, McAdam and Willman, 2012; León-Ledesma and Satchi, 2019).

pattern, following the three main phases of ideas growth earlier discussed (Figure 5). This may indicate that Bloom et al. (2020)'s celebrated result that "ideas are getting harder to find" could be reinterpreted in fact as "there is an increasing scarcity of R&D capital required to find new ideas", with a policy implication that R&D output could be increased by subsidizing and facilitating the accumulation of state-of-the-art R&D capital rather than necessarily increasing R&D employment.

Figure 5: Effective R&D capital-to-labor ratio (in logs)



In relation to the debate whether the observed slowdown in TFP growth over the last couple of decades is a sign of an upcoming secular stagnation (Jones, 2002; Gordon, 2016) or represents a transition phase to a digitally mature economy which would again grow faster once the transition period ends (Brynjolfsson and McAfee, 2014; Brynjolfsson, Rock and Syverson, 2019), our results are indicative of the latter option. According to our estimates, the current slowdown in R&D output is likely due to a relative shortage of R&D capital, rather than a sharp falling idea TFP.[21]

[21] Speculatively, the most promising kind of R&D capital required to achieve progress may be AI algorithms.

7 A Comparison with Alternative Measures of Ideas

In this section, we consider alternative measures of ideas and perform the same regression analysis. The first is to consider TFP growth as a measure of R&D output. The second is to stay within the area of patents, but to use measures of quality-adjusted patents.

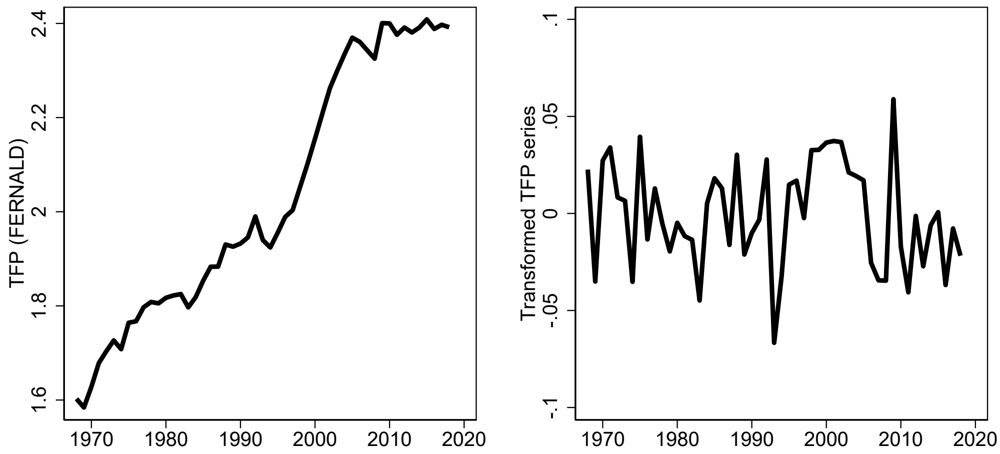
7.1 TFP as a measure of ideas

Let us now consider the growth of TFP as a measure of R&D output. We implement the same system of equations to estimate the parameters of interest, such as factor substitutability and productivity terms. We use the well-known *Fernald* TFP series for the US economy from 1968 to 2019 (consistent with the earlier sample). The Fernald series, here labeled $\log(TFP^F)$, has become the standard series for analyzing pure TFP in the US economy, since it corrects for factor utilization and changes in factor quality over time (features that would be missing from a conventional accounting measure).

To obtain the normalized TFP growth series, we exponentiate the original logged series, compute its gross growth rate, divide by the mean growth rate in the sample, and then take the log of the result. Specifically, we define $G_t = TFP_t^F / TFP_{t-1}^F \in \mathbb{R}^+$ and compute its sample mean G_0. Normalized TFP growth is obtained as $\widetilde{G}_t = G_t / G_0 > 0$. The logged variable $\log\{\widetilde{G}_t\} \in \mathbb{R}$ is used as log R&D output in the estimations. The original and transformed time series are shown in Figure 6.

This definition of the explained variable has three key characteristics. First, following Bloom et al. (2020), we use *relative* (percentage) rather than *absolute* (dollar value) increases in TFP as our current measure of R&D output. Second, following Klump and de La Grandville (2000); Klump, McAdam and Willman (2012), we use a *normalized* measure which has favorable properties for estimating the elasticity of substitution and factor-augmenting technical change. Third, we use *gross* rather than net growth rates (G_t, not $G_t - 1$) which eliminates negative TFP growth rates, problematic for the subsequent log transformation, without adding more noise to the data, e.g., through smoothing.

FIGURE 6: Log Total Factor Productivity (Fernald) and Its Transformation



Notes: The left side panel shows the log of the Fernald TFP series, whilst the right side shows its normalized growth rate dynamic.

Table 3 runs the same set of regression exercises as before. Again, we see that simple variants $(1)-(2)$ where R&D capital is excluded perform poorly in diagnostic tests (namely, residuals and likelihood). In these cases, moreover, it is not possible to reject the unit elasticity of substitution assumption, even though that is evidently rejected in subsequent, more data-congruent alternatives. Indeed, looking at the specifications $(3)-(6)$ we find an elasticity of substitution in a somewhat similar range to our baseline case at $0.6-0.8$. This again implies that R&D capital is an essential factor in the production of ideas (as understood by the TFP proxy) and complementary to R&D labor. This is the key takeaway from this exercise.

Table 3: Results using TFP

Parameter, Case	(1)	(2)	(3)	(4)	(5)	(6)
ξ	0.991^{***}	0.987^{***}	1.001^{***}	0.603^{***}	0.609^{***}	0.601^{***}
	(0.023)	(0.025)	(0.002)	(0.074)	(0.053)	(0.061)
$\gamma_{\mathcal{R}}$	-0.041^{***}	-0.040^{***}	0.405	-0.025^{***}	-0.025^{***}	-0.025^{***}
	(0.001)	(0.001)	(0.471)	(0.001)	(0.001)	(0.001)
$\lambda_{\mathcal{R}}$		0.861^{***}		1.074^{***}		
		(0.075)		(0.134)		
$\gamma_{\mathcal{K}}$			-0.892	-0.029^{***}	-0.035^{***}	-0.034^{***}
			(0.942)	(0.002)	(0.001)	(0.001)
$\lambda_{\mathcal{K}}$				0.423^{***}		
				(0.142)		
$\gamma_{\mathcal{K}}^{sin}$					0.035^{*}	0.025
					(0.020)	(0.021)
$\gamma_{\mathcal{K}}^{cos}$					-0.126^{***}	-0.108^{***}
					(0.015)	(0.027)
η						0.562^{***}
						(0.080)
R&D Labor Productivity	Exp.	**B**	Exp.	**B**	Exp.	Exp.
R&D Capital Productivity	no	no	Exp.	**B**	**F**	**F**
η	fixed	fixed	fixed	fixed	fixed	estimated
$\xi = 1$	[0.706]	[0.611]	[0.650]	[0.000]	[0.000]	[0.000]
$\lambda_{\mathcal{R}} = 1$		[0.063]		[0.579]		
$\gamma_{\mathcal{R}} = \gamma_{\mathcal{K}}$			[0.359]	[0.132]	[0.000]	[0.000]
$\kappa_{cos}^{\mathcal{K}} = \kappa_{sin}^{\mathcal{K}} = 0$					[0.000]	[0.000]
$\lambda_{\mathcal{K}} = 1$				[0.000]		
res_3	[0.075]	[0.026]	[0.062]	[0.032]	[0.000]	[0.000]
res_4	[0.057]	[0.053]	[0.030]	[0.004]	[0.004]	[0.005]
ll	103.7	105.4	108.7	116.9	134.4	134.5
bic	-191.6	-191.2	-197.7	-206.3	-241.3	-237.5
$rmse_4$	0.140	0.140	0.140	0.104	0.110	0.114
$rmse_3$	0.062	0.055	0.050	0.057	0.040	0.038

Notes: The numbers in parentheses are robust standard errors, where the significance stars are to be read as $^{*} < 0.1, ^{**} < 0.05, ^{***} < 0.01$. Probability values are in brackets. Symbols **B** and **F** denote the Box-Cox and Fourier forms respectively, and "Exp." denotes exponential (i.e., $\lambda = 1$). In the second section of the table, we present Wald tests of various parameter restrictions. In the (final) diagnostic section of the table, the first two rows refer to ADF test of the unit root null associated to the errors in equations (4) and the logged form of (3) (albeit with the TFP term as the dependent variable) and the p-values are obtained by bootstrapping distribution. Finally, terms ll, bic and $rmse$ denote, respectively, the Log Likelihood, and the Bayesian Information Criterion, and the Root Mean Square Error.

In contrast to our baseline results, however, the estimated growth rates of unit productivities of R&D capital and R&D labor in the production of ideas are now both negative. Specifically, we find a robust decline in the unit productivity of R&D labor over the sample, aligned with the claim that "ideas are getting harder to find" (Bloom et al., 2020). We can offer three comments in this regard. First, the explained variable in this robustness exercise is constructed as a growth rate and, therefore, it captures relative, not absolute increase in ideas. Thus, it rather answers the question: "Is exponential growth [rather than any growth] getting harder to achieve?". Second, the Fernald series itself makes quality adjustments to the TFP series, which may contribute to weak identification of the unit productivity terms. Third, the estimated decline in the unit productivity of R&D labor may also to a degree point to the presence of persistent bottlenecks in the real economy that prevent the effective adoption of newly invented technologies. If so, the implication would be that removing these bottlenecks, for example through implementation of advanced automation technologies such as AI algorithms, would strongly accelerate measured TFP growth.

7.2 Breakthrough patents as a measure of ideas

An issue with raw patents as a measure of ideas is that patents vary in significance. Some patents are widely cited, others hardly cited at all. Some patents well cited in one era, may be overlooked or superseded in another. Moreover, the degree to which a patent is cited is somewhat subjective, depending on the preferences and knowledge of inventors.

In an important paper, Kelly et al. (2021) construct a new indicator of US patent quality by analyzing the text of patent documents. Their main contribution is not merely to generate citation-weighted measures, but indicators of 'breakthrough' patents (i.e., a patent whose content is distinct from prior patents, but is *similar* to future patents).

Figure 7 shows three of their series.[22] The first (in black) is the number of breakthrough patents per capita. Specifically, breakthrough patents are those that fall in the top 10% of the unconditional distribution of Kelly et al.'s importance measure (defined as the ratio of the 10-year forward to the 5-year backward similarity, net of year fixed effects). The other series (in red) are the number of patents that fall in the top 10% of the unconditional distribution of forward citations measured either over the next 10 years (red dash) or over

[22] We plot their data from the 1960s to 2010, consistent with our existing results. However, their series extends back to 1836.

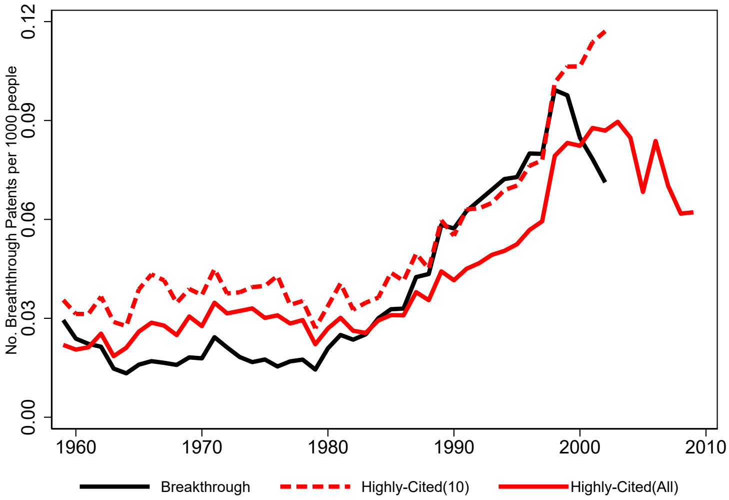
the entire sample (red solid), net of year fixed effects and population scaled. All series have relatively similar time series properties with a lift off from the mid-1980s (found to be driven by computers and communication innovations), and by no means dissimilar to our existing series (see Figure 1).[23]

Table 4 repeats our main exercise this time using the third of the plotted series, Highly-Cited All which is the most data-rich series and has the advantage of having the longest sample (extending to 2009). Again, we follow a progression of simple constrained cases to more complex ones. This, though, is still 10 years short of the previous cases of raw patents data and TFP.

Notwithstanding, a relatively similar pattern pertains to the breakthrough patents case as with the previous two. The initial cases where we either impose a unit elasticity or the absence of R&D capital yields weak results (i.e., easily dominated by subsequent cases, and exhibiting diagnostic issues). As we progress to more data-congruent cases, the elasticity settles in a $0.5 - 0.8$ range, with constant growth in R&D labor found to be consistent with the data and with cyclical dynamics for R&D capital. Compared to the latter columns of Table 2, we tend to find quite similar values and patterns for the productivity terms.

[23] Indeed, time-series analysis (presented in Appendix C) suggests a break in the early 1980s for our main patents series.

FIGURE 7: Breakthrough and Highly-Cited Patents



Notes: This figure plots the time series of breakthrough patents per capita. The series in black is the number of breakthrough patents per capita. Specifically, breakthrough patents are those that fall in the top 10% of the unconditional distribution of Kelly et al.'s importance measure (defined as the ratio of the 10-year forward to the 5-year backward similarity, net of year fixed effects). The other series (in red) are the number of patents that fall in the top 10% of the unconditional distribution of forward citations measured either over the next 10 years (red dash) or over the entire sample (red solid), net of year fixed effects and population scaled.

Source: Kelly et al. (2021).

TABLE 4: Results using Breakthrough Patents

Parameter, Case	(0)	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ξ	1.000	1.023^{***}	0.527^{***}	0.695^{***}	0.586^{***}	0.747^{***}	0.837^{***}	0.785^{***}
	(−)	(0.189)	(0.073)	(0.162)	(0.091)	(0.033)	(0.139)	(0.081)
$\gamma_{\mathcal{R}}$	0.001	0.001	0.006^{***}	0.007^{**}	0.003	0.008^{***}	0.007^{***}	−0.000
	(0.005)	(0.005)	(0.002)	(0.004)	(0.003)	(0.002)	(0.002)	(0.002)
$\lambda_{\mathcal{R}}$			−0.084		0.194			15.873^{**}
			(0.291)		(1.244)			(8.013)
$\gamma_{\mathcal{K}}$				-0.006^{*}	−0.003	-0.040^{***}	−0.066	-0.016^{*}
				(0.004)	(0.002)	(0.007)	(0.061)	(0.008)
$\lambda_{\mathcal{K}}$					−0.428			
					(0.398)			
$\kappa_{\mathcal{K}}^{sin}$						-0.701^{***}	−1.214	-0.482^{**}
						(0.109)	(1.199)	(0.202)
$\kappa_{\mathcal{K}}^{cos}$						0.128^{**}	0.288	0.217
						(0.065)	(0.383)	(0.160)
η							0.277	0.682^{**}
							(0.280)	(0.309)
R&D Labor Productivity	Exp.	Exp.	**B**	Exp.	**B**	Exp.	Exp.	**B**
R&D Capital Productivity	no	no	no	Exp.	**B**	**F**	**F**	**F**
η	fixed	fixed	fixed	fixed	fixed	fixed	estimated	estimated
$\xi = 1$		[0.905]	[0.000]	[0.060]	[0.000]	[0.000]	[0.241]	[0.008]
$\lambda_{\mathcal{R}} = 1$			[0.000]		[0.517]			[0.063]
$\lambda_{\mathcal{K}} = 1$					[0.000]	[0.000]		
$\gamma_{\mathcal{R}} = \gamma_{\mathcal{K}}$				[0.007]	[0.012]	[0.000]	[0.234]	[0.110]
$\kappa_{cos}^{\mathcal{K}} = \kappa_{sin}^{\mathcal{K}} = 0$						[0.000]	[0.336]	[0.031]
res_3	[0.555]	[0.900]	[0.070]	[0.020]	[0.030]	[0.011]	[0.011]	[0.009]
res_4	[0.231]	[0.776]	[0.032]	[0.044]	[0.031]	[0.004]	[0.001]	[0.007]
ll	40.721	40.759	39.615	41.250	46.707	67.387	67.850	66.345
bic	−70.230	−66.568	−60.542	−63.812	−67.250	−108.610	−105.798	−99.050
$rmse_4$	0.144	0.144	0.118	0.140	0.115	0.089	0.086	0.111
$rmse_3$	0.202	0.202	0.225	0.203	0.197	0.134	0.137	0.109

Notes: The numbers in parentheses are robust standard errors, where the significance stars are to be read as $^{*} < 0.1,^{**} < 0.05,^{***} < 0.01$. Probability values are in brackets. Symbols $\mathbf{B}$ and $\mathbf{F}$ denote the Box-Cox and Fourier forms respectively, and "Exp." denotes exponential (i.e., $\lambda = 1$). In the second section of the table, we present Wald tests of various parameter restrictions. In the (final) diagnostic section of the table, the first two rows refer to ADF test of the unit root null associated to the errors in equations (4) and the logged form of (3) (albeit with the TFP term as the dependent variable) and the p-values are obtained by bootstrapping distribution. Finally, terms ll, bic and $rmse$ denote, respectively, the Log Likelihood, and the Bayesian Information Criterion, and the Root Mean Square Error.

8 Conclusion

We introduced R&D capital alongside R&D labor into the Idea Production Function and estimated it using a flexible, non-neutral specification. We find that the elasticity of substitution between R&D inputs in the IPF is $0.7 - 0.8$ and significantly below unity. This implies that R&D capital should be considered as an essential factor in producing ideas, and complementary to R&D labor – in other words the marginal productivity of R&D labor will be enhanced by the presence of R&D capital.

We also identify a systematic positive trend in R&D labor productivity at about 1% per year on average and a cyclical trend in R&D capital productivity. Our results suggest that, aside from cyclical variability, the effective supply of R&D capital was systematically lagging behind R&D labor, constraining R&D output over the long run. These results were verified using raw patents, quality adjusted TFP, and breakthrough patents, suggesting a high degree of robustness.

Our results imply that ideas, rather than simply getting harder to find, in fact *require more sophisticated lab equipment* to be found (or implemented). This is a scarcity that can only be overcome by increased accumulation and development of R&D capital, not necessarily just by employing more R&D staff. Furthermore, because investments in R&D equipment are an endogenous variable that can be influenced by policy, our results suggest a weak case for future secular stagnation.

Our analysis could be extended in a number of dimensions. First, one could use international panel data or aggregated global-level data on R&D inputs and output to gauge whether our results hold more broadly. Second, one could consider using sectoral data to provide more granularity on the ideas production function, and thus whether there is marked sectoral heterogeneity. Finally, our results could be used in theoretical studies aimed at understanding the mechanisms of long-run growth in the presence of R&D capital.

References

Ang, James B., and Jakob B. Madsen. 2011. "Can Second-Generation Endogenous Growth Models Explain the Productivity Trends and Knowledge Production in the Asian Miracle Economies?" *Review of Economics and Statistics*, 93(4): 1360–1373.

Aum, Sangmin, Sang Yoon Lee, and Yongseok Shin. 2018. "Computerizing Industries and Routinizing Jobs: Explaining Trends in Aggregate Productivity." *Journal of Monetary Economics*, 97: 1–21.

Bai, Jushan, and Pierre Perron. 1998. "Estimating and Testing Linear Models with Multiple Structural Changes." *Econometrica*, 66(1): 47–78.

Bai, Jushan, and Pierre Perron. 2003. "Computation and analysis of multiple structural change models." *Journal of Applied Econometrics*, 18(1): 1–22.

Baqaee, David Rezza, and Emmanuel Farhi. 2020. "Productivity and Misallocation in General Equilibrium." *Quarterly Journal of Economics*, 135: 105–163.

Bernstein, Jeffrey I., and Theofanis P. Mamuneas. 2006. "R&D depreciation, stocks, user costs and productivity growth for US R&D intensive industries." *Structural Change and Economic Dynamics*, 17(1): 70–98.

Bloom, Nicholas, Charles I. Jones, John Van Reenen, and Michael Webb. 2020. "Are Ideas Getting Harder to Find?" *American Economic Review*, 110(4): 1104–1144.

Bresnahan, Timothy F., and Manuel Trajtenberg. 1995. "General Purpose Technologies: Engines of Growth?" *Journal of Econometrics*, 65: 83–108.

Brynjolfsson, Erik, and Andrew McAfee. 2014. *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies.* W. W. Norton & Co.

Brynjolfsson, Erik, Daniel Rock, and Chad Syverson. 2019. "Artificial Intelligence and the Modern Productivity Paradox: A Clash of Expectations and Statistics." In *The Economics of Artificial Intelligence: An Agenda.* , ed. Ajay Agrawal, Joshua Gans and Avi Goldfarb, 23–57. University of Chicago Press.

Christiano, Lawrence J., and Terry J. Fitzgerald. 2003. "The Band Pass Filter." *International Economic Review*, 44(2): 435–465.

Christopoulos, Dimitris K., and Miguel A. León-Ledesma. 2010. "Smooth breaks and non-linear mean reversion: Post-Bretton Woods real exchange rates." *Journal of International Money and Finance*, 29(6): 1076–1093.

Cowen, Tyler. 2011. *The Great Stagnation: How America Ate All the Low-hanging Fruit of Modern History, Got Sick, and Will (Eventually) Feel Better.* Dutton.

De Loecker, Jan, Jan Eeckhout, and Gabriel Unger. 2020. "The Rise of Market Power and the Macroeconomic Implications." *Quarterly Journal of Economics*, 135(2): 561–644.

Feenstra, Robert C., Robert Inklaar, and Marcel P. Timmer. 2015. "The Next Generation of the Penn World Table." *American Economic Review*, 105(10): 3150–3182.

Fernald, John G. 2015. "Productivity and Potential Output before, during, and after the Great Recession." *NBER Macroeconomics Annual*, 29(1): 1–51.

Fernald, John G. 2018. "A quarterly, utilization-adjusted series on total factor productivity." Federal Reserve Bank of San Francisco Working Paper Series 2012-19 (current update).

Fernald, John G., and Huiyu Li. 2022. "The Impact of COVID on Productivity and Potential Output." Presented at the 2022 Jackson Hole Symposium 'Reassessing Constraints on the Economy and Policy.', August 25-27, Jackson Hole, WY.

Fernald, John G., and J. Christina Wang. 2016. "Why Has the Cyclicality of Productivity Changed? What Does It Mean?" *Annual Review of Economics*, 8(1): 465–496.

Fernald, John G., Robert E. Hall, James H. Stock, and Mark W. Watson. 2017. "The Disappointing Recovery of Output after 2009." *Brookings Papers on Economic Activity*, 48(1 (Spring)): 1–81.

Gordon, Robert J. 2016. *The Rise and Fall of American Growth: The U.S. Standard of Living since the Civil War.* Princeton University Press.

Greenwood, Jeremy, Zvi Hercowitz, and Per Krusell. 1997. "Long-Run Implications of Investment-Specific Technological Change." *American Economic Review*, 87(3): 342–62.

Griliches, Zvi. 1998. *R & D and Productivity: The Econometric Evidence.* The University of Chicago Press, Chicago.

Grossman, Gene M., Elhanan Helpman, Ezra Oberfield, and Thomas Sampson. 2017. "The Productivity Slowdown and the Declining Labor Share: A Neoclassical Exploration." National Bureau of Economic Research, Inc NBER Working Papers 23853.

Growiec, Jakub. 2022*a*. *Accelerating Economic Growth: Lessons From 200,000 Years of Technological Progress and Human Development.* Springer. ISBN: 978-3-031-07194-2.

Growiec, Jakub. 2022*b*. "What Will Drive Global Economic Growth in the Digital Age?" SGH Warsaw School of Economics KAE Working Paper No. 2020/054.

Growiec, Jakub, Peter McAdam, and Jakub Mućk. 2018. "Endogenous labor share cycles: Theory and evidence." *Journal of Economic Dynamics and Control*, 87(C): 74–93.

Ha, Joonkyung, and Peter Howitt. 2007. "Accounting for Trends in Productivity and R&D: A Schumpeterian Critique of Semi-Endogenous Growth Theory." *Journal of Money, Credit and Banking*, 39(4): 733–774.

Helpman, Elhanan. 2010. *The Mystery of Economic Growth.* Harvard University Press.

Jones, Charles I. 1995. "Time Series Tests of Endogenous Growth Models." *Quarterly Journal of Economics*, 110(2): 495–525.

Jones, Charles I. 1999. "Growth: With or Without Scale Effects?" *American Economic Review*, 89(2): 139–144.

Jones, Charles I. 2002. "Sources of U.S. Economic Growth in a World of Ideas." *American Economic Review*, 92: 220–239.

Jones, Charles I. 2022*a*. "The End of Economic Growth? Unintended Consequences of a Declining Population." *American Economic Review*, 112(11): 3489–3527.

Jones, Charles I. 2022*b*. "The Past and Future of Economic Growth: A Semi-Endogenous Perspective." *Annual Review of Economics*, 14: 125–152.

Jorgenson, Dale W. 2005. "Accounting for Growth in the Information Age." In *Handbook of Economic Growth.* , ed. P. Aghion and S.N. Durlauf. North-Holland.

Kelly, Bryan, Dimitris Papanikolaou, Amit Seru, and Matt Taddy. 2021. "Measuring Technological Innovation over the Long Run." *American Economic Review: Insights*, 3(3): 303–20.

Kim, Hyeon-Wook, and Zia Qureshi. 2020. *Growth in a Time of Change: Global and Country Perspectives on a New Agenda.* Brookings Institution Press.

Klump, Rainer, and Olivier de La Grandville. 2000. "Economic growth and the elasticity of substitution: Two theorems and some suggestions." *American Economic Review*, 90(1): 282–291.

Klump, Rainer, Peter McAdam, and Alpo Willman. 2007. "Factor Substitution and Factor Augmenting Technical Progress in the US." *Review of Economics and Statistics*, 89: 183–192.

Klump, Rainer, Peter McAdam, and Alpo Willman. 2012. "The Normalized CES Production Function: Theory And Empirics." *Journal of Economic Surveys*, 26(5): 769–799.

Korinek, Anton. 2023. "Language Models and Cognitive Automation for Economic Research." National Bureau of Economic Research, Inc NBER Working Papers 30957.

Kruse-Andersen, Peter K. 2017. "Testing R&D-Based Endogenous Growth Models." University of Copenhagen Working paper.

La Grandville, Olivier de. 1989. "In Quest of the Slutzky Diamond." *American Economic Review*, 79: 468–481.

León-Ledesma, Miguel A, and Mathan Satchi. 2019. "Appropriate Technology and Balanced Growth." *Review of Economic Studies*, 86(2): 807–835.

León-Ledesma, Miguel A., Peter McAdam, and Alpo Willman. 2010. "Identifying the Elasticity of Substitution with Biased Technical Change." *American Economic Review*, 100(4): 1330–1357.

Li, Wendy C. Y., and Bronwyn H. Hall. 2019. "Depreciation of Business R&D Capital." *Review of Income and Wealth*, 66(1): 161–180.

Madsen, Jakob B. 2008. "Semi-endogenous versus Schumpeterian growth models: testing the knowledge production function using international data." *Journal of Economic Growth*, 13(1): 1–26.

Marco, Alan C., Michael Carley, Steven Jackson, and Amanda Myers. 2015. "The USPTO Historical Patent Data Files: Two Centuries of Innovation." USPTO Economic Working Paper 2015-1. Available at https://ssrn.com/abstract=2849514.

Oberfield, Ezra, and Devesh Raval. 2021. "Micro Data and Macro Technology." *Econometrica*, 89(2): 703–732.

OECD. 2009. *Measuring Capital OECD Manual, 2nd Edition.* OECD Publications.

OECD. 2015. *Frascati Manual 2015.* OECD Publications.

Ramey, Valerie A. 2020. "Secular stagnation or technological lull?" *Journal of Policy Modeling*, 42(4): 767–777.

Rivera-Batiz, Luis A., and Paul M. Romer. 1991. "Economic Integration and Endogenous Growth." *Quarterly Journal of Economics*, 106(2): 531–555.

Romer, Paul M. 1990. "Endogenous Technological Change." *Journal of Political Economy*, 98: S71–S102.

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek. 2019. "IPUMS USA: Version 9.0 [dataset]."

Sato, Ryuzo. 1970. "The Estimation of Biased Technical Progress and the Production Function." *International Economic Review*, 11(2): 179–208.

Solow, Robert M. 1957. "Technical Change and the Aggregate Production Function." *Review of Economics and Statistics*, 39(3): 312–320.

van Ark, Bart, and Kirsten Jäger. 2017. "Recent Trends in Europe's Output and Productivity Growth Performance at the Sector Level, 2002-2015." *International Productivity Monitor*, 33(Fall): 8–23.

Venturini, Francesco. 2012. "Looking into the black box of Schumpeterian growth theories: An empirical assessment of R&D races." *European Economic Review*, 56(8): 1530 – 1545.

Online Appendices

A Data Construction and Analysis

A fundamental problem in the current research project is the collection of long time series that would be acceptable proxies of variable of interest. We will discuss in detail all the specific problems, e.g. long-run trends in relative prices or changes in occupational systems, that affect precision of our measurement strategy.

A.1 Output

From an economic perspective, R&D activities aim to reduce unit cost of production or increase the variety of offered goods. At the aggregate level, the existing stock of knowledge/technology can be indirectly measures by the total factor productivity (TFP). Specifically one can use TFP estimates provided by Fernald (2018) which account for changing capacity utilization. Since the TFP is measured residually, adjustment by capacity utilization reduces unwarranted variation due to changes in short-run factors, e.g., demand fluctuation. Second, one could use the Multifactor Productivity (MFP) index provided by OECD. Thirdly, one could use the latest Penn World Table estimates of the TFP (Feenstra, Inklaar and Timmer, 2015).

However, using TFP as a proxy of the R&D output yields some problems. First, any changes in TFP might be driven by changes in the technology but also it could result from other processes. For instance, reduction in mis-allocation could improve TFP. Second, the TFP is a stock variable and, as a result, the output of the R&D sector is related to changes in existing technology so it could be measured by growth rates of the TFP. At the same time, functional form of the IPF requires positive values of the output. This condition makes the TFP growth a less applicable proxy as there could be some periods/events of decline in the TFP.[1]

Another strategy in measuring an aggregate macroeconomic outcome of the R&D sector is to use patent data. A common practice in related literature is to use the patent applications (Madsen, 2008; Ang and Madsen, 2011; Venturini, 2012). Since we are interested in long

[1] In the associated literature, the problem of negative TFP growth in estimating idea production function is overcome by considering approximation of idea production function (Ha and Howitt, 2007) or by taking 5-year averages (Ang and Madsen, 2011).

historical patent data our principal measure is taken from Marco et al. (2015). The series of interest – the flow of new patent applications – is portrayed on Figure A.1.

A.2 R&D Capital

We use Bureau of Economic Analysis (BEA) data to estimate R&D capital in the US economy. The key problem in measuring the total R&D capital stock is the fact that there is no available aggregate R&D capital stock, i.e., combining private and public sector. At the same time, the BEA does not publish long-run series on fixed-weights aggregates of R&D investment or R&D stock.[2] The reason for that is that there are long-run trends in the relative prices of inputs, long-run decline in relative prices of investment (Greenwood, Hercowitz and Krusell, 1997). Therefore, we consider two approaches. First, we use a direct measure of fixed assets. Second, the standard perpetual inventory method (PIM) is applied to the investment series in order to estimate the R&D capital.

In our first approach, we aggregate the available series on the R&D fixed assets. Although the BEA does not provide data on aggregate R&D assets it offers detailed series for the private and government sector. In particular, we use the BEA chain-type quantity indexes for R&D assets in both private (BEA code: `kcntotl1rd00`) and public sector (BEA code: `kcgtotl1rd00`). Since both series are indexes and, therefore, measure capital accumulation they do not contain information about capital level. Thus, we take the nominal value of net R&D capital stock to get an estimate of the real capital stock. Namely, we use current cost net R&D stock in 2012 in private (BEA code: `k1ntotl1rd00`) and public sector (BEA code: `k1gtotl1rd00`). Next, it is assumed that public and private R&D capital are perfect substitutes, i.e., the elasticity of substitution between these inputs tends to ∞, and we simply sum estimated aggregates.

In the second approach, we use the perpetual inventory method. The capital stock (K_t) is the sum of the capital stock in previous period reduced by depreciation and investment at period t:

$$\mathcal{K}_t = (1-\delta)\,\mathcal{K}_{t-1} + I_t^{rd} \tag{A.1}$$

where δ is the depreciation rate of R&D capital and I_t is the real investment in R&D. A

[2] The available data (in constant dollars) starts in 1999. This time span is too short to analyze the long-run patterns in R&D productivity.

key problem in calculating capital based on the PIM (A.1) formula is the initial condition problem. We follow OECD (2009) and apply the following formula:

$$\mathcal{K}_0 = I_0^{rd}/(g + \delta) \tag{A.2}$$

where g is the long-run (geometric) growth rate of R&D investment.

While the g parameter can be easily calculated from historical data there is a lot of uncertainty about the depreciation rate of the R&D capital. The standard choice in the literature is to fix depreciation rate at 15% (Venturini, 2012). More recently, there are several studies that provide empirical evidence that suggest a higher depreciation rate. Bernstein and Mamuneas (2006) find that depreciation rate is above 15% while Li and Hall (2019) document that the depreciation rate of R&D capital is even above 30%. In the 2019 KLEMS vintage, the depreciation rate for the R&D assets is fixed at 20%.[3]

In addition, the BEA publish historical series on depreciation of R&D capital. As previously, depreciation of R&D capital is also available separately for private and public sector. According to the BEA estimates, the implied depreciation rate is slightly above consensual value of 15%. However, the BEA estimates suggest that the depreciation rate has not been constant over time (Figure A.7). Before the WWII substantial short-run variation in the depreciation rate for both public and private R&D capital can be observed. This is due to approximation error related to the available BEA statistics. Namely, the BEA publish data expressed in billions of dollars and rounded to one digit. Therefore, in extreme case, i.e., for public capital, there is no depreciation of R&D capital in the 1920s because the reported value of depreciation is zero. Abstracting from this period, the implied depreciation rate has been stable since the WWII to the late 1950s. After that, there is unquestionable declining trend in depreciation rate. In our baseline setting, we use standard in the literature assumption that annual depreciation rate equals 15%. However, based on above discussion, we will carefully document a sensitivity of this choice.

To apply the PIM method we proceed as follows. Since there is no measures of real R&D investment expressed in chained dollars we estimate the series of interest based on available series, i.e, nominal R&D investment data as well as price indexes. For private sector, we divide nominal R&D investment (BEA code: `Y006RC`) by price index of this asset

[3] See https://euklems.eu/wp-content/uploads/2019/10/Methodology.pdf.

(BEA code: `Y006RG`). The same strategy is applied for public sector (BEA codes `Y057RC` and `Y057RG`, respectively). In addition, we also consider the following components of public investment: federal non-defense (BEA codes: Y069RC and Y069RG), defense (BEA codes: `Y076RC` and `Y076RG`) and state and local (BEA codes: `Y073RC` and `Y073RG`).

Based on the constructed series we can formulate the following stylized facts:

- The R&D capital stock has a unit root.[4] Non-stationarity of R&D capital is quite an intuitive feature as it can be expected that in the economy there is some accumulation of R&D capital. This implies that R&D capital should be rather an upward trending than a stationary variable. Technically speaking (see equation (A.1)), R&D capital would be stationary if investment in new capital (I_t) equals over the time depreciated capital stock. This case seems to be unrealistic.

- The dynamics of accumulation in aggregate R&D assets is complex. There are several time series features that can be simultaneously documented.

  - There is a downward (almost linear) trend in growth rate of R&D capital.

  - Even after differentiating the R&D stock are highly persistent. This is suggested by high persistence estimates obtained for the $AR(1)$ model.

  - There is a visible structural break in the R&D capital accumulation. Since the 1970s the growth rate of the R&D capital has dropped permanently. This can be observed for both the FAT and PIM based series. In addition, the above visual investigation is confirmed by a broad range of statistical tests.

  - The role of short-run fluctuations has been declining over the time. To evaluate the role of business cycles and medium-term variation we use a band pass filter proposed by Christiano and Fitzgerald (2003) and decompose all fluctuations in three groups: short-run/business cycles (frequency higher than 8 years), medium-term swings (from 8 to 50 years) and long-run oscillation and long-run trend (frequency below 50 years). Visual inspection of spectral decomposition suggests that the role of business cycles in shaping the R&D accumulation was significant only before the 1970s. Since the beginning of the 1970s the R&D

[4] We applied a battery of unit root and unit root under structural break tests. These are available on request.

accumulation is mostly driven by long-run trend and medium-run swings. Substantial magnitude of the long-run and medium-run fluctuation is consistent with high persistence that can be found for annual growth rates of R&D assets.

- The accumulation of R&D capital has been faster on average in comparison with total assets.
- The share of public/private assets in the total R&D capital has not been stable over the time. The following periods can be identified (see Figure A.12)
  - Sudden and substantial rise in public R&D capital during the WWII due to increasing role of defense R&D.
  - Slight upward trend in share of public R&D assets in total R&D assets due to rising role of defense sector as well as space programme.
  - Diminishing role of public assets in total R&D capital since the beginning of the 1970s.
- The properties of the PIM-based series of the R&D capital are slightly sensitive to a choice of (i) depreciation rate, and (ii) initial period.

  The characteristics of the PIM-based series depend on the depreciation rate as well as initial year (A.1). To check sensitivity of properties of the PIM based series of R&D capital we calculate the counterfactual PIM series (i) using various values of depreciation rate, (ii) truncating recursively available sample. To scrutinize an effect of these changes we calculate long-run averages. In addition, we consider two measure of comovement with the FAT-based measures of the R&D capital. First, the correlation between annual growth rates which measures short-run comovement. Second, we employ a long-run approach which is related to testing the cointegration.

  The long-run properties of the PIM-based series are not extremely sensitive to a choice of initial year and depreciation rate. Figure A.8 illustrates a dependence of geometric growth rate and the average annual growth rate of R&D capital on depreciation rate and initial period. There is a natural trade-off between the assumed depreciation rate of the R&D capital and its long-term growth rate. For a higher depreciation rate, more investment is required to replace obsolete R&D capital and this implies slower

R&D capital accumulation. However, this effect is not substantial. For extreme values of the depreciation rate, i.e., 0% and 40%, the average annual growth rate as well as geometric growth rate do not differ so much and range from 6% to 7% per annum.

Furthermore, the long-term average rate of accumulation of R&D capital depends substantially on the choice of initial period, but this relationship is consistent with the long-run slowdown in R&D accumulation. Both the PIM-based series and the FAT-based measure exhibit a visible decline in growth rates of available R&D stock (see Figure A.8, right panel). This fact is consistent with previous evidence in favor of the occurrence of structural break(s).

Finally, we look at comovement between the FAT-based measure and various PIM-based proxies that base on different values of δ. At first sight, there is an extremely high positive short-run correlation between the considered series (Figure A.9 and Figure A.10). In particular a choice of time invariant depreciation rate has no impact on the analyzed degree of comovement as the lowest estimated correlation coefficient is above 0.9. The short-run correlation is slightly lower for detected previously structural break (in the late 1960s/ early 1970s) nut it is still significantly positive. The analysis of the potential impact of our PIM assumptions on the long-run comovement with the FAT-based series is more puzzling. In our baseline case, that is, $\delta = 0.15$, there is no strong evidence in favor of cointegration between the time series analyzed. Abstracting for the low power of unit root and cointegration test, the reason for that is structural breaks in the analyzed series and its effect on the DGP (data generating process) could be not proportional.

– The constructed series of the R&D capital are comparable to measures in other databases. We use two additional data sources that offer data on the US R&D capital. First, we use capital the R&D capital stock from the KLEMS database (van Ark and Jäger, 2017). Second, we use an index of R&D capital services from the Multifactor Productivity (MFP) database provided by OECD.

All series are shown in Figure A.3. It is straightforward to observe that short-run comovement between those series is high. Moreover, the average growth rate is very similar among the considered sources.

– Finally, the share of the R&D assets in total capital stock has been systematically rising from the 1920s to the 1980s and has been roughly stable since the beginning of the 1980s. All in all, the average share of the R&D capital in nonresidential (total) available assets has fluctuated around 8% (5%) since the beginning of the 1980s. This empirical pattern is mostly determined by a rising role of R&D intensity in private sector. The share of the R&D assets in private capital stock has systematically increased since the 1920s. At the same time, the share of R&D in public assets exhibits hump-shaped trajectory, reaching the maximum in the 1980s.

A.3 R&D Labor

Estimation of the idea production function requires data on labor engaged in the research and development process. At the conceptual level and in line with the definition from the Frascati Manual (OECD, 2015) it refers to employees who undertake creative work that is aimed at general increase in an existing stock of knowledge.

In practice, an application of the above definition requires very detailed information about tasks that are related to R&D activity. However, according to the best of our knowledge such data are not available. As a result, R&D activity cannot be measured directly. Therefore, we will use two strategies.

First, we take advantage of publicly available data on R&D employment. Although statistical offices (Eurostat or OECD) publish estimates on R&D activity their availability is strongly limited. Namely, the Eurostat/OECD series starts in 1981. To overcome this problem, we use older vintages in order to extrapolate existing series. Before 1981 we use data collected within the IRIS (Industrial Research and Development Information System) program conducted by the NSF (the National Science Foundation). Moreover, based on historical data from Jones (1995) it is possible to get extrapolated observations earlier, i.e., in 1950s.

In our second approach, we estimate the labor input in R&D activity using microdata that contain information on the structure of occupations. An ideal strategy is to use detailed data on skills/abilities-content in occupations and merge them with the structure of labor force. The most important problem with this approach is that, according to our best knowledge, there is no longitudinal survey on research-intensity among occupations. For instance, the

O*NET data offers estimates on skill and abilities intensity but there is no direct measure of research intensity and the time span of this dataset is quite short since this survey started in 1998.

Thus, in our empirical part, we use IPUMS CPS data (Ruggles et al., 2019). This database offers harmonized micro data, namely the Current Population Survey (CPS), i.e., the monthly U.S. labor force survey. Based on the conceptual definition of R&D personal or the S&E groups we can identify the following occupational groups whose work could be classified as R&D activity:[5]

Scientists Agricultural and Food Scientists (*IPUMS code* 1600); Biological Scientists (1610); Conservation Scientists and Foresters (1640); Medical Scientists, and Life Scientists, All Other (1650); Astronomers and Physicists (1700); Atmospheric and Space Scientists (1710); Chemists and Materials Scientists (1720); Environmental Scientists and Geoscientists (1740); Physical Scientists, nec (1760).

Mathematical & Computer Occupations Actuaries (1200); Operations Research Analysts (1220); Statisticians (1230); Mathematical science occupations, nec (1240); Computer Scientists and Systems Analysts/Network systems Analysts/Web Developers (1000); Computer Programmers (1010); Software Developers, Applications and Systems Software (1020); Computer Support Specialists (1050); Database Administrators (1060); Network and Computer Systems Administrators (1100).

Engineers Architects, Except Naval (1300); Surveyors, Cartographers, and Photogrammetrists (1310); Aerospace Engineers (1320); Chemical Engineers (1350); Civil Engineers (1360); Computer Hardware Engineers (1400); Electrical and Electronics Engineers (1410); Environmental Engineers (1420); Industrial Engineers, including Health and Safety (1430); Marine Engineers and Naval Architects (1440); Materials Engineers (1450); Mechanical Engineers (1460); Petroleum, mining and geological engineers, including mining safety engineers (1520); Engineers, nec (1530); Drafters (1540); Engineering Technicians, Except Drafters (1550); Surveying and Mapping Technicians (1560).

Technicians Agricultural and Food Science Technicians (1900); Biological Technicians (1910); Chemical Technicians (1920); Geological and Petroleum Technicians, and Nuclear Technicians (1930); Life, Physical, and Social Science Technicians, nec (1960); Professional, Research, or Technical Workers, nec (1980).

Social Scientists Economists and market researchers (1800); Psychologists (1820); Urban and Regional Planners (1830); Social Scientists, nec (1840).

[5] In practice, we try to match the Eurostat definition. According to human resources in the science and technology approach, scientists and engineers (S&E) are workers who conduct research, improve or develop concepts, theories and operational methods and/or apply scientific knowledge related to fields. This definition can be covered by following group occupation (according to ISCO-08 classification): Science and engineering professionals (21), Health professional (22) and Information and communications technology professionals (25).

Health Professionals Chiropractors (3000); Dentists (3010); Dietitians and Nutritionists (3030); Optometrists (3040); Pharmacists (3050); Physicians and Surgeons (3060).

Based on the above available classification we define two aggregates of R&D labor. In our baseline definition, we include scientists, mathematical and computer occupations, and engineers. In addition, we consider a broader definition that includes also technicians, social scientists and health professionals.

The following set of stylized facts can be formulated:

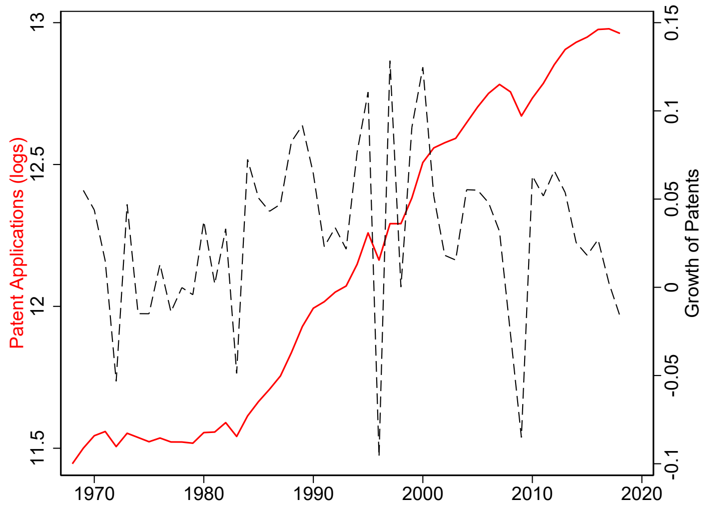
– Share of the R&D related workers i total employment is increasing over time. This is consistent with the previous empirical evidence in the literature (Jones, 1995; Ha and Howitt, 2007).

– There are substantial differences in the level of R&D employment. Even if we look at the series of the Scientists&Engineers the share of this group in total employment according to data provided by (Jones, 1995) is almost two times higher than the share estimated based on NSF/Eurostat data. By definition, this difference would be higher if we compared with the (IPSUM) share a broader group of occupations.

– Nevertheless, all proxies of the R&D employment suggest almost identical upward tendency. According to the merged series on Scientists&Engineers and the IPUMS-based group of scientists the share of R&D employment in total employment rose by around 80% between 1968 and 2017. For the baseline definition based on IPUMS, this increase has been even larger and exceeded 100%.

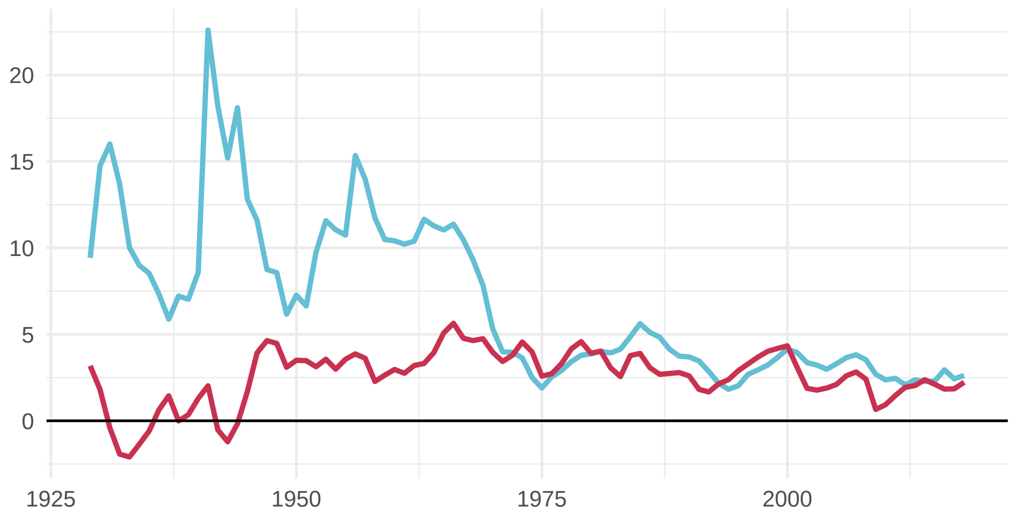
– A detailed decomposition of the measures based on IPUMS illustrates the key measurement problem. Ongoing technical change has created demand for new occupations that are closely related to new technologies. This is mostly observable for computer-related occupations. In the late 1960s this occupational group was almost absent on the labor market while in the late 2000s their share (together with mathematical occupation) was around 4%.

Figure A.1: New patent applications in the US, 1968-2019



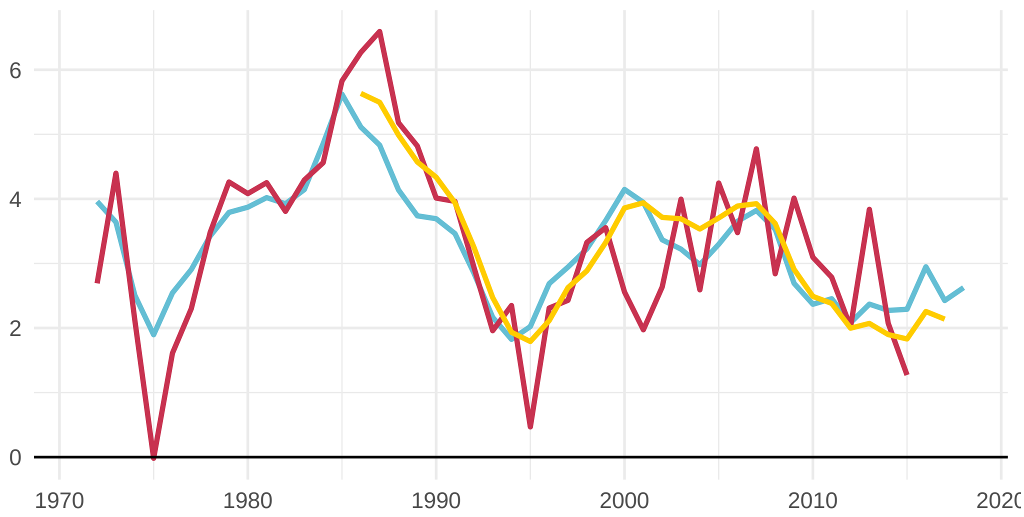
Notes: In this figure we plot on the lhs axis the log of patent applications and on the rhs the growth rate of patents. Data derived from Marco et al. (2015).

Figure A.2: Total R&D capital (annual growth rate) and Non-residential Private Fixed Assets



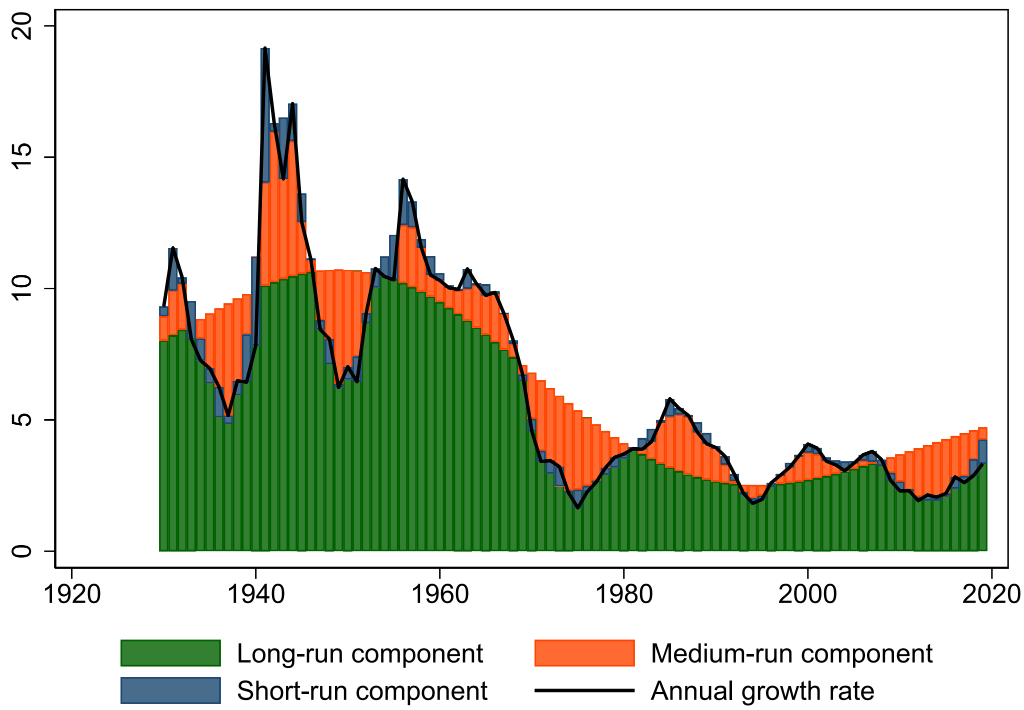
Notes: red color denotes nonresidential private capital while blue line stands for total R&D capital.

Figure A.3: The comparison of the R&D stock with other data sources (annual growth rate)



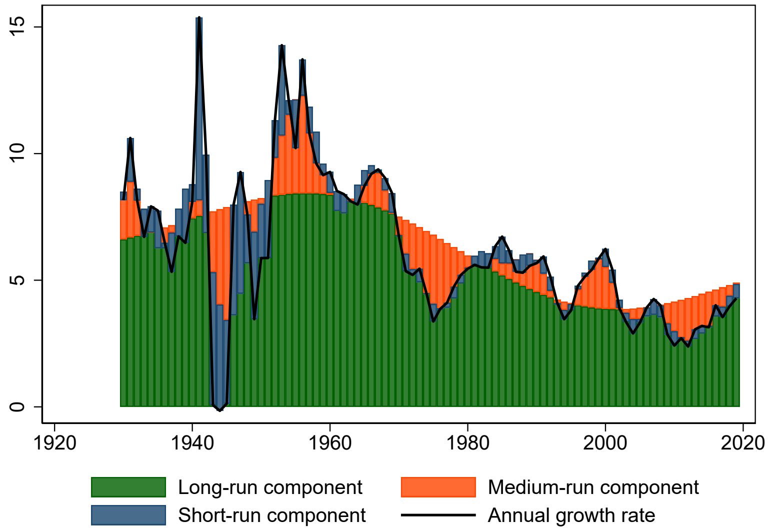
Notes: blue line is the BEA-based estimates of R&D stock, red color denotes the KLEMS estimate and yellow color is the OECD estimate of the R&D capital.

Figure A.4: Spectral decomposition of the total R&D capital (annual growth rate)



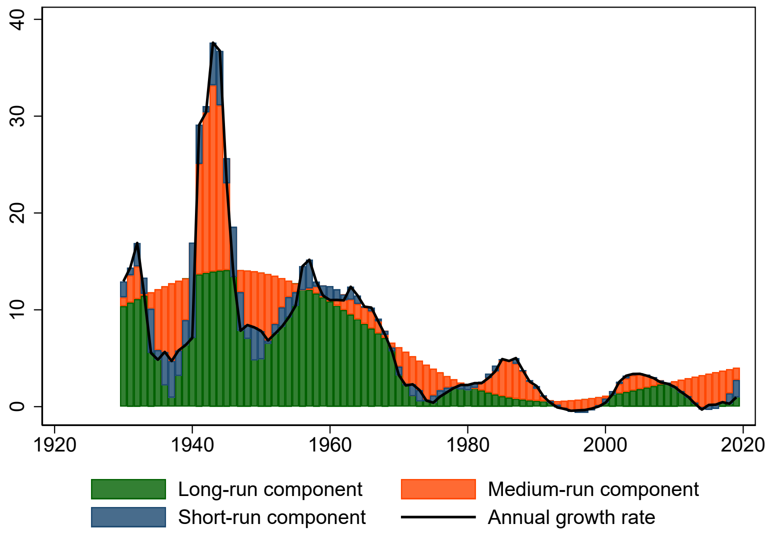
Notes: See main text for description.

FIGURE A.5: Spectral decomposition of the private R&D capital (annual growth rate)



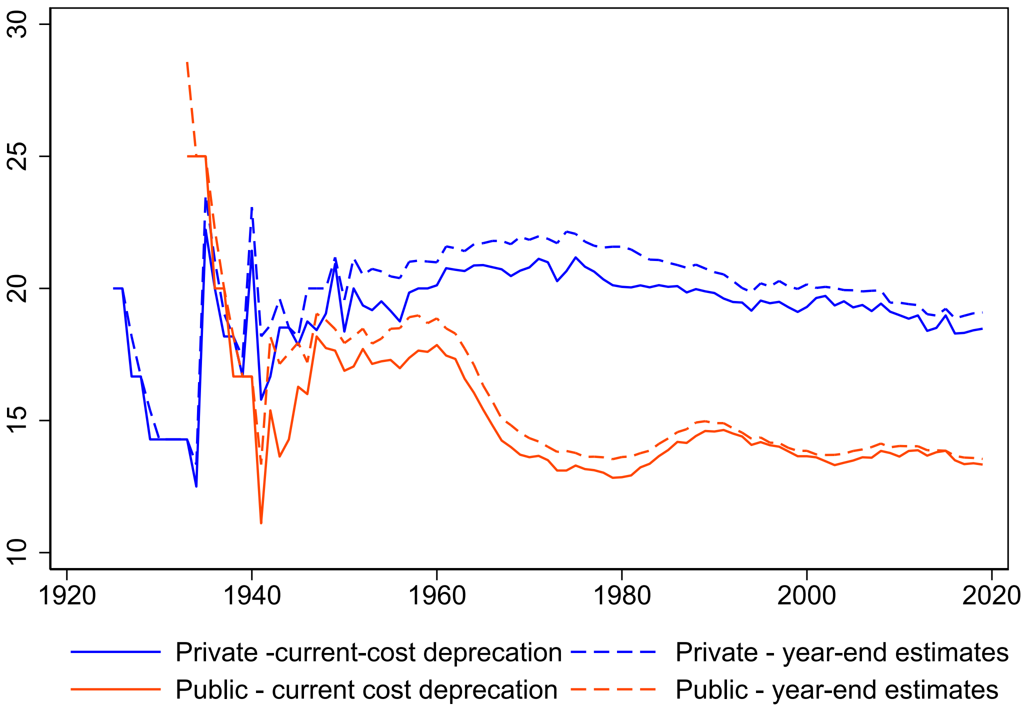
Notes: See main text for description.

FIGURE A.6: Spectral decomposition of the public R&D capital (annual growth rate)



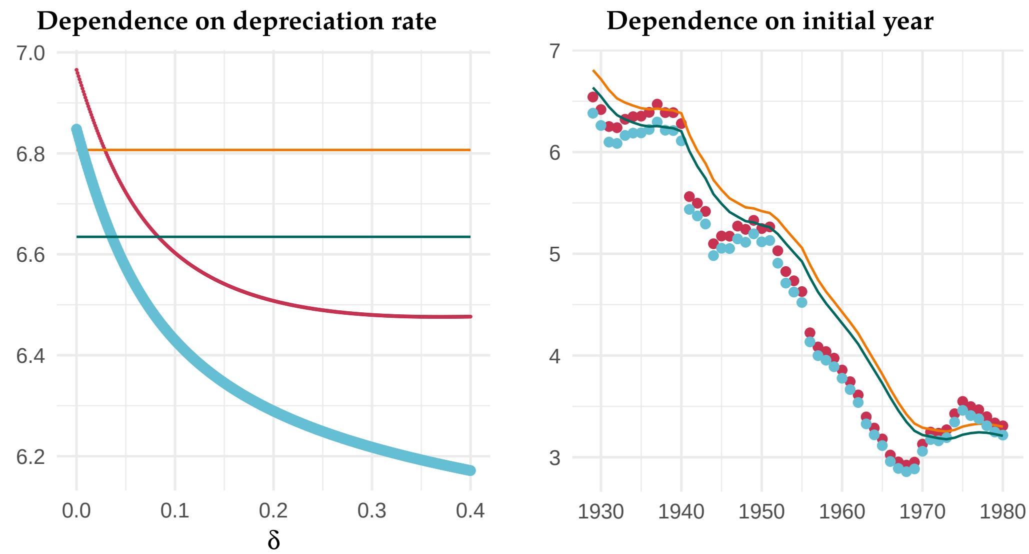
Notes: See main text for description.

Figure A.7: Implied depreciation rate of R&D capital based on BEA data



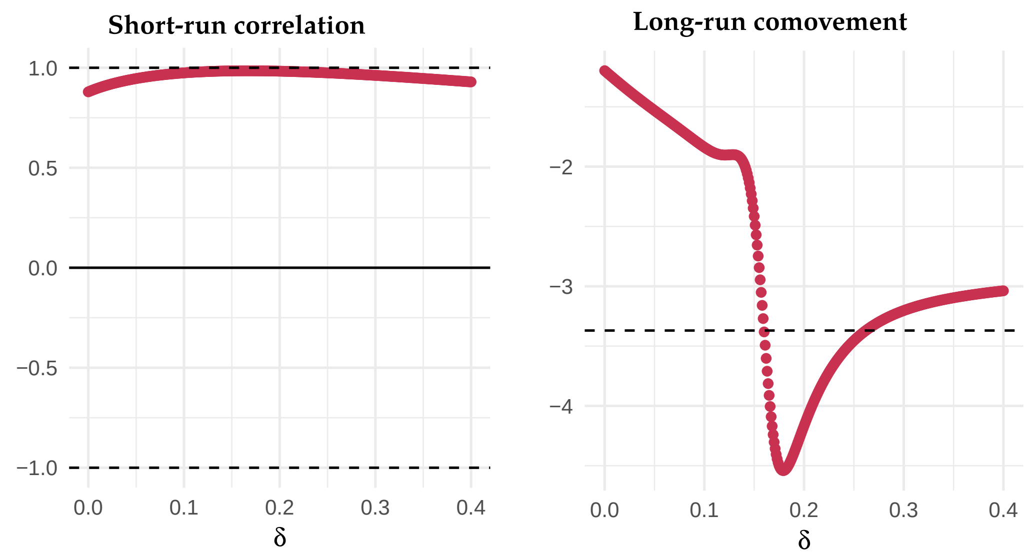
Notes: Implied depreciation rates for public and private R&D capital.

Figure A.8: Sensitivity analysis of the PIM-based series to depreciation rate values (δ) and choice of initial year



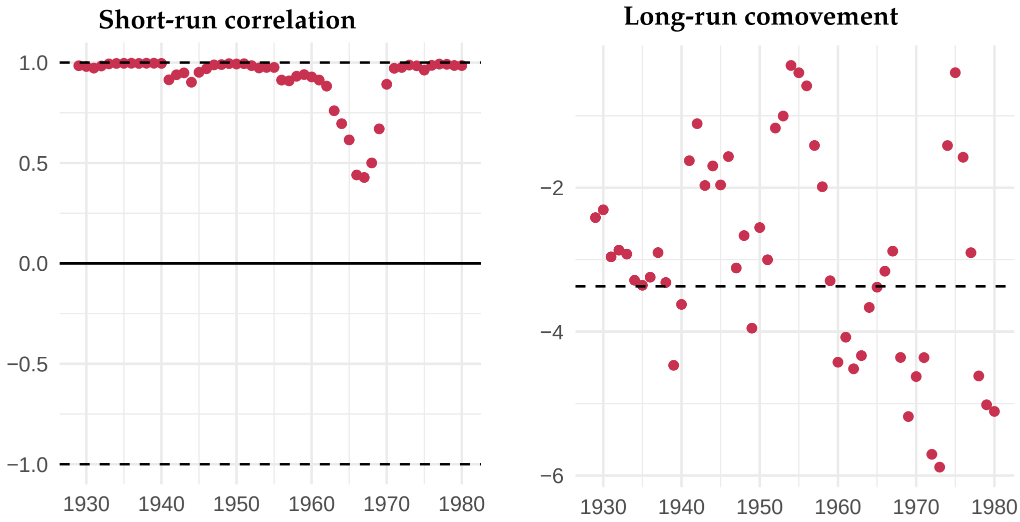
Note: blue color denotes average growth rate of the PIM-based measure, red color stands for geometric growth rate of the PIM-based measure, green color represents average average growth rate of FAT-based measure and orange color stands for the FAT-based geometric growth rate of FAT-based measure..

Figure A.9: Sensitivity analysis of the PIM-based series to depreciation rate value (δ)



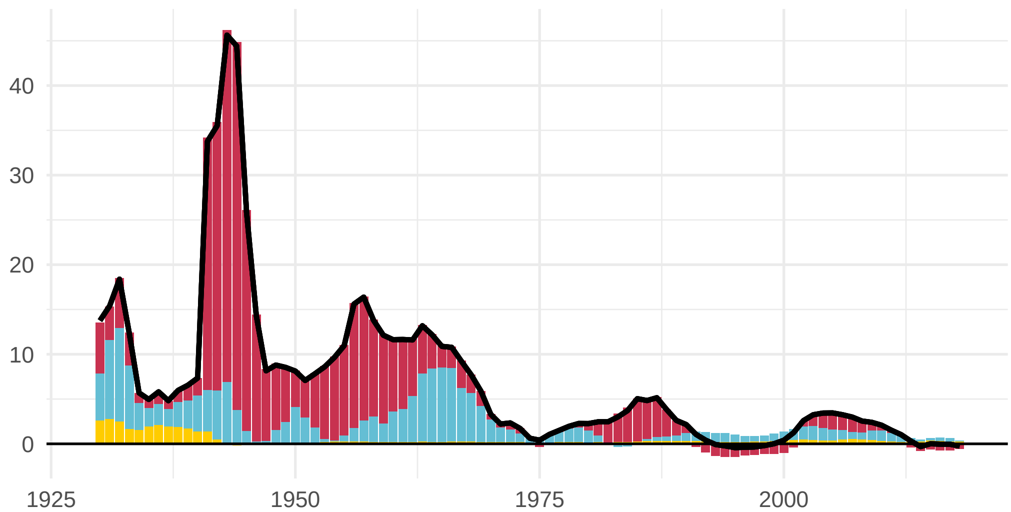
Notes: **Short-run correlation** portrays correlation between annual growth rates of the FAT R&D capital and the PIM-based R&D capital. **Long-run comovement** illustrates the ADF statistics from regression of the logged PIM-based R&D capital on the logged FAT-based R&D capital. The dashed horizontal line is the 5% critical value for testing co-integration.

FIGURE A.10: Sensitivity analysis of the PIM-based series to a choice of initial year



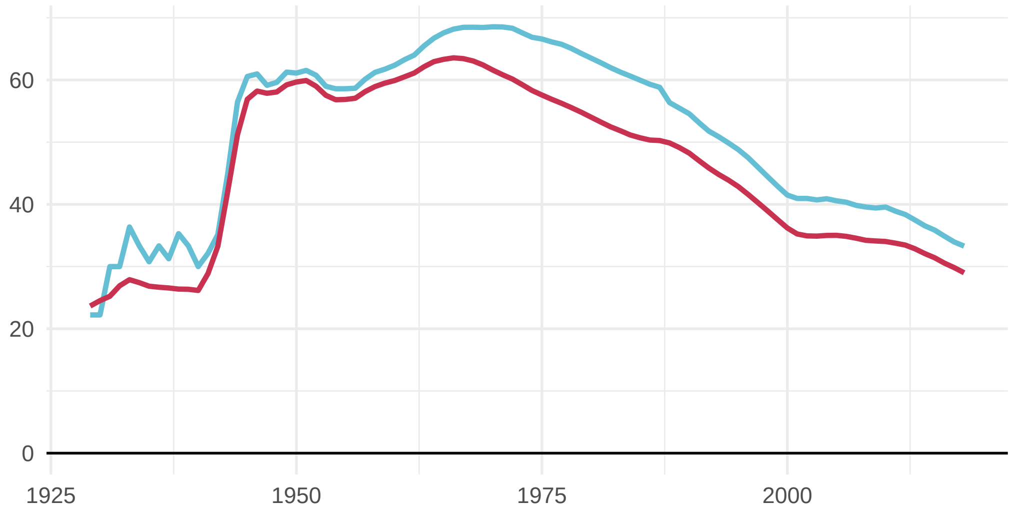
Notes: See Figure A.9.

FIGURE A.11: Public R&D capital (annual growth rate) and its components



Notes: red color denotes defense R&D capital, blue color stands for non-defense federal R&D capital while yellow refers to state and local R&D capital.

FIGURE A.12: Share of public R&D capital in total R&D capital (in %)



Notes: red color denotes the share calculated using PIM-based series while blue color represents share obtained from (nominal) BEA series.

FIGURE A.13: Share of R&D assets in total and non-residential capital stock (in %)

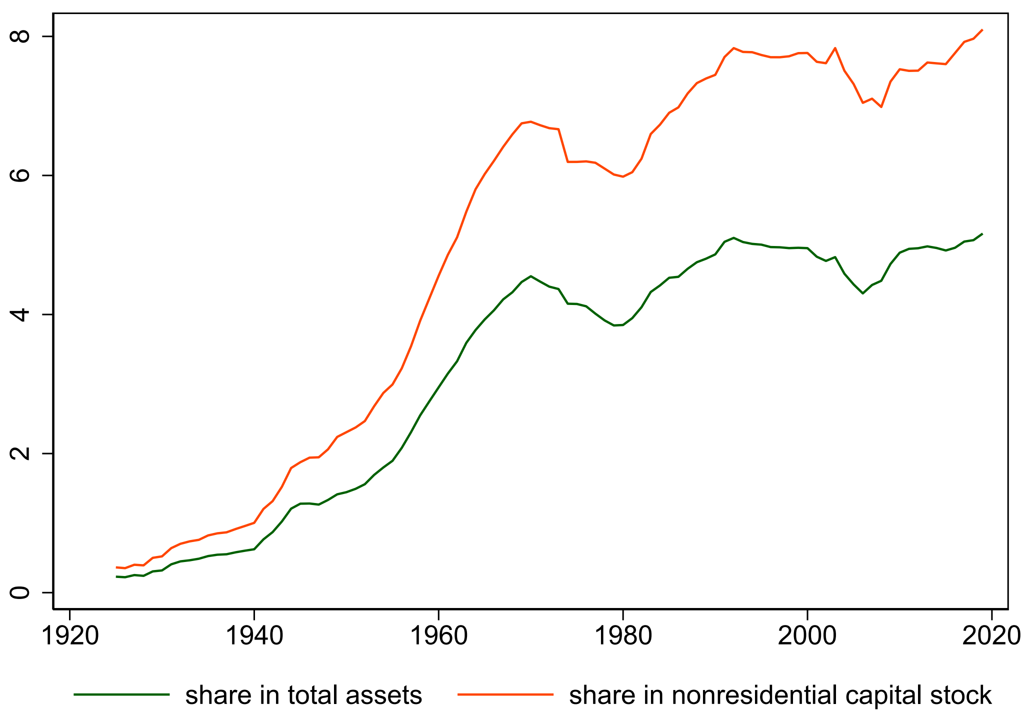
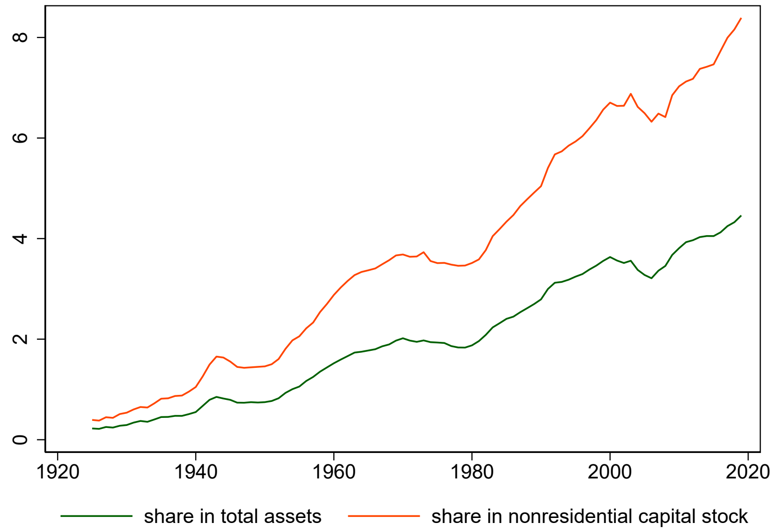
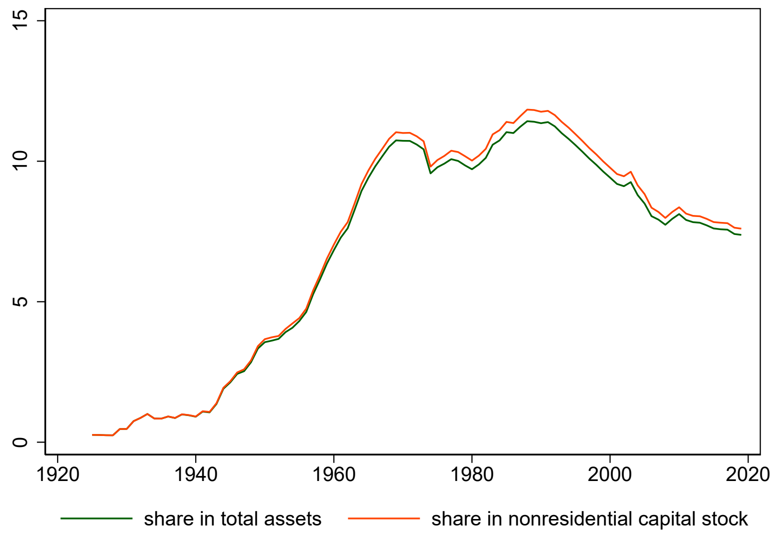


Figure A.14: Share of R&D assets in total private capital stock (in %)



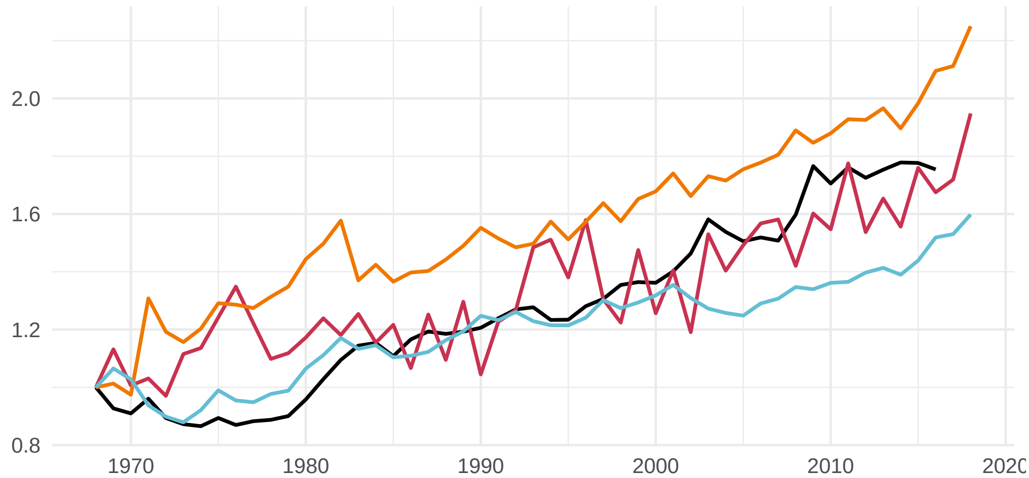
Notes: This figure shows private R&D assets by class.

Figure A.15: Share of R&D assets in total public capital stock (in %)



Notes: This figure shows public R&D assets by class.

Figure A.16: Share of the R&D employment (FTE, 1968=1)



Notes: The **black line is the merged (from various sources) series, red line is the IPUMS-based share of scientists, orange stands for the IPUMS based share of the R&D employees according to baseline definition while blue represents the IPUMS-based share of the R&D employees according to broader definition.**

Figure A.17: Share of the R&D related workers in total US employment and hours

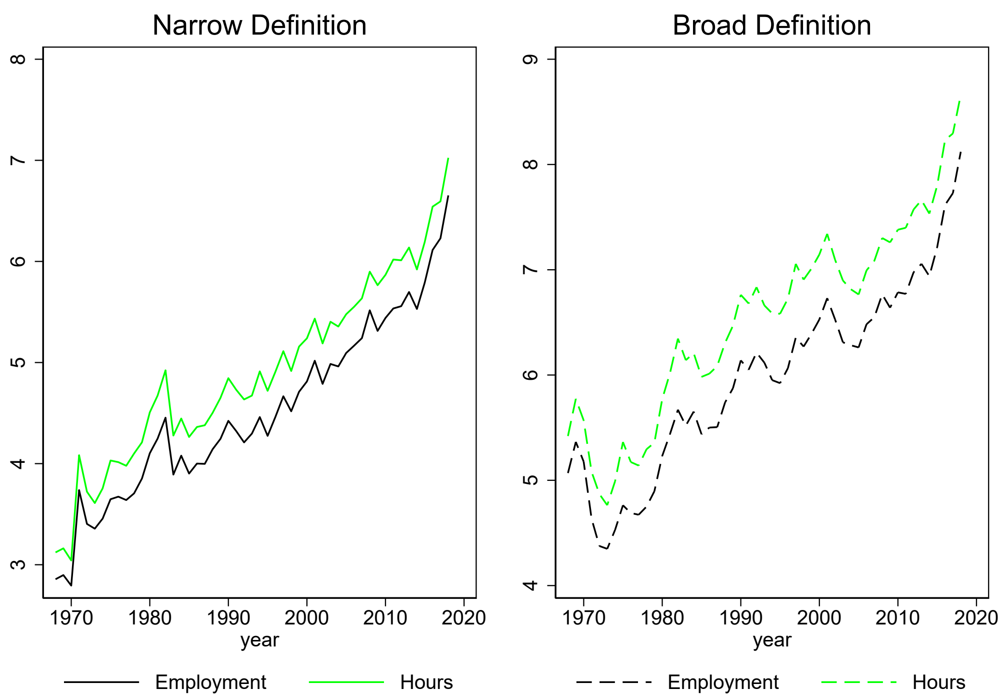
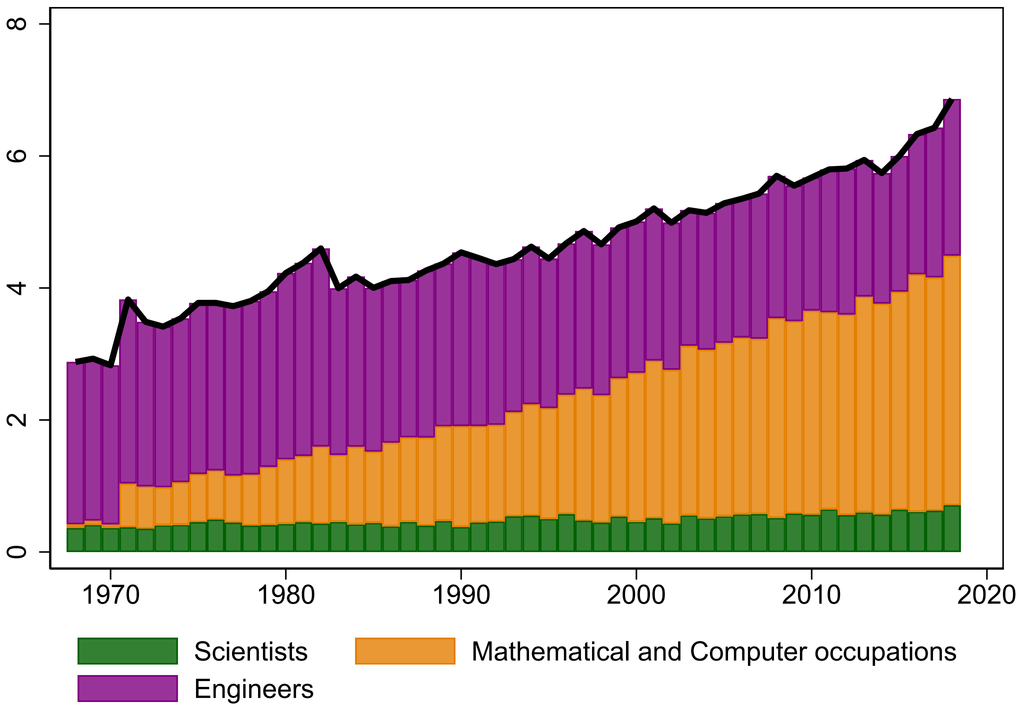
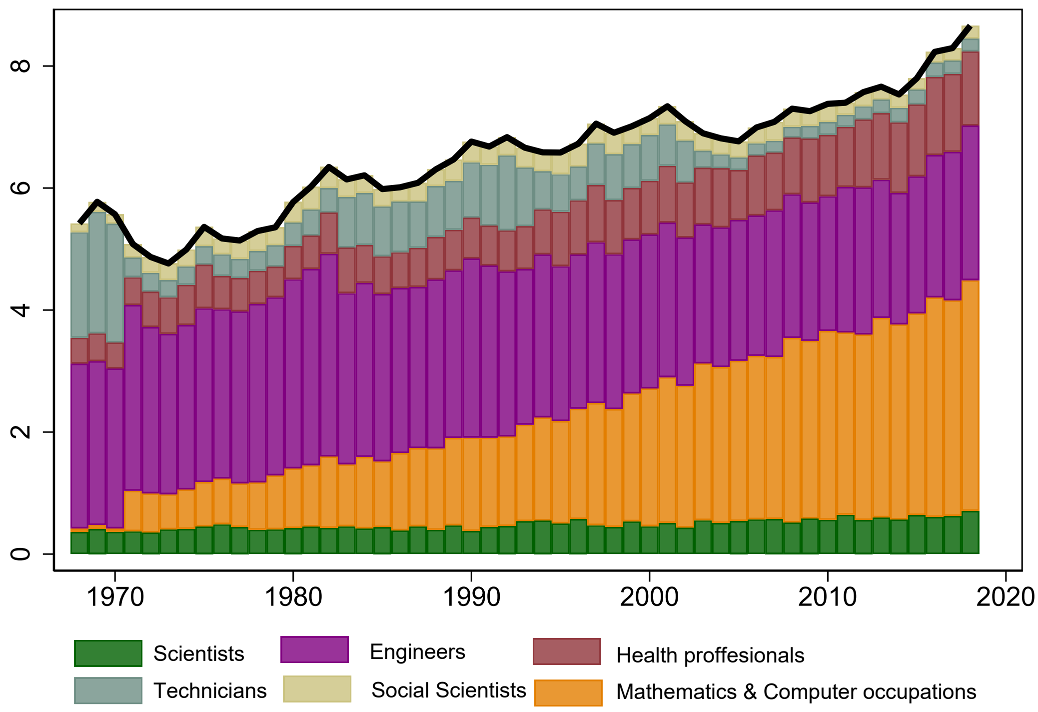


Figure A.18: Share of the R&D related occupations in aggregated hours (in %) – baseline definition



Notes: Narrower occupational definition of R&D occupations.

FIGURE A.19: Share of the R&D related occupations in aggregated hours (in %) – broader definition



Notes: Broader occupational definition of R&D occupations.

Figure A.20: Share of the R&D related occupation groups in total US employment and hours

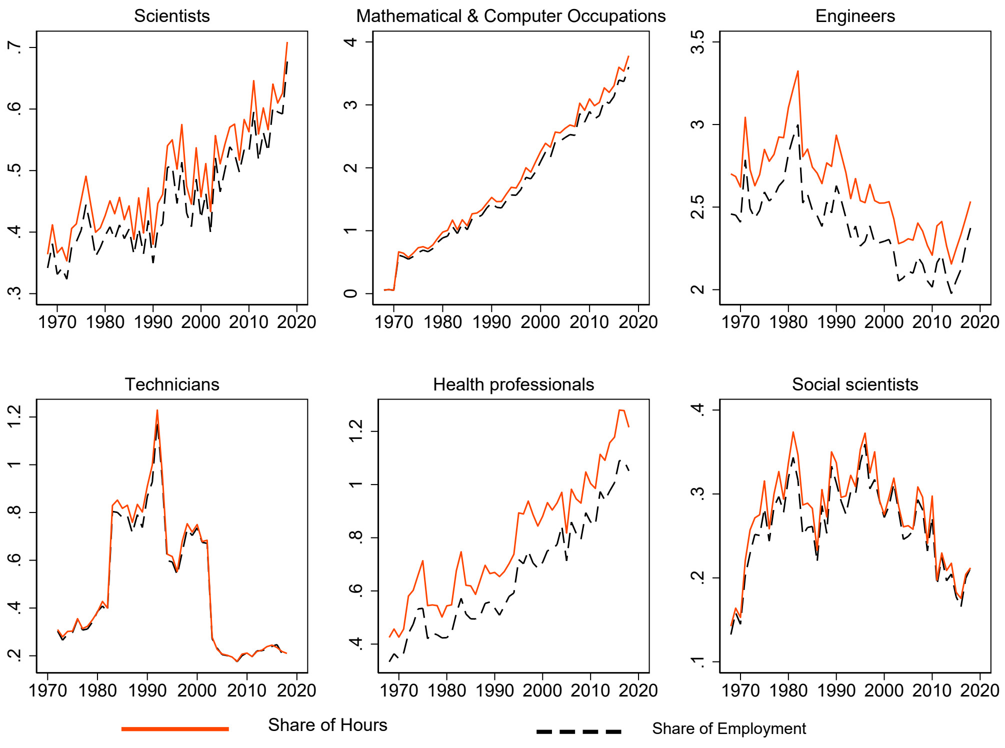
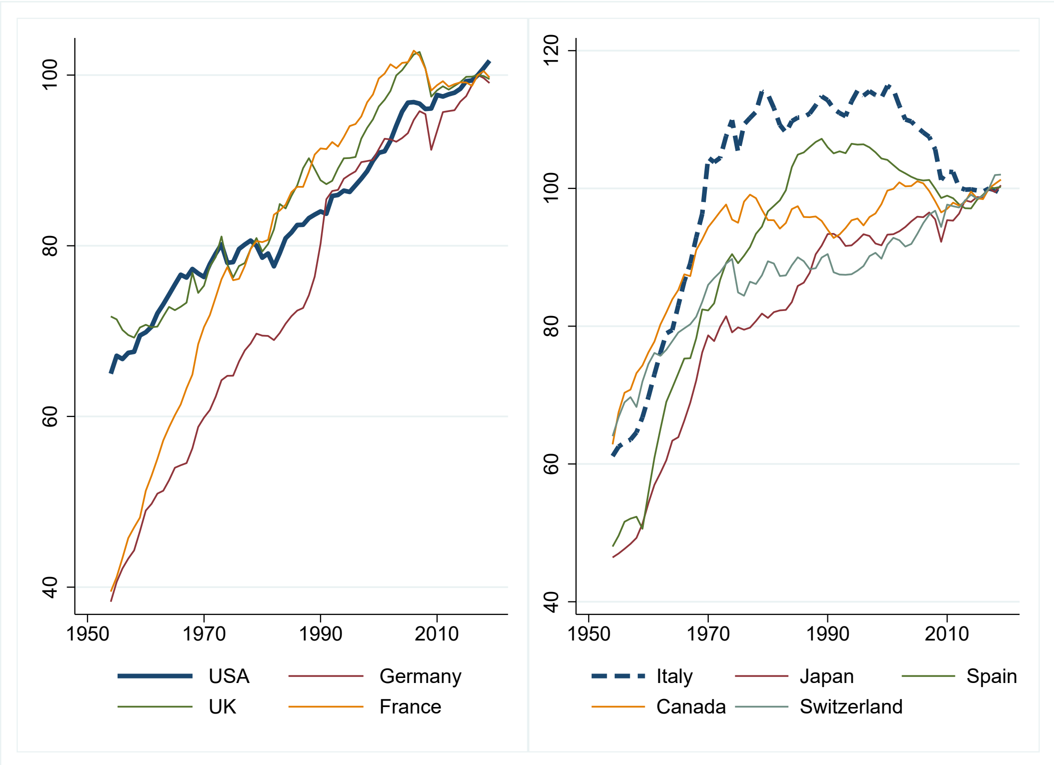


Figure A.21: Total Factor Productivity, 1954-2019: Constant National Prices, 2017=100

(G7 + Switzerland + Spain)



Source: Feenstra, Robert C., Robert Inklaar and Marcel P. Timmer (2015) "The Next Generation of the Penn World Table." *American Economic Review*, 105(10), 3150-3182. Indexed: 2017 = 100. The series were downloaded from FRED with the mnemonic `RTFPNAXXA632NRUG` where XX denotes the relevant country code, e.g. US.

B Unit Productivity Forms

B.1 Box Cox

Following Klump, McAdam and Willman (2007), we model time-varying technological progress terms using a Box-Cox transformation (specified in normalized form). This allows deterministic but time-varying technological progress terms where curvature or decay terms could be uncovered from the data in economically meaningful ways.

$$\Gamma_t^i = e^{g_t^i} \tag{B.1}$$

$$g_t^i = \gamma_i \times \left[\frac{\tilde{t}^{\lambda_i} - 1}{\lambda_i}\right] \times t_z, \tag{B.2}$$

The growth rate of technical change associated to factor i is therefore given by,

$$\gamma_t^i = \frac{dg_t^i}{dt} = \gamma_i \times \tilde{t}^{(\lambda_i - 1)} \tag{B.3}$$

where $\tilde{t} = t/t_z$ and curvature parameter $\lambda \in \mathbb{R}$ determines the shape of the technical progress function. Note, the re-scaling of γ and t by the fixed point value t_z in (B.2) allows us to interpret γ_i directly as the rates of i factor-specific unit productivity improvements at the fixed-point period ($t = t_z$).

For $\lambda = 1$, the technical progress functions are exponential; otherwise they are subexponential / superlinear ($\lambda \in (0,1)$), linear ($\lambda = 0$) or hyperbolic functions of time ($\lambda < 0$). Finally, if $\lambda > 1$, then technical progress witnesses a superexponential explosion.

Asymptotically, the function (B.2) would behave as follows in levels and growth rates, respectively:

$$\begin{aligned} &\lim_{t\to\infty} g_t^i \to \infty \qquad &\lambda_i \geq 0 \\ &\lim_{t\to\infty} g_t^i = -\tfrac{\gamma_i}{\lambda_i} t_z \quad &\lambda_i < 0 \end{aligned} \tag{B.4}$$

$$\gamma_t^i = \frac{dg_t^i}{dt} = \gamma_i \times \tilde{t}^{\lambda_i - 1} \Rightarrow \begin{cases} \infty \text{ (as } t \to \infty) & \lambda_i > 1 \\ \gamma_i & \lambda_i = 1 \\ 0 & \lambda_i < 1 \end{cases} \tag{B.5}$$

B.2 Fourier

Our second case uses a trigonometric trajectory which is a special case of a Fourier expansion:[6]

$$\log \Gamma_t^j = \exp\left[(t - t_z)\left(\gamma_j + \kappa_j^{sin} \sin\left(\frac{2\pi\kappa t}{T}\right) + \kappa_j^{cos} \cos\left(\frac{2\pi\kappa t}{T}\right)\right)\right],$$

where $\pi = 3.14$ and $j = K, L$. Any possible structural breaks will be captured by the $\kappa \in \mathbb{R}$ parameters, where $\kappa_j^{sin} = \kappa_j^{cos}$ retrieves the simple linear case. As regards the appropriate number of frequencies $\kappa \geq 1$ to include, we follow Ludlow and Enders (2000) who showed that a single frequency is invariably sufficient to approximate the Fourier expansion in the bulk of empirical applications.[7] Indeed with higher values of κ one might able capture rather low-frequency fluctuations in factor-biased technical change.

[6] See Christopoulos and León-Ledesma (2010) for a discussion of Fourier forms in economics.

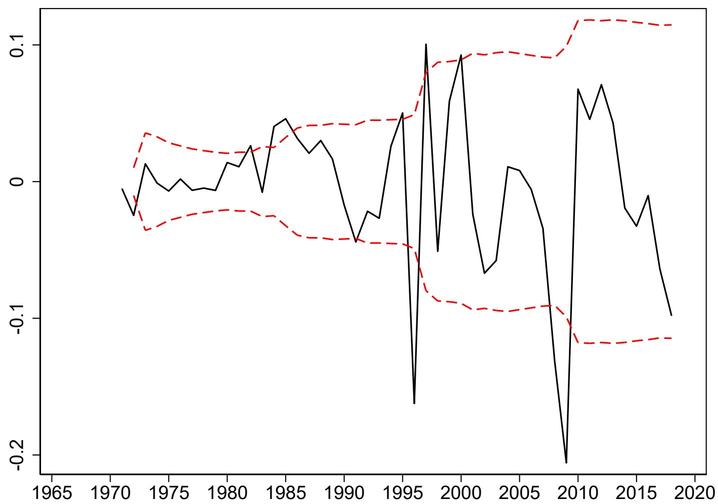
[7] Moreover, according to Becker et al. (2004) the Fourier expansion has more power to detect several smooth breaks of unknown form in the intercept than, say, the Bai and Perron (1998, 2003) multi-break tests.

C Stability Analysis

In this section, we perform a simple exploratory analysis of structural breaks in the patent growth series. We model it as a simple AR(1) process, which should capture reasonably well its time series path.

We then estimate that form recursively over time and plot the persistence parameters and its associated standard errors. Values of the residuals outside of the standard error bands are indicative of structural breaks, large movements or cyclical swings. Looking at Figure C.1, we can see some suggestive evidence for structural instability for the periods around the mid 1980s, mid 1990s, and around the Great recession.

Figure C.1: Recursive Residual Stability Analysis for Patents



Notes: In this figure we derive the recursive residual (in black) plus/minus their two standard errors (in red dash) for an $AR(1)$ regression in $\Delta\widetilde{A}_t$. The interpretation of those recursive exercise being that residuals outside the standard error bands suggest instability in the parameters of the equation.

D Robustness

As a robustness check we consider alternative empirical measures of R&D capital and R&D labor. The additional estimates for fixed and estimated η are presented in Table D.1. Qualitatively, all results replicate our previous preferred findings: the elasticity of substitution ξ is below unity; the average growth rate of R&D labor productivity ranges from $0.1\% - 2.6\%$ per annum; there is evidence in favor of presence of a cyclical dynamic / multiple structural breaks in R&D capital productivity.

Table D.1: Robustness

	(1),(2) Baseline[†]		(3),(4) Narrow R&D Labor		(5),(6) Merged R&D Labor		(7),(8) Private R&D capital	
ξ	0.793^{***}	0.760^{***}	0.836^{***}	0.810^{***}	0.687^{***}	0.639^{***}	0.789^{***}	0.815^{***}
	(0.019)	(0.062)	(0.034)	(0.044)	(0.025)	(0.057)	(0.017)	(0.085)
$\gamma_{\mathcal{K}}$	-0.016^{***}	-0.013^{***}	-0.007	-0.006	-0.005^{***}	-0.005^{***}	-0.062^{***}	-0.074^{*}
	(0.003)	(0.004)	(0.007)	(0.005)	(0.002)	(0.001)	(0.005)	(0.042)
$\gamma_{\mathcal{R}}$	0.011^{***}	0.011^{***}	0.001	0.002	0.003^{**}	0.002^{**}	0.026^{***}	0.026^{***}
	(0.001)	(0.001)	(0.003)	(0.003)	(0.001)	(0.001)	(0.002)	(0.002)
$\gamma_{\mathcal{K}}^{sin}$	0.556^{***}	0.438^{***}	0.650^{***}	0.510^{***}	0.401^{***}	0.329^{***}	0.537^{***}	0.638^{*}
	(0.045)	(0.137)	(0.041)	(0.134)	(0.039)	(0.064)	(0.047)	(0.356)
$\gamma_{\mathcal{K}}^{cos}$	-0.427^{***}	-0.337^{***}	-0.406^{***}	-0.345^{***}	-0.398^{***}	-0.316^{***}	-0.548^{***}	-0.664^{*}
	(0.028)	(0.109)	(0.03)	(0.086)	(0.033)	(0.07)	(0.032)	(0.365)
η		0.418^{***}		0.405^{***}		0.426^{***}		0.280^{*}
		(0.121)		(0.100)		(0.104)		(0.149)
R&D Labor Productivity	Exp.	Exp.	Exp.	Exp.	Exp.	Exp.	Exp.	Exp.
R&D Capital Productivity	**F**	**F**	**F**	**F**	**F**	**F**	**F**	**F**
η	fixed	estimated	fixed	estimated	fixed	estimated	fixed	estimated
$\xi = 1$	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.031]
$\gamma_{\mathcal{R}} = \gamma_{\mathcal{K}}$	[0.000]	[0.000]	[0.388]	[0.332]	[0.000]	[0.000]	[0.000]	[0.017]
$\kappa_{cos}^{\mathcal{K}} = \kappa_{sin}^{\mathcal{K}} = 0$	[0.000]	[0.006]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.192]
res_3	[0.004]	[0.006]	[0.008]	[0.011]	[0.000]	[0.000]	[0.008]	[0.009]
res_4	[0.001]	[0.001]	[0.000]	[0.000]	[0.000]	[0.000]	[0.002]	[0.003]
ll	133.2	134.2	115.6	116.2	134.2	134.7	127.9	127.4
bic	-239	-237	-203.7	-201	-241.1	-238.2	-228.2	-223.4

Notes: The numbers in parentheses are robust standard errors, where the significance stars are to be read as $^{*} < 0.1,^{**} < 0.05,^{***} < 0.01$. Probability values are in brackets. Symbols **B** and **F** denote the Box-Cox and Fourier forms respectively, and "Exp." denotes exponential (i.e., $\lambda = 1$). In the second section of the table, we present Wald tests of various parameter restrictions. In the (final) diagnostic section of the table, the first two rows refer to ADF test of the unit root null associated with the errors in equations (4) and the log form of (3) and the p-values are obtained by bootstrapping distribution. Finally terms ll and bic denote, respectively, the Log Likelihood, and the Bayesian Information Criterion.

†: the baseline columns replicate the final two columns of Table 2.