Boysen-Hogrefe, Jens

Article — Published Version

Is the supermultiplier currently nil? - A replication study of Deleidi and Mazzucato (2021)

Research Policy

Provided in Cooperation with:
Kiel Institute for the World Economy – Leibniz Center for Research on Global Economic Challenges

Suggested Citation: Boysen-Hogrefe, Jens (2025) : Is the supermultiplier currently nil? - A replication study of Deleidi and Mazzucato (2021), Research Policy, ISSN 1873-7625, Elsevier, Amsterdam, Vol. 54, Iss. 3, pp. 1-12, https://doi.org/10.1016/j.respol.2025.105176

This Version is available at:
https://hdl.handle.net/10419/318260

Research Policy 54 (2025) 105176

Contents lists available at ScienceDirect

Research Policy

journal homepage: www.elsevier.com/locate/respol

Short communication

Is the supermultiplier currently nil? - A replication study of Deleidi and Mazzucato (2021)☆

Jens Boysen-Hogrefe

Kiel Institute for the World Economy, Germany

ARTICLE INFO

Keywords
Replication study
Mission oriented innovation policies
Fiscal multiplier
Sraffian supermultiplier

ABSTRACT

Analyzing US macro data via a structural vector-autoregressive model (SVAR), Deleidi and Mazzucato (2021) find strong positive spillovers from mission-oriented government spending on private research and development, as well as on overall economic activity ("crowding in"). Deleidi and Mazzucato apply the SVAR to first-differenced data despite the possibility of cointegration. The replication shows that the result hinges on the transformation of the data and the choice of the sample period. The time variation of the estimation results is substantial. When estimating the model with data starting after 1985, the results point to a temporary "crowding out" of private research and development spending.

1. Introduction

In their article, Deleidi and Mazzucato (2021) aim to empirically test the proposition that mission-oriented government spending can greatly benefit the overall economy. They utilize an econometric model, specifically a structural vector autoregression (SVAR), to analyze data from the United States. The results of their study show substantial positive spillovers of mission-oriented government spending on private research and development and overall economic dynamism ("crowding in"). The authors refer to this phenomenon as a "supermultiplier." However, upon closer examination of the paper, concerns arise regarding the robustness of their findings. Consequently, I intend to replicate Deleidi and Mazzucato's work while making plausible adjustments to the initial article's specifications. The replication study reveals that the results of Deleidi and Mazzucato are not robust against variations in the sample's time span and the assumptions about the time series properties. In an early part of the sample, the SVAR shows counterintuitive, strongly negative effects of mission-oriented government spending on GDP. For the periods between 1966 and 1984, the "supermultiplier" is even larger than in the original paper, and for data from 1985 onwards, the results suggest a temporary "crowding out" of private research and development spending. While the time variation of the impact of government spending on research and development is noted in Kantor and Whalley (2023), who apply an alternative methodological approach, the time variation found here is rather extreme. This may indicate that, in general, the simple SVAR approach applied by Deleidi and Mazzucato faces methodological limitations and may be less well-suited to reach reliable quantitative results.

The replication study is organized as follows. Section 2 describes the empirical model applied in Deleidi and Mazzucato (2021), and Section 3 provides the corresponding results for the data vintage used in this replication study. Section 4 highlights possible causes that undermine the robustness of the results in the original paper. Section 5 presents the results for modified specifications. These modifications pertain to the transformation of the data, specifically how the variables are deflated and whether first-differenced data is used. To address possible structural instabilities, Section 6 provides results for subsamples identified via the breakpoint detection procedure proposed by Qu and Perron (2007). Finally, Section 7 concludes.

2. The empirical model

In their analysis, Deleidi and Mazzucato employ a structural vector autoregressive model (SVAR). The model is applied to four time series: Gross Domestic Product (1: GDP), Private Expenditures on Research and Development (2: R&D), Government Expenditures on Military Research and Development (3: G_MO), and Other Government Expenditures on Consumption and Investment (4: G_R). To identify the shocks, they utilize the Cholesky decomposition, which breaks down the symmetric variance-covariance matrix of the error terms into the product of a lower

☆ I am very grateful for extremely valuable and constructive comments on former versions of this replication study from an anonymous referee and the editor. All remaining errors are mine.

E-mail address: Jens.hogrefe@ifw.kiel.de.

https://doi.org/10.1016/j.respol.2025.105176
Received 6 March 2023; Received in revised form 13 June 2024; Accepted 12 January 2025
Available online 17 January 2025

triangular matrix. The shape of this matrix, or the associated zero restrictions, means that the direction of the effect of the error terms is specified, thus giving the covariance a causal interpretation. In the arrangement chosen by Deleidi and Mazzucato, the errors in the equation for G_MO affect all others, while they are unaffected by the other error terms. Second in the series, Deleidi and Mazzucato place G_R, then R&D, and finally GDP, whose error terms can thus be contemporaneously affected by all others, but contemporaneously have no effect on the other error terms. Deleidi and Mazzucato employ nominal data, all divided by the GDP deflator. In the original sense of the term, only GDP is price-adjusted. The other variables are adjusted by a deflator other than their own. Moreover, the model incorporates logarithmized data in the form of first differences.

3. Estimation results for the original specification

The impulse responses to the one-unit shocks in Fig. 1 stem from the estimation of the original specification by Deleidi and Mazzucato. The one-unit shock specification with logarithmized data allows the interpretation of the responses as elasticities. In contrast to the original article, a more recent data vintage is employed. Results can be directly compared to those presented in Annex B provided in Deleidi and Mazzucato (2021). Since deviations are only minor, it can be concluded that the results are not severely affected by real-time data issues. In the following, the results based on the more recent vintage will be used as the baseline.

Deleidi and Mazzucato's results suggest a large increase in private R&D spending due to a boost from G_MO (see Fig. 1). There is also a positive effect on overall economic performance (GDP). Although the statistical significance of this finding may not be particularly strong, the point estimates indicate substantial economic significance. It is important to keep in mind that 1 % in G_MO represents a relatively small amount of money compared to 1 % of economic output. Given the huge differences in size, even modest elasticities suggest that G_MO could stimulate a multiple of private R&D and especially of GDP. In their article, Deleidi and Mazzucato report that one additional real dollar of G_MO induces 40 to 60 real dollars of GDP after a few quarters, likely using ratios at the current edge. This is why Deleidi and Mazzucato refer to it as a "supermultiplier." However, for the sake of ensuring comparability across different time periods, this replication study only includes responses to one-unit shocks.

4. Criticisms

Instead of using price-adjusted variables, Deleidi and Mazzucato deflate the nominal variables by the GDP deflator. This means that price developments occurring in R&D that deviate from aggregate price developments are considered "price-adjusted" or real in Deleidi and Mazzucato's approach. Considering the identification approach employed in the SVAR, this choice might impact the results. Joint price movements in both research and development variables would be attributed to the government. Since originally price-adjusted data are available for all variables in Deleidi and Mazzucato's empirical model, these data are used to check whether the way the variables are deflated matters. However, the calculation of government spending on consumption and investment adjusted for governmental research and development spending is more involved to account for the properties of chain indices. Details can be found in Annex 1.

Another criticism of the modeling in Deleidi and Mazzucato (2021) is the use of first-differenced data. This implicitly assumes that there are no cointegration relationships between the variables. The issue of cointegration can be easily remedied by estimating the SVAR with level data, as suggested by Sims et al. (1990). The estimation results in response to the first two aspects are discussed in Section 5.

The identifying assumptions that Deleidi and Mazzucato make by applying the Cholesky decomposition are quite strict. An extensive literature on alternative identification methods has developed over the past decades, compare e.g. Kilian and Lütkepohl (2017). There is no extensive discussion of the identification strategy in the paper, and this aspect will not be the primary focus of the replication study. All replication results will also be based on the Cholesky decomposition. Nevertheless, in Annex 2, results for an alternative ordering of the data are presented to provide some insight into the potential relevance of this issue.

Additionally, the estimation sample encompasses quite heterogeneous time periods. In the 1950s, the government spent more money on military research than the private sector as a whole spent on R&D. It is also important to acknowledge that the early 1950s were influenced by the Korean War, which may have impacted the economic dynamics during that period. Furthermore, private spending on R&D as a share of

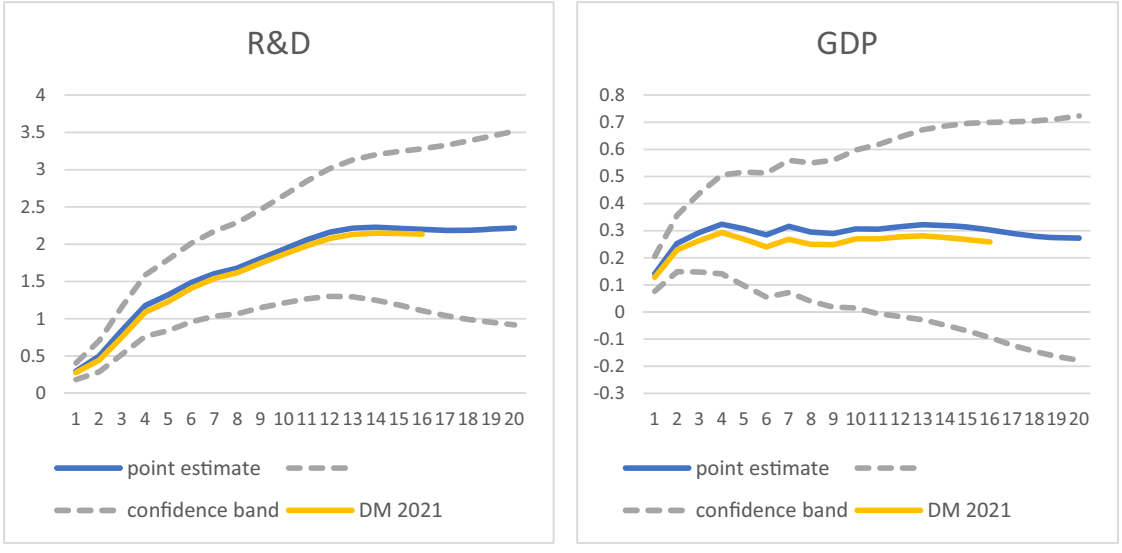

Fig. 1. Impulse response sequence for the specification proposed by Deleidi and Mazzucato.
Quarterly response to a one-unit shock of the G_MO. Confidence bands correspond to a confidence level of 68 % (one standard deviation). Data sample: 1947Q1 through 2018Q4. DM 2021 corresponds to the values presented in Deleidi and Mazzucato (2021), Annex B.
Source: Fred Economic Data St. Louis FED, Deleidi and Mazzucato (2021), own calculations.

GDP has trended upward over the years (see Fig. 2). Additionally, there is literature indicating a structural break in U.S. macroeconomic data around the year 1984, suggesting that the economic landscape and relationships between variables may have experienced significant changes at that time. This debate took place under the heading of the "Great Moderation," as discussed by Stock and Watson (2002). For this reason, methods to time structural breaks in multivariate time series models related to the approach in Qu and Perron (2007) are applied. The multiplier of G_MO varies substantially between the periods identified via the structural break procedure. Section 6 shows the corresponding results and discussion.

5. Estimation results for modified specifications

In this section and the subsequent section, various Structural Vector Autoregressions (SVARs) are estimated to empirically address the criticism concerning the baseline results and assess their robustness. Certain elements of the baseline specification used by Deleidi and Mazzucato are retained in these estimations. Specifically, each SVAR includes the same four variables mentioned earlier. Furthermore, all modified specifications, as discussed in the main text, are estimated using a lag order of 5. In Annex 3, the issue of lag length is discussed in some detail. In this section, the specifications are modified with respect to the way the variables are deflated and with respect to the transformation of the time series, i.e., results of the baseline specification are compared to results from SVARs in levels.

When the SVAR is estimated using price-adjusted data, it is observed that the response of R&D remains nearly unchanged compared to the baseline. As for the GDP response, the point estimate is slightly lower compared to the initial results (Fig. 3). Therefore, it can be concluded that the specific method of deflating the data used by Deleidi and Mazzucato does not appear to significantly affect the empirical findings.

Regarding the choice to estimate the SVAR in first differences, first, the Johansen test is performed for both the data deflated by the GDP deflator and the real data. Here, it should be stressed that the estimation of an SVAR with first-differenced data only provides valid results when cointegration of the level data can be ruled out. The results of the cointegration tests indicate at least one cointegration relation (Annex 4). This also holds true for various subsamples. Only for the maximum eigenvalue test with data after 1984 is the null of no cointegration not rejected at the 1 % level.

When the SVAR is estimated in levels, results change substantially in both cases, with variables deflated as in Deleidi and Mazzucato as well as with price-adjusted data (Fig. 4). The elasticities for private R&D are still sizeable but lower compared with the baseline, and the elasticities of GDP drop to zero after a few periods. At least qualitatively, the latter finding corresponds to the results in Fieldhouse and Mertens (2023). According to their research, there is no long-lasting impact of military R&D spending on the business sector's total factor productivity.

Since the results between level data and first differenced data vary and since there is evidence in favor of cointegration via the cointegration tests, I stick to level data in the following analysis of the subsamples.

6. The impact of time variation

The sample covers a rather long period of time with different economic and political conditions. It might be assumed that the link between G_MO and R&D or GDP varies over time. To analyze this aspect, a break point estimation procedure for multivariate time series models is chosen related to the one in Qu and Perron (2007). Details about this approach are given in Annex 5. Allowing for two breaks, the likelihood is maximized with breaks in the first quarter of 1966 and 1985. The latter break may correspond to the notion of the start of the "Great Moderation," which the literature assumes for 1984, as discussed in Stock and Watson (2002).

6.1. Early sample 1948–1966

In the first subsample identified according to the procedure of Qu and Perron (2007), the positive spillover on R&D is only temporary (Fig. 5). After two and a half to three years, the response drops to zero. The response of GDP is extremely negative. This result seems rather implausible. It's difficult to imagine any channel that would translate additional government spending on R&D into a severe decline in GDP that is a multitude of the means spent before. Here, I'd like to stress that

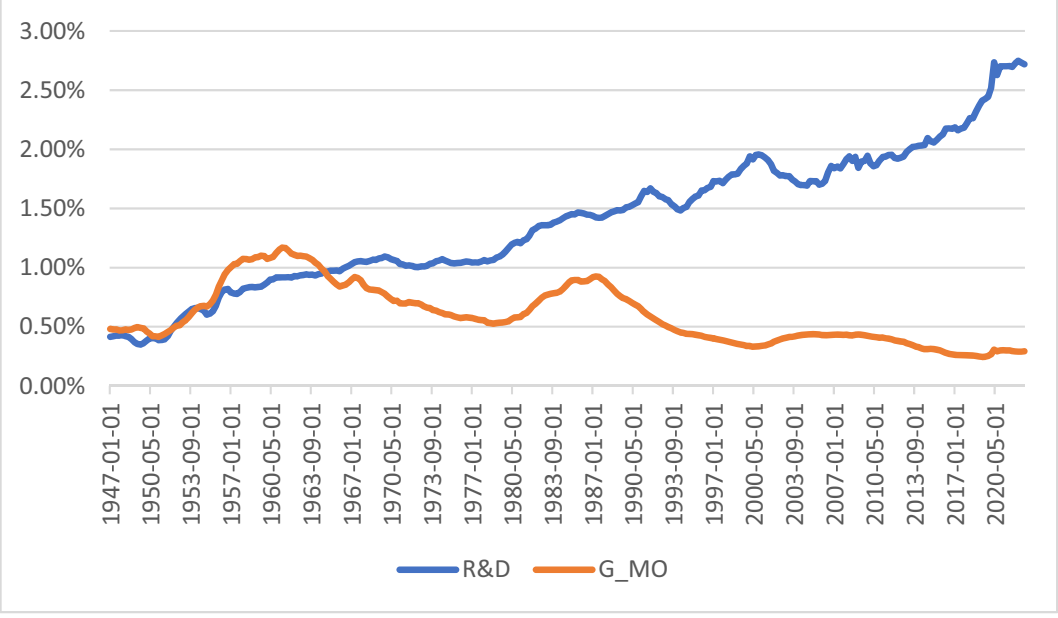

Fig. 2. Share of research and development expenditure in gross domestic product in %.
Nominal values.
Source: Fred Economic Data St. Louis FED, own calculations.

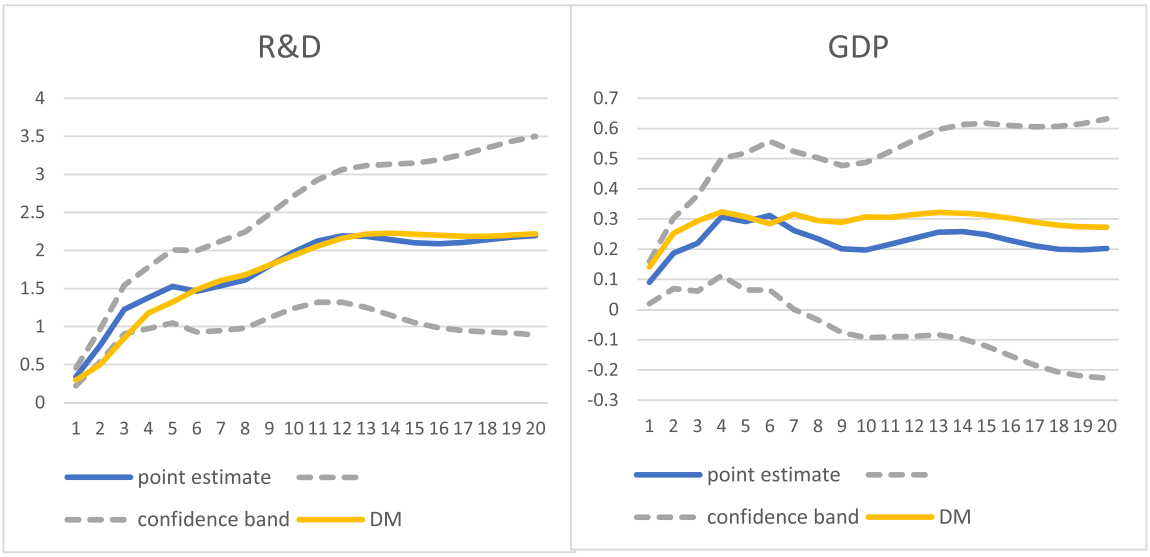

Fig. 3. Impulse-response sequence for price-adjusted data (first differences).
Quarterly response to a one-unit shock of the G_MO. Confidence bands correspond to a confidence level of 68 % (one standard deviation). Data sample: 1947Q1 through 2018Q4. DM refers to the specification of the SVAR in Deleidi and Mazzucato (2021) estimated with the recent data vintage.
Source: Fred Economic Data St. Louis FED, own calculations.

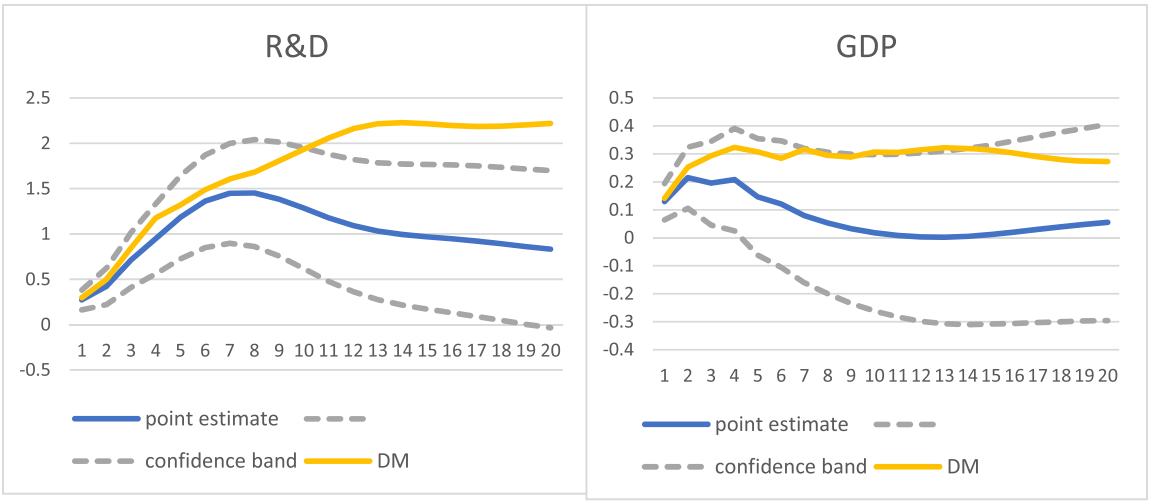

Fig. 4. Impulse-response sequence for level data.
Quarterly response to a one-unit shock of the G_MO. Confidence bands correspond to a confidence level of 68 % (one standard deviation). Data sample: 1947Q1 through 2018Q4. Price adjustment via GDP deflator. DM refers to the specification of the SVAR in Deleidi and Mazzucato (2021) estimated with the recent data vintage.
Source: Fred Economic Data St. Louis FED, own calculations.

the SVAR method has its methodological shortcomings. The VAR is relatively small, and omitted variables may matter. The sample covers the period after World War II and the Korean War, episodes where joint factors may have driven G_MO and GDP in opposing directions. Overall, the rather small sample covers very volatile times, partly reflected in wide confidence bands.

6.2. *The years 1967 through 1984*

In contrast, the results for the subsample between 1967 and 1984 are particularly strong (Fig. 6). The empirical relation between G_MO and R&D, as well as with GDP, exceeds the baseline results. Even when the SVAR is estimated using an alternative ordering of the data, the results remain noteworthy, especially in the case of R&D, as shown in Annex 3. The very strong response of GDP is even stronger than in the original paper of Deleidi and Mazzucato. During the years 1967 through 1984, a response of 1 percentage point in GDP to a change of 1 % in G_MO would imply a multiplier above 100. This is substantially higher than the positive effects found in other studies (compare Jones, 2021, Table 1) as well. It should be stressed that the confidence bands around the responses of GDP are relatively wide, and that the SVAR faces the methodological limitations mentioned before. At the 95 % confidence level, the null is included in almost all versions of the SVAR estimated in this replication study. For R&D, the estimation is more precise. Even at the high confidence level, significant results can be found.

Qualitatively, the evidence for positive GDP responses between the years 1967 and 1984 is in line with the findings in Kantor and Whalley (2023), who use cross-sectional data from various years spanning between 1947 and 1992. The response of value added in manufacturing to investments of space programs is particularly strong in the 60s, 70s, and

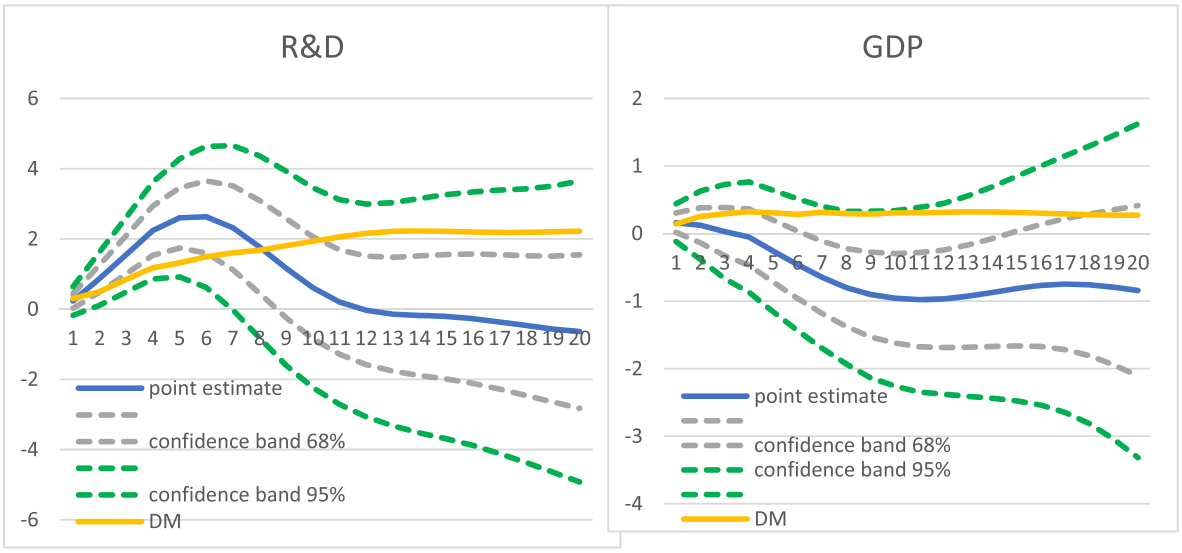

Fig. 5. Impulse-response sequence for level data spanning between 1948 and 1966.
Quarterly response to a one-unit shock of the G_MO. Confidence bands correspond to a confidence level of 68 % (one standard deviation). Data sample: 1947Q1 through 1966Q4.DM refers to the specification of the SVAR in Deleidi and Mazzucato (2021) estimated with the recent data vintage.
Source: Fred Economic Data St. Louis FED, own calculations.

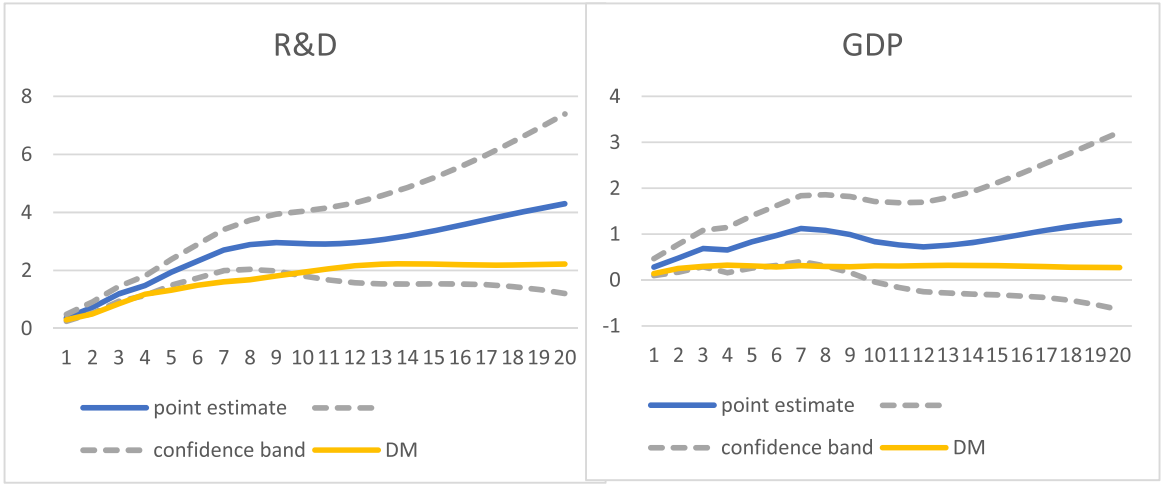

Fig. 6. Impulse-response sequence for level data spanning between 1967 and 1984.
Quarterly response to a one-unit shock of the G_MO. Confidence bands correspond to a confidence level of 68 % (one standard deviation). Data sample: 1967Q1 through 1984Q4. DM refers to the specification of the SVAR in Deleidi and Mazzucato (2021) estimated with the recent data vintage.
Source: Fred Economic Data St. Louis FED, own calculations.

80s. Furthermore, the results in Kantor and Whalley (2023) already point to time variation in the responses to G_MO as well.

6.3. *The years 1985 through 2018*

Accounting for two structural breaks, the second one aligns with the first quarter of 1985, coinciding with the onset of the "Great Moderation." Recent subset analyses reveal a transient "crowding out" impact on R&D, diminishing substantially after a few quarters. In contrast, GDP registers a positive response in initial quarters, gradually dipping below zero thereafter. Long-term effects, observable after 3 to 4 years, demonstrate minimal impact on both R&D and GDP (Fig. 7).

The observed "crowding out" phenomenon in private R&D spending, when focusing solely on post-1985 data, diverges from the outcomes in the overall dataset and the previous subset. This difference may suggest a shift in the relation between the private sector and the government sector. It is possible that R&D resources, such as graduate students, are now less inclined toward government-related tasks, and that the private sector and the government sector compete for these resources nowadays. Public funding for R&D in the military sector may have provided a temporary boost to the US economy, as highlighted by Gross and Sampat (2023), who underscore the role of public spending in initiating R&D activities from relatively low levels. Nowadays, this may have changed. However, the temporary "crowding out" effect on private R&D, while conflicting with the initial findings of Deleidi and Mazzucato, also contrasts with the recent outcomes in Moretti et al. (2021).

Moretti et al. (2021) present evidence from a recent sample using OECD country panel data and French firm data. The timeframe in Moretti et al. (2021) matches the one considered in this last subset here. Discrepancies may arise from differences in the identification methodology. Another conjecture for these mixed findings is that variations in the results for the US may not be indicative of other countries, possibly

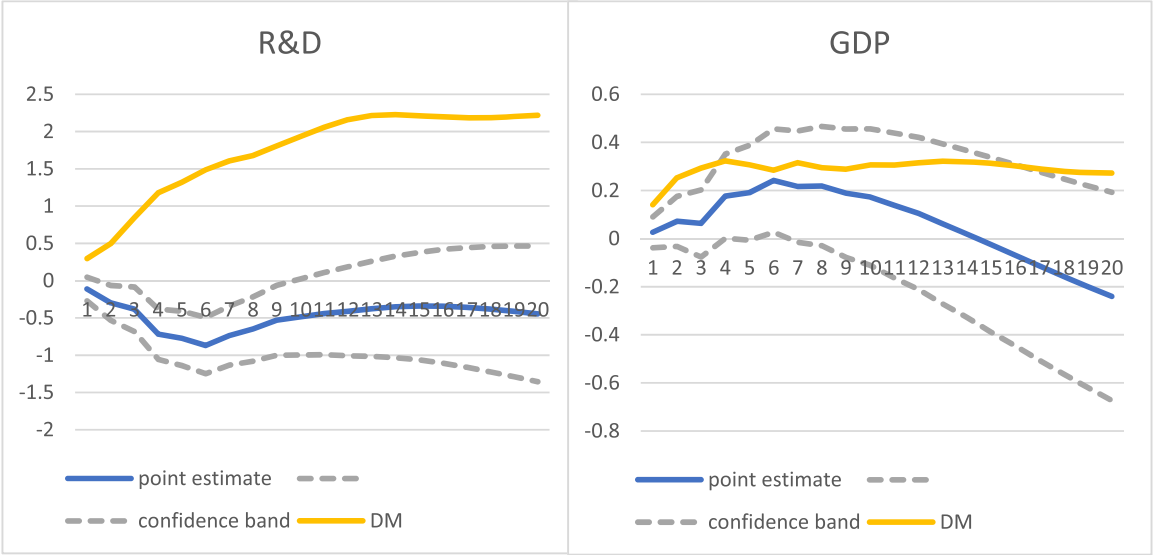

Fig. 7. Impulse-response sequence for level data since 1985.
Quarterly response to a one-unit shock of the G_MO. Confidence bands correspond to a confidence level of 68 % (one standard deviation). Data sample: 1984Q1 through 2018Q4. DM refers to the specification of the SVAR in Deleidi and Mazzucato (2021) estimated with the recent data vintage.
Source: Fred Economic Data St. Louis FED, own calculations.

due to varying stages of development. Since Moretti et al. (2021) do not report poolability tests for the OECD panel regression, this possibility cannot be discounted.

However, for the "crowding out" finding present in data after 1984, the same caveats apply as for the other subsamples. The methodological limitations of a small SVAR are also present for the results after 1984. And finally, to interpret the temporary crowding-out effect identified here, it is also important to recognize that this observation pertains specifically to a particular type of R&D expenditure. Notably, Fieldhouse and Mertens find stronger effects for non-defense compared to defense R&D on total factor productivity. Azoulay et al. (2019) highlight positive outcomes from public research funding in pharmaceutical and biotechnology firms, covering the period from 1980 to 2005. Additionally, Akcigit et al. (2021) emphasize the importance of funding basic research. Thus, even if the results of the SVAR analysis are trustworthy, the temporary crowding-out effect identified here does not inherently negate the government's capacity to stimulate R&D in the US, nowadays.

7. Conclusion

The findings of Deleidi and Mazzucato appear to be influenced by the data transformation, neglecting the potential for cointegration. Additionally, results are not stable over time. Support for the "super-multiplier" is notably present in a subset spanning from 1966 to 1985, while data from 1985 onwards indicates a temporary "crowding-out" effect in private research and development spending. However, uncertainty of the estimates is rather large, and the time variation of the point estimates is extreme. Therefore, it is crucial to recognize that the results in Deleidi and Mazzucato, but also in this replication study, are subject to methodological limitations and should not be overemphasized. Nevertheless, the outcomes of the replication study may signal temporal variations in the relevance of public spending on military R&D for overall R&D and the broader economy. Future research should consider that the impact of public R&D on the economy could exhibit time-varying patterns, contingent upon economic and political conditions.

CRediT authorship contribution statement

Jens Boysen-Hogrefe: Writing – review & editing, Writing – original draft, Investigation, Formal analysis, Conceptualization.

Declaration of competing interest

I declare that I have no conflicts of interest.

Annex 1. Data source and preparation

For the baseline variant, the following variables were downloaded from the Fred Economic Data St. Louis FED database.

- GDP: Gross Domestic Product, Billions of Dollars, Quarterly, Seasonally Adjusted Annual Rate (→GDP)
- GCE: Government Consumption Expenditures and Gross Investment, Billions of Dollars, Quarterly, Seasonally Adjusted Annual Rate (→G_MO + G_R)
- Y006RC1Q027SBEA: Gross Private Domestic Investment: Fixed Investment: Nonresidential: Intellectual Property Products: Research and Development, Billions of Dollars, Quarterly, Seasonally Adjusted Annual Rate (→R&D)
- Y076RC1Q027SBEA: Government Gross Investment: Federal: National Defense: Gross Investment: Intellectual Property Products: Research and Development, Billions of Dollars, Quarterly, Seasonally Adjusted Annual Rate (→G_MO)
- GDPDEF: Gross Domestic Product: Implicit Price Deflator, Index 2012 = 100, Quarterly, Seasonally Adjusted.

First, G_R is calculated by subtracting Y076RC1Q027SBEA from GCE. Then, all four variables are deflated using GDPDEF.
At the same time, price-adjusted chain indices are available for all variables and are used in the alternative estimates.

- GDPC1: Real Gross Domestic Product, Billions of Chained 2012 Dollars, Quarterly, Seasonally Adjusted Annual Rate (→GDP)
- GCEC1: Real Government Consumption Expenditures and Gross Investment, Billions of Chained 2012 Dollars, Quarterly, Seasonally Adjusted Annual Rate (→G_MO + G_R)
- Y006RA3Q086SBEA: Real Gross Private Domestic Investment: Fixed Investment: Nonresidential: Intellectual Property Products: Research and Development (chain-type quantity index), Index 2012 = 100, Quarterly, Seasonally Adjusted (→R&D)
- Y076RA3Q086SBEA: Real Government Gross Investment: Federal: National Defense: Gross Investment: Intellectual Property Products: Research and Development (chain-type quantity index), Index 2012 = 100, Quarterly, Seasonally Adjusted (→G_MO).

To calculate price-adjusted government spending excluding military research and development (G_R) spending, the following calculation steps are performed. The prior quarter changes in GCEC1 and G_MO are calculated. Since GCEC1 is the aggregate, this is the weighted sum of the prior quarter changes in G_R and G_MO. The weights correspond to the nominal values of the respective previous year. Since the previous quarter changes of GCEC1 are available, the equation is solved for the previous quarter changes of G_R. These are thus calculated, and from them, the price-adjusted index can be obtained, which is rebased to 2012.

Table A1.1
Descriptive statistics.

		G_MO	G&R	R&D	Y
Price adjusted data					
Log level	Mean	4.1187	7.4754	3.2512	8.8538
	Std	0.5637	0.4729	1.1665	0.6598
First differences	Mean	0.0062	0.0063	0.0143	0.0077
	Std	0.0207	0.0167	0.0204	0.0093
Adjusted via GDP deflator					
Log level	Mean	3.6861	7.2385	4.4342	8.8538
	Std	0.5097	0.6236	1.1050	0.6598
First differences	Mean	0.0055	0.0080	0.0137	0.0077
	Std	0.0206	0.0170	0.0207	0.0093

Source: Fred Economic Data St. Louis FED, own calculations.

Annex 2. Impact of the ordering of the data (Cholesky decomposition)

The ordering of the data in the SVAR, as determined by the Cholesky decomposition, plays a crucial role in the identification of causal relationships. In both Deleidi and Mazzucato's original study and throughout the replication study, it is assumed that shocks to G_MO have contemporaneous effects on all other variables, but there is no contemporaneous impact of the other error terms on G_MO. This specific ordering of the variables is motivated by the assumption that G_MO is an exogenous policy variable. As a result, according to this ordering, the correlation of the G_MO error term with the other error terms is interpreted as reflecting a causal relationship. To provide insight into the extent to which the results may be influenced by the identifying assumption, the following section presents results based on an alternative ordering of the data. In this alternative ordering, while the ordering of the other variables remains the same, G_MO's position is reversed. In this scenario, all shocks have contemporaneous effects on G_MO, while the G_MO shock does not have a contemporaneous impact on the other error terms.

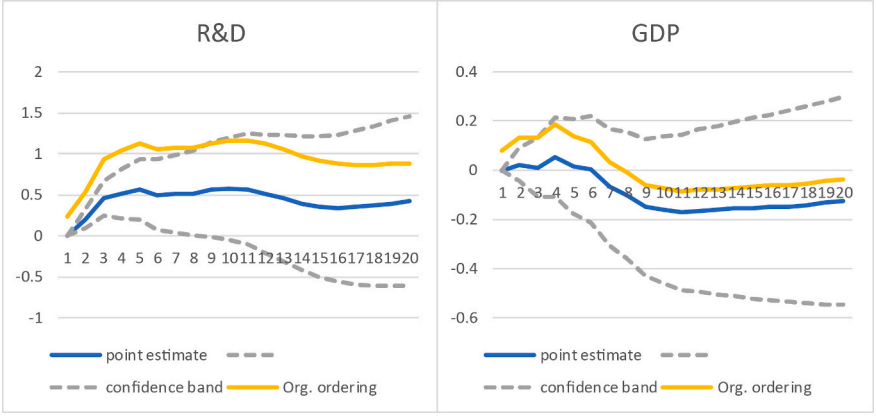

Fig. A.1. Impulse-response sequence with different ordering of the data.
Quarterly response to a one-unit shock of the G_MO. Confidence bands correspond to a confidence level of 68 % (one standard deviation). Data sample: 1947Q1 through 2018Q4. Org. ordering corresponds to the results for the ordering of the data as in Deleidi and Mazzucato (2021).
Source: Fred Economic Data St. Louis FED, own calculations.

For the full sample, there is a downward level shift in the responses of R&D and GDP to G_MO. However, there is still a positive and long-lasting effect for R&D. For the results after 1985, the change in the ordering of the variables has almost no impact.

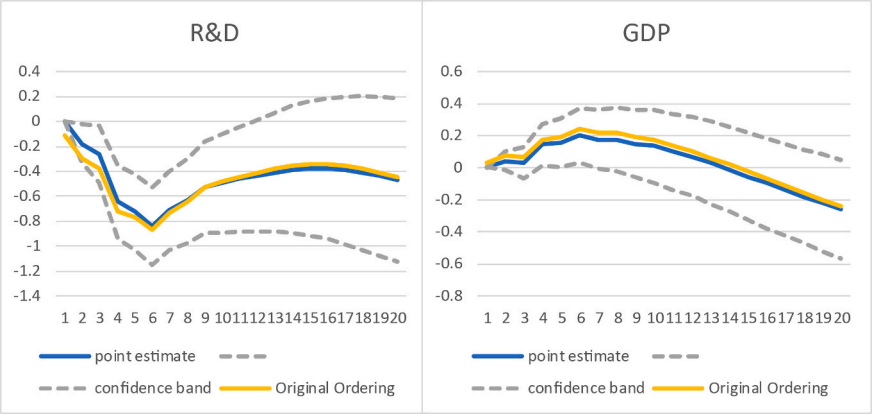

Fig. A.2. Impulse-response sequence for data since 1985 with different ordering of the data.
Quarterly response to a one-unit shock of the G_MO. Confidence bands correspond to a confidence level of 68 % (one standard deviation). Data sample: 1984Q1 through 2018Q4. Org. ordering corresponds to the results for the ordering of the data as in Deleidi and Mazzucato (2021).
Source: Fred Economic Data St. Louis FED, own calculations.

Annex 3. Impact of the lag order

The VARs in the main text are all estimated with 5 lags. To illustrate the impact of the number of lags, some specifications are re-estimated with numbers of lags suggested by information criteria, namely AIC and BIC.

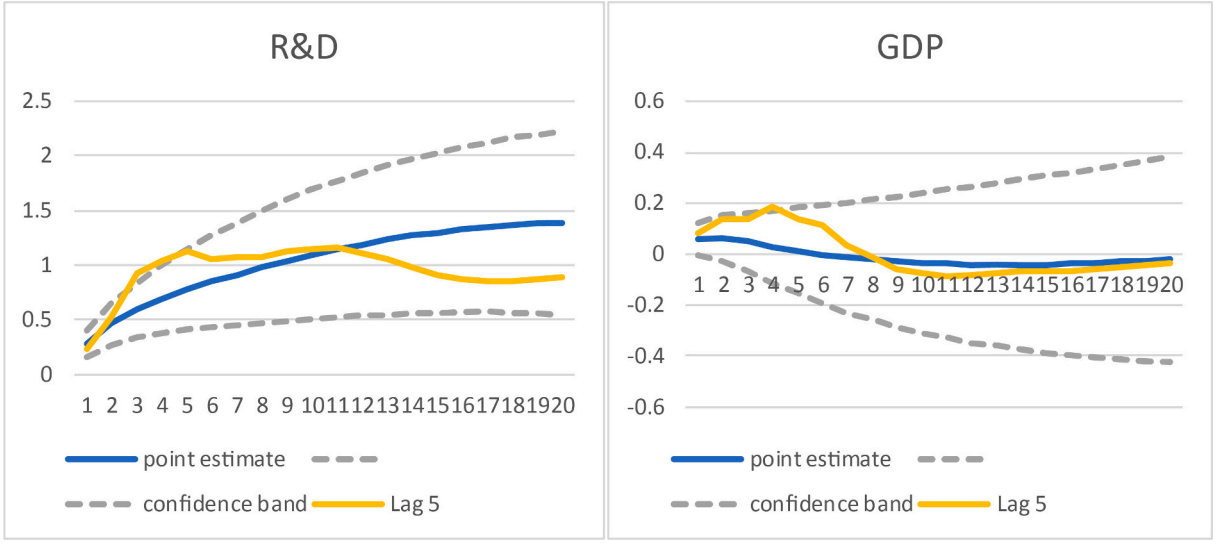

Fig. A.3. Impulse-response sequence for price adjusted level data with a VAR(2) for the full sample.
Quarterly response to a one-unit shock of the G_MO. Confidence bands correspond to a confidence level of 68 % (one standard deviation). Data sample: 1947Q1 through 2018Q4. Lag 5 corresponds to the results for the main specification with five lags.
Source: Fred Economic Data St. Louis FED, own calculations.

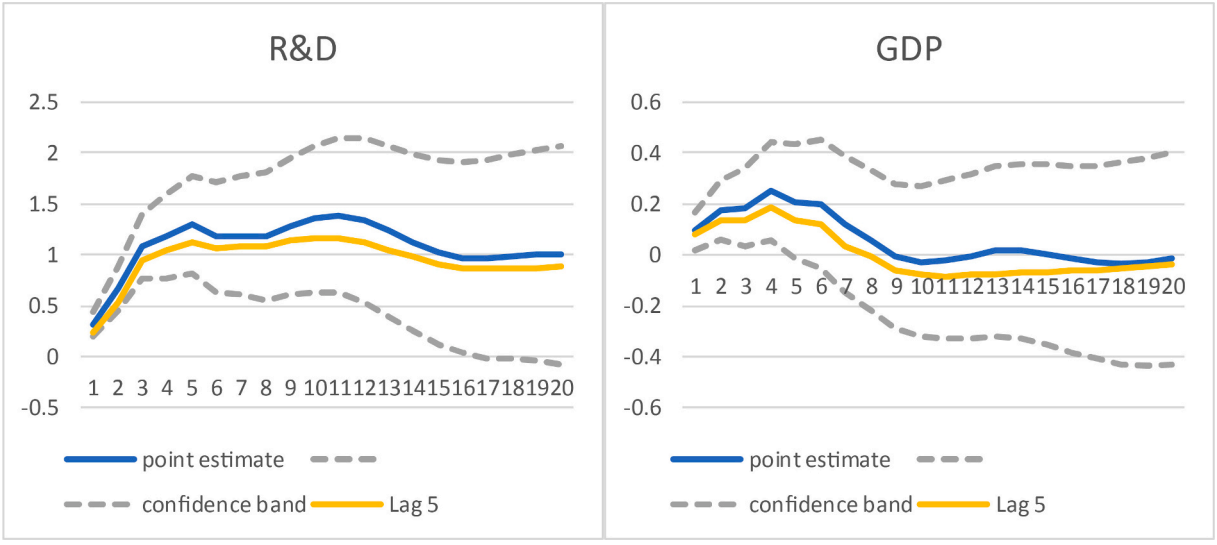

Fig. A.4. Impulse-response sequence for price adjusted level data with a VAR(6) for the full sample.
Quarterly response to a one-unit shock of the G_MO. Confidence bands correspond to a confidence level of 68 % (one standard deviation). Data sample: 1947Q1 through 2018Q4. Lag 5 corresponds to the results for the main specification with five lags.
Source: Fred Economic Data St. Louis FED, own calculations.

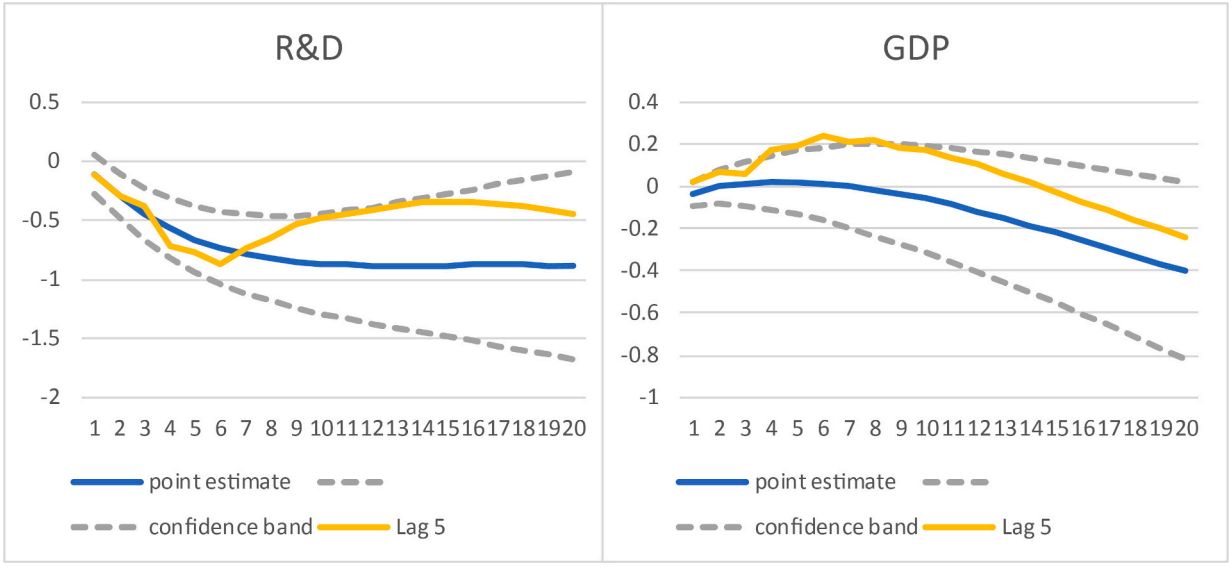

Fig. A.5. Impulse-response sequence for price adjusted level data since 1985 with a VAR(2).
Quarterly response to a one-unit shock of the G_MO. Confidence bands correspond to a confidence level of 68 % (one standard deviation). Data sample: 1985Q1 through 2018Q4. Lag 5 corresponds to the results for the main specification with five lags.
Source: Fred Economic Data St. Louis FED, own calculations.

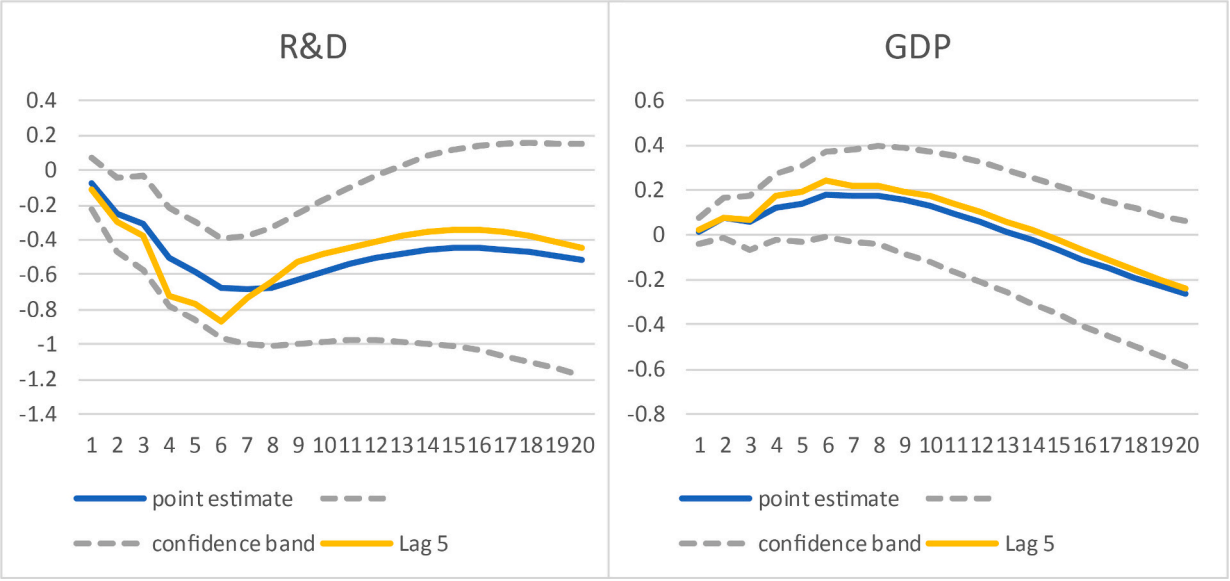

Fig. A.6. Impulse-response sequence for price adjusted level data since 1985 with a VAR(4).
Quarterly response to a one-unit shock of the G_MO. Confidence bands correspond to a confidence level of 68 % (one standard deviation). Data sample: 1985Q1 through 2018Q4. Lag 5 corresponds to the results for the main specification with five lags.
Source: Fred Economic Data St. Louis FED, own calculations.

The results obtained from the stronger parameterized models, where the number of lags is selected based on the AIC, generally align with the choice of using five lags. Even for the full sample, the smaller model with two lags doesn't deviate significantly from the one with five lags. In the post-1985 situation, the responses of R&D in the model with two lags show more severe differences. However, the lag 5 point estimates are still within the 68 % confidence band.

Annex 4. Cointegration test results

Table A.2
Cointegration test results (Johansen test).

# of cointegration relations	Full sample				Since 1984			
	Trace test		Max. eigenvalue test		Trace test		Max. eigenvalue test	
	Test stat.	p-Value	Test stat.	p-Value	Test stat.	p-Value	Test stat.	p-Value
GDP deflated data								
None	65.15	0.00	41.85	0.00	58.93	0.00	27.97	0.04
At most 1	23.30	0.23	10.85	0.66	30.96	0.04	15.79	0.24
At most 2	12.45	0.14	10.22	0.20	15.17	0.06	11.86	0.12
At most 3	2.23	0.14	2.23	0.14	3.31	0.07	3.31	0.07
Price adjust. data								
None	72.42	0.00	47.24	0.00	54.55	0.01	29.23	0.03
At most 1	25.18	0.16	11.21	0.63	25.32	0.15	14.42	0.33
At most 2	13.97	0.08	8.85	0.30	10.90	0.22	8.71	0.31
At most 3	5.12	0.02	5.12	0.02	2.18	0.14	2.18	0.14

Trend assumption for the alternative: linear deterministic trend. Intercept in cointegration relation and in VAR.
Source: Fred Economic Data St. Louis FED, own calculations.

Annex 5. Structural breaks in multivariate regressions

The procedure proposed by Qu and Perron (2007) estimates sequentially a multivariate regression for all possible subsamples that correspond to the a priori requirements specified with a maximal number of breaks and the minimum length of the subsample. The break points are identified according to the maximum likelihood principle. For this exercise, full breaks in all parameters of the VAR including the variance-covariance matrix are allowed. Furthermore, I assume a maximum of two breaks and the minimum length of a subsample is 25 % of the sample.

Unfortunately, Qu and Perron do not derive the limiting distribution for non-stationary regressors. Thus, to my knowledge so far, no valid test exists for multiple structural breaks in a VAR with trending data. To underline the relevance of structural breaks in the VAR for the given data set, I simulated data based on the OLS estimates for the whole sample and ran the estimation procedure for two breaks. The histogram in Fig. A.7 shows the differences in the conditional log-likelihood between the estimates with two freely chosen breaks and the estimation for the full sample. The log-likelihood difference for the empirical data surpasses all differences of the 2000 simulation runs.

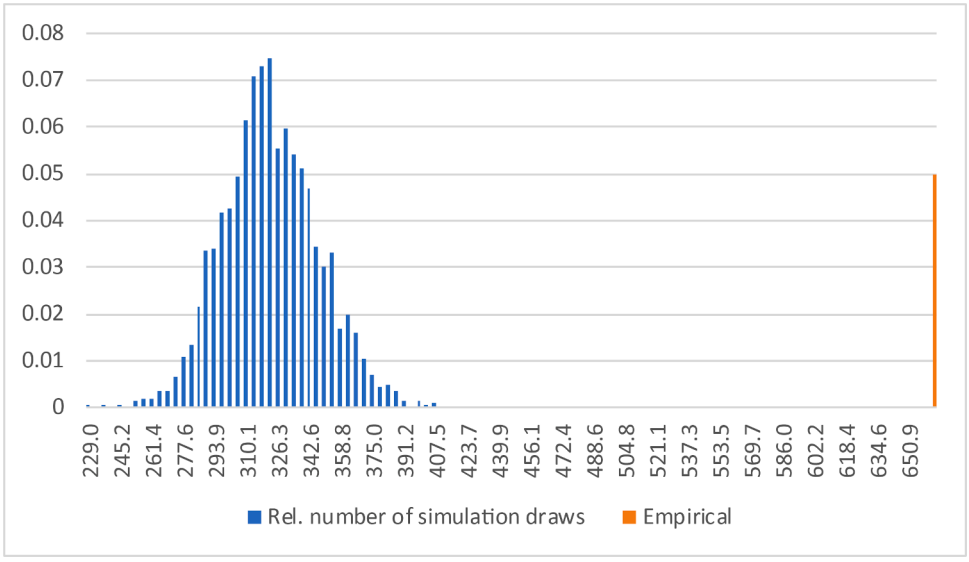

Fig. A.7. Simulated distribution for the likelihood with two unknown breaks.
Simulations based on the OLS estimates for the full sample. Likelihood conditional on the two break dates timed via the supremum. Empirical: Likelihood for the VAR with full breaks in 1966 and 1984.
Source: own calculation.

The results in the main text are based on the assumption of two breaks. Assuming three breaks shifts the second break into the year 1983 (instead of the beginning of 1985) and shows a further break in the year 2000 (Fig. A.8). The main finding that more recent data points to lower responses prevails.

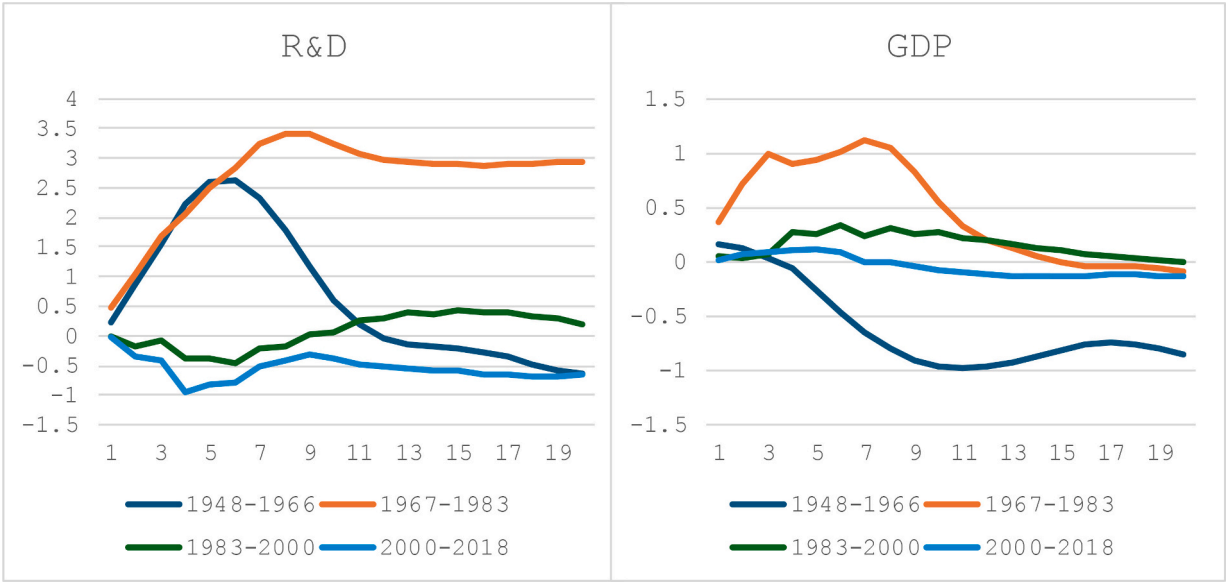

Fig. A.8. Impulse-response sequence for price-adjusted level data from various data samples.
Quarterly response to a one-unit shock of the G_MO. Specification names correspond to the time the data sample spans.
Source: Fred Economic Data St. Louis FED, own calculation.

Data availability

Data will be made available on request.

References

Akcigit, U., Hanley, D., Serrano-Velarde, N., 2021. Back to basics: basic research spillovers, innovation policy, and growth. Rev. Econ. Stud. 88 (1), 1–43.
Azoulay, P., Graff Zivin, J.S., Li, D., Sampat, B.N., 2019. Public R&D investments and private-sector patenting: evidence from NIH funding rules. Rev. Econ. Stud. 86, 117–152.
Deleidi, M., Mazzucato, M., 2021. Directed innovation policies and the supermultiplier: an empirical assessment of mission-oriented policies in the US economy. Res. Policy 50 (2), 104–151.
Fieldhouse, A.J., Mertens, K., 2023. The returns to government R&D: evidence from U.S. appropriations shocks. In: Research Department Working Papers No. 2305. Federal Reserve Bank of Dallas.
Gross, D., Sampat, B.N., 2023. America, jump-started: World War II R&D and the takeoff of the US innovation system. Am. Econ. Rev. 113 (12), 3323–3356.
Jones, B. F., 2021. Science and Innovation: The Under-Fueled Engine of Prosperity, in Rebuilding the Post-Pandemic Economy, ed. Melissa S. Kearney and Amy Ganz (Washington D.C.: Aspen Institute Press).
Kantor, S., Whalley, A.T., 2023. Moonshot: public R&D and growth. In: NBER Working Paper No. 31471.

Kilian, L., Lütkepohl, H., 2017. Structural Vector Autoregressive Analysis. Themes in Modern Econometrics. Cambridge University Press.
Moretti, E., Steinwender, C., Van Reenen, J., 2021. The intellectual spoils of war? Defense R&D, productivity and international spillovers. In: NBER Working Paper No. 26483.
Qu, Z., Perron, P., 2007. Estimating and testing structural changes in multivariate regressions. Econometrica 75 (2), 459–502.
Sims, C.A., Stock, J.H., Watson, M.W., 1990. Inference in linear time series models with some unit roots. Econometrica 58 (1), 113–144.
Stock, J.H., Watson, M.W., 2002. Has the business cycle changed and why?. In: NBER Working Paper Series Nr. 9127.