Measuring the Causal Economic Effects of Scientific Research — Evidence from the Staggered Foundation of the SENAI Innovation Institutes in Brazil

Torben Schubert[1,2], Denilton Darold[1,2], Markus Will[3]

[1]*Fraunhofer Institute for Systems and Innovation Research ISI*

[2]*CIRCLE — Centre for Innovation Research, Department of Design Sciences (LTH), Lund University*

[3]*Fraunhofer Institute for Production Systems and Design Technology IPK*

Abstract: How to estimate the economic returns of public science is a longstanding but equally challenging topic in quantitative science studies. In this paper, we exploit the staggered foundation of the SENAI Innovation Institutes (ISI) in Brazil since 2012 to estimate their effects on GDP using a difference-in-differences (DiD) approach. Building on historical and institutional insights from interviews on the foundation process, we unravel the conditions under which the parallel trends assumption is likely to hold. Our analysis reveals that these institutes significantly contribute to GDP per capita, with an average treatment effect of 985 BRL (approximately €160). Moreover, by relying on detailed project-level data, we were able to show that the effects come almost exclusively from genuine research projects and not from the provision of scientific services, such as metrology. Finally, tentative calculations suggest that the SENAI ISI institutes may account for about 0.66% of Brazil's overall GDP, emphasising the importance of applied science in regional economic development and providing insights into effective collaboration between research and industry.

Keywords: SENAI ISI, public research, economic effects, GDP per capita

Acknowledgments: The authors gratefully acknowledge the organisational and data support from SENAI. Specifically, we express our thanks to Gustavo Leal Sales Filho, SENAI's General Director, for his valuable insights into the foundation process of the ISI institutes. We would like to thank Alberto Pavim for his substantial contribution in providing the administrative data used in this study and for numerous insightful discussions. We would also like to thank Gabriel Bertholdo Vargas for his assistance in data provision. The corresponding author also acknowledges financial support from the German Ministry for Education and Research (BMBF) within the project PHOEBUS (grant number: 16NISTA04A).

Declaration of Conflict of Interest: The authors declare that this research paper is part of a research project commissioned by SENAI to Fraunhofer IPK and Fraunhofer ISI. The research was conducted independently, and the commissioning party did not influence the scientific work of the authors or the study results in any way.

1 Introduction

The analysis of the economic returns and effects of science has a long tradition in quantitative science studies. Almost 70 years ago, Zvi Griliches (1958) analysed the social returns of hybrid corn-related innovations by comparing the costs of their development with market values. Extending this idea, many studies since the 1970s have sought ways to measure the economic returns of science with various methodologies, including multiplier analyses (Glückler et al. 2015), regression approaches (Schubert and Kroll 2016b; Bertoletti et al. 2022; Agasisti and Bertoletti 2022) or macroeconomic simulation models (Allan et al. 2022). Yet, while these methods come with their specific advantages and disadvantages, they all face the pressing need to establish causality. In the context of the economic effects of science, credibly claiming causal links is, however, notoriously difficult because scientific organisations are usually not randomly located across space; they are attracted to economic or cultural centres, become embedded in their regions and then evolve with them, often over the course of centuries. Disentangling such deeply ingrained relationships to cleanly identify the causal economic effects of science is, irrespective of the level of statistical sophistication, usually rendered almost impossible simply due to the blurring of what is cause and what is effect. Thus, most of the existing analyses remain subject to the criticism that the estimated effects may, to a non-substantial degree, reflect associations rather than causation.

In this paper, we seek to address this gap by exploiting evidence from a unique case: the SENAI Innovation Institutes (ISI) in Brazil. The ISI institutes are private, not-for-profit research institutes focusing on applied sciences, mostly in natural sciences, engineering and life sciences, which make revenues by cooperating with contracting firms, for which they provide various types of research-related services. 27 ISI institutes have been founded sequentially since 2011 (26 up to 2019) and they are spread over the country. The staggered foundation offers a unique opportunity for a quasi-experimental setting usually unavailable in studies analysing the effects of public science. Specifically, we will make a case that the characteristics of the regional selection process were driven by fixed-effects-like selection based on long-term regional characteristics and therefore satisfy conditions under which parallel trends will hold (Ghanem et al. 2022). Thus, the characteristics of the foundation process —specifically, how regions were selected — allow us to estimate the causal effects of the SENAI ISI institutes on regional GDP. Moreover, our detailed data allows us to separate the effects based on the types of collaboration projects with firms, which gives us unique insights into which types engender the strongest economic effects.

Our results suggest that the staggered foundation of the ISI institutes had substantial economic effects in terms of GDP. Overall, we estimate the average treatment effect on the treated regions (ATT) in terms of GDP per capita to be 985 BRL. This effect comes almost exclusively from genuine research projects involving R&D, while more service-oriented projects, such as metrology or technical consultancy, appeared to cause little or no effect. Finally, we provide some tentative macroeconomic projects, which suggest that the SENAI institutes may be causally linked to about 0.66% of the overall Brazilian GDP.

2 Background

Studies measuring the economic value of public science can be subdivided into microeconomic studies that measure the effects on firm performance and macroeconomic studies that measure the effects on whole economies, such as regions or countries. While the former group of studies is broad (Robin and Schubert 2013a; Maietta 2015; Comin et al. 2019), they rely on very different kinds of datasets and thus on different identifying assumptions. We will therefore omit this group from our review and focus on studies estimating macroeconomic returns in the next section.

2.1 The Macroeconomic Returns of Scientific Research

Most analyses estimating the macroeconomic effects of public research have relied on multiplier analysis, using either Keynesian or input-output multipliers (Glückler et al. 2015). These studies rely on taking into account clearly identifiable and attributable expenditure streams, such as induced investments or consumption by staff or students, to which, in turn, multipliers are applied. While estimating rates of returns, typically defined as GDP increase divided by costs, these studies contain numerous sources of bias that seriously limit their economic interpretability (Schubert and Kroll 2015). The measurement of expenditure that flows into the region is, in many cases, economically problematic for the purpose of measuring social returns. For example, the increase induced by student consumption in the host region seems hardly interpretable. The students would have consumption expenditures irrespective of their student status — in many cases, even higher ones had they begun to work instead. Moreover, multiplier analyses ignore the value of knowledge (Glückler et al. 2015), which is arguably the most genuine contribution of scientific organisations. In particular, the latter problem is conceptually hard to defend and, within the methodological framework, impossible to tackle. A number of studies have specifically tried to address both problems by resorting to econometric estimations of the macroeconomic effects of public science organisations. In this setting, some sort of regression approach is used to let an economic outcome variable in a given region (often GDP) to be explained by variables measuring the activity of public science organisations in the same regression. Studies falling into this category include Schubert and Kroll (2013), Schubert and Kroll (2016b), Bertoletti et al. (2022), Agasisti and Bertoletti (2022) and Allan et al. (2022). A common finding of these studies is that estimated returns are much larger than those resulting from multiplier analyses, underscoring the paramount importance of knowledge generation. For example, Schubert and Kroll (2013) — using fixed-effects panel regressions for German NUTS 3 regions — estimate that about 8% of German GDP is attributable to its university system.

Nonetheless, while conceptually clearer in their interpretation, these studies are also subject to methodological criticisms. Notably, like all regression approaches, the key challenge is to establish causality — a non-trivial task with field data, which is subject to various sorts of selection effects. While these studies make considerable econometric efforts to ensure causality

— e.g., in their study of the economic effects of the Fraunhofer Society in Germany, Allan et al. (2022) provide a series of confirming placebo tests — ultimately, the ability to provide causal effects rather than mere associations requires the identification of some sort of exogenous variation. Some recent studies have tried tackling the causality issue by exploiting the foundations of specific types of organization, for example the applied science universities in Switzerland (Schlegel et al. 2022; Pfister et al. 2021) arguing that the event of a foundation is at least partly exogenous. Yet again, if locational choice for the new establishments is selective favouring e.g. a priori more central and economically stronger regions, even DiD-estimators may be biased. Thus, the ability to estimate estimate causal effects rests less on the choice of econometric estimator but more on whether the specific case is likely to provide a source of exogenous variation. In the next section, we will explain the institutional and historical details of the foundation of the SENAI institutes in Brazil to argue that this case appropriately provides such exogenous variation.

2.2 History and Role of the SENAI Innovation Institutes (ISI)

The SENAI ISIs (*SENAI Innovation Institutes, or Institutos SENAI de Inovação*) were established by SENAI (*Serviço Nacional de Aprendizagem Industrial, or National Service for Industrial Training*) as a response to the evolving needs of the Brazilian industrial sector with regard to advanced research and development capabilities. The foundation of these institutes marked a significant shift in the country's approach to industrial innovation, aiming to bolster competitiveness through cutting-edge technology and applied research (Kohl et al. 2020).

SENAI itself is the largest private vocational education complex in Latin America. Since its creation in 1942, it has trained more than 73 million workers in 28 industry areas. It is present in more than 2000 Brazilian municipalities and offers a wide variety of courses at all levels of professional and technological education. Additionally, it offers technological services, metrology, consulting through SENAI ISTs (SENAI Technology Institutes, or Institutos SENAI de Tecnologia) and, more recently, applied research and innovation through the establishment of SENAI ISIs.

The establishment of the SENAI ISI institutes represented an important evolution of SENAI's mission. It started in 2011, driven by a comprehensive dialogue with over 50 business leaders, organised by the MEI (*Mobilização Empresarial pela Inovação, or Business Mobilization for Innovation*), which highlighted the need for a more robust infrastructure dedicated to industrial innovation. The National Confederation of Industry (CNI) led this initiative, selecting SENAI as the primary institution due to its extensive experience, established credibility and deep-rooted connections within the Brazilian industrial sector. With its 70-year legacy, SENAI was well-positioned to lead the creation of a national network of applied research institutes through the SENAI ISIs.

The development of the SENAI ISI network was not merely an infrastructural endeavour but one that required significant intellectual capital, including business processes and human and relational resources. SENAI partnered with the Fraunhofer Society, a global leader in industrial research and development, to plan, implement and operationalise the ISI network. This collaboration facilitated the determination of actual industrial demand through workshops with over 300 companies across 12 Brazilian states.

When comparing the SENAI ISI network to other research and technology organisations globally, key similarities and distinctions emerge. Modelled after Germany's Fraunhofer Society, SENAI ISIs focus on high-impact industrial projects and emphasise collaboration with both large and small enterprises. Similarly, the Dutch TNO, the Canadian National Research Council (NRC) and the Finnish Technical Research Centre (VTT) highlight the importance of broad scope, national impact and transforming economies through innovation. The SENAI ISI network, with its well-planned establishment and strategic focus, mirrors these successful international models, with the aim of positioning itself as a vital component of Brazil's national innovation system.

The strategic focus was on creating a demand-driven network, ensuring sustainability and broad national coverage. Financially, the infrastructure part of the project was supported by investments from SENAI and loans from the Brazilian Development Bank (BNDES), worth approximately R$ 1 billion in initial investments and R$ 2.2 billion up to now. By 2024, 27 of the 28 planned innovation institutes were operational, employing over 1500 researchers. These institutes have collectively executed projects worth R$ 1.9 billion in partnership with more than 800 industrial companies by 2021.

The ISIs are also integrated into broader national and international research frameworks, with 18 institutes accredited as EMBRAPII units, 23 recognised by the ANP and 14 accredited by CATI. Notably, 56% of the R&D&I projects involve startups and small to medium enterprises, fostering a robust ecosystem of innovation. The network has connected over 185 startups with 90 larger companies through technological challenges.

The impact of the SENAI ISI network extends beyond immediate project outcomes. It has facilitated the practical application of scientific research, translating theoretical knowledge into tangible industrial solutions. This role as an intermediary between universities, research centres and the industry is crucial for maintaining the flow of innovation within Brazil's economic framework. Therefore, the SENAI ISI network is an exemplary case for measuring the economic effects of scientific research due to its unique setup. The staggered establishment of the institutes across diverse regions provides exogenous variation, which is ideal for applying the difference-in-differences (DiD) method to estimate causal impacts. This setting minimises pre-existing regional differences, ensuring credible results. Additionally, the institutes' focus on applied research and industry collaboration directly ties economic effects to scientific advancements. The comprehensive data available from their inception further strengthens the analysis. Thus, the SENAI ISI network offers a robust framework for evaluating the true economic returns of scientific research and innovation.

2.3 Locational Choice in the SENAI ISI Case and the Parallel Trends Assumption

Measuring the causal effects of the foundation of SENAI ISI institutes using a DiD approach depends crucially on the validity of the parallel-trends assumption, which means that in the absence of treatment, both treated and non-treated regions would have displayed parallel trends in the outcome variable in the post-treatment period. The parallel trends assumption is obviously defined partly in terms of unobservable counterfactuals, since all treated regions in the post-treatment period are only observed in their treated and not in their hypothetical non-treated state. This unobservability makes it impossible to test the parallel trends assumption empirically, but we can theoretically assess its plausibility. Whether it holds or not obviously depends on the locational choice. For example, if the SENAI ISI institutes had been randomly allocated to regions, then randomisation of the locational choice would imply that treatment and non-treatment regions would probabilistically not differ in any respect except for the treatment status. Needless to say, randomisation is an unreasonable assumption in our case, as the locations for establishing the institutes were strategically chosen. However, while at a trivial level, randomisation implies the validity of DiD, the parallel trend can also hold under weaker assumptions, too, as shown by Ghanem et al. (2022).

One of the conditions under which it can hold is that the locational choice depended only on fixed (potentially unobservable within the context of the study) effects of the regions. I.e., locational choices were made based on (roughly) time-constant structural characteristics of the regions. Examples of such characteristics include economic power, established presence of potential collaboration partners, presence of other scientific organisations, local demand/economic structures or human capital endowments. Based on desk research and insights from SENAI management, we will now make a case that the locational choice was, as expected, non-random and made in a way to a) strategically maximise the prospective success of the institutes and b) achieve a certain level of regional balance. Both mechanisms worked in a way where the key decision-makers relied on stable regional characteristics, which they used as proxies. Thus, our key argument is that the selection was based on fixed effects and consistency conditions, as defined by Ghanem et al. (2022), can be credibly assumed to apply.

Qualitative Evidence on Locational Choice for SENAI ISI Institutes

To better understand the context and circumstances of SENAI ISI’s conception and the selection criteria for branch establishment across the country, an interview was conducted in September

2024 with the current SENAI general director, who has been directly involved in the whole foundation process since its inception.

According to him, the selection of institutes followed a structured process that combined both technical and political considerations. Initially, the primary criterion for selecting locations was the technical capacity of pre-existing local SENAI institutes. The goal was to ensure that these institutes, focused on vocational training, had the necessary infrastructure, expertise and administrative capability to support R&D activities in specific technological areas. In his words, “institutes with high performance willing to implement R&D.” Next, thematic competence was assessed, where competencies that have already been developed converge with priority technologies. SENAI sought to align the institutes with fields deemed as national priorities, or “promising technologies,” based on international benchmarks, ensuring that emerging technologies would be the focus.

The first institutes were initially concentrated in Brazil’s coastal regions, where most of the population and major cities are located. The existing institutes had more developed technical capabilities and administrative infrastructure, making them natural candidates. However, aligned with the National Confederation of Industry (CNI), the assessment was that the geographic distribution was imbalanced. It was, in his words, “*muito litorâneo*,” or “too coastal.” This imbalance did neither reflect the internal political structure of SENAI as a federated organisation nor the underlying policy goals of regional development.

To address this geographic imbalance, a political element was introduced into the selection process. This shift led to the prioritisation of inland regions when the technical criteria had been exhausted. The rationale behind this decision was both pragmatic and political. From a development perspective, investing in economically weaker regions was seen as a way to foster innovation across the entire country, not just in areas that were already well-established. Politically, it was important to maintain equity within SENAI’s federative system, which included balancing investments among the various state industry federations to ensure regional representation and support.

Thus, the second phase of the selection process involved directing funds to inland states like Amazonas, Pará and Mato Grosso do Sul. This phase sought to ensure that the benefits of the SENAI ISI network extended beyond the country’s more developed coastal regions, helping to promote R&D activities in underdeveloped areas. It is important to note that the institutes were thematically oriented, meaning their focus was on specific technological fields rather than serving only local industries. As a result, their client base was often scattered across the country, and their location was largely independent of where their industrial partners were based.

In sum, the location of SENAI ISI institutes was driven by two main factors: pre-existing technical capabilities/proximity to economic centres and the political will for geographic distribution. Both selection mechanisms were implemented by the key decision-makers’ reliance on stable regional characteristics, which gives us a good a priori confidence that parallel trends (potentially conditional on covariates) are likely to hold in our setting. In the next Section, we will continue by presenting the data and the specifics of our identification strategy.

3 Data and Methodology

We utilised two primary data sources: regional macroeconomic statistics and administrative data from SENAI. The regional macroeconomic statistics, which include GDP per micro-region and population data, were sourced from the Brazilian Institute of Geography and Statistics (IBGE). The IBGE is the official government agency responsible for the collection of statistical, geographic, cartographic, geodetic and environmental information in Brazil. The second data source was provided by the SENAI administration and includes detailed administrative data of all SENAI institutes, their projects, respective categories, and client locations. This comprehensive data source encompassed various aspects of SENAI's operations, offering insights into the specific competencies of each institute and the geographical distribution of their clients.

To construct the dataset for econometric analysis, we merged the macroeconomic statistics with the SENAI administrative data using the micro-region as the key.[1] This approach allowed us to analyse the economic impact at the regional level, taking into account both the locations of the institutes and their clients, who are spread across the country. This widespread distribution is attributed to each institute's specialisation, which focuses on specific areas of competence rather than regional allocation or concession.

3.1 Overview of the Identification Strategy

Our econometric identification strategy aimed to estimate the causal effects of SENAI ISI's establishment on regional economic outcomes. To achieve this, we employed a difference-in-differences (DiD) methodology, leveraging the staggered introduction of SENAI ISI institutes across various micro-regions in Brazil.

Our identification strategy benefits from the quasi-experimental setting provided by the non-random and staggered placement of the institutes. To rule out the possibility that any short-term dynamics in the GDP growth path would bias the estimates, we also included one-year leads and lags of growth in GDP per capita. Moreover, since we know the SENAI ISI institutes, despite being founded equitably across the Brazilian states, were located in urban areas in the states, we also included the population of the hosting municipality. Thus, we ultimately employ a conditional DiD approach, where the central identifying assumption is that parallel trends hold conditionally on leads and lags in GDP per capita and population growth.

3.2 Estimation Framework in the SENAI ISI Case

While we have already discussed that parallel trends is likely a plausible assumption in our case, which makes DiD a reasonable choice, there are a number of subtleties in the ISI case, which we

[1] A table with summary statistics for all variables can be found in the appendix.

will now discuss. These include the staggered design, the concrete definition of treatment and its implications as well as the choice of estimator. Specifically, we will make use of two alternative treatment definitions, the first of which is based on Irreversible Treatment, while in the second version, regions may switch between treated and untreated states freely.

3.2.1 Staggered Irreversible Treatment

DiD was originally developed within a two-period-two-groups design, our design is staggered as the ISI institutes have been founded in many different time periods. This distinction is important because the most common two-way-fixed-effects (TWFE) estimator is consistent in staggered designs only if the treatment effects do not differ between the periods in which treatment first occurred (Goodman-Bacon et al. 2019). With time-heterogeneous treatment effects, estimators specifically tailored to staggered designs should be used (see below).

Exploiting the foundation dates, it is straightforward to assign the treatment status to all regions that are home to at least one SENAI institute from the year of foundation of the earliest institute in the region. Suppose a region is home to a SENAI ISI institute, which was founded in 2014, that region would be considered as treated from 2014 onwards and untreated before. To define what we consider the home region, we draw on insights from the literature on the geography of knowledge spillovers, which suggests that knowledge spillovers are often localised but may well reach beyond very narrow geographic boundaries. For example, Anselin et al. (1997) provide evidence for the US that academic spillovers will often reach beyond the boundaries of metropolitan statistical areas (defined as at least one metropolitan agglomeration and at least 50,000 inhabitants). Similar findings for Germany suggest that up to 90% of the economic effects for universities originate outside the boundaries of the local community (Schubert and Kroll 2016a). Thus, it seems necessary to define the home region in sufficiently broad terms. In our case, we opted for the state level, of which there are 26 in Brazil. Thus, we consider a region treated if there is at least one SENAI institute in the same state.

To formalize our design, we briefly explain the fundamental staggered DiD-design based on the exposition by Callaway and Sant'Anna (2021). Assume there are T periods and define G_{ig} as a dummy variable, which indicates that a unit i is first treated in a period $g \in \{2, \dots, T\}$. Following the standard staggered design, for now, we also assume irreversibility of treatment, i.e., units do not switch back to non-treatment states. Let $Y_{i,t}(0)$ be the potential outcome of unit i if it is untreated in period t. Furthermore, define $Y_{it}(g)$ the potential outcome that unit i would experience in period t if were first treated in period g. Then, the potential outcome for each unit and time period is given by:

$$Y_{i,t} = Y_{i,t}(0) - \sum_{t=2}^{T} \left(Y_{i,g}(g) - Y_{i,g}(0) \right) G_{i,g} \qquad (1)$$

In DiD approaches, the interest usually lies in the average treated effect of the treated (ATT). What is different in staggered designs is that there is not only one ATT, but many individual ones for each onset of treatment. I.e., for units which share the same first treatment period g, we can define an ATT in each period t as follows:

$$ATT(g,t) = E(Y_t(g) - Y_t(0)|G_g = 1) \quad (2)$$

As one can see, in this framework, the ATTs are allowed to differ over time. Usually, the primary policy interest is not in these individual treatment-onset-specific ATT but in summarizing aggregations thereof, a very natural one of which is the following:

$$\theta = \sum_g \sum_{t=2}^{T} w(g,t) ATT(g,t) \quad (3)$$

with $w(g,t)$ group-time-specific aggregation weights. As already said, unless the individual ATT is constant over time, it can be shown that TWFE does not consistently identify θ. It is, however, possible to derive estimators that provide consistent estimates. Among these are the ones by Callaway and Sant'Anna (2021) or the one by Chaisemartin and d'Haultfoeuille (2024).

3.2.2 Staggered reversible Treatment

The above staggered design rests on the assumption of irreversibility of treatment, i.e., once units are treated, they remain so forever. This is true for our treatment definition based on the opening/location of SENAI ISI institutes, as long as no institutes are closed down (which was not the case in our data). However, relying on the institute locations to define treatment rests on the assumption that the economic effects occur on the basis of geographic proximity as reflected by administrative borders. While this assumption is reasonable, it neglects the fact that the geographic patterns of interaction between SENAI ISI and firms that give rise to economic effects can be substantially more complex. An important reason for complex geographic interaction patterns is that the ISI institutes are very specialised in terms of topics. Thus, for any firm with a certain activity profile, the institutes are not interchangeable. If the firm sees value in collaborating with ISI institutes, it can reasonably be expected to select a thematically appropriate institute as a partner, rather than just selecting the closest. While there is certainly a significant distance decay component, because firms may be reluctant to enter collaborations with partners located very far away, we would expect cooperation patterns to be more spread out than for public science organisations covering a much broader range of activities, notably universities. One way to address this point is to use the location of the cooperating firms rather than the location of the institute. Based on this information, we also use an alternative treatment definition and consider a region as treated if, in a given year, it was home to at least one firm that had a contractual relationship with an ISI institute (irrespective of where the latter was located). Because our database covers more than 25,000 different contracts, we define the home region of the firm relatively narrowly by the 5-digit regions, of which there are 510 in Brazil. This regional treatment definition, based on partner firm location, gives a very fine-grained account based on actual interaction patterns that are not biased by ad-hoc assumptions regarding regional proximity.

While our treatment measures actual collaboration patterns and may be considered more accurate, it comes at the price that it does not result in a standard staggered DiD design. Obviously, regions can depending on the start and end of individual projects switch arbitrarily back and forth between treated and untreated states. The estimator provided by Callaway and Sant'Anna (2021) is not fit for such a design. However, the estimator provided by Chaisemartin and d'Haultfoeuille (2024) is more general and can also handle such complex intertemporal treatment.[2] For the sake of comparability, we used the estimator by Chaisemartin and d'Haultfoeuille (2024) as the baseline estimator and reported the one by Callaway and Sant'Anna (2021) only in the robustness section. The choice of estimator, however, did not seem to have much influence in the cases where both are applicable. Before we turn to the estimation of the treatment effects, we now briefly present some descriptive analyses of the evolution of the SENAI ISI institutes as well as the geography of their operations.

3.3 The Evolution and the Development of SENAI Innovation Institutes

At the writing of this article, the SENAI ISI network consisted of 27 operational institutes,the first of which was founded in 2012.[3] These institutes are distributed across 13 federal states, with some states, such as Bahia and Rio de Janeiro, hosting multiple institutes established in different years during the study period. Initially, the institutes were concentrated in densely populated coastal regions, where the largest vocational training centres were located, before gradually expanding to inland areas, which is in line with the two phases of the foundation process we described in Section 2.3. Figure 1 illustrates Brazil's population density along with the locations of the institutes, labelled with their respective foundation years.

[2] We do not introduce their design here because it is similar to the one described in Eqs. (1)-(3), where however the additional generality comes at the price of a more complex notation.

[3] A further one was in the planning phase.

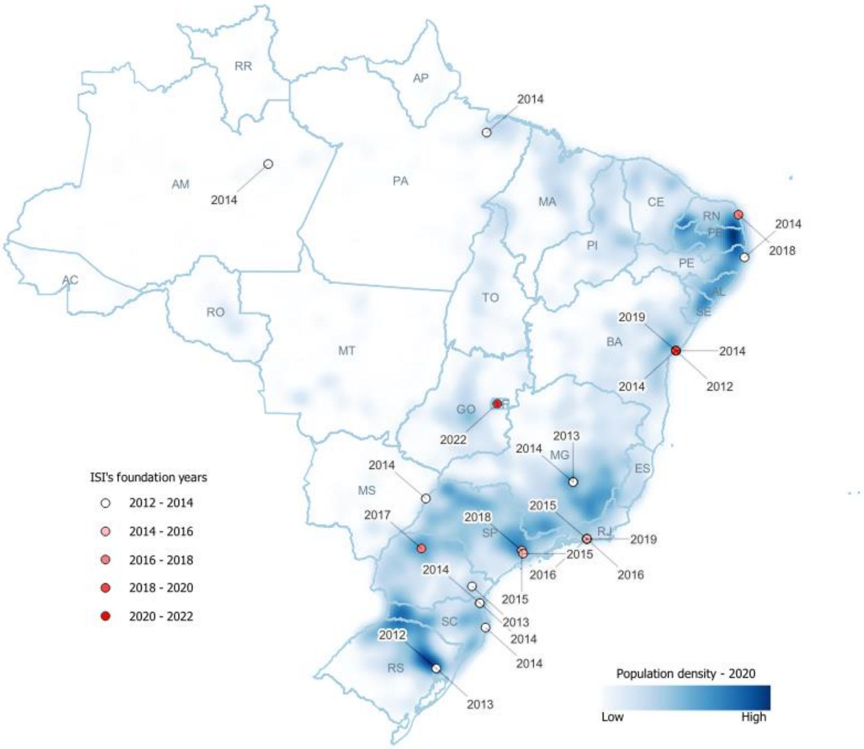


Figure 1: Population density and ISI institutes

To better understand our definition of treatment, it would be useful to take a closer look at the official regional classification in Brazil. While there are 26 states at the top level, these are hierarchically divided into 126 intermediate regions, 510 immediate (or micro) regions, and 5570 municipalities. While the 26 states were used to define treatment using the location of institutes, we opted to draw on the 510 micro-regions to define treatment based on project locations with ISI clients, i.e., companies that have ordered or commissioned projects with ISI institutes. Of the 510 immediate regions, 313 have at least one project listed, as can be seen in Figure 2, showing the geographic outreach of the institutes' client base.

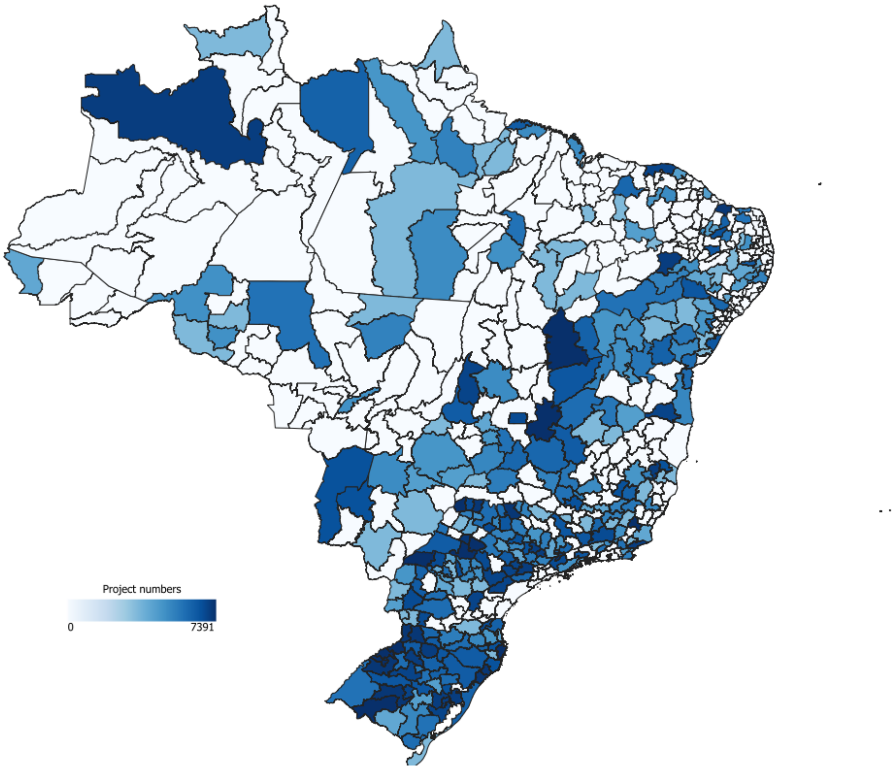


Figure 2: Project numbers per region

Figure 2 reveals a high regional concentration of the projects. While most regions have only a small number of projects, a few regions exhibit disproportionately high project counts and revenues. The top region has 7391 projects, and the leading client accounts for 915 projects. Similarly, revenue per project varies widely, with most generating modest amounts, though a few outliers greatly inflate the average, with one project earning over R$ 25 million, as Table 1 shows. This highlights the fact that a small number of high-impact projects and regions account for a disproportionate share of overall activity and financial returns. The figures are interesting because they can imply challenges for DiD estimation, which implicitly assumes that treatment is equally sized. We come back to this issue later with additional robustness checks allowing for continuous treatment.

	Regions	Clients	Revenue per Project
Count	313	7402	43.64
Mean	147.00	6.22	23.17
Std Dev	692.61	23.76	329.80
Min	1	1	-1181.14
25%	3	1	0.25
Median	11	2	0.6
75%	56	4	1.54
Max	7391	915	25594.35

Table 1: Summary of project statistics
Note: Revenue in thousand BRL

Examining the revenue grouped by service types provides us with further details, as seen in Table 2. R&D stands out with the highest average revenue by far, confirming that R&D projects tend to be the most financially impactful, though the variation is substantial, as seen in the large standard deviation and a maximum value of R$25.5 million. Technology Consulting shows the highest outlier at over R$2.5 million, though it also has a wide range of outcomes. Metrology, with the most entries, has a lower average revenue but a broad range, while specialised services show notable extremes, with a maximum of over R$2.3 million. Overall, a small number of high-revenue projects, particularly in R&D, drive much of the financial impact.

	Technology Consultancy	Metrology	R&D	Complementary Services	Specialised Technical Services
Count	1380.00	38435.00	2080.00	39.00	1709.00
Mean	22204.52	1963.12	422883.19	52582.76	13844.97
Std	109713.31	22559.62	1445190.35	135812.16	103146.88
Min	-1181149.90	-960.00	0.00	480.00	0.00
25%	1980.00	229.50	11660.00	3630.00	450.00
Median	5515.00	510.00	78079.46	19199.81	1250.00
75%	20000.00	1196.50	320660.49	51653.47	3600.00
Max	2527547.10	2190876.00	25594357.26	843579.87	2375021.00

Table 2: Project revenue per service type

Regarding the evolution over time, there is a consistent upward trend of projects across all service types from 2013 to 2019 and then stabilisation until 2023, as seen in Figure 3. The majority are metrology services, followed by R&D, consultancy and specialised services. The slight dip in 2020 was driven by a decrease in metrology projects, whereas R&D presented a consistent increase over the whole period, peaking in 2022.

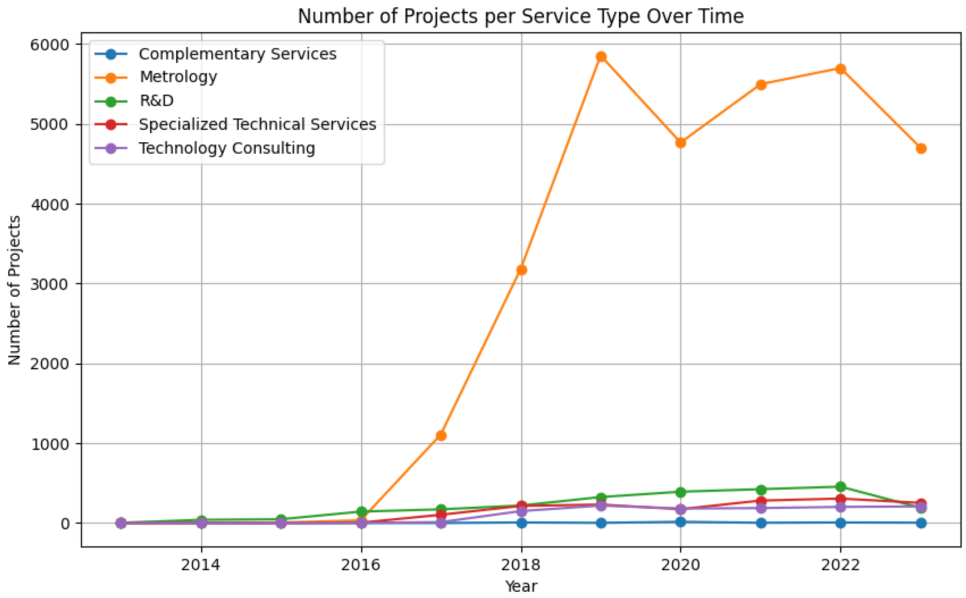


Figure 3: Project numbers over time

These descriptive statistics lay the groundwork for the subsequent analysis of the economic effects of SENAI ISI institutes.

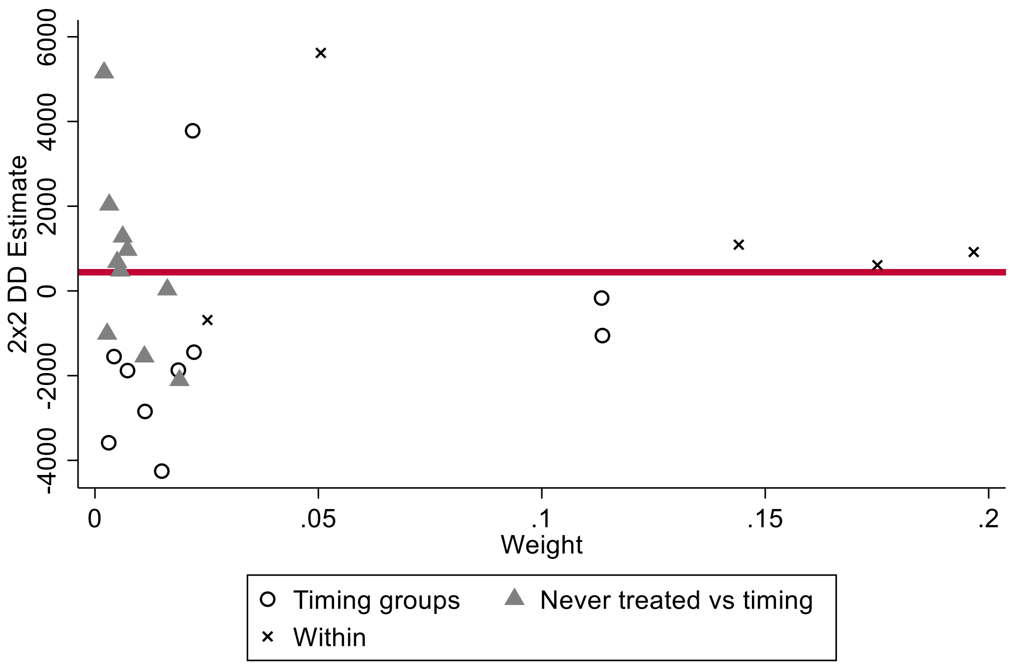
4 Results

4.1 The Economic Effects of SENAI

Before we present our main key results, we now briefly provide some evidence of the central role of the staggered roll-out. An important conclusion is that indeed, treatment effects are time heterogeneous implying that TWFE will be biased and the estimators discussed in the previous section should be used instead.

4.1.1 The Role of Staggered Roll-out for Estimation

Goodman-Bacon et al. (2019) have shown that the regular TWFE estimator of the average treatment effect of the treated group (ATT) is a weighted sum of all constituent 2x2 DiD designs. Out of these 2x2 designs, the particularly problematic cases are those where later treatment groups are compared to earlier treatment groups. Bias arises unless treatment effects are constant over time. The decomposition results by Goodman-Bacon et al. (2019) make it possible to analyse the significance of these potentially bias-inducing later-vs.-earlier comparisons by calculating the ATT for each individual 2x2 design alongside the aggregation weight. Problems are likely if i) treatment effects differ strongly over the 2x2 designs and ii) if the problematic later-vs.-earlier comparisons have high weights.



Overall DD Estimate = 442.35799
Within component = -684.2442 (weight = .0251527)

Figure 4: Bacon decomposition for DiD design based on institute location

The results from the Bacon decomposition are shown in Figure 4, where the label "Timing groups" refers to a set of potentially problematic later-vs.-earlier comparisons. As we can see, the overall TWFE DiD estimate corresponds to a 442 BRL increase in GDP per capita following the foundation of an institute, which, from a normal TWFE regression, would obtain a p-value of 6.7%. We should note that this effect does not include any control variables and is merely a point of reference. Moreover, we can see that the individual ATTs differ substantially as there are, besides many positive ATTs, also a larger number of negative effects. Most importantly, however, the potentially bias-inducing later-vs.-earlier comparisons are — with one exception — all negative, and two of them have very high weights of about 11% each. Thus, the Bacon decomposition suggests that the treatment effects are heterogeneous over time and that they have most likely induced a negative bias. Indeed, when using only the later-vs.-earlier comparisons to estimate the overall treatment effect, we obtain an estimate of 691 BRL, which is thus likely to be more significant than the baseline TWFE estimate, including the biasing comparisons. Overall, the results from the Bacon decomposition suggest that it is highly important to rely on estimators that are robust to heterogeneous treatment effects in staggered designs. We present the results of these estimations in the next subsection.

4.1.2 Estimating the Baseline Effects on GDP per Capita

As discussed in Section 3.2, our data allowed us to define treatment either by relying on foundation dates in combination with institute location or alternatively based on the location of the cooperating firms. The baseline results of these estimations are presented in Table 3.

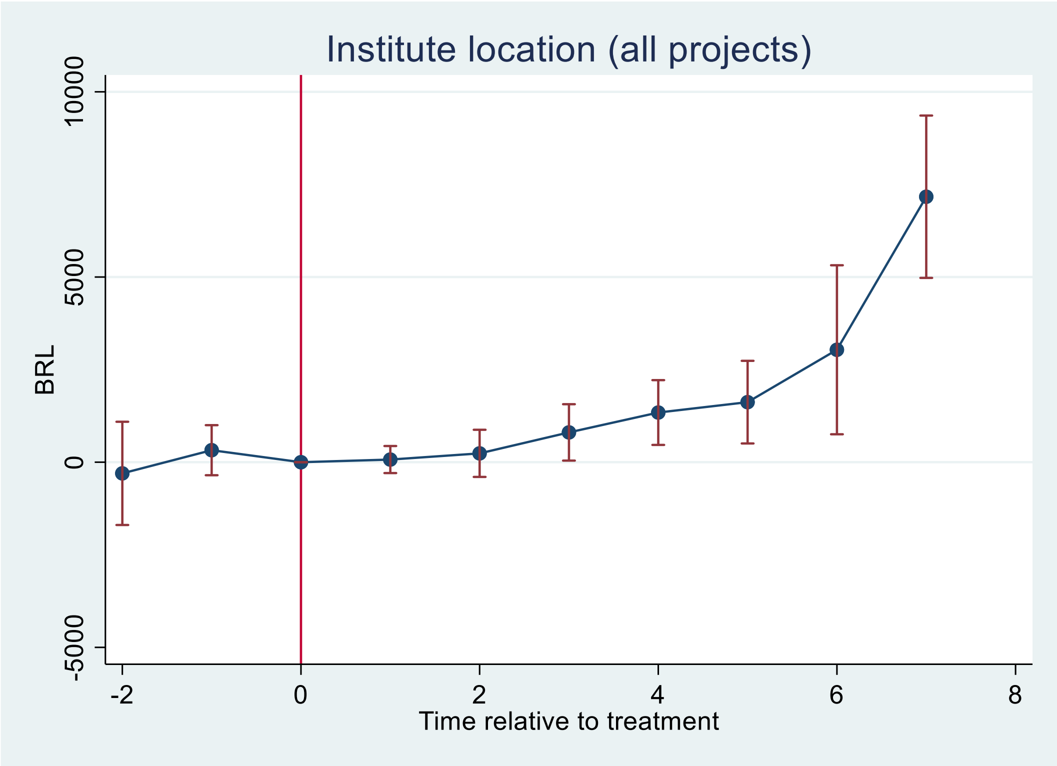
In the first row, we show the results for the treatment definition based on the institute location using the estimators. We can see that the estimates are positive and significant (ATT: R$ 1210, pval<0.01). It is, as suggested by the Bacon decomposition, also larger than the naive TWFE estimate of 442 BRL, which confirms the downward bias of the time-heterogeneous treatment effects in our case. Most importantly, when looking at the placebo test of the joint common trend and non-anticipation assumption, there is no strong evidence of the violation with a p-value of the joint nullity of the pre-treatment effects of 0.12. When looking at the alternative treatment definition based on the firm location, the results are quite similar both in terms of size and significance level (C&H firm: R$ 985, pval<0.01). It should be noted that in this latter case, the test of the common trend assumption does not reach the 5% significance level, even though it is close, and indeed, the individual placebo treatments reach significance. This suggests that although not overwhelming, in particular for the treatment based on firm location selection, issues will remain. While we will dig further into this issue in the robustness section, we nonetheless conclude that most evidence is in favour of positive treatment effects, while the concern of selection issues is limited.

Treatment	ATT	SE	z-val.	p-val. placebo (joint)	Coef. placebo (t-2)	Coef. placebo (t-1)
Institute location	1210.49**	381.86	3.17	0.12	-301.79	324.20
Firm location	985.08**	339.39	2.90	0.06	-534.30*	-355.40*

Table 3: Staggered DiD estimation

* p<0.05, ** p<0.01, *** p<0.001

Event-study representations of the two DiD estimations are presented in Figure 5. What is interesting about these figures is that the effects are increasing over time, with the first two treatments being small and insignificant, and the latter being the largest, at more than 5000 BRL. This is well in line with the intuitive principle that the effects of scientific research on the economy do not occur instantaneously and will take time to unfold.



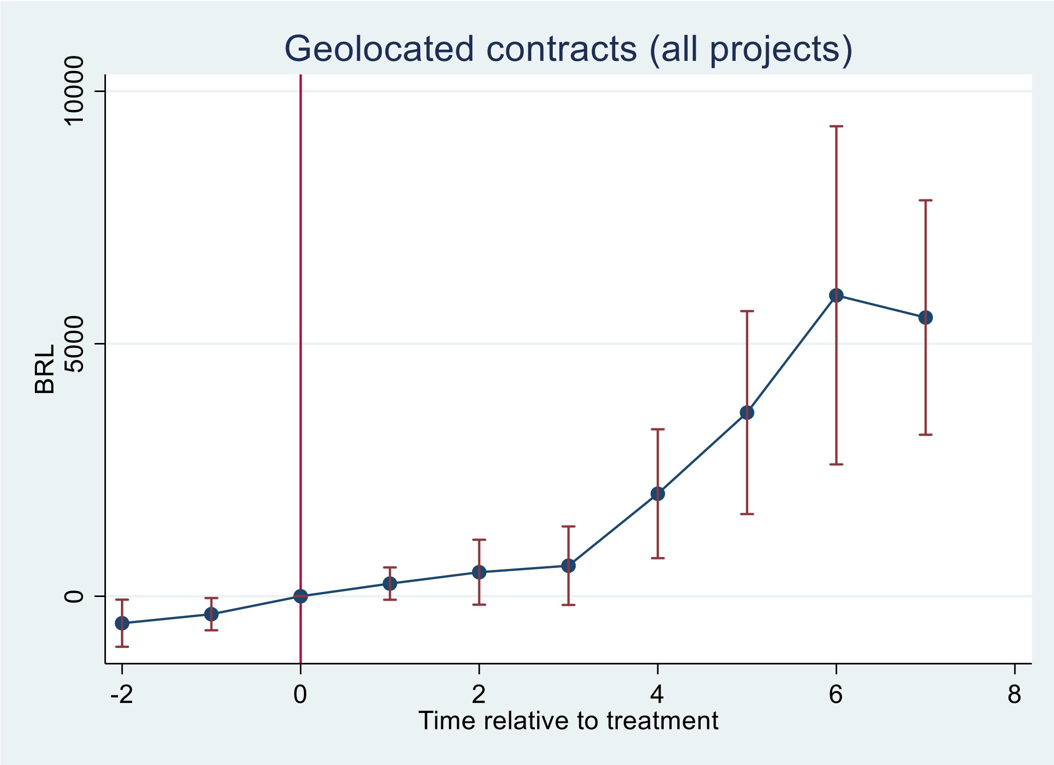


Figure 5: Event-study plots of the baseline DiD estimations

4.1.3 Robustness Checks

4.1.3.1 Further Assessments of Parallel Trends

Although the existing estimates of the last subsection provided useful evidence in favour of parallel trends, we now provide a series of more detailed placebo tests. One concern that is pertinent to the so far used methodology is that it allowed in our sample only for very few pre-treatment placebos, which is the result specific formulation allowing for several switches between treated and untreated states in the same unit. This generality is necessary for treatment definition based on firm location, but it is not necessary for the institute location. In that case, we can also resort to simpler estimators, which consider units as consistently treated from the first treatment onwards. One estimator that is suitable in this simple case, while still allowing for staggered treatments with potentially time-heterogeneous treatment, is the one by Callaway and Sant'Anna (2021). Indeed, when this estimator with the doubly robust option based on inverse probability weighting of tilting (Sant'Anna and Zhao 2020) is used, we obtain the following event study plot.

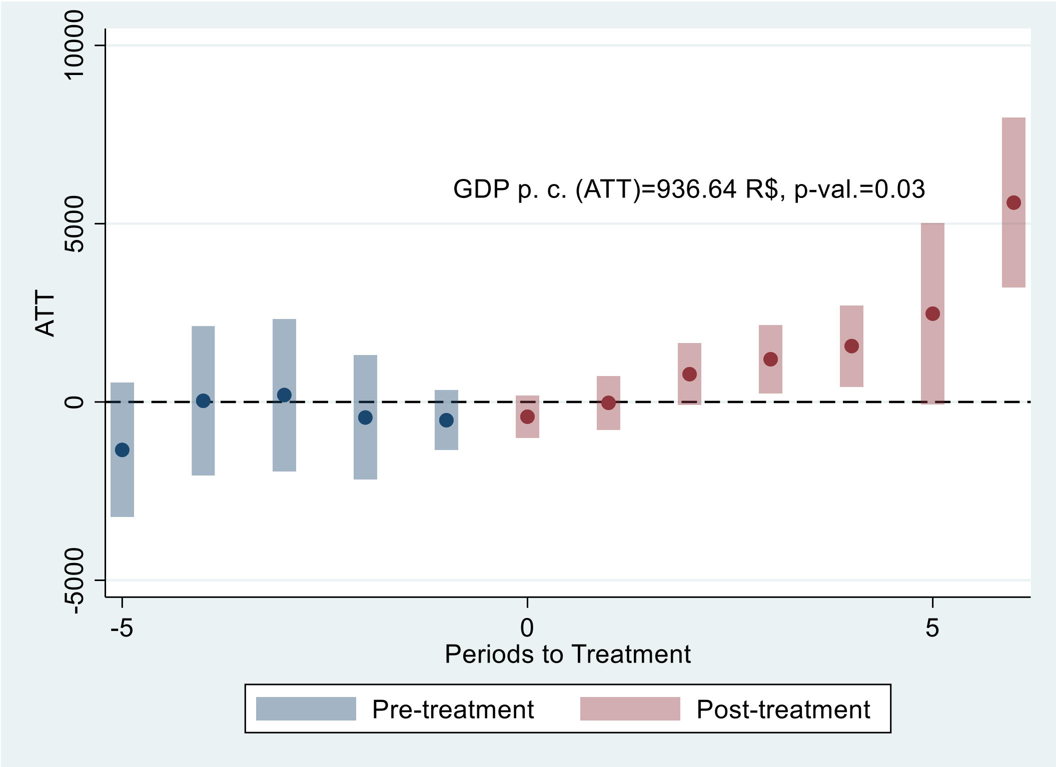


Figure 6: Event-study plot of the baseline DiD for the institute location based on the Callaway and Sant'Anna estimator

The effect is with 936.64 BRL a bit lower than those resulting from the method by Chaisemartin and d'Haultfoeuille (2024). Yet, it is still significant. Moreover, the method allowed in our sample for up to five pre-treatment placebos, which are all non-significant. The overall p-value of joint significance was accordingly not showing signs of significance with a p-value of 0.19. Thus, even based on estimators allowing for the identification of more pre-treatment placebos, we do not see obvious violations of the parallel trends assumption.

In our empirical staggered setting, it is also possible to derive further assessments of the parallel trends assumption by exploiting the fact that we have two different kinds of control group units. These are firstly the never-treated regions and secondly the treated regions before treatment. By default, DiD estimators use both types of control units, assuming that they are similar. This assumption may or may not hold in practice. It can, however, be assessed by a placebo-type of test. Specifically, we restrict the sample to the control units, irrespective of their type. Then, we define group-wise placebo-treatments for each of the two groups in a way that the other group is treated as the control group. Obviously, if the assumption is correct that both groups are valid control groups, all placebo-treatments for every year and group should be insignificant. To implement this idea, we run a regular two-way fixed effects (TWFE) model taking the placebo treatments as explanatory variables. Under the Null-hypothesis of the equivalence of the two control groups, TWFE should be consistent even in a staggered design, because the treatment effects are consistently zero. The results are presented in Table 4

	(1) F1.GDP per capita (treatment group: never treated)	(2) F1.GDP per capita (treatment group: treated before treatment)
Treatment 2012	NA	-764.58180*
	NA	(-2.09)
Treatment 2013	-674.10303	1391.80446**
	(-1.38)	(2.86)
Treatment 2014	-29.30634	-377.22131
	(-0.06)	(-0.64)
Treatment 2015	82.91136	-150.70043
	(0.17)	(-0.12)
Treatment 2016	-517.63430	-337.49234
	(-1.06)	(-0.21)
Treatment 2017	-591.00649	-515.33476
	(-1.22)	(-0.33)
Treatment 2018	-239.51290	NA
	(-0.57)	NA
L1.GDP growth	10904.76901***	10968.36380***
	(15.83)	(15.89)
F1.GDP growth	16420.90583***	16316.85214***
	(24.19)	(24.04)
Population	0.00196	0.00216
	(1.02)	(1.11)
Constant	-499.32465	-1193.06037
	(-0.36)	(-0.86)
Year dummies	Yes	Yes
Observations	4080	4080

Table 4: Placebo-treatment tests on the equivalence of never-treated and treated before treatment as control units (TWFE-estimation)

t-values in parantheses, * p<0.05, ** p<0.01, *** p<0.001

Indeed, our results do not provide strong evidence of a non-equivalence. In the first column, we take the never-treated units as treated and the before-treatment treated units as controls. In this setting, all placebo-treatments are individually insignificant. In the second column, we take the treated before treatment groups as the treatment units and take the never treated as controls. While we do observe that the first two placebo treatments are significant, the remaining are not and indeed the placebo treatments are jointly insignificant. Thus, overall, evidence of a non-equivalence of both control groups is at best marginal, reassuring us of this crucial assumption of the staggered design of the previous section.

A final assessment of robustness towards the failure of parallel trends results from the notion that parallel trends in practice will not hold exactly but only to varying degrees. An important question that results from this notion is how strong violations of parallel trends would need to be to invalidate the finding of significant treatment effects. Assessing this question is possible by the method proposed by Rambachan and Roth (2019), which is implemented in the honestdid package for STATA. This method first estimates the existing violations of pre-treatment parallel trends. Then, it determines a proportionality factor M giving the percentage of post-treatment violations of parallel trends that would invalidate the claim that the treatment effects are significantly different from zero. In short, this method allows for a continuous assessment of how robust the results are to violations of the parallel trends assumption. It is common to take a proportionality factor of M at least 100% as an indication of high robustness. This would mean that the post-treatment violations are robust to violations of parallel trends and it would need to be larger than the maximum detected in the pre-treatment period. When applying this method, the results indicate that for the last estimated treatment effect in post-treatment period 7, M is bound between 105 and 110%. This implies that our findings are robust even to increases in violations of the parallel trends over what was maximally observed before treatment.

4.1.3.2 Differences in Dosage

We know from the descriptive statistics that the intensity of treatment differs considerably between the treated units. In some regions, the treatment is assigned on the basis of just one institute. In other treated regions, such as São Paulo, there are several institutes. Also, there are differences concerning the contract location. In some regions that have been assigned a positive treatment status, there is only one or a few contracts, while the observed maximum is 1209. In simple terms, the treatment "dosage" differs substantially. A solution that has been established so far to address this issue is to define treatment status based on varying cut-offs. Yet this approach remains, of course, arbitrary. However, the estimator by Chaisemartin and d'Haultfoeuille (2024) allows for an extension of DiD's dichotomous treatment logic to the continuous case. We see in Table 5 that the overall results are confirmed, documenting a positive and significant effect overall, with the estimates indicating that an additional institute, on average, raises GDP per capita by 356 BRL and an additional contract by 51 BRL. The p-values on the placebo tests are, as before, not significant and therefore do not provide strong evidence against the common trend and non-anticipation assumptions.

Treatment	ATT	SE	z-val.$^{+}$	p-val. placebo (all)	Coef. placebo (t-2)	Coef. placebo (t-1)

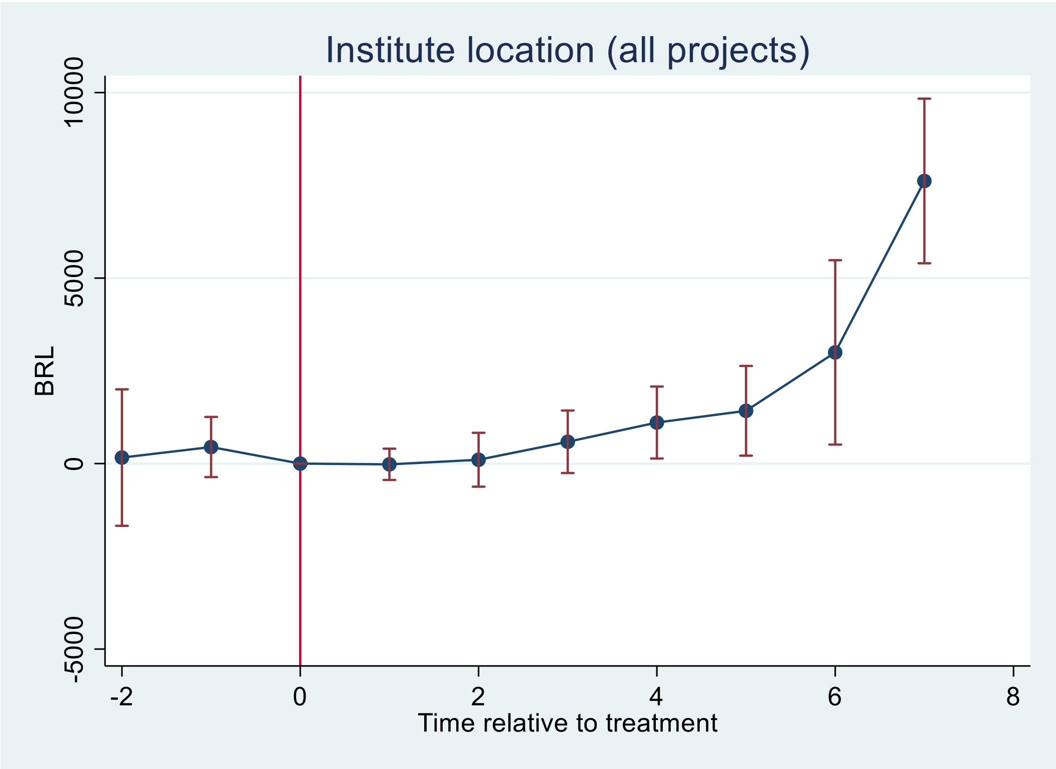
Number of institutes	356.67*	162.83	149.39	-301.79	324.20	299.17
Number of contracts	51.84**	238.19	20.51	-534.30*	-355.40*	-355.40**

Table 5: Staggered DiD estimation (continuous treatment)

* p<0.05, ** p<0.01, *** p<0.001

+ There is currently no proven asymptotic normality result for the continuous DiD estimator of the ATT. The z-values and resulting significance levels are reported only as points of reference.

Although the continuous estimator suggests that effects remain significant once we allow for continuous treatment, a further validation of the previously used binary treatments seems reasonable. In particular, our results indicated that although one of joint placebo-tests in Table 3 was close to significance, identifying the source of potential selection issues is crucial. To find out more, we have rerun the binary treatment models but excluding regions with very large treatments. The reason is that selection may be more pronounced for regions where the strategic investments were large. If our reasoning is correct, we would expect that the p-values of the placebo tests decline. Moreover, to warrant that the estimated effects are not just the result of residual uncontrolled selection, in the restricted sample, the treatment effects should be smaller (because large treatments are excluded) but still significant (because also small treatments should have an effect). The results of this rerun estimation are presented in Figure 7.


Institute location (all projects)
BRL
10000
5000
0
-5000
-2
0
2
4
6
8
Time relative to treatment

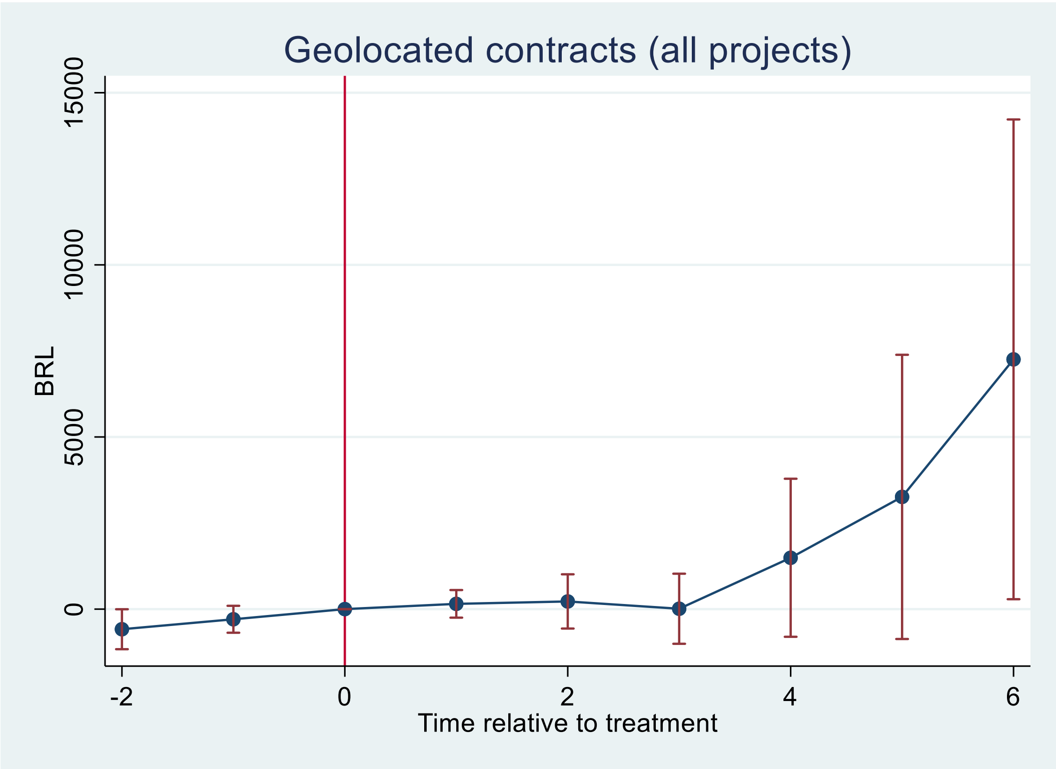

Geolocated contracts (all projects)
BRL
15000
10000
5000
0
-2
0
2
4
6
Time relative to treatment

Figure 7: Event-study plots of the baseline DiD estimations excluding regions with large investments

Indeed, the results confirm our expectations: first, we still observe significant effects. Most pronounced are these for the specification building on the institute location, where the results appear not to differ a lot from the unrestricted sample. For the contract location, we see that despite an increasing trend, there is a significant spike only for the last estimated period. Second, we find that the estimated treatment effects are smaller overall. For the institute location, it is 1098.48 BRL on average (p-val.<0.05). For the firm-location, the average effect dropped to 398.58 BRL, which is still positive but not significant anymore (p-val.>0.05). Again, this is expected to some degree because excluding large treatments should, of course, also reduce average treatment effects in the restricted population, which is treated at lower dosages. This implies that these estimates can be considered to be (potentially overly) conservative lower bounds, where, at least for later years, the effects are still positive and significant. Third, when looking at the p-values of the pre-treatment placebos, we indeed see a decline. For the institute location, in the restricted sample, the p-value declined from originally 0.11 to 0.29. In the firm-location specification, it dropped from 0.06 to 0.14, which suggests that indeed, selection effects may play a role, particularly for regions with very large investments. Nonetheless, overall, the general claim that treatments are even in conservative settings positive is corroborated.

4.1.4 Heterogeneity across Project Type

As we have seen, the SENAI institutes cover quite a large portfolio of activities that range from consultancy projects and technical services such as metrology to genuine R&D projects. The bulk of these activities corresponded to the second category of metrology. An important question from the policy as well as from the managerial side is whether the effects differ by type of project. The treatment definition based on firm location allows for an easy test of differential effects by splitting the treatment definition by project type; i.e., in one variant, the definition of treatment would be based only on consultancy projects, in another, only on metrology projects, and so on. If treatment effects were homogeneous, all estimated ATTs would not differ substantially.

Figure 8 presents the associated event-study plots by project type. Although there are a few significant spikes, for consultancy and metrology projects at least, only R&D projects show a clear positive pattern. Indeed, the overall ATTs are only significant for this group of projects (ATT=1800 BRL, pval<0.001), while the ATTs of the other project categories are not. Thus, the results suggest that the overall positive effects of the SENAI institutes are, by far and large, driven by more basic R&D projects rather than consultancy or technical services such as metrology. Such a finding is also in line with findings that the value of scientific organisations comes from their ability to create and make available fundamental advanced knowledge and technologies rather than to generate income from side hustles that, in principle, could also be provided by commercial firms.

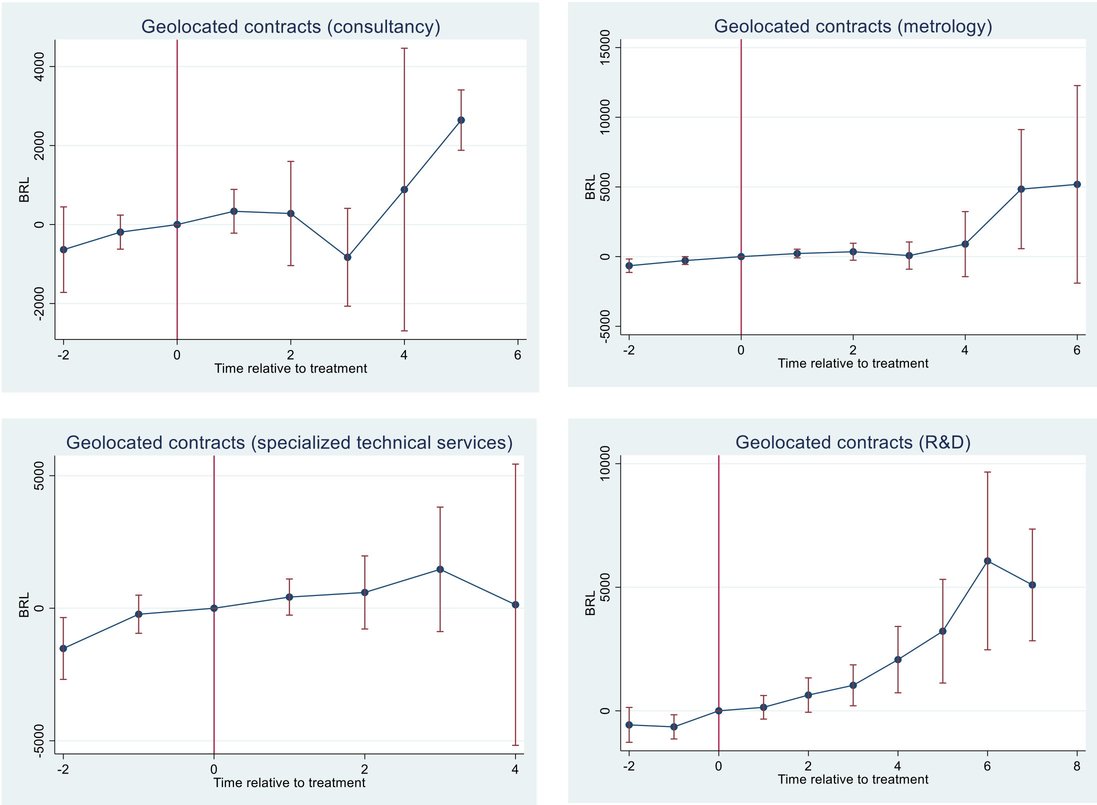


Figure 8: Event-study plots of the baseline DiD estimations by project type

4.2 The Overall Macroeconomic Value of SENAI

Like the study by Schubert and Kroll (2016a), our data represents a regional breakdown of the whole total economy and thus does not represent a mere sample. This feature is desirable in order to derive simple estimates of the overall value of SENAI for the Brazilian economy as opposed to just the treated regions. It is, for example, straightforward to provide results for the share of Brazilian GDP explained by the presence of SENAI institutes. It is, however, important to emphasise that these calculations are tentative and should be seen as indicative of the range of the effect size rather than exact estimates.

	2011-2020

ATT: GDP per capita	985.00
Average share of treated regions	13.45%
Average population	399,705
Average population in treated regions	1,060,365
Population-weighted share of treated regions	35.68%
ATE (Brazil)	351.48
Average GDP per capita in Brazil 2011-2022	52,894.74
Share of explained GDP per capita	0.66%

Table 6: Share of Brazilian GDP explained by SENAI

Sources: own calculations, Worldbank

Such a calculation is presented in Table 6. We assume the lowest estimated ATT resulting from contractor locations of 985 BRL. Over the course of the observation period, an average of 13% of the regions were treated each year. However, since these regions were, on average, larger by a factor of about 2.65, the population-weighted share of treated regions was 36%. Assuming that the treatment effect of the untreated is zero,[4] we can calculate an estimate of the average treatment effect for Brazil (ATE) by multiplying the ATT by the population-weighted share of treated regions. If we compare the ATE to the overall average GDP of 52,894 BRL, we find that approximately 0.66% of the Brazilian GDP can be attributed to SENAI. This estimate is substantial, particularly in comparison to the overall budget. However, it falls well in line with, for example, the results for the (admittedly much larger) Fraunhofer Society in Germany, to which 1.6% of total GDP (Allan et al. 2022) and 0.55% of labour productivity may be attributed (Comin et al. 2019). It is also in line with research indicating that the returns on innovation-related expenditures may induce multipliers that are often in the double digits (Jones and Summers 2020).[5]

[4] This assumption is an approximation and may be violated if, for example, treatment leads to relocation of economic activities to treated regions (negative regional spillovers) or if induced economic activities in treated regions also drive economic activities in untreated regions (positive regional spillovers).

5 We refrain, however, from reporting such multipliers because of quality concerns about cost data in the internal bookkeeping system.

5 Discussion and Conclusion

The literature on the economic effects of science is well established and has mostly pointed to substantial effects for a wide array of variables, including GDP (Schubert and Kroll 2016b; Agasisti and Bertoletti 2022; Glückler et al. 2015), productivity (Comin et al. 2019; Fritsch and Wyrwich 2018), and innovativeness (Robin and Schubert 2013b). However, generating credible causal evidence has always been a non-trivial challenge because scientific organisations tend to have long histories, often coevolving with or sometimes even creating their environments. This coevolution tends to blur clear cause-and-effect relationships and makes it difficult to understand the direction of causality. Because of this, many studies have, arguably, struggled to make compelling arguments as to why the observed associations can validly be interpreted as causal effects.

Recently, a few studies have employed the foundation of scientific organisations in difference-in-differences approaches (Pfister et al. 2021; Schlegel et al. 2022), which provides a promising framework to address the endogeneity challenges. These studies are, however, still limited in number and naturally focus on very specific cases, all of them in Western economies. Our work adds to this stream of literature in two important ways. Firstly, our unique institutional insights into the historic locational selection processes allow us to make a very credible case for the validity of the parallel trends assumption, going beyond mere empirical corroboration by placebo tests. So far, existing studies have focused on broader types of organisations, such as the Universities of Applied Sciences in Switzerland, which, despite sharing a common organisational form, are heterogeneous and highly independent of each other from a legal perspective. Thus, in these studies, an institutional assessment of locational selection criteria naturally faces limitations. Secondly, our specific setting of the SENAI ISI institutes is one of the very few instances where evidence is provided for a catch-up economy.

Overall, our results confirm that public research organisations can have profound impacts on the economy. Focusing on productivity, we estimated the average treatment effect on the treated regions (ATT) in terms of GDP per capita to be at least 985 BRL. This effect was robust in the light of various specification alternatives. It is also economically substantial, given that it implies that about 0.66% of the total Brazilian GDP can be attributed to the SENAI ISI institutes. Indeed, this effect may appear very large given the fact that investment levels are still rather limited, leading to enormous returns on investment. Yet, at the same time, it is known that the returns on innovation and knowledge-related activities are very large. Jones and Summers (2020) report high multipliers of 10–30. In an emerging economy such as Brazil, these effects may even be amplified because, up to now, organisations closing the gap between universities and firms, which are an established part of innovation systems in most advanced economies, have been largely absent. Thus, it is not unreasonable to assume that the SENAI ISI institutes closed a systemic gap in the Brazilian innovation system, which explains the considerable returns. Needless to say, decreasing returns may, of course, set in once the ISI activities are sufficiently expanded. However, at least while spending remains at its current limited level, our results suggest that increased investments appear to be economically warranted.

A highly important observation with regard to high-relevance public research organisations in general concerns the question of which types of projects cause the effects. Most public research organisations — including those in other countries, e.g., Fraunhofer in Germany, SINTEF in Norway, TNO in the Netherlands, or RISE in Sweden — engage in a wide variety of activities ranging from genuine research and technological development projects to the provision of technical services, such as metrology, or consultancy. All of these activities can potentially provide benefits for industrial partners. However, our results clearly suggest that the economic effects almost exclusively result from genuine research projects. Thus, the social or economic value of public research organisations seems to be associated with their capacity to contribute to research, technology development, and innovation. It is apparently much less related to their capacity to provide technical services or consultancy, which commercial actors could also provide.

This is not to say that public/private research organisations, such as SENAI, should not engage in these activities. Indeed, in many cases, organisations may be critically reliant on revenue streams coming from the provision of services to maintain their operations. However, it does have important implications for management and policy. Concerning the managerial aspects, the provision of technical services should not be a goal in itself but a means to sustain a scientific organisation that sees its ultimate objective as providing collaborative applied research with and for industrial partners. Regarding policy, our results suggest that public funding should primarily be linked to the performance of research-oriented activities. The case for funding technical services or consultancy, which can essentially also be provided by commercial suppliers, is probably less compelling.

Needless to say, even if we provided an extensive list of robustness and validity checks, our results may depend on modelling assumptions as well as selectivity issues. Crucial is, for example, the modelling decision on how to define treatment. While the per-region ATT does not seem to be affected strongly, the macroeconomic value of SENAI-ISI is, because our treatment definitions differ considerably in terms of the share of treated regions. In our study, we have opted to report only the lower macroeconomic estimates, which therefore represent conservative bounds. However, a more in-depth understanding of treatment would help to better understand the reach of treatment. The firm-location definition of treatment effects considers only direct treatments administered to the location of the firm. The implicit assumption is that all potential spillover effects running from the treated firm are limited to the region in which the firm is located. Yet, if effects spill over to neighbouring regions, too, the reach may be underestimated. That this is likely the case is suggested by the institute-location based treatment definition. Yet, overall, the geography of the reach of the effects needs to be better understood in order to make statements with great confidence. On the side of selectivity: Although formal tests provided in no case formally present significant evidence against the validity of the parallel trends assumption, in certain subgroups (e.g., regions with particularly high SENAI investments) or for specific treatment definitions (treatment based on firm-location), the issues may be stronger. Thus, even this unusually clean case exploiting the staggered foundation of the SENAI-ISI institutes is not beyond all methodological doubts, which suggests that more evidence from a wider variety of settings would help to strengthen further the confidence in the economic effects of scientific research.

6 Appendix

Variable	N	T	Mean	Std. Dev.	Min	Max
GDP per capita (BRL)	5100	10	22489.00	14601.50	3450.25	203328.00
Treatment: Institute location (all projects)	5100	10	0.51	0.50	0.00	1.00
Treatment: Geolocated contracts (all projects)	5100	10	0.13	0.34	0.00	1.00
Treatment: Geolocated contracts (R&D)	5100	10	0.05	0.21	0.00	1.00
Treatment: Geolocated contracts (metrology)	5100	10	0.11	0.31	0.00	1.00
Treatment: Geolocated contracts (specialized technical services)	5100	10	0.02	0.14	0.00	1.00
Treatment: Number geolocated contracts	5100	10	3.45	39.42	0.00	1209.00
Treatment: Number institutes	5100	10	1.30	1.45	0.00	4.00
GDP growth	4590	9	1.08	0.09	0.54	2.56
Population	5100	10	399705.00	1200000.00	27711.00	22000000.00

Table 7: Descriptive Statistics of the main Variables

7 Publication bibliography

Agasisti, Tommaso; Bertoletti, Alice (2022): Higher education and economic growth: A longitudinal study of European regions 2000–2017. In *Socio-Economic Planning Sciences* 81, p. 100940.

Allan, Grant; Figus, Gioele; Schubert, Torben (2022): Understanding the Macroeconomic Effects of Public Research: An Application of a Regression Microfounded CGE-Model to the Case of the Fraunhofer-Gesellschaft in Germany.

Anselin, Luc; Varga, Attila; Acs, Zoltan (1997): Local geographic spillovers between university research and high technology innovations. In *Journal of urban economics* 42 (3), pp. 422–448.

Bertoletti, Alice; Berbegal-Mirabent, Jasmina; Agasisti, Tommaso (2022): Higher education systems and regional economic development in Europe: A combined approach using econometric and machine learning methods. In *Socio-Economic Planning Sciences* 82, p. 101231.

Callaway, Brantly; Sant’Anna, Pedro H. C. (2021): Difference-in-differences with multiple time periods. In *Journal of econometrics* 225 (2), pp. 200–230.

Chaisemartin, Clément de; d'Haultfoeuille, Xavier (2024): Difference-in-differences estimators of intertemporal treatment effects. In *Review of Economics and Statistics*, pp. 1–45.

Comin, Diego; Licht, Georg; Pellens, Maikel; Schubert, Torben (2019): Do companies benefit from public research organizations? The impact of the Fraunhofer Society in Germany. In *The Impact of the Fraunhofer Society in Germany*, 19-006.

Fritsch, Michael; Wyrwich, Michael (2018): Regional knowledge, entrepreneurial culture, and innovative start-ups over time and space—an empirical investigation. In *Small Business Economics* 51, pp. 337–353.

Ghanem, Dalia; Sant'Anna, Pedro H. C.; Wüthrich, Kaspar (2022): Selection and parallel trends. In *arXiv preprint arXiv:2203.09001*.

Glückler, Glückler; Panitz, Robert; Wuttke, Christian (2015): Die wirtschaftliche Wirkung der Universitäten im Land Baden-Württemberg. In *Raumforschung und Raumordnung| Spatial Research and Planning* 73 (5), pp. 327–342.

Goodman-Bacon, Andrew; Nichols, Austin; Goldring, Thomas (2019): Bacon decomposition for understanding differences-in-differences with variation in treatment timing. In *NBER Working Paper* (25018).

Griliches, Zvi (1958): Research costs and social returns: Hybrid corn and related innovations. In *Journal of political economy* 66 (5), pp. 419–431.

Jones, Benjamin F.; Summers, Lawrence H. (2020): A calculation of the social returns to innovation: National Bureau of Economic Research.

Kohl, Holger; Will, Markus; Prim, Marcelo Fabricio; Pavim, Alberto Xavier (2020): Building up a national network of applied R&D institutes in an emerging innovation system. In *Production* 30, e20190151.

Maietta, Ornella Wanda (2015): Determinants of university–firm R&D collaboration and its impact on innovation: A perspective from a low-tech industry. In *Research Policy* 44 (7), pp. 1341–1359.

Pfister, Curdin; Koomen, Miriam; Harhoff, Dietmar; Backes-Gellner, Uschi (2021): Regional innovation effects of applied research institutions. In *Research Policy* 50 (4), p. 104197.

Rambachan, Ashesh; Roth, Jonathan (2019): An honest approach to parallel trends. In *Unpublished manuscript, Harvard University*.

Robin, Stéphane; Schubert, Torben (2013a): Cooperation with public research institutions and success in innovation: Evidence from France and Germany. In *Research Policy* 42 (1), pp. 149–166.

Robin, Stéphane; Schubert, Torben (2013b): Cooperation with public research institutions and success in innovation: Evidence from France and Germany. In *Research Policy* 42 (1), pp. 149–166.

Sant'Anna, Pedro H. C.; Zhao, Jun (2020): Doubly robust difference-in-differences estimators. In *Journal of econometrics* 219 (1), pp. 101–122.

Schlegel, Tobias; Pfister, Curdin; Harhoff, Dietmar; Backes-Gellner, Uschi (2022): Innovation effects of universities of applied sciences: an assessment of regional heterogeneity. In *The Journal of Technology Transfer* 47 (1), pp. 63–118.

Schubert, Torben; Kroll, Henning (2013): Endbericht zum Projekt „Hochschulen als regionaler Wirtschaftsfaktor ". In *Im Auftrag von Stifterverband für die Deutsche Wissenschaft. Karlsruhe*.

Schubert, Torben; Kroll, Henning (2015): Ökonomischer Wert der Relativitätstheorie? Nur Papier und Bleistift? In *Wissenschaftsmanagement*.

Schubert, Torben; Kroll, Henning (2016a): Universities' effects on regional GDP and unemployment: The case of Germany. In *Papers in Regional Science* 95 (3), pp. 467–490.

Schubert, Torben; Kroll, Henning (2016b): Universities' effects on regional GDP and unemployment: The case of Germany. In *Papers in Regional Science* 95 (3), pp. 467–490.