Department of Economics

Athens University of Economics and Business

WORKING PAPER no. 07-2025

The heterogeneous causal effects of the EU’s Cohesion Fund

Angelos Alexopoulos , Ilias Kostarakos,
Christos Mylonakis, and Petros Varthalitis

March 2025

The heterogeneous causal effects of the EU's Cohesion Fund*

Angelos Alexopoulos †1, Ilias Kostarakos‡2, Christos Mylonakis§1, and Petros Varthalitis¶1

[1]Athens University of Economics and Business, Department of Economics, Athens, Greece
[2]European Commission, Joint Research Centre

March 14, 2025

Abstract

This paper quantifies the causal effect of cohesion policy on EU regional output and investment focusing on one of its least studied instruments, i.e., the Cohesion Fund (CF). We employ modern causal inference methods to estimate not only the local average treatment effect but also its time-varying and heterogeneous effects across regions. Utilizing this method, we propose a novel framework for evaluating the effectiveness of CF as an EU cohesion policy tool. Specifically, we estimate the time varying distribution of the CF's causal effects across EU regions and derive key distribution metrics useful for policy evaluation. Our analysis shows that relying solely on average treatment effects masks significant heterogeneity and can lead to misleading conclusions about the effectiveness of the EU's cohesion policy. We find that the impact of the CF is frontloaded, peaking within the first seven years after a region's initial inclusion in the program. The distribution of the effects during this first seven-year cycle of funding is right skewed with relatively thick tails. This indicates positive effects but unevenly distributed across regions. Moreover, the magnitude of the CF effect is inversely related to a region's relative position in the initial distribution of output, i.e., relatively poorer recipient regions experience higher effects compared to relatively richer regions. Finally, we find a non-linear relationship with diminishing returns, whereby the impact of CF declines as the ratio of CF funds received to a region's gross value added (GVA) increases.

Keywords: Cohesion policy; Growth; Regional transfers
JEL Classification: O18; O40; R11; R58

*The views expressed in this paper are those of the authors alone and do not represent the official views of the European Commission. All remaining errors are our own.

†angelos@aueb.gr
‡ilias.kostarakos@ec.europa.eu
§mylonakisc@gmail.com
¶pvarthalitis@aueb.gr

1 Introduction

Cohesion policy is one of the most important fiscal policy initiatives at the European Union (EU) level. Over the last three decades, nearly €960 billion (in 2015 prices) have been disbursed to the EU regions as part of this sub-national transfers program. To put things in a historical perspective, cohesion-related expenditures account for almost one-third of the EU's budget (see Benedetto (2024) and Bachtler (2022)), making it one of the largest redistribution programs ever implemented in the EU. In the current seven-year budget period,[1] €392 billion have been dedicated to cohesion policy objectives, reinforcing its status as a top spending priority of the EU. Naturally, given its scale and overarching objectives, the effectiveness of cohesion policy has been debated over time by both academics and policymakers (see European Commission (2024)).

This paper quantifies the causal effect of cohesion policy on EU regional output and investment. By adopting recent developments in the causal inference and potential outcomes literature, i.e., Xu (2017), Athey et al. (2021) and Borusyak et al. (2024),[2] we extend beyond the existing literature (e.g., Becker et al. (2010, 2013, 2018)) which has primarily focused on estimating a common effect across all treated regions, i.e. the (local) average treatment effect. Our empirical strategy allows us to capture not only the average impact of cohesion policy but also its time-varying and heterogeneous effects across regions, providing a more nuanced understanding of its effectiveness.

We focus on one of the least studied instruments of cohesion policy: the Cohesion Fund (CF). This policy instrument is solely available to the relatively poorer Member States and has disbursed, over time, almost 20% of total cohesion policy expenditures. Our paper introduces a novel approach to evaluate its core objectives, which go beyond simply boosting income across recipient regions to actively address economic disparities among EU regions. To do this, we utilize the strengths of our methodology, which allows us to estimate the impact of each region for every year following its inclusion in the CF program. This enables us to estimate the time-varying distribution of CF effects across regions and derive key distributional metrics that are useful for policy evaluation. Specifically, we compute the percentage of regions that experience a positive (or negative) impact. Moreover, we examine the skewness of the distribution which indicates whether the CF program has predominantly positive or negative effects, along with the kurtosis which measures the degree of uneven dispersion of CF effects across regions. We do this over consecutive funding cycles aligning with the EU's institutional seven-year budgeting framework. In the context of the CF's policy evaluation, a decline in the percentage of regions experiencing a positive impact over time may suggest a diminishing effectiveness of the CF. However, this should not always be interpreted at face value but, rather, should be assessed alongside additional distributional metrics. Specifically, a simultaneous decrease in skewness and/or kurtosis—indicating a more symmetric distribution—may instead suggest that the CF effect becomes more symmetric as a result of reducing regional disparities over time.

The identification of the causal effects of such regional transfer programs has mainly been conducted by relying either on panel data methods related to the Differences-in-Differences (DiD) approach (see for example Gobillon and Magnac (2016) for a discussion) or by employing a Regression Discontinuity Design (RDD) to assess, for example, the impact of the cohesion policy expenditures across the regions of the EU. These papers leverage on the institutional rule that determines whether a region is eligible or not for the relevant EU objectives (e.g. the so-called Objective I or Convergence Objective), which in turn determines the

[1] The so-called Multiannual Financial Framework.

[2] For a recent review see Xu (2023)

corresponding size of the cohesion transfers - e.g. see Becker et al. (2010, 2012) - more details on this are provided in section 2.

However, in the analysis of regional policies there are specific issues which require careful consideration. The first issue, which has received quite a lot of attention, is that of the spatial dependence across units. These dependencies arise mainly due to geographical reasons and are local in nature. A number of studies have employed standard spatial econometrics techniques to account for these issues (see, among others, Amendolagine et al. (2024), Bourdin (2018), Crescenzi and Giua (2020), Di Caro and Fratesi (2022) and Mohl and Hagen (2010)) to account for these factors. To our knowledge, only Crescenzi and Giua (2020) have attempted to account for these dependencies in a causal inference contaxt, by employing a spatial RDD in order to estimate the causal effects of cohesion policy.

A second important issue is that of cross-sectional dependence, highlighted by Gobillon and Magnac (2016). Specifically, cross-sectional dependence is a form of time-varying heterogeneity which arises from the presence of unobserved common factors whose impact extends well beyond a local geographical area. Examples of such factors that are prominent in our sample include the Global Financial Crisis of 2008 and the debt crisis of the southern periphery of the EU. These unobserved common factors can lead to severe biases if they are correlated with the treatment variable.

Lastly, it is important to recognize that regional policies are more likely than not to exert region-specific effects. To our knowledge, the relevant empirical literature on the policy evaluation of Cohesion transfers has largely assumed that the impact of the EU's regional policy is identical across all regions. Essentially, this amounts to disregarding any heterogeneity that arises due to differences in the relative position of regions across the income distribution along with a wide range of other idiosyncratic characteristics.

In this paper, we overcome these challenges by drawing on recent developments from the causal inference and potential outcomes literature in order to estimate the causal effects of the Cohesion Fund payments. Specifically, our analysis builds on the matrix completion technique implemented in the context of factor models - see, Athey et al. (2021) and Xu (2017). The factor or interactive fixed effects models (see Bai (2009a), Pesaran (2006)) nest the standard Two-Way Fixed Effects model and can directly account for the presence of cross-sectional dependence and the potential correlation between the treatment and the outcome (see Gobillon and Magnac (2016)).

Our approach is a generalization of the synthetic control method of Abadie et al. (2010), which relaxes key assumptions of standard causal inference methods. In particular, by employing the synthetic control methodology we avoid the imposition of the so-called *parallel trends* assumption which is central to all the techniques that rely on the Differences-in-Differences approach -see, Angrist and Pischke (2009), De Chaisemartin and d'Haultfoeuille (2024), Sun and Abraham (2021). It is worth mentioning that the presence of the time-varying unobserved common factors, which are an integral part of our approach, would invalidate this assumption (see Xu (2017)). In our setup, the interactive fixed effects will directly model and account for these confounders. Moreover, under this generalized approach, the staggered adoption of the treatment from more than one units is allowed.

Our main findings are as follows. We find that, on average, the CF has a persistent positive effect on both output (Gross Value Added) and investment (Gross Fixed Capital Formation) of the recipient regions, compared to a counterfactual scenario where no funds are provided by the EU. The impact of the CF is frontloaded, peaking within the first seven years after a region's initial inclusion in the program. For a hypothetical average EU region, the CF led to an increase in GVA per capita of €468 compared to the counterfactual scenario. In the second seven-year cycle, i.e. after 14 years of inclusion in the CF program, the positive effect persists but diminishes in magnitude, with GVA per capita increasing by €330 compared to the counterfactual. In other words, a citizen in this hypothetical EU region will be wealthier by €468 in the 7th year after the inclusion to the program and by €330 in the 14th year. Similarly, the hypothetical average EU region, experiences an increase in GFCF per capita of €385 and €415 after 7 and 14 years, respectively, compared to the counterfactual.

However, our analysis shows that relying solely on average treatment effects masks significant heterogeneity and can lead to misleading conclusions about the effectiveness of this EU policy tool. We provide novel evidence on the distribution of the impact of the Cohesion Funds across EU regions and its evolution over consecutive seven-year cycles (mimicking the EU budgeting cycles). Specifically, the distribution of the GVA effect during the first seven-year cycle of funding is right skewed with relatively thick tails. This indicates positive GVA effects which are unevenly distributed across regions. In the second cycle (after 14 years of treatment) both the skewness and the kurtosis of the distribution decrease. This indicates that the magnitude of the CF effect becomes smaller, while its impact is more symmetric across regions. We quantify the distributional effects and, specifically, we find that in the first seven-year cycle 64.5% of treated regions experience a positive effect, while this percentage reduces to 49% in the second seven-year cycle. When we categorize regions into percentiles based on their initial GVA, we find that the poorest regions – those in the bottom 0-10th percentile– experience a significantly larger in magnitude positive effect from receiving Cohesion Funds. Specifically, their GVA after the first seven-years of receiving the fund is estimated to grow by 21.5 percentage points more compared to the counterfactual scenario. For regions in the 10th-25th and 50th-75th percentile, the gains are 18.8 and 12.3 percentage points higher, respectively, compared to the counterfactual. However, the benefits diminish for the relatively wealthier regions and even turn to negative for the wealthiest EU regions. Another important finding is that the positive effect persists in the second seven-year cycle, i.e., after 14 years under treatment; however, it becomes significantly smaller in magnitude. Turning to the GFCF distribution, most of the patterns observed for GVA hold. However, we observe two key differences. First, a higher percentage of the treated regions experience a positive effect within the first seven years under treatment, namely 90%. Second, the distribution is relatively more right-skewed and leptokurtic (higher tails) which implies that the positive effects are larger and more unevenly distributed across region. For the poorest regions, i.e., those in the 0-10th percentile in terms of initial GVA per capita, the investment effect is 57.5 percentage points higher compared to the counterfactual. This impact is more than 2.5 larger than the estimated effect on GVA growth. This finding provides evidence that the CF payments have a regional impact that adheres to the policy objectives, namely, to stimulate regional investment and infrastructure, and, through these channels, enhance regional GVA growth, especially for the poorest regions in the EU.

Having these granular effects of the Cohesion Fund payments by region and over time, we use them to visualize the effects on a map of the European regions. A distinct spatial pattern emerges from visual inspection of such a map. The smallest effects of the Cohesion Fund are concentrated in Southern Europe. In contrast, the largest effects are observed in Eastern Europe (note that countries in Northern Europe were not eligible for CF support). This pattern reveals a new geographical division within Europe. Specifically, Southern Europe, which consists of relatively wealthier regions, experiences weaker CF effects, whereas Eastern Europe, which consists of poorer regions that became members of the EU in the early 2000s, benefits the most from such a policy. Moreover, when we aggregate these results at the country level, most of the findings observed at the regional level are confirmed. Specifically, the impact of the Cohesion Fund is frontloaded, with the largest effects estimated within the first seven-year cycle. Additionally, the magnitude of the impact is negatively correlated with a country's initial GVA per capita, meaning poorer countries tend to experience stronger positive effects. These differences in GVA growth stem from the causal impact of regions within each country receiving Cohesion Fund support. Finally, we find a nonlinear relationship and diminishing returns for the CF impact on GVA as a function of the CF to GVA ratio. That is, for relatively low funding as a share of GVA, the effect is negligible or even negative. As funding increases, the effect grows, peaking when the cohesion fund reaches 0.58% and 0.94% of the region's GVA for the seven- and fourteen-year period respectively. Beyond a point, we observe diminishing returns, where the effect of the Cohesion Fund gradually declines and eventually becomes zero or even negative at higher funding as a share of GVA.

The rest of our paper is laid out as follows: section 2 presents an overview of the related literature while section 3 presents some additional information on the Cohesion Fund. Section 4 presents our dataset, while section 5 analyses our methodology and presents the empirical model. Section 6 presents the results and section 7 concludes.

2 Related literature

The present paper contributes to the empirical literature studying the effects of the EU's cohesion policy expenditures on regional economic performance. Based on the approach utilized, this literature can be split into two broad categories: one that employs causal inference techniques and one that focuses on more traditional (spatial) panel econometrics approaches. A schematic review is presented in Table 1.

Starting with the first branch of the relevant literature, the focus is on disentangling the causal impact that cohesion transfers exert on the economic performance of the recipient EU regions. These papers leverage on an institutional rule that determines whether a region is eligible or not for EU transfers. In particular, the rule stipulates that if a region's GDP per capita (in PPP terms) is less than 75% of the EU average prior to the commencement of the programming period, then it is awarded the 'Objective I' or 'Convergence criterion' or 'less-developed' region status and is thus eligible for the EU transfers.[3]

By focusing on a binary treatment indicator based on the GDP per capita rule, the literature examined whether being awarded the 'less-developed' status had a causal impact on economic performance. In this context, Becker et al. (2010, 2013, 2018) and Crescenzi and Giua (2020) focus on an interesting observation and develop a 'fuzzy' RDD in order to assess the impact of cohesion transfers. In particular, they observe that there are cases of non-compliance with the eligibility rule which led to a distinction between a region being *eligible for* and a *recipient of* EU transfers; that is, there exist cases were regions were awarded the Objective 1 status despite having a larger level of GDP per capita as well as cases were the regions were not awarded the 'less-developed' status despite meeting the criteria.[4] Becker et al. (2010) show that cohesion policy transfers to the lagging regions increase real per capita GDP growth by 1.6% within the same programming period. This translates to a multiplier of 1.2, implying that the policy yields net benefits. Becker et al. (2013) develop a heterogeneous variant of the RDD in order to examine if absorptive capacity –proxied by an index of the quality of government and human capital– is the driving force behind heterogeneous treatment effects. They conclude that heterogeneity in absorptive capacity is crucial, with only one-third of the regions in their sample exhibiting a high enough level that allows them to attain large, positive economic effects. Lastly, Becker et al. (2018) use an updated dataset that

[3]It should be clarified here that all EU NUTS2 regions are eligible for cohesion policy transfers. However, the bulk of the transfers is directed to the 'less developed' regions mentioned above; that is, it is the intensity of the transfers that changes across regions.

[4]Interestingly, Pellegrini et al. (2013) use a 'sharp' RDD approach.

allows them to account for the impact of the financial crisis of 2008. They document a weaker impact of the EU's cohesion policy transfers during the programming period containing the crisis, pointing to a downward pressure that was exerted by the cyclical effects. Additionally, they document that the effects of the policy seem to be contained within one programming period.

A separate branch of this literature shifted the focus from the broad question of whether cohesion transfers have a causal impact on growth to the more granular issue of whether the level of the transfer intensity that regions receive is what matters for economic performance. Technically, these papers exploit the variation in the intensity of the treatment. In this case, the question at hand is whether higher levels of the treatment lead to a higher or lower impact on the outcome compared to lower levels of the treatment. In this context, Hagen and Mohl (2008), Becker et al. (2012) and Cerqua and Pellegrini (2018) employ a generalized propensity score approach following Hirano and Imbens (2004) and estimate the so-called *dose-response* function. This allows the examination of whether there exists either a *minimum necessary* level or a *maximum desirable* level of regional transfer intensity necessary to achieve positive effects from cohesion policy expenditures (e.g. see Becker et al. (2012)). This has important policy implications, as the existence of either bound implies that a redistribution of the EU's transfer budget could be undertaken, leading to positive effects for more regions. In particular, Becker et al. (2012) identify that the maximum desirable level of transfer intensity for their sample is equal to 1.3% of GDP (and, specifically, of the beginning-of-period GDP). Moreover, they calculate that the optimal transfer intensity –the level of transfers beyond which one additional euro of transfers leads to the generation of less than one euro of additional GDP– is equal to 0.4% of GDP. Similar results are obtained by Cerqua and Pellegrini (2018). On the contrary, Hagen and Mohl (2008) show that for a sample focusing on earlier programming periods, Cohesion transfers have a positive yet statistically insignificant effect thus rendering void the discussion of an 'optimal dose' of EU transfers. Overall, the literature utilizing counterfactual policy evaluation techniques seems to provide evidence pointing to a clear positive effect of cohesion policy funds on regional economic growth and employment.

Our work extends this literature by applying a recently developed technique that is able to disentangle region- and time-specific effects for the EU regions that were recipients of the Cohesion Fund payments. This approach allows us to assess the causal effects of the Cohesion Fund under a more general, heterogeneous setting compared to what was implemented in the relevant literature.

The second broad category of the relevant empirical studies focuses on the use of panel data techniques for the assessment of the impact of cohesion policy expenditures on economic performance. As the approach followed in these articles goes beyond the scope of the present paper, we only briefly discuss it here. Interested readers are referred to the papers included in the upper panel of Table 1 for an entry point to this literature. This rather extensive literature employs various panel econometrics techniques and aims to disentangle the (long-run) effects of these sub-national transfers on regional economic performance using mainly instrumental variables and spatial approaches (see, for example, Di Caro and Fratesi (2022) and Mohl and Hagen (2010)). Recent contributions have incorporated into the analysis institutional factors - e.g. see Rodríguez-Pose and Garcilazo (2015). Lastly, a rather recent set of contributions focuses on the estimation of the multiplier effects of Cohesion expenditures - e.g. see Canova and Pappa (2024).

Table 1: Schematic Literature Review

Reference	Sample	Period	Data	Dep. Variable	Estimator	Homogeneity	Spatial dimension
Panel data							
Amendolagine et al. (2024)	245 NUTS2	2000-2018	Cohesion database	GDP pc growth	HSAR	Heterogeneous panel	Weighting matrix
Bourdin (2018)	147 NUTS3	2000-2016	Payments	GDP pc growth	GWR, SDM	Heterogeneous panel	Weighting matrix
Di Caro and Fratesi (2022)	250 NUTS2	1990-2015	Cohesion database	GDP pc	ECM	Heterogeneous panel	–
Fiaschi et al. (2018)	175 NUTS2	1991-2008	Payments, commitments	GDP pw growth	SDM	Homogeneous panel	Weighting matrix
Fidrmuc et al. (2024)	272 NUTS2	1997-2014	Cohesion database	GDP pc growth	2SLS, SDM	Homogeneous panel	Weighting matrix
Mohl and Hagen (2010)	126 NUTS2	1995-2005	Payments	GDP pc growth	GMM, Spatial Lag	Homogeneous panel	Weighting matrix
Rodríguez-Pose and Garcilazo (2015)	169 NUTS2	1996-2007	Payments	GDP pc growth	FE, GMM	Homogeneous panel	–
Policy evaluation for EU funding impact assessment							
Becker et al. (2010)	674 NUTS2(p)	1989-2006	Binary	GDP pc growth	Fuzzy RDD	Homogeneous	–
Becker et al. (2012)	2078 NUTS3(p)	1994-2006	Payments	GDP pc growth	Dose-response (GPS)	Homogeneous	–
Becker et al. (2013)	186-251 NUTS2(p)	1989-2006	Binary	GDP pc growth	Fuzzy RDD	Homogeneous	–
Becker et al. (2018)	1989-2013	259 NUTS2	Binary	GDP pc growth	Fuzzy RDD	Homogeneous	–
Cerqua and Pellegrini (2018)	208 NUTS2	1994-2006	Payments	GDP pc growth	Dose-response (GPS)	Homogeneous	–
Crescenzi and Giua (2020)		2000-2014	Binary	GDP pc growth	Spatial RDD	partially Heterogeneous	–
Hagen and Mohl (2008)	122 NUTS1/NUTS2	1995-2005	Payments	GDP pc growth	Dose-response (GPS)	Homogeneous	–
Pellegrini et al. (2013)	213 NUTS2	1994-2006	Binary	GDP pc growth	Sharp RDD	Homogeneous	–

Note: *p* denotes that the regions are pooled for the estimation, *binary* denotes a dummy variable specifying whether a region was treated, i.e. whether it receives cohesion transfers and/or is awarded the 'Objective I' or 'Convergence' or 'less-developed' status.
Estimators: *2SLS* denotes the two-stage least squares, *ECM* denotes an error-correction model, *GMM* denotes a Generalized Method of Moments estimator, *GPS* denotes a generalized propensity score, *GWR* denotes a geographically weighted regression estimator, *FE* denotes the fixed effects estimator, *HSAR* denotes the heterogeneous spatial autoregressive estimator, *SDM* denotes the Spatial Durbin Model
Data: Cohesion data denotes the Historic EU payments database of DG REGIO mentioned in the main text, payments denotes data received by the authors directly from DG REGIO or by ESPON.
Homogeneity: Partial means that the paper includes results based on groupings of regions.
Weighting matrix: In order to preserve the readability of the table, the various weighting schemes employed are mentioned here: k-nearest neighbor, Gaussian distance decay function,

3 Background: the Cohesion Fund

The European Structural and Investment Funds (henceforth, ESIF) constitute a major component of a system of (sub)national transfers made by the EU to the Member States to foster regional economic performance. ESIF is primarily composed by the European Regional Development Fund (ERDF), the European Social Fund (ESF) and the Cohesion Fund (CF) – which are collectively known as the cohesion policy instruments– and the European Agricultural Fund for Rural Development (EAFRD) (which, together with the European Agricultural Guarantee Fund (EAGF) are the two components of the Common Agricultural Policy (CAP)).

The vast majority of the relevant literature has focused either on the impact of all cohesion policy instruments combined or on the two largest instruments, namely the ERDF and the ESF. In this paper we deviate from this trend and focus our attention on the largely neglected CF instrument.

The CF was established in 1994 in order to further strengthen the cohesion of the EU via funding investment projects related to the fields of environment and transport infrastructure. A key characteristic of the Cohesion Fund is that the rule based on which the European Commission allocates the budget across the Member States differs compared to the rest of the ESIF instruments. In particular, the fund is available to countries that exhibit per capita Gross National Income (GNI) lower than 90% of the EU27 average. As a result, the CF instrument is targeting, by design, the poorest European territories and offers them additional assistance to achieve higher levels of output. This is made evident in the following Table 2 which breaks down CF expenditures by programming period and also provides further details about the fund's recipients.

Table 2: Cohesion Fund recipients and transfers

Period	Countries	Regions	GDP pc <75%	Population (%)	CF amount	Share of CP(%)
1994-1999	4	42	16	14.9	18078.27	25.04
2000-2006	16	100	74	37.3	30618.79	17.7
2007-2013	16	102	71	39.6	68939.32	27.1
2014-2020	15	83	79	38.9	47953.53	22.3

Notes: Values are in €millions.
Notes: The dates in the table depict the beginning of the respective programming period.
The number of regions with GDP per capita lower than 75% of the EU average is calculated based on data corresponding to the first year of each programming period.
Share of CP refers to CF payments as a share of the total cohesion policy payments over the respective period.

As can be observed in the Table, the scope of the CF increased significantly over time. In particular, when the Fund first became operational, payments were made toward four countries only –namely, Greece, Ireland, Portugal and Spain– that accounted for 15% of the EU's population. Moreover, only 40% of the beneficiaries exhibited a level of GDP per capita lower than 75% of the EU average, effectively classifying them as being 'less developed' regions. During the 2000-2006 period, which saw the enlargement of the Union with the accession of 10 Central and Eastern European countries, the CF beneficiaries significantly increased. In particular, during that programming period, the fund was available to 16 countries, with the number of regions and the affected population more than doubling. It is interesting to note that of the 100 recipient regions, 74 exhibited a level of GDP per capita less than 75% of the EU average, a significant increase compared to the previous period (essentially, 3 out of 4 beneficiaries were less developed regions). These numbers remained stable for the 2007-2013 period; however, in the 2014-2020 period with the number of eligible countries decreasing, 79 out of the 83 treated regions were classified as less developed.

It is also worth noting that the enlargement of the EU brought about a significant increase in the budget allocated to the CF. More specifically, during the third programming period (2007-2013), the CF budget more than doubled reaching almost €70 billion or 27% of the overall cohesion policy budget. The 2014-2020 period saw a decrease in the CF budget, which however still accounted for more than 20% of the total cohesion policy budget.

This set of evidence highlights the fact that the CF, with its expanding reach over time that covers almost exclusively less developed regions and the increasing budget that aimed to fund environment and transport investment projects –which are crucial for achieving sustainable growth– is of central importance for ensuring increased cohesion across the EU and merits an in depth analysis of its potential effects.

4 Data and Descriptive Statistics

We analyze data from two sources: (i) the Cohesion portal[5] from which we source data on the EU's Cohesion Fund expenditures (in current prices) and (ii) the Annual Regional Database (ARDECO) of the European Commission[6] for macroeconomic data.

Starting with the Cohesion Fund data, we note that they are generally available for the 1993-2022 period.[7] However, the vast majority of regions became eligible for the CF payments in 2000 (as already mentioned, prior to 2000, only four countries received Cohesion Fund payments). Given that the data are provided in nominal terms, we convert them to 2015 prices using the regional GVA deflator. We note here that prior to using the data for the empirical analysis, we had to account for two issues: first, the fact that the data are not reported under a single NUTS2 classification; and, second, the issue of the time of the recording of Cohesion Fund expenditures. Further details regarding these issues, along with our approach, are available in Appendix B.

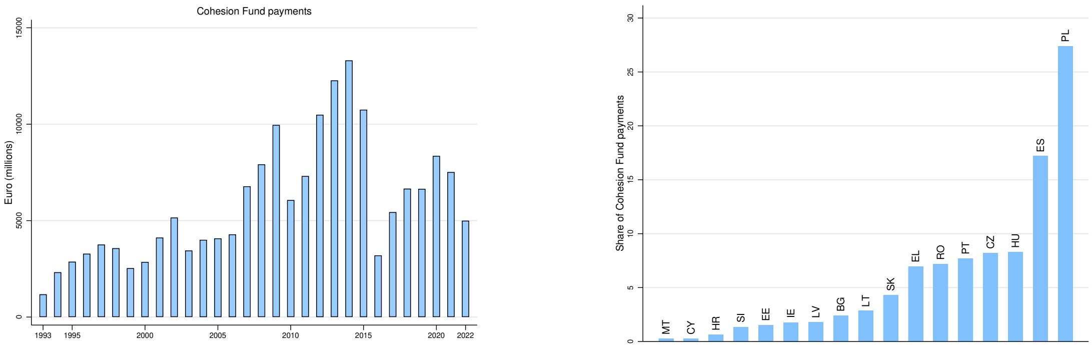
Panel (a) of Figure 1 depicts the evolution of the CF payments over time for all recipient countries, while panel (b) of the Figure depicts the share of disbursements received by the eligible countries. As is evident from the Figure, during the 2007-2013 period there was a significant increase in the CF payments as a means of accommodating the accession of the newest Member States. These increased payments continued up to 2015, due to the well-known delays in the disbursement of EU funds (for more details on this, see Dicharry (2023)). In total, for the period covered by our sample, €165.15 billion were disbursed to the recipient countries, with Poland and Spain receiving around 27% and 17% of total payments, respectively.

[5] `https://cohesiondata.ec.europa.eu/Other/Historic-EU-payments-regionalised-and-modelled/tc55-7ysv/about_data`

[6] `https://knowledge4policy.ec.europa.eu/territorial/ardeco-database_en`

[7] Although the Cohesion Fund was officially launched in 1994, there were some payments in advance towards the eligible countries in 1993.

Figure 1: Cohesion Fund payments



(a) Yearly Cohesion Fund payments

(b) Share of Cohesion Fund payments

Notes: Panel (a) depicts the total expenditures for the Cohesion Fund by year, while panel (b) presents the average shares of expenditures received by the 17 recipient countries. Note that not all countries received Cohesion Fund payments throughout the sample, e.g., Spain was ineligible for the 2014-2020 period.

Table 3 provides a disaggregation of the CF disbursements by beneficiary, along with some additional data on the relative importance of these payments in terms of GVA and public investment. The Table covers the 2000-2022 period, which includes the last three programming periods, in order to ensure maximum coverage in terms of recipient countries. We note here that Ireland is excluded from the analysis for two reasons: firstly, Ireland was eligible for the Cohesion Fund for only one programming period (2000-2006). Secondly, due to the well-documented issues with the Irish National Accounts stemming from the distortionary impact of large multinational enterprises operating within the country it would be increasingly difficult to disentangle the effects of the Cohesion Fund on economic performance. Overall, it is evident from the Table that Poland has received the bulk of the payments over the last 23 years, with a total amount that is more than twice in magnitude compared to the one received by Spain, which ranks second. It is worth highlighting here that Spain was not eligible for funding from the CF during the last programming period (2014-2020), implying that the country benefited significantly from these disbursements within a narrower time frame.

Turning to the relative size of the CF payments in terms of GVA, we observe that across beneficiaries the CF ranges from 0.08 to 0.68%, with a median value of 0.39%. As such, the contribution of the EU's CF transfers constitute a significant boost of liquidity for the implementation of investment projects. This is further exemplified in the last column of the Table, which depicts the CF transfers as a share of public investment. We observe that, with a median value of 11.7%, the CF transfers emerge as a significant source of additional funds for the financing of investment, in a complementary fashion to public investment. In particular, in Eastern countries like Lithuania, Latvia, Hungary, Slovakia and Poland, the CF disbursements amounted to at least 15% of public investment.

Table 3: Cohesion Fund expenditures, 2000-2022

Country	Regions	Programming Period	CF amount	CF(% GVA)	CF (% Gov. Investment)
BG	6	Periods 1-3	4206.9	0.48	12.46
CY	1	Periods 1-3	459.3	0.12	3.9
CZ	8	Periods 1-3	14365.4	0.42	11.7
EE	1	Periods 1-3	2634.8	0.65	13.3
EL	13	Periods 1-3	8884.4	0.22	5.8
ES	19	Periods 1-2	20079.3	0.13	3.3
HR	4	Periods 1-3	1110.9	0.17	4.1
HU	8	Periods 1-3	14537.8	0.67	17.1
LT	2	Periods 1-3	4975.9	0.68	15.89
LV	1	Periods 1-3	3169.9	0.67	15.13
MT	1	Periods 1-3	459.3	0.26	8.3
PL	17	Periods 1-3	47988.7	0.56	27.1
PT	7	Periods 1-3	9186.6	0.25	6.9
RO	8	Periods 1-3	12592.7	0.39	14.4
SI	2	Periods 1-3	2348.5	0.3	6.5
SK	4	Periods 1-3	7495.9	0.48	15.5

Notes: The Table covers the 2000-2022 period as in the first programming period (1994-1999) during which the CF was operational only 4 countries were eligible for the funds.
Values are in €millions, 2015 prices. Country-level GVA deflators were used.
Our proxy for general government investment is the Gross Fixed Capital Formation of the public sector (NACE codes: O to Q) from Eurostat's industry accounts (series: nama_10_a64_p5).
Column [3] refers to the programming periods in which each country was eligible for CF transfers. The 2000-2006 period is defined as Period 1, 2007-2013 as Period 2 and 2014-2020 as Period 3.

Our source for macro-related data is ARDECO. We obtain data for real Gross Value Added (GVA), at constant 2015 prices (series code: SOVGE) and for the average annual population at the regional level (series code: SNPTD) in order to construct the corresponding per capita measure.

Our dataset originally included data from 27 European countries and $N = 242$ regions in the 2021 NUTS2 classification for a maximum of $T = 43$. The treated sample covers 96 EU NUTS2 regions[8] over the 1993-2022 period for a total of 2297 observations and an average of 23.9 years under treatment. These regions belong to 16 countries, the majority of which are the newest Member States of the EU, i.e. the Central and Eastern European countries that entered the Union starting with the 2004 enlargement. Only 18 regions (located in Greece and Portugal) are under treatment for the whole sample period, with 78 regions being treated for up to 23 years.

5 Empirical Strategy

5.1 Notation and preliminaries

Let y_{it} denote the outcome of interest[9], for the i-th EU region at year t, $i = 1,\dots,N$ and $t = 1,\dots,T$. Throughout the rest of the paper we refer to regions that were eligible for CF payments as the *treated regions*, we call the regions that did not receive this type of funding

[8]We exclude from the sample of treated regions the so-called outermost regions of Portugal, namely Azores and Madeira, and Spain, namely the Canary Islands, along with the autonomous cities of Ceuta and Melilla.

[9]In this paper we consider two outcomes of interest; the logarithm of real GVA per capita and the logarithm of Gross Fixed Capital Formation.

the *control regions* and we use the term *treatment* when a region receives funding from the CF program. Let N_c and $N_{tr} = N - N_c$ be the numbers of control and treated units, respectively. We also denote by T_i the last time period before unit i received the intervention, i.e, $T_i = T$ for control units, while $\mathcal{N}$ denotes the set of all treated regions.

We work under the framework of potential outcomes (see Holland (1986)) to assess the impact of the policy intervention implied by the CF program by relying on observational time-series. For each region and at each time period two potential outcomes are considered; the outcome that region i would have if it was not treated at time t, denoted by $y_{it}(0)$ and typically referred as the *potential untreated outcome*, and the outcome that the i-th region would have at time t if was treated, namely the *potential treated outcome* denoted by $y_{it}(1)$. Notice that only one of these outcomes is observed, while the other is known as the *counterfactual* outcome. Specifically, for a treated region $i \in \mathcal{N}$ at time $t > T_i$ we observe $y_{it} = y_{it}(1)$ whereas for $t \leq T_i$ as well as for $i \notin \mathcal{N}$ and for each $t = 1, \ldots, T$ we observe $y_{it} = y_{it}(0)$. The causal effect τ_{it} of the policy intervention is defined in terms of the difference between the observed outcome and the one that region i would have if it was not treated at time t, i.e. the counterfactual, so that

$$\begin{aligned} \tau_{it} &= y_{it}(1) - y_{it}(0) \\ &= y_{it} - y_{it}(0). \end{aligned} \tag{1}$$

Finally, let D_{it} denote the treatment indicator, $D_{it} = 1$ if region i received the intervention at time t and $D_{it} = 0$ otherwise.

5.2 Causal inference by using matrix completion

To assess the effect of the policy intervention implied by the CF program we follow recent directions in the literature and particularly the factor augmented approach (Xu, 2023) where unobserved confounders are modeled through interactive fixed effects (Bai, 2009b). Specifically, we employ the matrix completion method (Athey et al., 2021, Borusyak et al., 2024), considering the unobserved potential untreated outcomes $y_{it}(0)$, $i \in \mathcal{N}$ and $t > T_i$, as missing values to be estimated as $\widehat{y}_{it}(0)$. This allows us to compute the causal effect τ_{it} as

$$\widehat{\tau}_{it} = y_{it} - \widehat{y}_{it}(0). \tag{2}$$

In particular, we assume for each $i = 1, \ldots, N$ and $t = 1, \ldots, T$ the following factor model for both observed and potential untreated outcomes

$$\begin{aligned} y_{it} &= \tau_{it} D_{it} + \lambda_i^\top f_t + \epsilon_{it} \\ &= \lambda_i^\top f_t + \epsilon_{it} \end{aligned} \tag{3}$$

where f_t is a $K \times 1$ vector with unobserved factors, λ_i is $K \times 1$ vector with factor loadings and ϵ_{it} is a random variable that accounts for measurement error and is assumed to follow a sub-Gaussian distribution; the second line in (3) emphasizes that for potential untreated outcomes we have $D_{it} = 0$ for all $i = 1, \ldots, N$ and $t = 1, \ldots, T$, i.e,. (3) can also be interpreted as our assumption for the data generation mechanism of untreated outcomes. Moreover, the model in (3) is a generalization of the Two Way Fixed Effects (henceforth, TWFE) model which is the standard tool for estimation of causal effects under the DiD approach; see Appendix A for more details. Crucially, causal estimation using the TWFE model aims to estimate the parameter $\tau = \mathbb{E}[\tau_{it} | D_{it} = 1]$ which is typically known as the average treatment effect on the treated (ATT) while disregarding the estimation of the unit- and time- specific effects τ_{it}. Additionally, the estimation of τ by using the TWFE modelling approach relies on the following set of strong assumptions: (i) additive separability which states that any unobserved heterogeneity can be decomposed into additive unit- and time- specific fixed effects, (ii) the parallel trends assumption, which states that the treated and control outcomes would follow the same paths in the absence of the treatment, (iii) strict exogeneity of treatment, which rules out the possibility of the treatment being assigned based on future outcomes, (iv) no dynamic treatment effects, i.e., treatment effects are assumed to be homogeneous and constant over time and (v) the stable unit treatment value assumption (SUTVA) which states that the potential outcomes of any given unit are not affected by the treatment status of other units.

Importantly, the model specification described by (3) mitigates or rules out most of the aforementioned assumptions. First, the parallel trends assumption is relaxed since unobserved time-varying confounders that cause its failure are modelled by the interactive fixed effects term $\lambda_i^\top f_t$. In fact, the latent factors f_t can be understood as time-varying parameters that model unobserved confounders which are common across the EU regions but exert a heterogeneous impact on each region; this region-specific effect of the unobserved confounders is the parameter λ_i, also known as the factor-loading parameter. Furthermore, the interactive fixed effects term rules out the additive separability accounting for richer cross-sectional dependence. Notably, this is particularly useful in regional policy evaluation, where units considered to be in close proximity (either geographical or based on economic considerations, e.g. due to trade) may respond differently to unobserved shocks. Furthermore, as already noted, our considered approach results in the estimation of heterogeneous causal effects τ_{it} across time and regions, with the ATT for each time t being estimated as

$$\widehat{ATT}_t = \frac{1}{N_{tr}} \sum_{i \in \mathcal{N}} \hat{\tau}_{it}. \tag{4}$$

Overall, the identification of causal effects provided by the factor modelling approach in (3) relies on the following two assumptions: (i) strict exogeneity of treatment, i.e., treatment assignment is assumed to be uncorrelated with the idiosyncratic error after controlling for the latent factors and thus rules out endogenous treatments (ii) the potential outcomes of one unit should not be affected by the treatment status of another unit. Notice, that (ii) is a relaxed form of SUTVA since the interactive effects in (3) can capture unobserved common shocks, such as sectoral linkages, trade flows, or institutional similarities that may lead to correlated outcomes.

5.2.1 Estimation of causal effects

To estimate $y_{it}(0)$, for $i \in \mathcal{N}$ and $t > T_i$, we fit the model in (3) by discarding all the post-intervention data. In particular, following Athey et al. (2021) we consider the partially observed outcome matrix $Y \in \mathbb{R}^{N\times T}$ in which

$$y_{it} = \begin{cases} y_{it}(0), & \text{if } i \notin \mathcal{N} \text{ and } t = 1,\dots,T, \\ y_{it}(0), & \text{if } i \in \mathcal{N} \text{ and } t \leq T_i, \\ \text{missing}, & i \in \mathcal{N} \text{ and } t > T_i. \end{cases}$$

Then, the model in (3) implies that

$$Y = L + E, \tag{5}$$

where L is a $N \times T$ low-rank matrix representing the underlying structure of the data with elements $L_{it} = \lambda_i^\top f_t$ and E is also a $N \times T$ matrix with elements ϵ_{it}. Clearly, from (5) we have both for the observed and the missing entries in Y that

$$y_{it} = L_{it} + \epsilon_{it},$$

The goal is to estimate the entries L_{it} that correspond to the missing entries in Y. To achieve this we employ the matrix completion with nuclear-norm regularization method developed recently by Athey et al. (2021). In particular, given the observed entries in Y indexed by $\mathcal{O} \subset \{1,\dots,N\} \times \{1,\dots,T\}$, the estimation of L is formulated as a convex optimization problem of the form

$$\hat{L} = \arg\min_{L} \|P_{\mathcal{O}}(Y-L)\|_F^2 + \lambda\|L\|_*, \tag{6}$$

where $P_{\mathcal{O}}((Y-L))$ denotes the projection operator that retains entries in $\mathcal{O}$ and sets others to zero, $\|P_{\mathcal{O}}(Y-L)\|_F$ is the Frobenius norm of the matrix $P_{\mathcal{O}}(Y-L)$, $\|L\|_*$ is the nuclear norm (sum of singular values), which promotes a low-rank solution and λ is a tuning parameter controlling the trade-off between fit and complexity; see Appendix A for the definitions of the Frobenius and nuclear norms as well as for more details on solving the minimization problem in (6). Finally, having at hand the estimators $\hat{L}_{it}$ we estimate the causal effects for the treated units, i.e., for $i \in \mathcal{N}$ and $t \geq T_i$, as

$$\hat{\tau}_{it} = y_{it} - \hat{y}_{it}(0) = y_{it} - \hat{L}_{it}.$$

6 Results

In this section we present the empirical estimates of the impact of Cohesion Fund payments on regional economic performance.[10] The section is organized as follows: firstly, in section 6.1 we present the average treatment effect on the treated. Next, in section 6.2, we delve into the heterogeneous nature of the effects on the level of regional GVA per capita, focusing both on the sub-national (regional) and the national level. Then, in section 6.2.3 we turn our attention on the growth effects of the Cohesion Fund expenditures. Lastly, section 6.3 provides some graphical evidence on the nonlinear nature of the relationship between funding intensity and its GVA effects.

6.1 Average Treatment Effect

We start by presenting the average treatment effect across all regions as estimated in section 5.2.1; specifically, see equation (4). Figures 2 and 3 present the average treatment effect on the y-axis against years relative to treatment for GVA per capita and Gross Fixed Capital Formation (GFCF), respectively. The vertical red line that crosses the x-axis in year 0 represents the initial year in which each region received CF payments. Note that this year might differ for each region as EU countries became eligible for the CF payments at different years and, as such, the regions in our sample received payments at different points in time. Thus, the treatment effect is shown by the black line for each year that follows the zero year. The grey shaded area represents the upper and lower bound of the 95% confidence interval calculated using a parametric bootstrap technique as in Xu (2017) (see Appendix A for more details).
Figures 2 and 3 show that on average CF payments have a persistent positive effect on both the GVA and GFCF of a region that receives them (treated) compared to the counterfactual scenario where these funds are not provided by the EU.[11] Our methodology allows us to provide a quantification of the average GVA effect in the regions which were eligible for Cohesion

[10] The estimation is conducted by using the gsynth library (Xu and Liu, 2021) in the statistical software R (R Core Team, 2021)

[11] Figure C.1 in Appendix C plots both the observed and the respective estimated counterfactual average GVA in logs as well as in growth rates. Their difference is the average treatment effect presented in 2. Additionally, Figure C.2 depicts the same series for the growth rate of GVA.

Fund payments. Specifically, on the seventh year under treatment the average treatment effect is equal to 0.086 (in logs), while for the entire period it averages to 0.05. Similarly, for the 14th year under treatment, the average treatment effect is equal to 0.091, while for the second 7-year period it is equal to 0.097 on average. This implies that, for the hypothetical EU27 average region, GVA per capita increases by €468 more than it would under the counterfactual scenario (i.e., without CF support).[12] After 14 years the average treatment effect is equal to 0.075 in logs. This implies that the hypothetical EU27 average region, GVA per capita increases by €330 more than it would under the counterfactual scenario. Following similar calculation, the hypothetical EU27 average region, GFCF per capita increases by €385 and €415 more than it would under the counterfactual scenario after 7 and 14 years under treatment, respectively. Lastly, as can be seen from the Figures, the majority of this rise occurs in the first seven years, while further gains beyond that period are relatively smaller. This is a first indication that the impact of the Cohesion Fund is, on average, frontloaded.

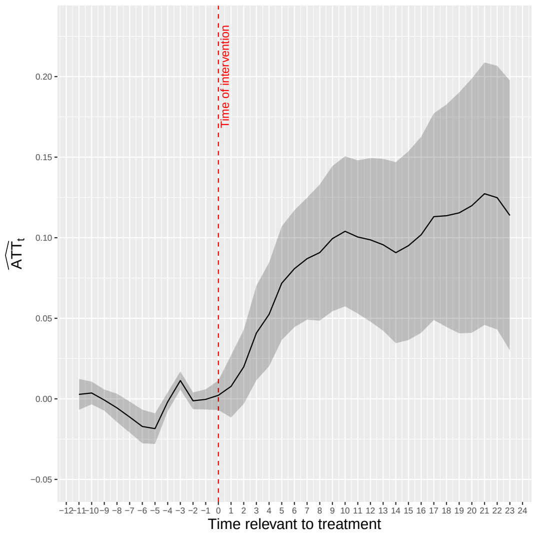


Figure 2: $\widehat{ATT}_t$ for GVA

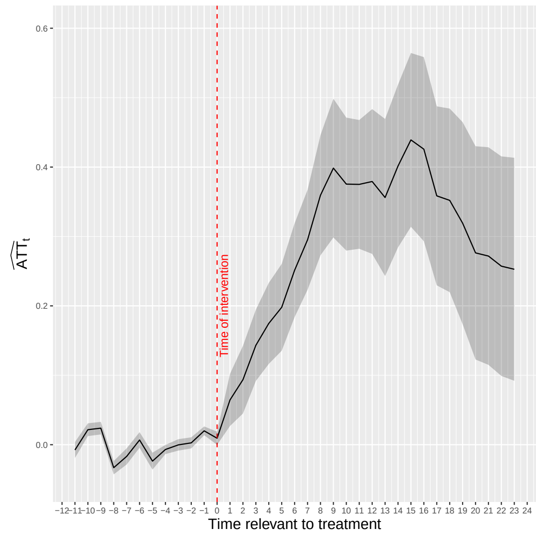


Figure 3: $\widehat{ATT}_t$ for GFCF

Note: Average Treatment Effect on the Treated (ATT) over time (black solid line) and 95% credible intervals (gray shaded area) for the log of GVA per capita (left panel) and the log of GFCF (right panel).

6.2 Heterogeneous effects

6.2.1 Quantitative analysis of the heterogeneous effects

The previous section focused on the average treatment effect across all regions that have received CF payments. However, a key advantage of our empirical strategy is that it allows us to explore the heterogeneous effects of CF across the different regions in our sample. Specifically, Figure 4 presents the entire distribution of the GVA effects across treated regions at 7 and 14 years after receiving CF for the first time along with the average effect across all years. Similarly, Figure 5 displays the analogous distributions for the GFCF effect.

A visual inspection of Figures 4 and 5 shows significant heterogeneity in the quantitative effect of CF in EU regions. This finding is novel, as most existing papers usually limit their

[12] In this hypothetical EU27 average region, GVA per capita rises by €2,168 over the first seven years of treatment—from €9,064 in the first year to €11,232 in year seven. Under the counterfactual scenario, the corresponding increase is €1,700 —from €8,928 to €10,628 in year seven.

analysis to the estimation of an average effect across regions, ignoring any potential heterogeneity. Both Figures suggest that the heterogeneous impacts should be closely accounted for and studied, as they can provide valuable information to EU policymakers regarding the optimal allocation of EU resources. Indicatively, the observed heterogeneity could be related to the intensity of the treatment, i.e. the size of the Cohesion Fund expenditures relative to GVA. If a maximum desired intensity is identified, this provides flexibility to re-allocate the resources and achieve better outcomes. We touch upon this issue in more detail in secton 6.3.

Figure 4: Distribution of GVA effects

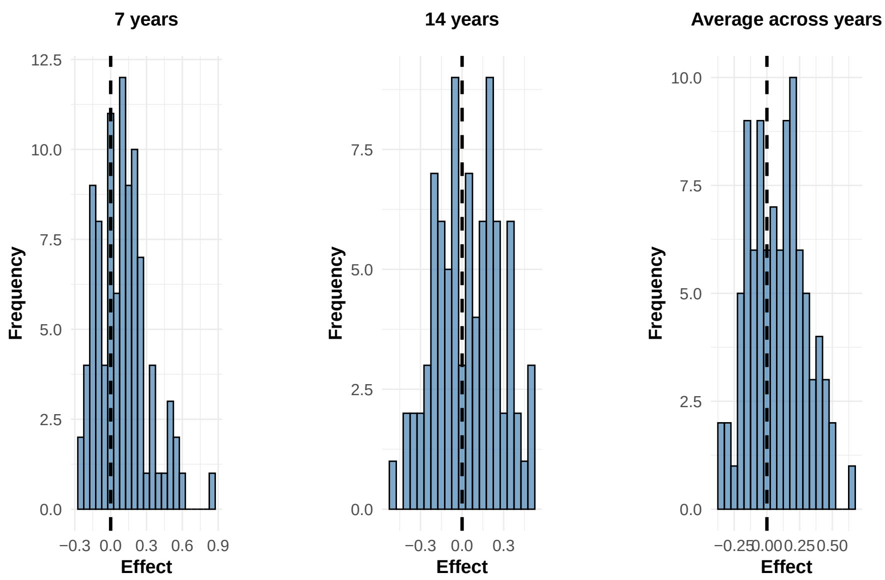


Figure 5: Distribution of GFCF effects

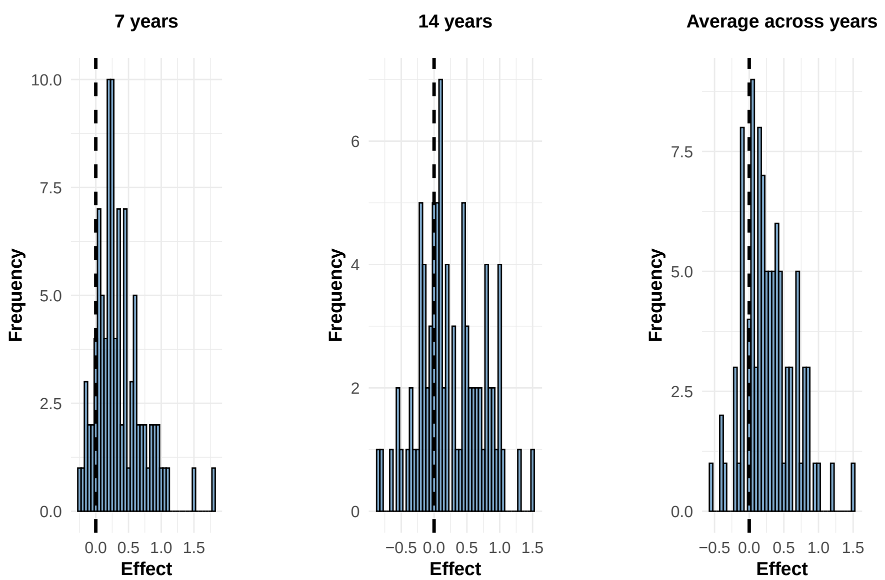


Table 4 presents summary statistics for the distribution of the CF effects on regional GVA. These results indicate that the inclusion in the CF program has a predominantly positive impact on the majority of treated regions (see first row). However, this positive effect is front-loaded, meaning its intensity concentrates in the first 7 years. Specifically, during the first seven-year cycle, 64.5% of treated regions experience positive effects, but this reduces to 49% after 14 years. On average, a 60% of treated regions benefit from the CF. The skewness of the distribution (see second row) reveals that the treatment effect is positively skewed in the first seven-year cycle. This indicates a longer right tail, implying that some regions experience significantly higher positive effects. However, over time, skewness decreases during the second seven-year cycle (see 14 years). This suggests that large positive effects become less pronounced. The evolution of kurtosis further support the above evidence (see third row). During the first seven-year cycle, the distribution exhibit higher kurtosis which implies relatively thicker tails. As can be seen, in our case, the right side (positive effect) is thicker. Over time, however, kurtosis declines which indicates that the distribution becomes flatter and more symmetric. This suggests that the effects of the CF program become more evenly distributed and milder across regions.

Table 4: Summary statistics for the distribution of the regional effect on GVA per capita

	7 years	14 years	Average
% of regions with $\hat{\tau}_i > 0$	64.5	48.9	60.4
Skewness	0.81	-0.06	0.15
Kurtosis	3.81	2.33	2.42

Table 5 presents summary statistics for the distribution of CF effects on gross fixed capital formation (GFCF). Although most of the patterns observed for GVA effects hold, some interesting differences emerge. Specifically, the impact of the CF on regional investment is significantly larger in magnitude, affecting a higher percentage of treated regions. Almost 90% of the treated regions experience a positive effect on their investment within the first 7 years compared to 64.5% for GVA. This proportion decreases over time, and 62. 5% of the treated regions still have a positive impact after 14 years. Both skewness and kurtosis are higher than the respective metrics of the GVA distribution. This indicates a more pronounced right-skewed shape with fatter tails. In other words, the distribution is more concentrated in the right tail which indicates positive effects on GFCF. These findings align with the core objective of the CF program, which is designed to stimulate investment and infrastructure development in less developed EU regions. Our results confirm that CF funds play a particularly strong role in boosting capital formation, with a frontloaded effect on investment. Additionally, this seems to affect the GVA over a longer horizon. Thus, these findings suggest that the increase in GVA occurs through the stimulation of investment.

Table 5: Summary statistics for the distribution of the regional effect on gross fixed capital formation

	7 years	14 years	Average
% of regions with $\hat{\tau}_i > 0$	89.6	62.5	81.2
Skewness	1.16	0.15	0.51
Kurtosis	5.12	2.62	3.58

6.2.2 Geographical analysis of the heterogeneous level effects

This section illustrates the geographical dispersion of EU Cohesion Fund's policy impact across Europe. To do this, we categorize regions into five percentiles based on their effects on GVA and GFCF ranging from the lowest (light colors—yellow) to the highest (dark colors—red). Figure 6 maps the GVA effect across geographical regions.

A distinct spatial pattern emerges from the visual inspection of the map. The smallest effects of the Cohesion Fund, represented by lighter colors (yellow), are concentrated in Southern Europe. In contrast, the largest effects are observed in Eastern Europe (red), while gray-shaded regions, corresponding to Northern Europe, were not eligible for CF support.

This pattern reveals a new geographical division within Europe. In this case, Southern Europe, which consists of relatively wealthier regions, experiences weaker CF effects compared to Eastern Europe, where poorer regions benefit the most from such a policy. As such, the newest Member States that gained their EU status post-2000 are the ones that seem to reap the benefits of the Cohesion Fund assistance. A similar pattern though less pronounced is apparent when focusing on the effects on GFCF. The largest effects are again observed in Eastern Europe, while Southern Europe experiences milder impacts. It is interesting to note here that in the 14-year case, Spanish regions exhibit positive effects on investment.

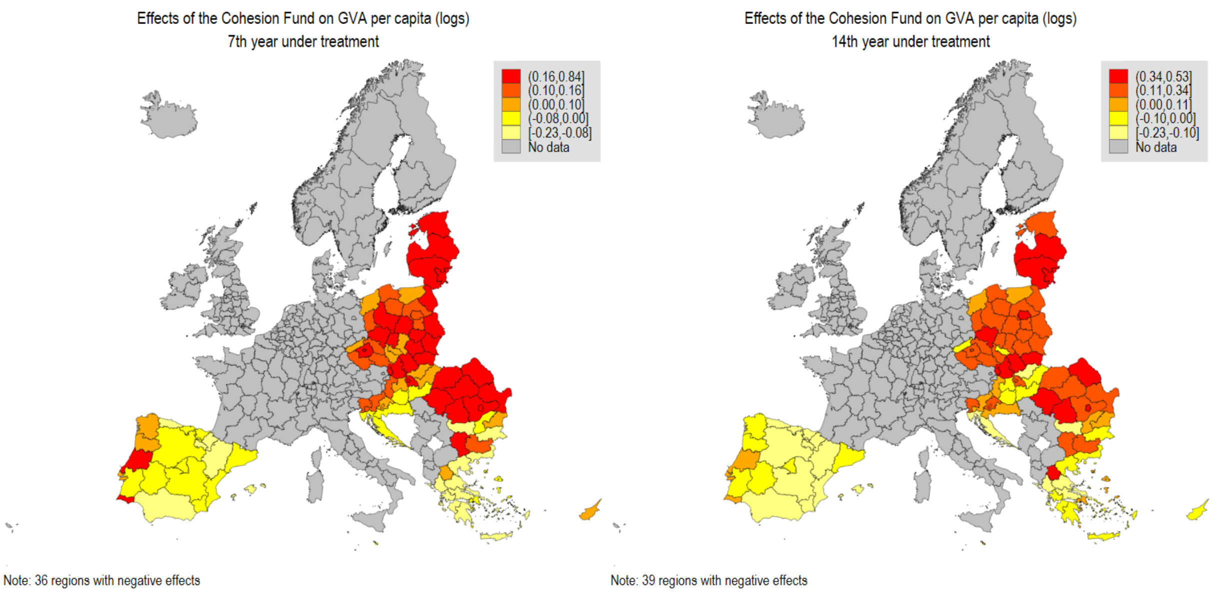


Figure 6: Geographical Dispersion of GVA effects

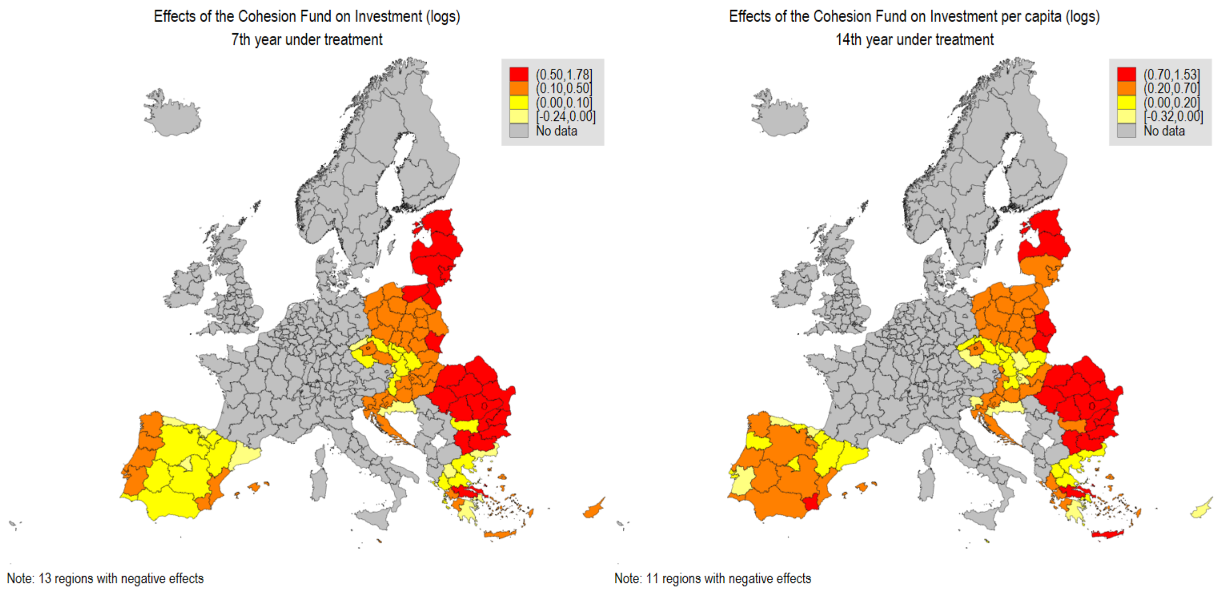


Figure 7: Geographical Dispersion of GFCF effects

6.2.3 Heterogeneous impact of the Cohesion Fund on GVA growth

In this section we focus our attention on the regional growth effects of the Cohesion Fund expenditures. Specifically, Table 6 compares the observed GVA growth, g_t^y, from the year of initial treatment with its counterfactual GVA growth as estimated in Section 5.2.1. That is the growth that a region would have experienced had it not received the CF. The counterfactual GVA growth for the t-th year after treatment, $\widehat{g}_t^y$, is computed as:

$$\widehat{g}_t^y = \left[\log(\widehat{\mathrm{GVA}}_t) - \log(\mathrm{GVA}_0)\right] \times 100 \tag{7}$$

where GVA_0 represents the GVA at the beginning of the programming period and $\widehat{\mathrm{GVA}}_t$ is the estimated counterfactual GVA at the end of the t-th year after the treatment.[13] To better assess the heterogeneous effects, we categorize the treated regions into five percentiles (column 1) based on their initial GVA per capita (column 2). The observed and counterfactual GVA growth rates, shown in columns four and five, respectively, are the average growth across regions of each percentile. We do this over two consecutive 7-year cycles, mimicking the actual 7-year budgeting programming periods of the EU.

[13] Table C.1 in Appendix C repeats the same analysis, excluding outlier regions based on the pre-treatment RMSE

Table 6: GVA per capita growth rates by years under treatment (group averages)

Percentile	Initial GVA	g^y, in %	$\widehat{g}^y$, in %	$g^y - \widehat{g}^y$
Effect After First 7-Year Treatment Period				
0-10th	4137.21	32.7	11.2	21.5
10th-25th	5273.65	28.8	10.01	18.79
25th-50th	8416.783	24.9	12.55	12.35
50th-75th	11928.78	15.57	14.01	1.56
75th-90th	17153.59	11.58	15.67	-4.09
Effect After Second 7-Year Treatment Period				
0-10th	4137.21	7.5	5.3	2.2
10th-25th	5273.65	9.4	11.3	-1.9
25th-50th	8416.783	8.09	5.57	2.52
50th-75th	11928.78	9.99	9.89	0.10
75th-90th	17153.59	11.89	13.56	-1.67

Notes: Initial GVA refers to the GVA per capita in the first year of the treatment. The last two columns depict the overall period growth rate for the cases in which the region is treated and the counterfactual, respectively.

Table 6 yields some interesting findings. The impact of EU's Cohesion Fund is frontloaded, peaking within the first seven years of inclusion in the program. However, the effects exhibit significant heterogeneity, with a clear negative relationship between the initial GVA per capita and the estimated impact. That is, the regions that have benefited the most are those in the lower percentiles of the initial GVA distribution - specifically, those below the 50th percentile. Regions belonging in the 0-10th, 10th-25th, and 25th-50th percentiles experienced an overall GVA growth that was 21.5, 19 and 12 percentage points higher than their respective counterfactual scenario (i.e., in the absence of EU funding), respectively. In contrast, the benefits in terms of GVA growth diminish for the relatively wealthier regions, such as those in the 50th-75th percentile. For the highest-income regions, the effect turns to negative. This suggests that the Cohesion Fund may have had limited or even adverse effects in the wealthiest regions. When comparing across different funding cycles, we observe that the growth effects persist but diminish significantly over time. However, it is important to note that the second cycle, i.e., the 14 years under treatment, overlaps with the Global Financial Crisis, which may have affected the outcomes. Nonetheless, even during the Crisis period, it seems that the CF payments seem to have assisted a large number of recipient regions across the initial GVA distribution to attain an average GVA growth rate larger than the counterfactual case.

Table 7 is analogous to Table 6 but it computes overall investment growth. The counterfactual investment growth, $\widehat{g}_t^I$, is defined similarly to equation (7). For comparability with the above discussion, we categorize treated regions into the same five percentiles based on their initial GVA per capita. The results reinforce the discussion on the heterogeneous GVA effects discussed above. As with GVA growth, the impact of the EU's Cohesion Fund on investment overall growth is frontloaded, peaking within the first seven years. The effect is also more pronounced in relatively poorer regions, with the higher impact estimated for the poorest 0-10th percentile, followed by the 10th-25th and 25th-50th percentiles. Compared to the effect on GVA growth, the effect on overall investment growth is significantly larger in magnitude. For instance, the poorest percentile experienced an investment growth rate that was 57.45%

percentage points higher than in the counterfactual scenario (i.e., without EU funding). This impact is more than 2.5 larger than the estimated effect on GVA growth. This is intuitive as the CF is designed and targeted to stimulate investment and infrastructure in relatively poorest EU regions. Another key difference is that while the effects on GVA growth persist, the benefit to investment growth dissipate faster in the second-year cycle.

Table 7: Investment growth rates by years under treatment (group averages)

Percentile	**Initial GVA**	$\widehat{g}^I$**, in %**	g^I**, in %**	$\widehat{g}^I - g^I$
Effect After First 7-Year Treatment Period				
0-10th	4137.21	66.67	9.22	57.45
10th-25th	5273.65	48.38	6.22	42.16
25th-50th	8416.783	25.99	7.74	18.25
50th-75th	11928.78	32.48	17.45	15.03
75th-90th	17153.59	32.87	21.47	11.40
Effect After Second 7-Year Treatment Period				
0-10th	4137.21	-19.09	-11.14	-7.95
10th-25th	5273.65	-0.80	-5.26	4.46
25th-50th	8416.783	-7.47	-4.89	-2.58
50th-75th	11928.78	13.66	4.52	9.14
75th-90th	17153.59	25.71	12.68	13.03

Notes: Initial GVA refers to the GVA per capita in the first year of the treatment. The last two columns depict the overall period growth rate for the cases in which the region is treated and the counterfactual, respectively.

6.2.4 Analysis of the effects by country

Given that the Cohesion Fund is, by design, allocated to countries (rather than to regions, as with the rest of the EU's Structural Funds) it is informative to examine the economic impact of the Cohesion Fund at a more aggregate level. In particular, given that we have estimated region- and time-specific treatment effects we can use equation (4) in order to obtain country-specific treatment effects (i.e. in this case the set $\mathcal{N}$ contains the NUTS2 regions of each of the eligible countries).

The following Tables 8 and 9 present the country-specific treatment effects for GVA per capita and investment, respectively, in the 7th and 14th year under treatment. We rank the countries in ascending order based on their initial GVA per capita, starting with the lowest initial GVA per capita, i.e., Bulgaria, and ending with the highest, i.e., Cyprus.

As expected, most of the results observed at the regional level are confirmed when aggregated at the country level. Specifically, the impact of the EU Cohesion Fund is frontloaded, with the largest effects estimated within the first seven-year cycle. Additionally, the magnitude of the impact is negatively correlated with a country's initial GVA per capita, meaning poorer countries tend to experience stronger positive effects. These differences in GVA growth stem from the causal impact of regions within each country receiving Cohesion Fund support. The eight poorest countries in terms of initial GVA —Bulgaria, Romania, Lithuania, Hungary, Latvia, Poland, Slovakia, and Estonia—generally exhibit positive effects on both GVA and investment growth. As before, the effect on investment growth is more pronounced compared to the effect on GVA growth. Romania, Latvia, and Lithuania stand out with particularly high GVA growth rates over the first seven-year cycle, although some of these gains diminish over time. In contrast, Hungary and Slovakia display more moderate yet stable positive effects. The effect on investment growth is more volatile than GVA growth. While Bulgaria, Romania, Lithuania, and Latvia experience significant investment growth within the first seven years, their long-term effects tend to decline and/or even turning negative. Poland and Greece, in contrast, show more sustained investment effects over the two cycles. Cyprus experiences a dramatic investment decline from the first seven-year to the second seven-year cycle. Wealthier countries such as Spain and Portugal show weaker or even negative effects, suggesting diminishing returns to EU funding in economies with higher GVA per capita.

Table 8: Country-specific GVA growth effect, $\widehat{g}^y - g^y$

	GVA pc	Years under treatment	
Country		7	14
Bulgaria (BG)	2991	14.1	3.36
Romania (RO)	4099	32.5	-2.17
Lithuania (LT)	5204	27.43	2.8
Hungary (HU)	5254	8.96	-4.29
Latvia (LV)	5300	30.93	2.52
Poland (PL)	5898	10.1	7.46
Slovakia (SK)	7352	22.44	7.81
Estonia (EE)	7775	16.42	-7.07
Croatia (HR)	9434	-6.25	3.24
Czech Republic (CZ)	9925	14.2	-0.61
Portugal (PT)	11954	-0.35	-5.47
Slovenia (SI)	12334	9.08	-2.92
Greece (EL)	12585	-1.26	8.14
Malta (MT)	13159	-4.64	9.47
Spain (ES)	15817	-4.94	-5.92
Cyprus (CY)	18192	0.12	-10.06

Notes: Cyprus, Estonia, Latvia, Malta constitute of only one region; as such, the average treatment effect is the same irrespective of the level of aggregation. g^y denotes is the growth rate in the treated case and $\widehat{g}^y$ is the growth rate in the counterfactual scenario. The initial year of treatment differs across countries.

Table 9: Country-specific investment growth effect, $\hat{g}^I - g^I$

	GVA pc	Years under treatment	
Country		7	14
Bulgaria (BG)	2991	66.3	11.24
Romania (RO)	4099	66.98	-12.97
Lithuania (LT)	5204	61.63	-19.45
Hungary (HU)	5254	16.73	6.19
Latvia (LV)	5300	57.91	-23.62
Poland (PL)	5898	6.65	16.37
Slovakia (SK)	7352	9.78	4.02
Estonia (EE)	7775	65.93	-8.21
Croatia (HR)	9434	-0.06	2.11
Czech Republic (CZ)	9925	15.11	-0.51
Portugal (PT)	11954	17.08	-11.84
Slovenia (SI)	12334	15.92	-26.23
Greece (EL)	12585	19.26	12.62
Malta (MT)	13159	22.08	7.54
Spain (ES)	15817	3.63	15.65
Cyprus (CY)	18192	29.03	-49.71

Notes: See notes of Table 9.

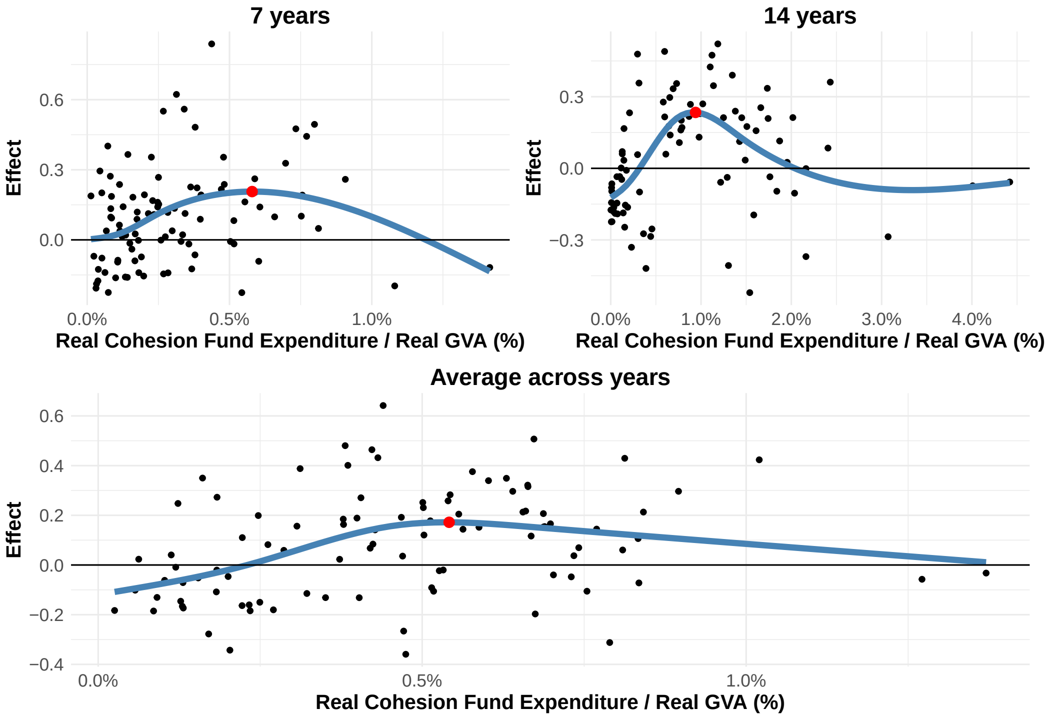
6.3 Non-linearity

In Figure 8, we plot the effect on GVA (y-axis) against the amount received from the Cohesion Fund as a share of GVA for each treated region, i.e. the treatment intensity. Each black dot represents the average "treatment effect" for each treated region. A spline is fitted to these dots represented by the blue line. Interestingly, a nonlinear relationship between the impact on GVA and the cohesion funds received as a share of GVA is revealed. This echoes the concept of a *maximum desirable level of transfers* of Becker et al. (2012), based on which beyond a certain level of treatment intensity the impact declines.

At relatively low funding as a share of GVA, the effect on GVA appears small, negligible or even negative. However, as funding increases, the effect grows, reaching its peak when the Cohesion Fund accounts for 0.58% of the region's GVA over the first seven years. Beyond this point, we observe diminishing returns, where the effect of cohesion funds gradually declines and eventually becomes zero or even negative at higher funding shares of GVA. Similarly, the treatment effect peaks when the CF share reaches 0.94% of a region's GVA for the fourteen year period On average the maximum effect is achieved when the CF share is equal to 0.54% of a region's GVA.

This nonlinear relationship suggests that the EU policymaker should allocate this type of funding with caution. On the one hand, increased funding can generate positive effects, but only up to a certain upper limit where the impact reaches its peak. Beyond this point, additional funding leads to diminishing returns and may even become counterproductive above a certain threshold. Our analysis enables us to quantify both the upper bound and the threshold beyond which further funding results in negative effects.

Figure 8: Nonlinear effect of Cohesion Funds



Notes: The red bullet points correspond to the 0.58%, 0.94%, 0.63% and 0.54% of the Real CF expenditure/ Real GVA for 7, 14, 21 and average across years, respectively.

7 Conclusions

This paper studies the impact of Cohesion Fund payments on regional economic performance, focusing on the heterogeneous responses of the EU's regions to this sub-national transfers programme. In particular, relying on the matrix completion technique, it disentangles both the region- and the time-specific effects of the policy. Our main results are that these transfers exert, on average, a positive and persistent effect on both output and investment, with the majority of the gains materializing in the first seven years under treatment. Moreover, the relatively poorer regions are the ones reaping the benefits of the programme. Lastly, we uncover a nonlinear relationship between funding intensity and the size of the effect.

Our results give rise to some interesting conclusions regarding the evaluation of such programmes. First and foremost, focusing on a (local) average treatment effect as the single parameter of interest masks significant information from the policymakers. Specifically, our finding that the Cohesion Fund transfers exert a larger impact on the poorer regions highlights the importance of strengthening its place-based approach. This will allow a more tailored design of the policy, focusing on the core needs of the regions. However, whether this tailored design should be the result of Member State-based designs is an issue that is left for future research. Second, the apparent existence of a threshold beyond which the policy exhibits diminishing returns implies that there is scope for a 'budget-neutral' redistribution of the Cohesion Fund payments. Specifically, as also shown in Becker et al. (2012), there is scope for a redistribution of the budget that could increase the efficiency of the Cohesion Fund in terms of regional growth performance and, potentially, enhance the convergence process.

References

Abadie, A., A. Diamond, and J. Hainmueller (2010). Synthetic control methods for comparative case studies: Estimating the effect of california's tobacco control program. *Journal of the American statistical Association 105*(490), 493–505.

Amendolagine, V., F. Prota, and L. Serlenga (2024). The Impact of European Cohesion Policy: A Spatial Perspective. *Journal of Economic Geography 24(4)*, 477–494.

Angrist, J. D. and J.-S. Pischke (2009). *Mostly harmless econometrics: An empiricist's companion.* Princeton university press.

Athey, S., M. Bayati, N. Doudchenko, G. Imbens, and K. Khosravi (2021). Matrix completion methods for causal panel data models. *Journal of the American Statistical Association 116*(536), 1716–1730.

Athey, S. and G. W. Imbens (2017). The state of applied econometrics: Causality and policy evaluation. *Journal of Economic perspectives 31*(2), 3–32.

Bachtler, J. (2022). Cohesion policy - where has it come from? where is it going? *European Court of Auditors Journal 1*, 7–12.

Bai, J. (2009a). Panel data models with interactive fixed effects. *Econometrica 77*, 1229–1279.

Bai, J. (2009b). Panel data models with interactive fixed effects. *Econometrica 77*(4), 1229–1279.

Becker, S., P. H. Egger, and M. von Ehrlich (2010). Going NUTS: The effect of EU structural funds on regional performance. *Journal of Public Economics 94(9-10)*, 578–590.

Becker, S., P. H. Egger, and M. von Ehrlich (2012). Too much of a good thing? on the growth effects of the EU's regional policy. *European Economic Review 56(4)*, 648–668.

Becker, S., P. H. Egger, and M. von Ehrlich (2013). Absorptive capacity and the growth effects of regional transfers: A regression discontinuity design with heterogeneous treatment effects. *American Economic Journal: Economic Policy 5(4)*, 29–77.

Becker, S., P. H. Egger, and M. von Ehrlich (2018). Effects of EU regional policy: 1989-2013. *Regional Science and Urban Economics 69*, 143–152.

Benedetto, G. (2024). The history of the eu budget. *European Parliament, Policy Department for Budgetary Affairs PE 636.475 - May 2024.*

Borusyak, K., X. Jaravel, and J. Spiess (2024). Revisiting event-study designs: robust and efficient estimation. *Review of Economic Studies*, rdae007.

Bourdin, S. (2018). Does the cohesion policy have the same influence on growth everywhere? a geographically weighted regression approach in central and eastern europe. *Economic Geography 95(3)*, 256–287.

Canova, F. and E. Pappa (2024). The macroeconomic effects of eu structural funds. *Journal of the European Economic Association.*

Cerqua, A. and G. Pellegrini (2018). Are we spending too much to grow? the case of structural funds. *Journal of Regional Science 58*, 535–563.

Crescenzi, R. and M. Giua (2020). One or many cohesion policies of the european union? on the differential economic impacts of cohesion policy across member states. *Regional Studies 54(1)*, 10–20.

De Chaisemartin, C. and X. d'Haultfoeuille (2024). Difference-in-differences estimators of intertemporal treatment effects. *Review of Economics and Statistics*, 1–45.

Di Caro, P. and U. Fratesi (2022). One policy, different effects: Estimating the region-specific impacts of eu cohesion policy. *Journal of Regional Science 62*, 307–330.

Dicharry, B. (2023). Regional growth and absorption speed of eu funds: when time isn't money. *Regional Studies 57(3)*, 511–524.

European Commission (2017). Regionalisation of ESIF payments 1989-2015. *Directorate-General for Regional and Urban Policy, Publications Office.*

European Commission (2024). Forging a sustainable future together – cohesion for a competitive and inclusive europe – report of the high-level group on the future of cohesion policy. *Directorate-General for Regional and Urban Policy, Publications Office.*

Fiaschi, D., A. Lavezzi, and A. Parenti (2018). Does eu cohesion policy work? theory and evidence. *Journal of Regional Science 58*, 386–423.

Fidrmuc, J., M. Hulényi, and O. Zajkowska (2024). The long and winding road to find the impact of eu funds on regional growth: Iv and spatial analyses. *Regional Studies 58(3)*, 583–600.

Fiuratti, F. I., D. E. Nikolova, S. M. Pennings, and M. T. Schiffbauer (2024). Are regional fiscal multipliers on eu structural and investment fund spending large? a reassessment of the evidence. *Policy Research Working Paper Series 10658.*

Gobillon, L. and T. Magnac (2016). Regional policy evaluation: Interactive fixed effects and synthetic controls. *Review of Economics and Statistics 98*(3), 535–551.

Hagen, T. and P. Mohl (2008). Which is the right dose of eu cohesion policy for economic growth? *ZEW Discussion Papers, ZEW - Leibniz Centre for European Economic Research 08-104.*

Hirano, K. and G. W. Imbens (2004). The propensity score with continuous treatments. In A. Gelman and X. Meng (Eds.), *Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives*, Wiley Series in Probability and Statistics, pp. 123–130. Wiley.

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American statistical Association 81*(396), 945–960.

Mazumder, R., T. Hastie, and R. Tibshirani (2010). Spectral regularization algorithms for learning large incomplete matrices. *The Journal of Machine Learning Research 11*, 2287–2322.

Mohl, P. and T. Hagen (2010). Do eu structural funds promote regional growth? new evidence from various panel data approaches. *Regional Science and Urban Economics 40(5)*, 353–365.

Pellegrini, G., F. Terribile, O. Tarola, T. Muccigrosso, and F. Busillo (2013). Does EU regional policy enhance growth? *Papers in Regional Science 92*, 217–233.

Pesaran, M. (2006). Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica 74*, 967–1012.

R Core Team (2021). *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing.

Rodríguez-Pose, A. and E. Garcilazo (2015). Quality of government and the returns of investment: Examining the impact of cohesion expenditure in european regions. *Regional Studies 49(8)*, 1274–1290.

Sun, L. and S. Abraham (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics 225*(2), 175–199.

Xu, Y. (2017). Generalized synthetic control method: Causal inference with interactive fixed effects models. *Political Analysis 25*(1), 57–76.

Xu, Y. (2023). Causal inference with time-series cross-sectional data: a reflection. *Available at SSRN 3979613.*

Xu, Y. and L. Liu (2021). *gsynth: Generalized Synthetic Control Method.* R package version 1.2.1.

Appendix

A Technical details of empirical strategy

Definitions

Here, we provide definitions for terms used in Section 5. A real-valued random variable $epsilon$ is σ-sub-Gaussian if for all $s \in \mathbb{R}$,

$$\mathbb{E}[\exp(s\epsilon)] \leq \exp(\sigma^2 s^2/2).$$

For any matrix **A** and a set of index pairs $\mathcal{O}$, define the projection matrix

$$(P_{\mathcal{O}} A)_{it} = \begin{cases} A_{it}, & \text{if } (i,t) \in \mathcal{O}, \\ 0, & \text{otherwise.} \end{cases}$$

The singular value decomposition (SVD) of a $N \times T$ matrix A writes $A = S\Sigma R^\top$, where S and U are $N \times N$ and $T \times T$ matrices respectively and Σ is $N \times T$ rectangular diagonal matrix with elements denoted by $\sigma_i(A)$, $i = 1, \ldots, \min(N, T)$. The Frobenius norm of a matrix A is defined as

$$\|A\|_F = \left(\sum_i \sigma_i(A)^2\right)^{1/2},$$

and the nuclear norm

$$\|A\|_* = \sum_i \sigma_i(A).$$

Relation of the TWFE with the interactive fixed effects models

We assume $K = 2$ latent factors in (3) and $\lambda_{i1} = 1$, $\lambda_{i2} \in \mathbb{R}$, $f_{1t} \in \mathbb{R}$ and $f_{2t} \in \mathbb{R}$ as well as $\tau_{it} = \tau$ for each $i = 1, \ldots, N$ and $t = 1, \ldots, T$ then (3) becomes

$$y_{it} = \tau D_{it} + f_{1t} + \lambda_{2i} + \epsilon_{it},$$

which is the TWFE discussed in detail, among others, by Angrist and Pischke (2009).

Estimation of the low rank matrix L

Here, we describe briefly the methodology for solving the minimization problem defined by equation (6) in order to estimate the matrix L; see Athey et al. (2021) for more details.

First we note an estimator of L could be also calculated by minimizing the sum of squared differences,

$$\min_{\mathbf{L}} \frac{1}{|\mathcal{O}|} \sum_{(i,t)\in\mathcal{O}} (Y_{it} - L_{it})^2 = \min_{\mathbf{L}} \frac{1}{|\mathcal{O}|} \|\mathbf{P}_{\mathcal{O}}(Y - L)\|_F^2, \tag{8}$$

where $\mathcal{O} \subset \{1, \ldots, N\} \times \{1, \ldots, T\}$ is the set of pairs (i, t) such that y_{it} is an observed entry in the matrix Y with potential untreated outcomes. However, the objective in (8) is not useful, as the estimator simply sets $L_{it} = y_{it}$ for observed entries. To address this, Athey and Imbens (2017) introduce the regularization term $\lambda\|L\|$ in (8) and even more to avoid the regularization of fixed effects the modify the minimization in (6) to obtain an estimator of the form

$$(\hat{L}, \hat{\Gamma}, \hat{\Delta}) = \arg\min_{L,\Gamma,\Delta} \left\{ \frac{1}{|\mathcal{O}|} \|P_{\mathcal{O}}(Y - L - \Gamma \mathbf{1}_T^\top - \mathbf{1}_N \Delta^\top)\|_F^2 + \lambda \|L\|_* \right\}, \tag{9}$$

where $\mathbf{1}_N$ and $\mathbf{1}_T$ are N- and T- dimensional vector with ones.

Importantly, the resulting estimator can be computed using fast convex optimization programs, such as the soft-impute algorithm (Mazumder et al., 2010) which is briefly described as follows. For any given matrix A with SVD $A = S\Sigma R^\top$ Athey et al. (2021) define the matrix shrinkage operator

$$\text{shrink}_\lambda(\mathbf{A}) = S\tilde{\Sigma}R^\top, \tag{10}$$

where in $\tilde{\Sigma}$ is equal to Σ but each singular value $\sigma_i(A)$ of the latter is replaced by $\max(\sigma_i(A) - \lambda, 0)$. Using this, we initialize

$$L_1(\lambda, \mathcal{O}) = \mathbf{P}_{\mathcal{O}}(\mathbf{Y}),$$

and iteratively update

$$L_{k+1}(\lambda, \mathcal{O}) = \text{shrink}_{\frac{\lambda|\mathcal{O}|}{2}}\left\{P_{\mathcal{O}}(Y) + P_{\mathcal{O}}^{\perp}(L_k(\lambda, \mathcal{O}))\right\}. \tag{11}$$

This iteration continues until convergence. Finally, the optimal λ is selected via cross-validation where $\mathcal{O}$ is randomly splitted into K subsets, ensuring that each subset preserves the observed fraction of data. A sequence of candidate values $\lambda_1 > \lambda_2 > \cdots > \lambda_L = 0$ is tested, and the value minimizing the average squared error on hold-out data is chosen. To expedite computation, each $\hat{L}(\lambda_i, \mathcal{O}_k)$ can serve as a warm-start for computing $\hat{L}(\lambda_{i+1}, \mathcal{O}_k)$.

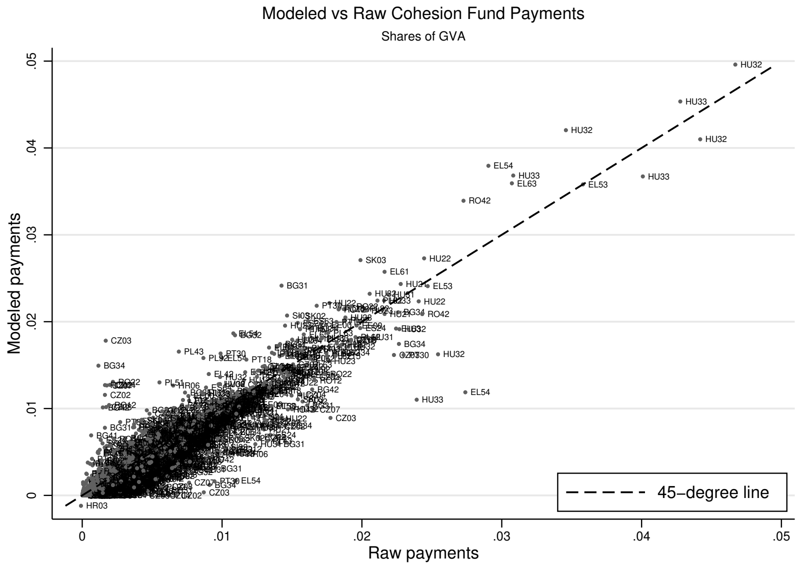
B Data

As already mentioned in the main text, there are two important issues related to the dataset of the Cohesion Fund expenditures.

The first concern is related to the timing between the moment when these disbursements became financial inflows, i.e. the moment when the expenditures actually took place on the ground and the reimbursement of the funds by the European Commission to the Member States (for more on this, see European Commission (2017)). In practical terms, this means that the dataset includes many gaps as the reimbursement of expenditures by the European Commission is not made in a timely manner. To circumvent this issue, we make use of the 'modeled' expenditures series. This is a series constructed using Monte Carlo simulations that aim to approximate what the actual expenditure series may look like.

As can be gleaned from Figure B.1, the correlation between the modeled and the actual payments, as shares of GVA, is –on average– very high, with a correlation coefficient equal to 0.93 (see, also, Fiuratti et al. (2024) for additional analysis).

Figure B.1: Correlation between modelled and raw Cohesion Fund payments



Notes: Scatter plot of modeled and raw Cohesion Fund expenditures. The Figure is based on a sample of 105 regions that are recipients of the Cohesion Fund and 2439 observations.

Another issue is that the data are not reported under a single NUTS2 version[14]; rather, we observe that a mix of NUTS2 versions is included in the data. We proceed with converting all the data to the NUTS2 2021 version by making use of the NUTS converter developed by the European Commission's Joint Research Centre. The converter can be located in the following link: `https://urban.jrc.ec.europa.eu/tools/nuts-converter?lng=en`.

[14]The Nomenclature of territorial units for statistics (NUTS), introduced in 2003, is a geographical nomenclature dividing the European Union territory into regions at three different levels. In this paper, we utilize data for the NUTS2 level given that this is the level with which EU cohesion policy is concerned. There have been 5 updates to the NUTS versions, the most recent one being introduced in 2021.

C Tables and figures

C.1 Observed and counterfactual GVA levels and growth rates

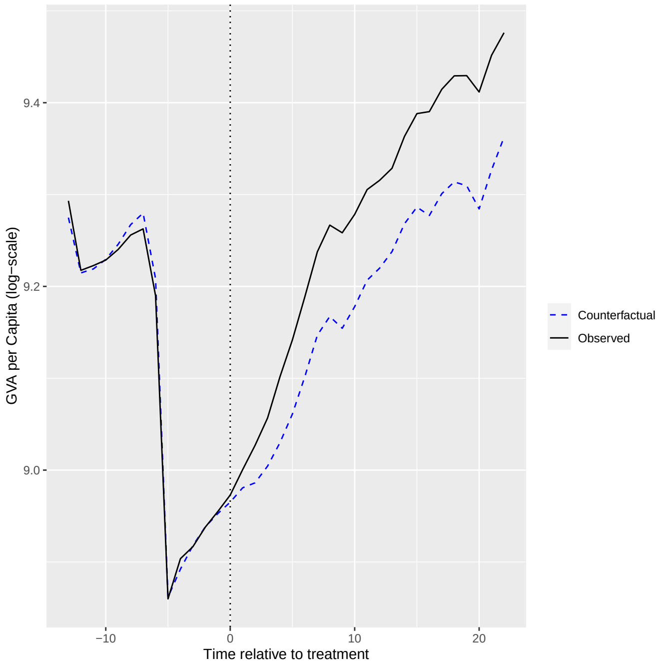


Figure C.1: Observed and estimated counterfactual average GVA in logs.

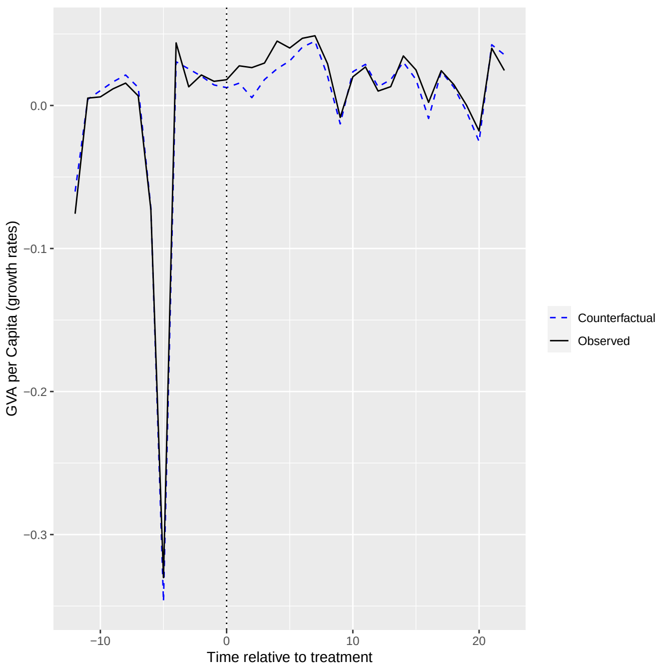


Figure C.2: Observed and estimated counterfactual average GVA growth rates.

C.2 Tables of growth rates excluding outliers

Table C.1: Log-difference of GVA per capita by programming period (%) excluding outliers

Period	Regions	Observed	Counterfactual
1994-1999	32	13.84	18.01
2000-2006	87	22.2	13.8
2007-2013	87	-1.66	0.69
2014-2020	71	22.4	15.11

Notes: The table depicts the difference between the end and beginning of period log of GVA per capita. As such, it corresponds to the overall period growth rate (not the annual average).

The 1994-1999 period actually corresponds to the 1993-1999 calendar year, as all the regions received a small proportion of the Cohesion Fund envelope in 1993.

The decline in regions in the 2014-2020 period is due to Spain not being eligible for the Cohesion Fund.

Table C.2: Log-difference of GVA per capita, years under treatment (%) excluding outliers

Years	Regions	Observed	Counterfactual
7	87	20.9	14.2
14	87	9.87	9.33
21	83	4.31	1.78

Notes: For additional notes, see Table

Table C.3: GVA growth rates by years under treatment

Percentile	**Initial GVA**	**Years under treatment**	$g_y^{\textbf{treated}}$**, in %**	$g_y^{\textbf{counter}}$**, in %**
10th	4137.21	**7 Years**	25.7	-9.03
25th	5273.65		23.22	9.33
50th	8416.783		18.38	9.54
75th	11928.78		8.06	13.27
90th	17153.59		13.29	22.15
10th	4137.21	**14 Years**	-22.33	1.79
25th	5273.65		17.14	8.98
50th	8416.783		-6.01	1.47
75th	11928.78		20.04	11.52
90th	17153.59		11.66	15.88
10th	4137.21	**21 Years**	-3.23	-6.56
25th	5273.65		21.83	8.97
50th	8416.783		5.71	2.06
75th	11928.78		-27.8	-5.27
90th	17153.59		-9.15	-8.78

Notes: The last two columns depict the overall period growth rate for the cases in which the region is treated and the counterfactual, respectively.
Initial GVA refers to the GVA per capita in the first year of the treatment

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS
ΣΧΟΛΗ ΟΙΚΟΝΟΜΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
SCHOOL OF ECONOMIC SCIENCES
ΤΜΗΜΑ ΟΙΚΟΝΟΜΙΚΗΣ ΕΠΙΣΤΗΜΗΣ
DEPARTMENT OF ECONOMICS

Department of Economics
Athens University of Economics and Business

List of Recent Working Papers

2023

01-23 Real interest rate and monetary policy in the post Bretton Woods United States, George C. Bitros and Mara Vidali
02-23 Debt targets and fiscal consolidation in a two-country HANK model: the case of Euro Area, Xiaoshan Chen, Spyridon Lazarakis and Petros Varthalitis
03-23 Central bank digital currencies: Foundational issues and prospects looking forward, George C. Bitros and Anastasios G. Malliaris
04-23 The State and the Economy of Modern Greece. Key Drivers from 1821 to the Present, George Alogoskoufis
05-23 Sparse spanning portfolios and under-diversification with second-order stochastic dominance, Stelios Arvanitis, Olivier Scaillet, Nikolas Topaloglou
06-23 What makes for survival? Key characteristics of Greek incubated early-stage startup(per)s during the Crisis: a multivariate and machine learning approach, Ioannis Besis, Ioanna Sapfo Pepelasis and Spiros Paraskevas
07-23 The Twin Deficits, Monetary Instability and Debt Crises in the History of Modern Greece, George Alogoskoufis
08-23 Dealing with endogenous regressors using copulas; on the problem of near multicollinearity, Dimitris Christopoulos, Dimitris Smyrnakis and Elias Tzavalis
09-23 A machine learning approach to construct quarterly data on intangible investment for Eurozone, Angelos Alexopoulos and Petros Varthalitis
10-23 Asymmetries in Post-War Monetary Arrangements in Europe: From Bretton Woods to the Euro Area, George Alogoskoufis, Konstantinos Gravas and Laurent Jacque
11-23 Unanticipated Inflation, Unemployment Persistence and the New Keynesian Phillips Curve, George Alogoskoufis and Stelios Giannoulakis
12-23 Threshold Endogeneity in Threshold VARs: An Application to Monetary State Dependence, Dimitris Christopoulos, Peter McAdam and Elias Tzavalis
13-23 A DSGE Model for the European Unemployment Persistence, Konstantinos Giakas
14-23 Binary public decisions with a status quo: undominated mechanisms without coercion, Efthymios Athanasiou and Giacomo Valletta
15-23 Does Agents' learning explain deviations in the Euro Area between the Core and the Periphery? George Economides, Konstantinos Mavrigiannakis and Vanghelis Vassilatos
16-23 Mild Explocivity, Persistent Homology and Cryptocurrencies' Bubbles: An Empirical Exercise, Stelios Arvanitis and Michalis Detsis
17-23 A network and machine learning approach to detect Value Added Tax fraud, Angelos Alexopoulos, Petros Dellaportas, Stanley Gyoshev, Christos Kotsogiannis, Sofia C. Olhede, Trifon Pavkov
18-23 Time Varying Three Pass Regression Filter, Yiannis Dendramis, George Kapetanios, Massimiliano Marcellino
19-23 From debt arithmetic to fiscal sustainability and fiscal rules: Taking stock, George Economides, Natasha Miouli and Apostolis Philippopoulos
20-23 Stochastic Arbitrage Opportunities: Set Estimation and Statistical Testing, Stelios Arvanitis and Thierry Post

21-23 **Behavioral Personae, Stochastic Dominance, and the Cryptocurrency Market, Stelios Arvanitis, Nikolas Topaloglou, and Georgios Tsomidis**
22-23 **Block Empirical Likelihood Inference for Stochastic Bounding: Large Deviations Asymptotics Under *m*-Dependence, Stelios Arvanitis and Nikolas Topaloglou**
23-23 **A Consolidation of the Neoclassical Macroeconomic Competitive General Equilibrium Theory via Keynesianism (Part 1 and Part 2), Angelos Angelopoulos**
24-23 **Limit Theory for Martingale Transforms with Heavy-Tailed Noise, Stelios Arvanitis and Alexandros Louka**

2024

01-24 **Market Timing & Predictive Complexity, Stelios Arvanitis, Foteini Kyriazi, Dimitrios Thomakos**
02-24 **Multi-Objective Frequentistic Model Averaging with an Application to Economic Growth, Stelios Arvanitis, Mehmet Pinar, Thanasis Stengos, Nikolas Topaloglou**
03-24 **State dependent fiscal multipliers in a Small Open Economy, Xiaoshan Chen, Jilei Huang, Petros Varthalitis**
04-24 **Public debt consolidation: Aggregate and distributional implications in a small open economy of the Euro Area, Eleftherios-Theodoros Roumpanis**
05-24 **Intangible investment during the Global Financial Crisis in the EU, Vassiliki Dimakopoulou, Stelios Sakkas and Petros Varthalitis**
06-24 **Time will tell! Towards the construction of instantaneous indicators of different agent-types, Iordanis Kalaitzoglou, Stelios Arvanitis**
07-24 **Norm Constrained Empirical Portfolio Optimization with Stochastic Dominance: Robust Optimization Non-Asymptotics, Stelios Arvanitis**
08-24 **Asymptotics of a QLR-type test for optimal predictive ability, Stelios Arvanitis**
09-24 **Strongly Equitable General Equilibrium Allocations, Angelos Angelopoulos**
10-24 **The Greek macroeconomy: A note on the current situation and future outlook, Apostolis Philippopoulos**
11-24 **Evolution of Greek Tax System, A Survey of Legislated Tax Changes from 1974 to 2018, Panagiotis Asimakopoulos**
12-24 **Macroeconomic Impact of Tax Changes, The case of Greece from 1974 to 2018, Panagiotis Asimakopoulos**
13-24 **`Pareto, Edgeworth, Walras, Shapley' Equivalence in a Small Economy, Angelos Angelopoulos**
14-24 **Stimulating long-term growth and welfare in the U.S, James Malley and Apostolis Philippopoulos**
15-24 **Distributionally Conservative Stochastic Dominance via Subsampling, Stelios Arvanitis**
16-24 **Before and After the Political Transition of 1974. Institutions, Politics, and the Economy of Post-War Greece, George Alogoskoufis**
17-24 **Gaussian Stochastic Volatility, Misspecified Volatility Filters and Indirect Inference Estimation, Stelios Arvanitis, Antonis Demos**
18-24 **Endogenous Realtor Intermediation and Housing Market Liquidity, Miroslav Gabrovski, Ioannis Kospentaris, Victor Ortego-Marti**
19-24 **Universal Choice Spaces and Expected Utility: A Banach-type Functorial Fixed Point, Stelios Arvanitis, Pantelis Argyropoulos, Spyros Vassilakis**

2025

01-25 **Democracy, redistribution, and economic growth: Some evidence from post-1974 Greece, George C. Bitros**
02-25 **Assessing Downside Public Debt Risks in an Environment of Negative Interest Rates Growth Differentials, Yiannis Dendramis, Georgios Dimitrakopoulos and Elias Tzavalis**
03-25 **Wicksellian General Equilibrium, Angelos Angelopoulos**

Department of Economics
Athens University of Economics and Business

The Department is the oldest Department of Economics in Greece with a pioneering role in organising postgraduate studies in Economics since 1978. Its priority has always been to bring together highly qualified academics and top quality students. Faculty members specialize in a wide range of topics in economics, with teaching and research experience in world-class universities and publications in top academic journals.

The Department constantly strives to maintain its high level of research and teaching standards. It covers a wide range of economic studies in micro-and macroeconomic analysis, banking and finance, public and monetary economics, international and rural economics, labour economics, industrial organization and strategy, economics of the environment and natural resources, economic history and relevant quantitative tools of mathematics, statistics and econometrics.

Its undergraduate program attracts high quality students who, after successful completion of their studies, have excellent prospects for employment in the private and public sector, including areas such as business, banking, finance and advisory services. Also, graduates of the program have solid foundations in economics and related tools and are regularly admitted to top graduate programs internationally. Three specializations are offered:1. Economic Theory and Policy, 2. Business Economics and Finance and 3. International and European Economics. The postgraduate programs of the Department (M.Sc and Ph.D) are highly regarded and attract a large number of quality candidates every year.

For more information:

https://www.dept.aueb.gr/en/econ/