Erasmus University Rotterdam
Erasmus School of Economics
Bachelor Thesis Econometrics and Operations Research

Unveiling the Instability of the Fiscal Multiplier Estimation

Alyde Bles (519018)

Supervisor:	Mikhail Zhelonkin
Second assessor:	Teresa Marreiros Bago d'Uva
Date final version:	2nd July 2023

Abstract

The fiscal multiplier, which estimates the impact of changes in fiscal spending on a nation's economic output, has been the subject of research for many years. There has been a wide range of findings, indicating the inherent instability of the fiscal multiplier. This paper addresses this issue by comparing the estimate of the fiscal multiplier calculated in a study by Acconcia et al. (2014), with an alternative methodology, namely the Local Projection (LP) method. The study by Acconcia et al. (2014) uses Instrumental Variable (IV) analysis to estimate the multiplier by exploiting a law related to Mafia activities in Italy. The first part of this paper examines the assumptions and pitfalls of the IV method. There are clear signs of sensitivity to outliers and the instruments have high leverage. Moreover, there is evidence of heteroskedasticity in the residuals. These pitfalls are a threat to the reliability of the results of the IV analysis. In the second part, the LP approach is examined. The results show that the dynamic multiplier obtained through LP is substantially smaller than the multiplier in the study of Acconcia et al. (2014). Additionally, this study highlights the difference in fiscal multiplier estimates between economic regimes. Overall, we demonstrate the instability of the fiscal multiplier estimate when a different methodology is used, raising concerns about the reliability of the findings of Acconcia et al. (2014). These results caution policymakers about the variability of the fiscal multiplier estimation and emphasise the need for careful consideration in decision-making processes.

1 Introduction

Over the years, numerous researchers have studied the fiscal multiplier, aiming to uncover the true nature of this multiplier. These studies use a variety of methods in different contexts to estimate the multiplier. Leading to a wide range of findings, with estimates ranging from negative values to estimates as high as 5. The fiscal multiplier estimates the effect of increased or decreased fiscal spending on a nation's economic output or Gross Domestic Product (GDP). This estimate can vary by the state of the economy (Baum et al., 2012; Auerbach & Gorodnichenko, 2012), the use of different identification strategies for the government spending shocks (Caldara & Kamps, 2017; Ramey, 2011) or perhaps a change in the monetary policy (Miyamoto et al., 2018).

A study conducted by Geli & S. Moura (2023) shows how methodological details can significantly affect the size, persistence, and precision of fiscal multiplier estimates. The issue of the instability of the results is not limited to the fiscal multiplier field alone. It is a broader problem within the empirical economic research field. The instability arises due to the reliance on limited data that is heavily influenced by the specific context in which it is collected. Ioannidis & Doucouliagos (2017) identified widespread bias and weak statistical power in the empirical economics literature, with the majority of the average effects being exaggerated by a factor of at least 2. They conclude that there is a need for improvements in study designs and use replication opportunities to reduce bias and enhance the robustness of the findings.

In this paper, we address this challenge - the instability of the fiscal multiplier. The paper by Acconcia et al. (2014) is used to compare and show this instability. In the paper by Acconcia et al. (2014), the fiscal multiplier is calculated by exploiting a law that mandates the mafia in Italy. The method is based on an Instrumental Variable (hereinafter referred to as IV) analysis with the use of Two Stage Least Squares (hereinafter referred to as 2SLS). Specifically, this paper aims to answer the following research question: Does the fiscal multiplier estimated by Acconcia et al. (2014) remain stable when an alternative methodology is employed? In order to answer this research question, the study is conducted in two parts. First, the paper by Acconcia et al. (2014) is replicated and reviewed. The assumptions of the 2SLS method are tested and several pitfalls are listed. Second, an alternative approach, the Local Projection (hereinafter, LP) method, for the estimation of the fiscal multiplier is examined.

To ensure the comparability of the results, the same data set is used for both parts. By doing so, the results are not influenced by varying environments, enabling a primary focus on the methods. The data set is assembled by Acconcia et al. (2014) and contains various macroeconomic variables from 1986 until 1999 of 95 provinces in Italy.

For the first part of the paper, the method and assumptions of Acconcia et al. (2014) are examined. Based on recommendations of Young (2022) two potential pitfalls are investigated. The sensitivity to outliers is researched and the findings suggest that there is strong evidence that the strength of the instruments relies on outliers. Second, the behaviour of the residuals is analysed and we find evidence of heteroskedasticity in the residuals of the first-stage regressions.

For the alternative approach, we use the LP method that is proposed by Jordà (2005). The government investment shocks are identified by applying a recursive identification strategy to a Structural Vector Auto Regressive (hereinafter, SVAR) model. The obtained Impulse Response Functions (hereinafter, IRFs) are used to calculate the multiplier. As an extension, the public investment multiplier is calculated in slack and good economic situations. This is estimated, by applying smooth-transition panel local projections to the data set. Lastly, we discuss the fiscal foresight problem. The results show that the dynamic multiplier is substantially smaller than the multiplier of Acconcia et al. (2014). In the LP method the 1-year dynamic multiplier has a value of 0.29 and in the method of Acconcia et al. (2014) it reaches a dynamic multiplier of 1.95. Furthermore, we provide evidence that the fiscal multiplier estimates can differ based on the economic state. These results suggest that the estimated fiscal multiplier in the study conducted by Acconcia et al. (2014) is unstable when a different method is used. There is evidence that the results of the study of Acconcia et al. (2014) are over-exaggerated, as well as that there is an overemphasis on the relevance of the instruments.

An important limitation of this study is the data set. The reliability of the results obtained through the LP approach tends to improve when a larger data set with quarterly observations is used. Nonetheless, the use of this particular data set allowed us to focus on the strengths and limitations of the 2SLS and LP methods primarily, as they were applied to the same data set.

The paper is organised in the following way. Section 2 provides a brief literature review of the different empirical approaches for calculating the fiscal multiplier. In Section 3, the study by Acconcia et al. (2014) is reviewed, starting with a detailed description of the data and the variables that are used. Moreover, within this section, the pitfalls of the IV approach are listed. Section 4 presents the methodology of the LP approach, with Section 5 presenting the results of the linear and state-dependent model calculated with the LP approach. Section 6 focuses on the comparison of the different approaches and addresses the fiscal foresight problem. Finally, Section 7 concludes the paper.

2 Literature Review

This section summarises the empirical approaches and estimates of the public spending multipliers. The three methods discussed in this section are the IV approach, the Vector Auto Regressive (hereinafter, VAR) approach, and the LP approach.

2.1 The IV Approach

The first approach that we discuss is the IV estimation approach. This is the method that Acconcia et al. (2014) use and is therefore deeply examined. In Acconcia et al. (2014) IV estimation with 2SLS is used. The basic idea behind IV estimation is to use the variation of the instrumental variables to identify the causal relationship between the exogenous variables and the outcome variable. This can be done through a two-stage process, known as the 2SLS technique. In the first-stage, the instrumental variables are used to estimate the relationship between the endogenous variables and the instrumental variables. The estimation provides the predicted values of the endogenous variables, which are now exogenous. In the second-stage, the predicted values obtained from the first-stage are used as new values for the endogenous variables. These predicted variables are then included in the regression model (Wooldridge, 2015). In our case, these predicted values are government spending and are used to estimate the effects on a nation's economic output. Several studies have applied this methodology to examine the impact of fiscal shocks. Nakamura & Steinsson (2014) use this technique, by exploiting an instrument based on military build-ups. They estimate an "open economy relative multiplier" of approximately 1.5. Corsetti et al. (2012) also use the 2SLS technique to estimate the multiplier. They identify exogenous spending shocks by estimating the difference between fiscal policy rules and their effects on the economy. The authors find an unconditional multiplier between 0.5 and 1. Other instruments that have been used in the literature are for example fiscal rules (Clemens & Miran, 2012) or predetermined factors for public spending (Kameda et al., 2021).

One of the major concerns in 2SLS is the issue of weak instrument bias. Weak instruments refer to instrumental variables with a weak or insignificant relationship with the endogenous regressor. As a result, the weak instruments can lead to biased estimates in the 2SLS regression and large size distortions of the hypothesis tests (Stock & Yogo, 2002). Moreover, Kiviet & Niemczyk (2007) highlight that 2SLS estimation may perform worse than the inconsistent Ordinary Least Squares (OLS) when weak instruments are available. In a study by Andrews et al. (2018) it is found that in a sample of 100 papers, there is a spike in first-stage F-statistics at a value of 10. This observation is intriguing as it corresponds to the threshold value for instruments to be considered strong. Young (2022) reviewed many published articles that applied the 2SLS method. In many of the 2SLS estimations that he examined, the results of instrument relevance were heavily influenced by one atypical observation or cluster. Furthermore, he showed that with clustered and heteroscedastic errors, in high-leverage papers the probability of an F-statistic greater than 10 rises to 60%. Thus, the benchmark of 10 that is now used may not be accurate anymore. Finally, Young (2022) analysis revealed that the 2SLS point estimates can rarely reject the OLS point estimates. He concludes that this is always true for high-leverage papers, but also in low-leverage papers, only 25% are able to reject the OLS point estimates.

2.2 The VAR Model

The VAR approach is the most common technique to estimate the fiscal multiplier. The method is popular because it has strong prediction power and it is easy to implement. The key challenge of estimating a VAR is to isolate the exogenous shocks. The identification of the exogenous shocks can be achieved through different identification schemes.

SVAR: in this identification strategy, one uses quarterly or annual data and makes the assumption that policy changes do not lead to immediate fluctuations in output within a single time lag. Blanchard & Perotti (2002) were the first who used this form of identification to estimate the fiscal multiplier. They identify the exogenous shock as they assume that government spending is not forecasted by lags of any of the variables included in the model. They estimated a government spending multiplier of 0.84. When they included anticipated shocks the multiplier reached a value of 2. Auerbach & Gorodnichenko (2012) adapted this approach by allowing for a regime-switching model to improve the flexibility of the fiscal multiplier and to study the multiplier under certain economic states. They conclude that fiscal policy is more effective when the economy is in a recession than in an expansion.

Sign Restriction: the second strategy for the identification of shocks in a VAR model is the sign restriction. This strategy is less dependent on a qualitative assessment of exo-or endogeneity. It identifies fiscal policy shocks by imposing restrictions directly on the sign of the impulse response functions. Mountford & Uhlig (2009) use this method and identified four shocks. Furthermore, they find clear negative effects of positive tax shocks. Other researches that used this identification strategy include: Pappa (2009) and Chian Koh (2017).

Narrative Approach: in this approach one calculates the multipliers using pre-identified events of exogenous increases in spending. A common example of this approach is to use war dates to identify fiscal spending shocks (Ramey & Shapiro, 1998).

The SVAR technique has been widely criticised. This technique uses numerous lagged endogenous variables to ensure that the error term is independent of the historical macroeconomic data. However, anticipated changes are not fully captured by these lagged values, and thus should be included. Ramey (2011) demonstrated this by showing that narrative strategies can better capture the difference between a policy announcement and its actual implementation. However, this narrative method can be very time consuming and is constrained to data availability, which limits the reliability of the results. Additionally, VAR models are impractical when dealing with panel data, due to their high dimensionality (Li et al., 2022). Finally, another limitation is that the results of a VAR analysis can be less reliable if one uses a small data set.

2.3 The LP Method

More recently, the LP method proposed by Jordà (2005), has become popular in the fiscal multiplier literature. This approach offers an alternative to VAR models by estimating IRFs through separate regressions for each time horizon. The local projection method allows for a dynamic analysis, capturing the effects of fiscal shocks over time. Auerbach & Gorodnichenko (2017) use this approach with annual data and estimated a government spending multiplier of 0.67 at impact. Similarly, Ramey & Zubairy (2018) used the identified military news shock from Ramey & Shapiro (1998) to substitute it in the LP equations and estimated a 2-year multiplier of 0.66. However, when they estimated the multiplier using shocks identified by Blanchard & Perotti (2002), the resulting estimate for the 2-year multiplier was 0.37.

The main advantage of the LP method compared to the VAR approach is that it is more robust to model misspecifications (Li et al., 2022). It can easily adapt non-linear specifications and can easily be estimated using standard regression packages (Restrepo-Ángel et al., 2022). However, the LP method may not necessarily outperform the VAR method in calculating IRFs. The LP method does not impose constraints on the impulse responses between different horizons, resulting in potentially erratic responses and reduced statistical efficiency. Moreover, the LP impulse responses may show abrupt oscillations when analysed over long horizons (Restrepo-Ángel et al., 2022). Several studies have calculated fiscal shocks through the LP framework using shocks that are identified through the narrative approach. Recently, researchers combine SVAR models with LPs by first identifying shocks in the SVAR framework and then substituting them into the LP equations. This approach has been used in several studies, such as in Deleidi et al. (2023).

3 Analysis of the IV Study

In this section, we first provide an explanation of the data set and give an introduction of the variables that are used in this study. Then, as we replicated the study by Acconcia et al. (2014), a review of the study is given and the assumptions are stated along with their importance. Finally, the assumptions of the methods used by Acconcia et al. (2014) are tested and the limitations are discussed.

3.1 Data and Variable Description

The data set used in this paper consists of province-level panel macroeconomic variables from 1986 until 1999 of 95 provinces in Italy. The variables are assembled by Acconcia et al. (2014) and are obtained from various data sources in Italy, including the Italian Institute of Statistics (ISTAT), Istituto Guglielmo Tagliacarne and the Ministry of Internal Affairs of Italy (Ministero dellí'Interno). The following variables are included in the analysis:

Value Added: the percentage growth rate of real per-capita total value added is defined as $Y_{i,t} = \frac{y_{i,t} - y_{i,t-1}}{y_{i,t-1}}$, where $y_{i,t}$ is the real per capita value added. The measurement unit is in millions of euros at current prices.

Government Investment Spending: this variable has two specifications. The first one is the year-on-year change as a ratio of lagged value added, denoted as $G^y_{i,t}$. It is computed as $\frac{g_{i,t} - g_{i,t-1}}{y_{i,t-1}}$, where $g_{i,t}$ is the real per capita public investment value added. The second variable is the growth rate of real per capita public investment, defined as $G_{i,t} = \frac{g_{i,t} - g_{i,t-1}}{g_{i,t-1}}$.

Council Dismissals (CD): the variable CD represents the number of municipalities placed under the administration of an external commissioner by the government due to evidence of mafia infiltration. Additionally, two specifications of CD are included in the analysis. The first one, denoted as $CDS1$, is based on the publication of the official degree in the first semester of the year. The second specification, $CDS2$, considers cases where the average number of days between the dismissals of the city council and the year-end is less than 180.

Employment: two specifications for employment are included in the analysis. The first one, denoted as $U1$, is the logarithm of per capita employment. The second specification, $U2$, captures the change in the logarithm of per capita hours of wage supplement available to employees of private firms in Italy through the unemployment insurance scheme.

The paper of Acconcia et al. (2014) uses various control variables, however as the control variables are not important for this analysis these are not discussed. For a specific explanation of the sources of the variables and the definitions of the control variables that are used, we refer to the replication package of Acconcia et al. (2014). Descriptive statistics for the variables can be found in Table 6 of the appendix. It is important to note that the exact same data set is used for both the replication part and the implementation of the LP method to ensure the comparability of the results.

3.2 Review of the IV Study

In this part, the method of Acconcia et al. (2014) is discussed. The study focuses on the fiscal multiplier of government spending at a local level using the IV approach with 2SLS.

The baseline empirical model is shown below, where the literature of Barro & Redlick (2011) is followed:

$$Y_{i,t} = \beta G^y_{i,t} + \alpha_i + \lambda_t + \gamma X_{i,t} + \epsilon_{i,t}, \tag{1}$$

here, β is the parameter of interest and measures the contemporaneous one-year government investment multiplier. The variables $Y_{i,t}$ and $G^y_{i,t}$ are specified as described in Section 3.1. They make use of two fixed effects, namely the province-fixed effect (α_i) and the year-fixed effect (λ_t). The province-fixed effect addresses the endogeneity issue as the province-specific characteristics are correlated with the government spending allocation criteria. The year-fixed effect controls for endogeneity for the cyclical developments and for the monetary and fiscal policy at a national level. The control matrix X has five variables measuring the number of people related to the judicial authority for different crimes. The authors acknowledged that their model has two potential problems that could bias the results. The first problem is that there may be anticipation effects because government spending on infrastructure is usually planned several years before it is actually implemented. The second problem is allocating bias, which occurs because the government allocates funds based on local developments. To address these issues, the researchers needed an exogenous shock that was unrelated to the local economy. They introduced two variables related to the compulsory administration of municipalities when there was evidence of mafia infiltration. The two instruments that were used are the $CDS1$ and the $CDS2$ variables, as described in Section 3.1. Government spending is then instrumented with $CDS1$ and the one-period lag of $CDS2$. This results in the following first-stage regression:

$$G^y_{i,t} = \delta_1 CDS1_{i,t} + \delta_1 CDS2_{i,t-1} + \alpha_i + \lambda_t + \gamma X_{i,t} + \epsilon_{i,t}. \tag{2}$$

To test the validity of the estimation of the fiscal multiplier using 2SLS, several tests were conducted. The first and most crucial test is the instrument relevance test, which aims to assess whether the instruments have a causal effect on the endogenous regressor. A violation of this test can result in biased estimates. When the instruments do not have a causal effect on the endogenous regressor, it becomes irrelevant, leading to zero correlation between the instrument and the outcome variable. In such cases, the 2SLS method loses its advantages in addressing the endogeneity problem and can introduce more bias as it uses two regressions. The instrument relevance assumption is typically evaluated with the use of the first-stage F-statistics. In our replication, the F-statistics for the headline results[1] reached a value of 12.00, exceeding the threshold value of 10 that is suggested by the "rule-of-thumb" for a single regressor (Staiger & Stock, 1994). This outcome indicates that the instruments are indeed relevant.

The second assumption of IV estimation is the exogeneity assumption, which requires the instruments to be uncorrelated with the error term and only affects the outcome variable through the endogenous variable. If this assumption fails, the correlation between the instrument and the outcome variable might just reflect some unobserved confounding effect rather than a true causal effect. To verify this assumption, two tests were conducted. First, the endogeneity of the potentially endogenous variable, government spending, was tested using the Wu-Hausman test, yielding a p-value of 0.00005. The null hypothesis of government spending being exogenous was rejected. Second, the Sargan J test (Sargan, 1958) was performed to assess the exogeneity of the instruments. This test can be applied when there are more instruments than endogenous variables. In our case, the p-value was high (0.97801), suggesting that the null hypothesis cannot be rejected and the instruments are not correlated with the errors in the 2SLS. Hence, they qualify as correctly identified instruments.

The last assumption that needs to hold is the exclusion restriction. This assumption requires that the instruments do not have a direct causal effect on the outcome variable. If there is an actual direct causal effect, it becomes challenging to separate the effect from the true effect of the endogenous regressor on the dependent variable. Validating this exclusion restriction requires a combination of statistical analysis and access to administrative documents, as used by the authors of the paper. For the statistical analysis, they controlled for the mafia activity by including measures of police investigations in their regressions model. On the administrative side, the researchers examined official documents and reports, to gain insight into the impact of the city council dismissals on economic activity. Due to the complexity and reliance on specific documentation, a further detailed discussion of this assumption and its validity is beyond the scope of this summary.

3.3 Concerns Regarding the IV Approach

The assumptions as stated above may not be valid, despite passing the tests. These tests are based on assumptions themselves, which can also be violated. In this part, the focus is on the violation of the weak instrument assumption. A study by Young (2022) raises concerns regarding this assumption, as he highlights the potential influence of heteroskedastic and robust errors on the reliability of results obtained through IV estimations, particularly in papers with high leverage. Moreover, the author emphasises the need for a sensitivity analysis to examine whether the results are heavily influenced by the outliers. In this part, we first discuss the sensitivity to outliers and thereafter the behaviour of the residuals.

[1] We define the headline results as the estimate the authors mention in the abstract, introduction and conclusion.

3.3.1 Sensitivity to Outliers

The first pitfall that is investigated, is the sensitivity to outliers. To assess this, we perform a sensitivity analysis by excluding one cluster iterative in each first-stage and second-stage regression of the paper. This analysis allows for a detailed examination of the impact of the outliers on the IV approach. This analysis is based on the methodology of Young (2022). Specifically, we removed one cluster and obtained the maximum p-values of the estimated coefficient of the instrumented variable, G^y, in the second-stage regression and the estimated coefficients of the instruments of the first-stage regression. Moreover, the minimum F-statistic is calculated iteratively after removing one cluster at a time.

Table 1 presents the results of the sensitivity analysis. We considered the headline result and the average results across the 2SLS regressions that are conducted throughout the paper. As seen in the table the p-value of the instrumented variable decreases in significance level. As the original p-value has a significance level of 0.01 and after deletion, the significance is 0.05. However, it can still be considered significant. Next, as seen in the table, the p-values of the instruments in the first-stage increase drastically. Making the instrument $CDS1$ insignificant for both the headline results and the average over the regressions. With a p-value of 0.112 and 0.113 respectively. Furthermore, the instrument $CDS2_{t-1}$ is after the deletion of one cluster only significant at a level of 0.1. Therefore, it is not surprising that the F-statistic falls below the "rule-of-thumb" of 10. Specifically, the average F-statistic across the regressions falls to a value of 9.64. These results suggest that there is evidence of weak instruments on an individual and a joint level.

Table 1: Sensitivity to outliers: deleting one cluster

	Headline		**Average over the Regressions**	
	Original	**Delete one cluster**	**Original**	**Delete one cluster**
$\boldsymbol{G^y}$	0.001	0.01	0.002	0.016
$\boldsymbol{CDS1}$	0.000	0.112	0.001	0.113
$\boldsymbol{CDS2_{t-1}}$	0.000	0.046	0.000	0.053
F-statistic	12.67	9.90	12.00	9.64

Notes: In this table, the sensitivity analysis is presented. The results for both the headline regressions and the average over 10 regression are shown. The p-values for the variables $CDS1$ and $CDS2_{t-1}$ in the first-stage regressions, as well as the p-value of the instrumented coefficient for G^y, are included. Additionally, the F-statistics of the first-stage regression are provided.

The results are sensitive to outliers. This sensitivity can reflect a concentration of leverage in a few clusters, which can be problematic for the reliability of the results. Specifically, it can influence the IV inference heavily, if a few clusters can generate the main variation in the instruments (Young, 2022). Leverage is defined as the diagonal elements of the hat matrix $H = X(X'X)^{-1}X'$. However, we want to calculate the leverage of each cluster of the instruments. This can be calculated as specified below:

$$Z'_{(i)}(Z'Z)^{-1}Z_{(i)}, \tag{3}$$

where $Z_{(i)}$ denotes the residual from the regression of the instruments on the covariates of the i-th cluster and Z represents the residuals when all clusters are included. For the headline results the maximum instrument leverage of instrument $CDS1$ is found to be 0.497, while for $CDS2_{t-1}$ it is 0.312. Following the literature of Young (2022), the average of these two leverages can be taken, which results in a maximum instrument leverage of 0.404. This value is considered "high" when compared to the values reported in the paper by Young (2022).

3.3.2 Behaviour of the Residuals

The second pitfall that we investigate is the deviation of residuals from the independent and identically distributed (i.i.d) normal ideal. As this poses a threat to the reliability of the IV estimation results. Specifically, we study the impact of heteroskedasticity, which is found to be the most influential factor leading to unreliable results in IV estimation. Andrews et al. (2018) discovered that heteroskedastic designs can lead to large coverage distortions in the standard 2SLS confidence intervals and the F-test which relies on the assumption of homoskedasticity, may also be invalid in such cases, which is also shown by Young (2022). If there is heteroskedasticity, the standard errors become more volatile, resulting in a greater dispersion in t-statistics while the degrees of freedom that are used to evaluate the distribution remain constant. In fact, an increase in heteroskedasticity raises the probability of obtaining an F-statistic greater than 10 when the instruments are actually irrelevant (Young, 2022).

However, for the sake of completeness, we test normality through a Q-Q plot, as shown in Figure 4 in the appendix. The Q-Q plot reveals that the residuals show symmetry, but there is evidence of heavy tails. However, since the residuals remain symmetric, this does not pose a significant problem for the reliability of the results. Moreover, in the results of Young (2022) it is seen that if the residuals are not normal, this does not increase the probability of obtaining a first-stage F-statistic of above 10 when actually the instruments are irrelevant.

To test the heteroskedasticity, the method formulated by Wooldridge (2015) is used. This is based on the test of Breusch-Pegan (Breusch & Pagan, 1979), where the Lagrange Multiplier is used as the test statistic (Koenker, 1981), as this test offers generally greater applicability than other tests (Wooldridge, 2015). In Table 2 the results are presented. We can reject the null hypothesis of homoskedasticity in every regression for the 0.05 significance level. This provides strong evidence of the presence of heteroskedasticity in the first-stage regressions. Moreover, two residual plots are shown, plotting the instruments against the residuals. Both plots show no signs of heteroskedasticity, as the points fluctuate around zero. This contradicts the statistical tests that were conducted before.

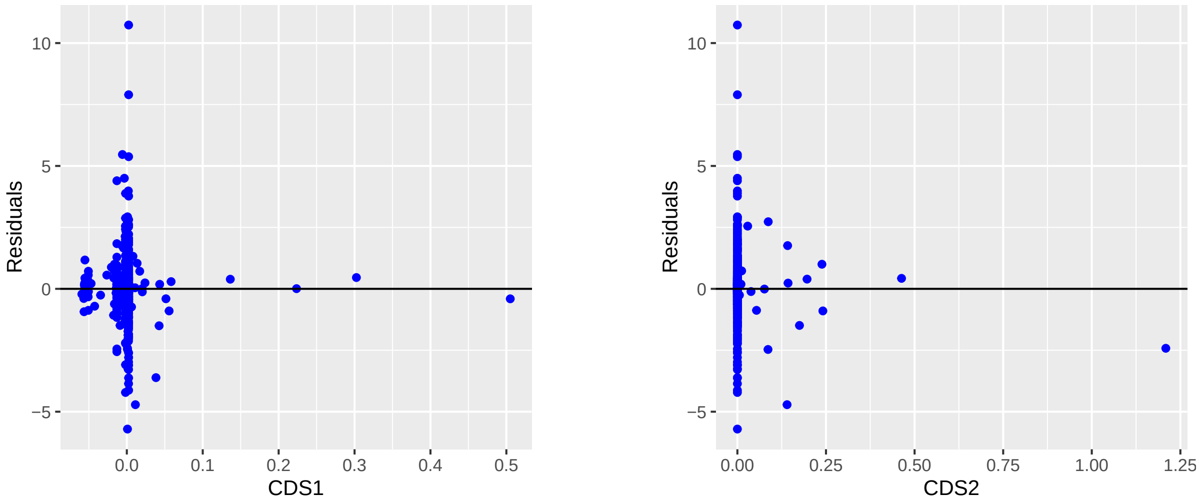
Concluding, there are signs of heteroskedasticity in the residuals of the first-stage, the results exhibit sensitivity to outliers and there is high leverage in the instrumental variables. These findings are potential limitations and could bias the result of the 2SLS.

Table 2: Test of Normality and Homoskedasticity

	0.05	**0.01**	**Average P-value**
Heteroskedasticity Test Koenker (1981)	1.00	0.67	0.018

Notes: 0.01/ 0.05 is the level of the test. This table represents the fraction of the regressions that reject the null of homoskedasticity for the given level of the test. There were 9 first-stage regressions tested. The test for heteroskedasticity is based on the R^2 of the residuals.

Figure 1: Residual Plots



Notes: The residual plots are presented in this figure. In the left panel the residuals of the first-stage are plotted against the instrument CDS1 and the right panel represents the residuals against the instrument CDS2.

4 The Alternative Approach: LP Method

In the previous section, concerns are raised about the reliability of the results. This section presents an alternative approach for estimating the fiscal multiplier, namely the LP Method. As explained in Section 3.1, the same data set is used for both parts of this study: in the review of the IV study and in the examination of the LP methodology. Therefore, in Section 3.1 you can find a description of the data set and the variables that are used within this section.

4.1 The Linear Model

Building on several recent contributions we estimated the public investment multiplier using the LP method (Jordà, 2005). The idea behind the LPs is to estimate the IRFs individually at different time horizons. There are several reasons why the LP method is used instead of the VAR approach to estimate the IRFs. First, the LP methodology is more robust to model misspecifications as it does not constrain the shape of the IRFs. Second, the VAR becomes impractical when one uses panel data as this results in high dimensionality. Concluding, the LP offers a natural and simple alternative for estimating the IRFs (Jordà, 2005).

The general expression as formulated by Jordà (2005) for the LP method is specified as:

$$\boldsymbol{y}_{t+h} = \boldsymbol{\alpha}^h + \boldsymbol{B}_1 \boldsymbol{y}_{t-1} + ... + \boldsymbol{B}_p^h \boldsymbol{y}_{t-p} + \boldsymbol{u}_{t+h}^h, \quad h = 0, ..., H-1, \tag{4}$$

where $\boldsymbol{\alpha}^h$ is a n x 1 vector of constants, and $\boldsymbol{B_h^p}$ are the parameter matrices for lag p and forecast horizon h. The structural impulse responses can be estimated from eq. (4) as:

$$\hat{IR}(t, s, \boldsymbol{d_i}) = \hat{\boldsymbol{B}}_1^s \boldsymbol{d}_i, \quad s = 0, ..., h, \tag{5}$$

where $\boldsymbol{d}_i$ is the shock matrix that can be identified from a linear VAR and with the normalisation $\boldsymbol{B_1^0} = \boldsymbol{I}$.

The first challenge of the LP approach is the identification of the shocks. The shocks need to be exogenous in order to obtain unbiased estimates. We follow the approach of Amendola (2022) and Deleidi et al. (2023) for the identification of the shocks and for the formulation of the LP equations. The fiscal multiplier can be calculated in two steps:

(i) Identify the shocks by applying a recursive identification strategy to an SVAR model (Blanchard & Perotti, 2002).

(ii) Estimate the multiplier through the IRFs, which can be obtained with the LP equations.

For the initial step, a recursive identification strategy is implemented, where the growth of per capita public investment is ordered first, and the growth of per capita value added is ordered as the second variable. Thus, assuming that public investment is not affected by the contemporaneous shock of the other variable. The structural shocks can be identified as the residuals of the first equation of the SVAR model:

$$G_{i,t} = aG_{i,t-1} + bY_{i,t-1} + e_{i,t}, \tag{6}$$

where $G_{i,t}$ is the growth rate of per capita public investment, $Y_{i,t}$ is the growth rate of per capita value added and $e_{i,t}$ is the identified shock.

The second step is to estimate the multiplier through the use of the IRFs. The dynamic effect of the public investment shock on a variable of interest can be easily retrieved with a panel LP by estimating the following series of LP equations:

$$(1) \quad Y_{i,t+h} = {\alpha_i}^h + {\lambda_t}^h + \beta_Y^h shock_{i,t} + \phi^h(L)X_{i,t} + \epsilon_{i,t+h}, \tag{7}$$

$$(2) \quad G_{i,t+h} = {\alpha_i}^h + {\lambda_t}^h + \beta_G^h shock_{i,t} + \phi^h(L)X_{i,t} + \epsilon_{i,t+h}, \tag{8}$$

where α_i and λ_t are the country and time fixed effects and $shock_{i,t}$ is the structural shock which is identified by eq. (6) as $e_{i,t} = shock_{i,t}$, and $X_{i,t}$ are the control variables with $\phi^h(L)$ being a polynomial lag of order 2. The control variables are the real growth of value added, the year-on-year change of public investment as a ratio of value added and two specifications of employment, $U1$ and $U2$[2]. As previously mentioned, β_y^h and β_G^h are the parameters of interest, and measure the dynamic effect of the IRFs.

The second challenge of the LP method is the calculation of the fiscal multiplier from the IRFs. The parameter β_y^h represents the value added elasticities to public investment, however,

[2]These variables are specified in Section 3.1.

this does not directly reveal the dynamic government investment multiplier. The estimated elasticities need to be converted to euro equivalents. This can be done by using an ex-post conversion factor that is based on the sample average of the ratio between the output and government investment ($\frac{\bar{Y}}{\bar{G}}$). According to Ramey & Zubairy (2018), the use of the ex-post conversion factor can introduce bias in estimating the multiplier. The ratio $\frac{\bar{Y}}{\bar{G}}$ can vary substantially over time in a historical sample, making it inappropriate to use as a constant value. However, in the context of this paper, where the sample size is only 10 years, the bias is not relevant. Therefore, we can assume a constant value for $\frac{\bar{Y}}{\bar{G}}$. Furthermore, as recommended by Ramey (2016), we estimate the cumulative fiscal multiplier. The cumulative multipliers are estimated as the integral response of value added relative to the integral of government investment change that occurred over a specific period. This allows us to examine potential long-lasting effects associated with permanent increases in public investment spending. In addition to the cumulative fiscal multiplier, we also calculated the dynamic multiplier for the purpose of comparison with the findings of Acconcia et al. (2014).

The estimation of the cumulative multiplier is shown below:

$$\text{Cumulative Multiplier}_k = \frac{\sum_{h=0}^{h=k} \beta_Y^h}{\sum_{h=0}^{h=k} \beta_G^h} \frac{\bar{Y}}{\bar{G}}, \quad 0 \leq k \leq H, \tag{9}$$

where $\beta^h(Y)$ and $\beta^h(G)$ are the coefficients corresponding to the growth of value added and growth of public investment spending.

4.2 The State-Dependent Model

Another advantage of the LP method is the flexibility of handling non-linearity and state dependency (Ramey & Zubairy, 2018). Therefore, as an extension, we analyse whether public investment multipliers vary with the state of the economy. The linear specification is extended into the following panel smooth transition LP model. The literature of Auerbach & Gorodnichenko (2012) and Ramey & Zubairy (2018) is followed. The state-dependent LP model is formulated as follows:

$$\begin{aligned} Y_{i,t+h} =& {\alpha_i}^h + {\lambda_t}^h + \beta_1^h shock_{i,t} \cdot F(z_{i,t-1}) \\ & + \beta_2^h shock_{i,t} \cdot (1 - F(z_{i,t-1}) + \gamma^h(L) X_{i,t} \\ & + \epsilon_{i,t+h}, \quad h = 0, ..., H, \end{aligned} \tag{10}$$

where indexes 1 and 2 represent the two states of the economy, namely the slack and good regimes. Variable z is an index of the business cycle, where a positive z indicates that the economy is in a good economic situation (expansion). $F(z)$ is the transition function, which is explained in more detail below. Furthermore, β_1 and β_2 capture the dynamic effect of the impulse response function when the economy is in a slack or a good state. The following transition function $F(z)$ is used (Auerbach & Gorodnichenko, 2012):

$$(1) \quad F(z_{i,t-i}) = \frac{\exp(-\gamma z_{i,t-1})}{1 + \exp(-\gamma z_{i,t-1})}, \quad \text{with} \quad \gamma > 0, \tag{11}$$

$$(2) \quad \text{Var}(z_t) = 1, \quad \text{and} \quad E(z_t) = 0.$$

The choice of index z is not trivial, as there is no theoretical prescription for what this value should be. In this research, two specifications are used. Namely, we use data on the "dynamic" of the value added growth and data on the rate of employment. The dynamics of value added are defined as the moving average of value added in Italy calculated with a rolling window. In Auerbach & Gorodnichenko (2012) a rolling window of 7 quarters is used to calculate the "dynamics", thus we used a rolling window of 2 years. The z is decomposed by the Hodrick Prescott (hereinafter, HP) filter (Hodrick & Prescott, 1997). For the HP filter, two different values are used, namely 10^6 and 4×10^3, to assess the robustness of the results obtained from the filtering process. Furthermore, the transition function, eq. (11), transforms the state variables into values between zero and one. Where an equal weight between the two regimes results in a $F(z)$ of 0.5. The parameter γ defines how fast the transition proceeds between the slack and good economic states. γ is calibrated in such a way that the economy spends about 20 percent of its time in the slack state, as we want the calibration of γ to be consistent with the duration of recessions (Auerbach & Gorodnichenko, 2012). This translates into a value of γ of 1.4 for the state variable the dynamic of value added and a value of γ of 1.8 for the state variable employment.

A complication associated with the Jordà (2005) method is the potential serial correlation of the error term with possible dependence across the provinces and the time. Therefore, the robust Driscoll & Kraay (1998) standard errors are used to account for this problem and make statistical inferences.

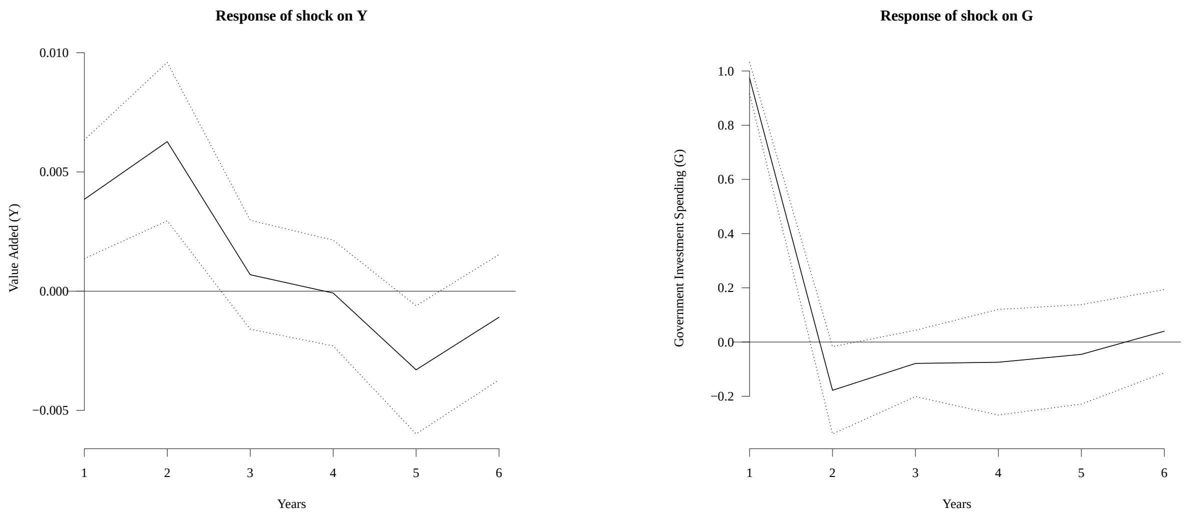
5 Results

In this section, the results of the local projections are discussed. Starting with the linear results and thereafter the state-dependent results.

5.1 The Linear Model

Figure 2 plots the IRFs of the linear specification. The left panel shows the effect of a government investment shock on value added, while the right panel shows the effect of the shock on government investment. The local projections are estimated six years ahead (h=6) with a confidence interval of 95% using standard errors from Driscoll & Kraay (1998). In the left panel of Figure 2 it is observed that the response to the government investment shock is positive up to a 4-year horizon. This indicates that the shock initially boots the value added, but after 4 years the effect turns negative. Additionally, the government investment shock is not persistent over time, as it starts to decline immediately. To further analyse the impact of the shocks the impulse response functions are transformed into multipliers, resulting in cumulative and dynamic public investment multipliers. This is shown in Table 3. By year 1, the size of the multiplier is 0.30 and increases to a value of 1.26 by year 4. Suggesting that the cumulative multiplier has an effect above unity over time. The dynamic multiplier exhibits a different pattern. By year 1 the dynamic multiplier has a value of 0.29 and rises to 0.47 by year 2. Thereafter, the effect disappears. Note that the cumulative multiplier and the dynamic multiplier are only significant in years 1 and 2. The lack of significance in the multipliers and the widening confidence intervals are not

Figure 2: The IRFs of the Linear Model



Notes: Linear impulse responses to a 1% shock of public investment spending. The dashed lines represent the 95% confidence intervals, based on the Driscoll & Kraay (1998) standard errors. The left panel shows the shock on value added (Y) and the right panel shows the shock on government investment spending (G).

surprising. As the LP method may exhibit erratic responses and reduced statistical efficiency over long horizons. Nevertheless, since our main focus lies in estimating the effects of the fiscal multiplier in the initial years, particularly for the comparison with the findings of Acconcia et al. (2014), these challenges will not significantly impede the comparison of the results.

Table 3: Fiscal Multiplier in the Linear LP Model

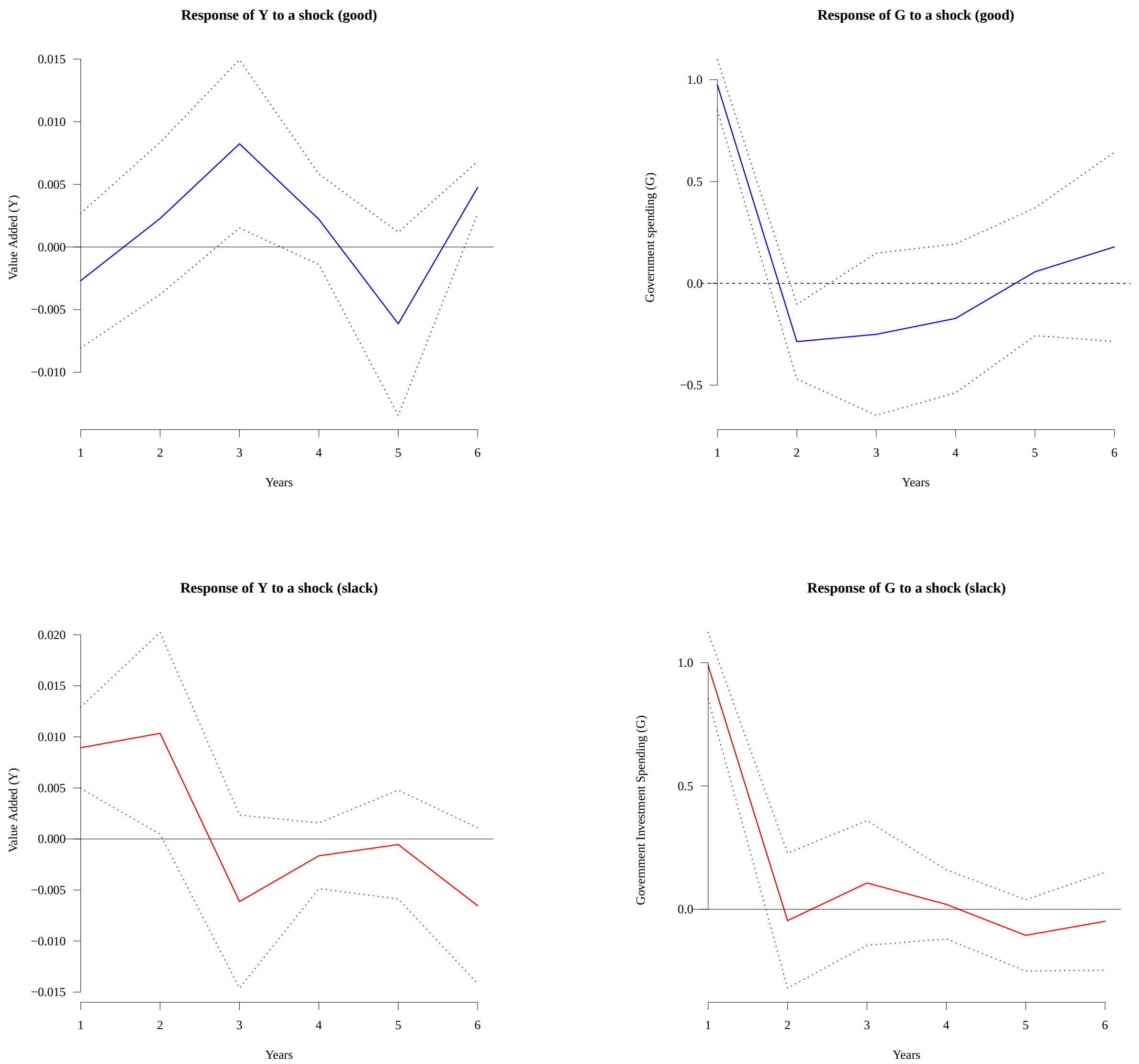
	1-year	**2-year**	**3-year**	**4-year**
Cumulative Multiplier	0.30**	0.96***	1.14	1.26
Dynamic Multiplier	0.29**	0.47**	0.05	-0.01

Notes: This table presents the cumulative and dynamic multiplier of the government investment shock over a 4-year horizon. The significance level is indicated with: *** $p < 0.01$, ** $p < 0.05$ and * $p < 0.1$.

5.2 The State-Dependent Model

As an extension, the impulse responses are investigated for different economic conditions, namely between slack and good economic situations. To further strengthen the findings, we tested the state-dependent model that incorporates two state variables and explore various configurations for these variables. The state variables investigated are the employment rate and the dynamics of value added. For the state variables, we looked at the deviations from the HP-filtered mean with a value of λ of 10^6 and 4 x 10^3. These specifications are based on the methodology of Amendola (2022). The benchmark state is the state defined as the standard deviations of the dynamics of value added from the HP-filtered mean with $\lambda = 10^6$. In Figure 3 the impulse responses in a good period and in a slack period for the benchmark state are presented. The

Figure 3: The IRFs for the state-dependent model for the state variable dynamics of value added



Notes: The state-dependent impulse responses to a 1% government investment shock. The dashed lines represent the 95% confidence intervals, based on Driscoll & Kraay (1998) standard errors. The IRFs in blue represent the functions in a good regime and the IRFs in red represent the responses in a slack regime. The state variable is the standard deviations of the dynamics of GDP (moving average of output growth) HP filtered with $\lambda = 10^6$.

first thing to notice is that in a slack period, the impulse response function has a similar shape as in the linear specification (see Figure 2). However, the response of Y in year 1 is more than two times bigger in a slack period compared to the linear model. In a good period, the shape of the response function of Y differs greatly compared to the slack period and thus also from the linear impulse responses. Value added has now a hump-shaped response with a peak at three years after the shock. The figures show that in a slack period, a government investment shock has a strong immediate effect and in a good period the shock has a delayed effect only reaching its peak at 3 years.

Table 4: The cumulative and dynamic multipliers for the state-dependent model

Dynamics of Value Added (HP-filtered with $\lambda = 10^6$)				
Multiplier	**1-year**	**2-year**	**3-year**	**4-year**
Cumulative Good	-0.21	-0.05	$1.36**$	2.86
Cumulative Slack	$0.68***$	$1.55**$	0.94	0.81
Dynamic Good	-0.20	0.17	0.62	0.17
Dynamic Slack	0.68	0.78	-0.46	-0.12
Dynamics of Value Added (HP-filtered with $\lambda = 4 \times 10^3$)				
Multiplier	**1-year**	**2-year**	**3-year**	**4-year**
Cumulative Good	-0.21	-0.05	$1.34**$	2.82
Cumulative Slack	$0.69***$	$1.55**$	0.95	0.82
Dynamic Good	-0.21	0.17	0.62	0.16
Dynamic Slack	0.68	0.78	-0.46	-0.12
Change in Employment (HP-filtered with $\lambda = 10^6$)				
Multiplier	**1-year**	**2-year**	**3-year**	**4-year**
Cumulative Good	-0.08	0.16	$0.54**$	$0.64*$
Cumulative Slack	$1.07**$	$5.09**$	$-6.22***$	-0.66
Dynamic Good	-0.11	0.32	0.66	0.36
Dynamic Slack	$0.70**$	$0.59***$	$-0.50***$	-0.46

Notes: The cumulative and dynamic multiplier effect of the government investment shock for the state-dependent model. It is calculated for two state variables: dynamics of value-added and the change of employment. Furthermore, two specifications for the HP trend are used. The significance level is indicated with: *** p <0.01, ** p < 0.05 and * p <0.1.

Table 4 represents the cumulative and dynamic multipliers for the different specifications of the state-dependent model. Focusing on the benchmark state, we observe that the cumulative multiplier during a slack period is positive and significant in the first two years following the shock. In particular, it reaches a value of 1.55 in year 2. On the other hand, the cumulative multiplier is not significant in the initial years during a good period, but shows a delayed effect. Which becomes significant only at year 3 with a cumulative multiplier of 1.36. The dynamic multiplier yields a similar pattern. In a period of slack, the multiplier is significant and high in the first years but turns negative and not significant in years 3 and 4. Furthermore, when comparing the benchmark state-dependent model with the linear model (Table 3), we find that in a slack period, the cumulative multiplier is twice as large in the first year and considerably higher in the second year. This suggests that a government investment shock has a stronger initial effect on value added in Italy during a slack period. These findings align with existing literature on fiscal multipliers in different economic conditions.

Additionally, when investigating the different specifications of the state variables in Table 3, we find that there is almost no difference when using a different value for the HP filter of the state variable dynamics of value added. This indicates that the results are robust to changes of λ. In the second state variable, the employment rate, the results are more volatile compared to the state-dependent model of the dynamics of value added and the linear model. Nevertheless, the cumulative multiplier still exhibits a strong initial effect during a slack period in the first two years and is statistically significant, reaching a value of 5.09 in year 2. However, the cumulative multiplier drops rapidly to a value of -6.22 in year 3. These results indicate more pronounced fluctuations and a faster decay in the employment rate-based specification. Lastly, which also aligns with the other state variable (dynamics of value added), in a good period the multiplier is not significant in the first years. The figures of the IRFs of the specification of the state variable employment can be seen in Figure 5 in the appendix.

Overall, this study provides empirical evidence that public investment multiplies vary depending on the state of the economy. Specifically, we find that the impact of a government investment shock is initially more pronounced during periods of economic slack.

6 Discussion of the IV and LP Approach

The results of the LP method are compared to the approach presented in the study by Acconcia et al. (2014), the IV approach. Thereafter, a discussion is given of the fiscal foresight problem and finally, the limitations of both approaches are stated.

6.1 Comparison of the Fiscal Multiplier

When comparing the fiscal multipliers across different methods, it is important to be cautious. The cumulative multiplier in this study cannot be directly compared with the multipliers reported in Acconcia et al. (2014). The multipliers in Acconcia et al. (2014) are in a static or dynamic form. Thus, the dynamic multiplier of Acconcia et al. (2014) can be compared with the dynamic multiplier of this study. In Table 4, the dynamic multiplier reaches its highest value at a 2-year horizon with a value of 0.47 and the one-year multiplier is 0.29. Both values are statistically significant. These values are lower than the one-year static and dynamic multiplier effect reported in the study by Acconcia et al. (2014), which were estimated to be 1.55 and 1.96, respectively. Furthermore, the analysis of the state-dependent model reveals that the multipliers can vary depending on the economic regime. During periods of economic slack, the multipliers initially have a stronger effect compared to periods of economic prosperity. This highlights the importance of considering different specifications and a longer time frame when estimating fiscal multipliers. As the results can depend heavily on the economic regime and may over-or underestimate the multiplier depending on the specific years of the data set. Concluding, the estimate of Acconcia et al. (2014) is not robust when a different methodology is employed and the multiplier can vary across different economic states.

6.2 The Fiscal Foresight Problem

For the identification of the shocks, the recursive identification strategy of Blanchard & Perotti (2002) is used. However, this approach is subject to criticism as it may not accurately capture the precise timing of the shocks. This issue, commonly referred to as the fiscal foresight problem, arises from the fact that economic agents often possess prior knowledge or expectations regarding government spending shocks and adjust their behaviour accordingly, even before the policy is implemented. Thus, empirical models that only consider changes in public spending to identify the shocks may lead to unreliable conclusions. This is because they overlook the information provided by policymakers. If the variables actually capture the fiscal foresight, but it is not included in the model, errors can occur due to the omission of relevant variables. As a result, the identified fiscal policy shocks may not be truly unexpected.

This issue has also been highlighted by Acconcia et al. (2014) as one of the problems that can bias their results. They argue that failing to account for the anticipations effects between the announcement and the realisation of projects can significantly bias the multiplier estimates in a downward direction. To account for this effect, Acconcia et al. (2014) used two instruments based on the council dismissals due to mafia infiltration.

Ramey (2011) validated the importance of incorporating fiscal foresight when calculating shocks. She demonstrated that fiscal shocks were predicted by professional fiscal forecasts and war dates with the use of a Granger Causality Test (Granger, 1969). This test can determine whether a time series is useful in forecasting another time series. Therefore, she showed that the identified VAR shocks were forecasted by the professional forecasts and the war dates, and thus not truly unexpected. As explained before, Acconcia et al. (2014) used the council dismissals to mitigate this problem. The approach of Ramey (2011) is followed and it is investigated whether the government spending shocks that are identified in this paper are foreseen by the council dismissals. To demonstrate this, the Granger Causality Test is used. If this is the case, the council dismissals should be incorporated in the estimation of the shocks. In this study, we use panel data and thus perform the test with the Pairwise Dumitrescu-Hurlin Panel Causality Test Dumitrescu & Hurlin (2012). The findings are represented in Table 5. It is seen that the p-value is above 0.05 and therefore the null hypothesis that the council dismissals Granger-cause the VAR shocks in any of the provinces can not be rejected. The identified shocks are not foreseen by the council dismissals and thus the council dismissals are not incorporated in our model.

Table 5: Pairwise Dumitrescu-Hurlin Panel Causality test between Council Dismissals and the identified shocks

Hypothesis Test	**Zbar-Statistic**	**P-value**
H0: Council Dismissals do not cause the shocks	0.842	0.400
H0: The shocks do not cause Council Dismissals	1.342	0.179

Notes: This table presents the Pairwise Dumitrescu-Hurlin Panel Causality test (Dumitrescu & Hurlin, 2012) to show if the identified shocks are foreseen by the council dismissals.

6.3 Limitations of the IV and LP Method

In this part, the limitations of the approaches are discussed. Starting with the IV approach for calculating the fiscal multiplier. As discussed in Section 2 and Section 3, the IV approach has some pitfalls. One major pitfall is that instruments can be considered weak, despite the F-statistic indicating that they are strong. Weak instruments can make the 2SLS results unreliable, as the first-stage regression estimates are biased and the t-test fails to control for the size. The instrument relevance can depend heavily on a few outliers, as demonstrated by Young (2022) and confirmed in our analysis. We demonstrated this in Section 3, where the F-statistic falls below the threshold value after the deletion of one outlier. Another pitfall is the behaviour of the errors also highlighted by Young (2022). The bias advantage that 2SLS has decreased with the presence of non-iid errors. In Section 3, we showed that the first-stage regression has evidence of having non-iid errors and thus the result should be looked at with caution. As mentioned by Andrews et al. (2018), it is important to acknowledge the challenges posed by heteroskedastic errors. However, to this day there has not been a definitive solution to address these challenges in 2SLS.

Despite the advantages of the LP approach, it has some limitations. The LP approach is usually conducted with quarterly data over a long time frame. To ensure the comparability with the results of Acconcia et al. (2014) the same data-set was used. However, this came at the cost of not having a big data set. The less preferable data set made our results less reliable, as is evident in Figures 2, 3 and 5. These figures demonstrate the presence of large confidence intervals and erratic responses. Furthermore, in general practices for the estimation of the fiscal multiplier in an LP or SVAR framework, additional control variables are included, such as the real interest rate or the overall government spending. The inclusion of such variables can enhance the reliability and precision of the results.

Lastly, it is important to note that the LP approach relies on OLS regressions, and the identification of the shocks is based on the recursive identification method of Blanchard & Perotti (2002). Therefore, the results are similar to the OLS estimation of Acconcia et al. (2014). The endogeneity problem is acknowledged with, among other things, the anticipated effects of fiscal shocks. However, using council dismissals as a proxy for this fiscal foresight is not appropriate in our analysis and can not make the shocks truly exogenous. Other possibilities such as a professional forecast could improve the results.

In summary, both approaches have their limitations and pitfalls. Given the current conditions and calculations, it is difficult to reach a definitive conclusion regarding which method could offer more reliable results, as both approaches face significant limitations.

7 Conclusion

The fiscal multiplier can differ widely depending on the chosen method and data set. Commonly used methods include 2SLS, SVAR, and the more recent LP approach. Each method possesses its own strengths and limitations. This study builds upon the research conducted by Acconcia et al. (2014) and raises concerns regarding the reliability of their findings. We use another approach, the LP method, to calculate the multiplier to see how robust the results are. In particular, the following main research question is answered: Does the fiscal multiplier estimated by Acconcia et al. (2014) remain stable when an alternative methodology is employed?

This study highlights several important findings. First, the presumed strength of the instruments in the analysis of Acconcia et al. (2014) is questionable. There is strong evidence that the strength of the instruments relies on outliers and the residuals show heteroskedastic behaviour, which can alter the significance of the first-stage. Second, the results obtained by Acconcia et al. (2014) are not robust when the LP method is used. The findings show that the dynamic multiplier is notably smaller than the multiplier of Acconcia et al. (2014), with respective estimates of 0.29 and 1.95. Lastly, the study contributes to the literature by showing the significance of considering different economic regimes, as it has a significant impact on the estimated results.

Concluding, the results of this study reveal that the findings of Acconcia et al. (2014) are not stable when a different method is used, leaving out the possibility that the difference between the estimates depends on the data set. It is important to note that this study does not claim its results to be superior or more resistant to misspecifications. Instead, it demonstrates the instability of the results when a different method is used and provides evidence of potential over-exaggeration of the fiscal multiplier and the relevance of the instruments used in the study by Acconcia et al. (2014). For policymakers, these results serve as a reminder to be cautious when they rely on fiscal multiplier estimation for important decision-making actions such as budget planning and resource allocations as the results can vary significantly.

Several limitations exist in this paper. A major limitation is the data set. While it has the advantage of leaving out the dependency of the results on a specific environment or data set, it is not the ideal data set for the LP method. To enhance the reliability of the results, further research could consider using a larger data set with quarterly observations and incorporating additional control variables such as the real interest rate. In addition, the issue of the endogeneity of the shocks remains unsolved. Although Acconcia et al. (2014) tried to address this use by using instruments, our analysis shows that these instruments do not solve the problem. The results of a Granger Causality test (Granger, 1969) indicate that the identified shocks are not foreseen by council dismissals. Further research could explore alternative approaches to overcome this problem, such as incorporating a professional forecast of the government investment shocks. Lastly, it would be valuable to explore other methods, such as an SVAR analysis and conduct a comprehensive comparison among the three methods: IV, LP and SVAR.

References

Acconcia, A., Corsetti, G. & Simonelli, S. (2014). Mafia and public spending: Evidence on the fiscal multiplier from a quasi-experiment. *American Economic Review*, *104*(7), 2185–2209.

Amendola, M. (2022). Public consumption multipliers in slack and good periods: Evidence from the euro area. *Macroeconomic Dynamics*, 1–25. doi: 10.1017/S136510052200058X

Andrews, I., Stock, J. H. & Sun, L. (2018). Weak instruments and what to do about them. *Lectures, NBER Summer Institute*.

Auerbach, A. & Gorodnichenko, Y. (2012). Measuring the output responses to fiscal policy. *American Economic Journal: Economic Policy*, *4*(2), 1–27.

Auerbach, A. & Gorodnichenko, Y. (2017). *Fiscal stimulus and fiscal sustainability* (Tech. Rep.). National Bureau of Economic Research, Inc.

Barro, R. J. & Redlick, C. J. (2011). Macroeconomic effects from government purchases and taxes. *The Quarterly Journal of Economics*, *126*(1), 51–102.

Baum, M. A., Poplawski-Ribeiro, M. M. & Weber, A. (2012). *Fiscal multipliers and the state of the economy.* International Monetary Fund.

Blanchard, O. & Perotti, R. (2002). An empirical characterization of the dynamic effects of changes in government spending and taxes on output. *The Quarterly Journal of Economics*, *117*(4), 1329–1368.

Breusch, T. S. & Pagan, A. R. (1979). A simple test for heteroscedasticity and random coefficient variation. *Econometrica: Journal of the Econometric Scoiety*, 1287–1294.

Caldara, D. & Kamps, C. (2017). The analytics of SVARs: A unified framework to measure fiscal multipliers. *The Review of Economic Studies*, *84*(3), 1015–1040.

Chian Koh, W. (2017). Fiscal multipliers: new evidence from a large panel of countries. *Oxford Economic Papers*, *69*(3), 569–590.

Clemens, J. & Miran, S. (2012). Fiscal policy multipliers on subnational government spending. *American Economic Journal: Economic Policy*, *4*(2), 46–68.

Corsetti, G., Meier, A. & Müller, G. J. (2012). What determines government spending multipliers? *Economic Policy*, *27*(72), 521–565.

Deleidi, M., Iafrate, F. & Levrero, E. S. (2023). Government investment fiscal multipliers: Evidence from euro-area countries. *Macroeconomic Dynamics*, *27*(2), 331–349.

Driscoll, J. C. & Kraay, A. C. (1998). Consistent covariance matrix estimation with spatially dependent panel data. *Review of Economics and Statistics*, *80*(4), 549–560.

Dumitrescu, E.-I. & Hurlin, C. (2012). Testing for Granger non-causality in heterogeneous panels. *Economic Modeling*, *29*(4), 1450–1460.

Geli, J. F. & S. Moura, A. (2023). *Getting into the nitty-gritty of fiscal multipliers: Small details, big impacts* (IMF Working Papers No. 2023/029). International Monetary Fund.

Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: journal of the Econometric Society*, 424–438.

Hodrick, R. J. & Prescott, E. C. (1997). Postwar US business cycles: an empirical investigation. *Journal of Money, Credit, and Banking*, 1–16.

Ioannidis, J. P. & Doucouliagos, H. (2017). *The power of bias in economics research.* Oxford University Press Oxford, UK.

Jordà, Ò. (2005). Estimation and inference of impulse responses by local projections. *American Economic Review*, *95*(1), 161–182.

Kameda, T., Namba, R. & Tsuruga, T. (2021). Decomposing local fiscal multipliers: Evidence from Japan. *Japan and the World Economy*, *57*, 101053.

Kiviet, J. F. & Niemczyk, J. (2007). The asymptotic and finite sample distributions of OLS and simple IV in simultaneous equations. *Computational Statistics & Data Analysis*, *51*(7), 3296–3318.

Koenker, R. (1981). A note on studentizing a test for heteroscedasticity. *Journal of Econometrics*, *17*(1), 107–112.

Li, D., Plagborg-Møller, M. & Wolf, C. K. (2022). *Local projections vs. VARs: Lessons from thousands of DGPs* (Tech. Rep.). National Bureau of Economic Research.

Miyamoto, W., Nguyen, T. L. & Sergeyev, D. (2018). Government spending multipliers under the zero lower bound: Evidence from Japan. *American Economic Journal: Macroeconomics*, *10*(3), 247–277.

Mountford, A. & Uhlig, H. (2009). What are the effects of fiscal policy shocks? *Journal of Applied Econometrics*, *24*(6), 960–992.

Nakamura, E. & Steinsson, J. (2014). Fiscal stimulus in a monetary union: Evidence from US regions. *American Economic Review*, *104*(3), 753–792.

Pappa, E. (2009). The effects of fiscal shocks on employment and the real wage. *International Economic Review*, *50*(1), 217–244.

Ramey, V. A. (2011). Identifying government spending shocks: It's all in the timing. *The Quarterly Journal of Economics*, *126*(1), 1–50.

Ramey, V. A. (2016). Macroeconomic shocks and their propagation. *Handbook of Macroeconomics*, *2*, 71–162.

Ramey, V. A. & Shapiro, M. D. (1998). Costly capital reallocation and the effects of government spending. In *Carnegie-Rochester Conference Series on Public Policy* (Vol. 48, pp. 145–194).

Ramey, V. A. & Zubairy, S. (2018). Government spending multipliers in good times and in bad: evidence from US historical data. *Journal of Political Economy*, *126*(2), 850–901.

Restrepo-Ángel, S., Rincón-Castro, H. & Ospina-Tejeiro, J. J. (2022). Multipliers of taxes and public spending in colombia: SVAR and local projections approaches. *Latin American Journal of Central Banking*, *3*(3), 100070.

Sargan, J. D. (1958). The estimation of economic relationships using instrumental variables. *Econometrica: Journal of the Econometric Society*, 393–415.

Staiger, D. O. & Stock, J. H. (1994). *Instrumental variables regression with weak instruments.* National Bureau of Economic Research Cambridge, Mass., USA.

Stock, J. H. & Yogo, M. (2002). *Testing for weak instruments in linear IV regression.* National Bureau of Economic Research Cambridge, Mass., USA.

Wooldridge, J. M. (2015). Introductory econometrics: A modern approach. In (p. 268-296). Cengage Learning, South-Western.

Young, A. (2022). Consistency without inference: Instrumental variables in practical application. *European Economic Review*, *147*, 104112.

A Appendix

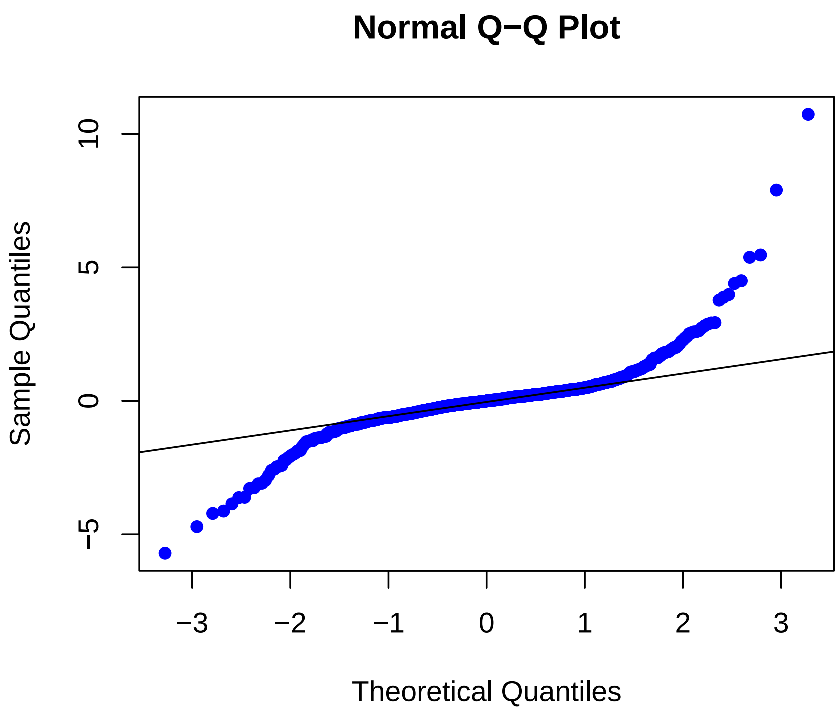
A.1 Figures and Tables

Table 6: Summary statistics for the data of 95 provinces in Italy from the year 1986 until 1999

Variable	Y	G^y	G	CD	CDS1	CDS2	U1	U2
Mean	1.050	-0.067	6.1946	0.005	0.002	0.003	-0.110	-0.005
St. Dev	2.414	1.277	48.396	0.051	0.023	0.039	0.390	0.036
Max	9.449	10.138	747.497	1.209	0.576	0.576	1.469	0.106
Min	-9.721	-9.920	-88.829	0.000	0.000	0.000	-2.137	-0.1688

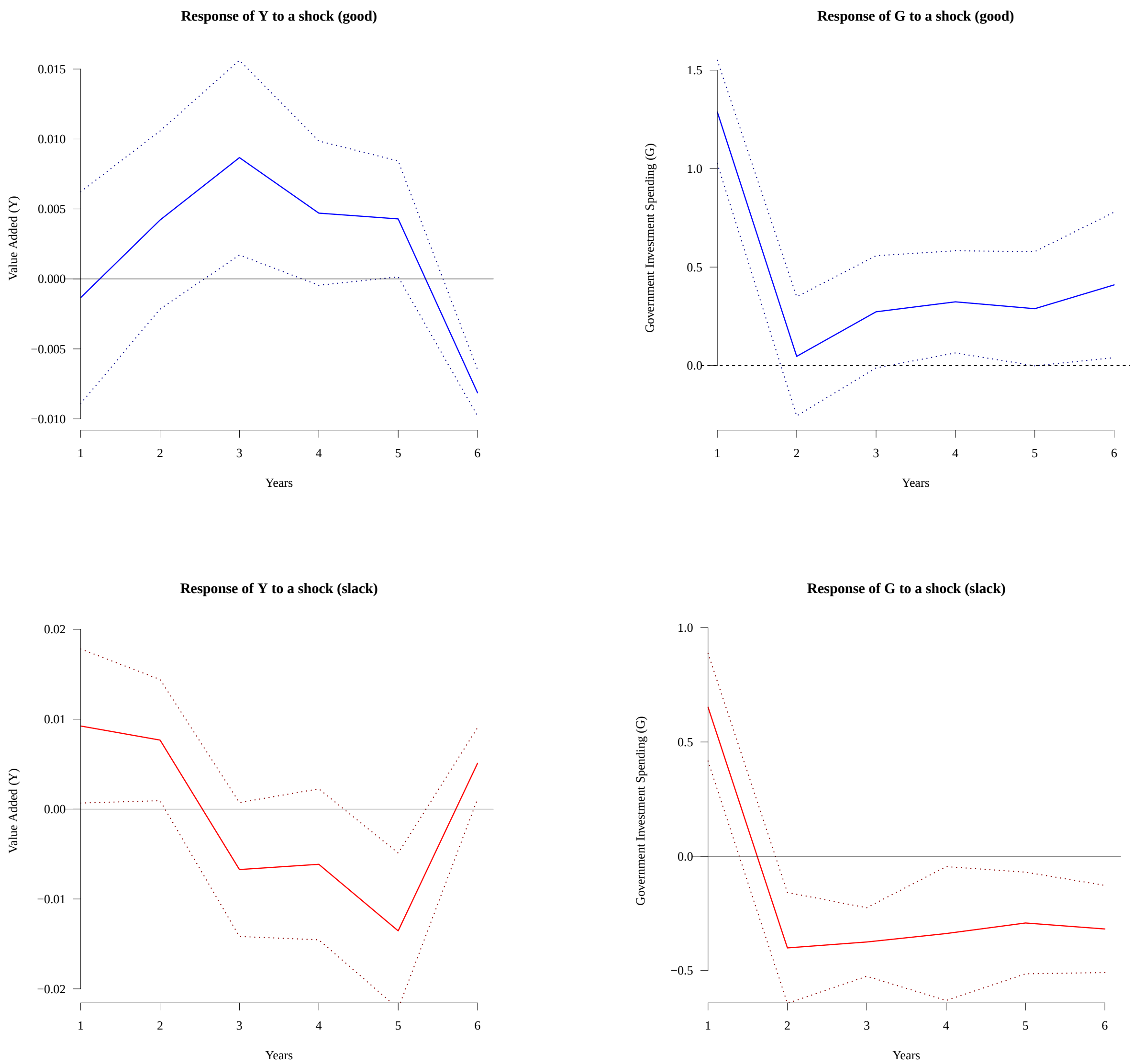
Notes: The data presented in this table shows summary statistics for the variables as described in Section 3 of the sample period 1986 until 1999. The data set consists of 1330 observations: 14 years of 95 provinces.

Figure 4: Normality Q-Q plot of the headline results



Notes: In this figure the Q-Q plot of the residuals of the first-stage regression of the headline results is presented.

Figure 5: The IRFs for the state-dependent model for the state variable employment



Notes: The state-dependent impulse responses to a 1% government investment shock. The dashed lines represent the 95% confidence intervals, based on Driscoll & Kraay (1998) standard errors. The IRFs in blue represent the functions in a good regime and the IRFs in red represent the responses in a slack regime. The state variable is the standard deviation of the logarithm of employment HP filtered with $\lambda = 10^6$.