INSTITUTO
UNIVERSITÁRIO
DE LISBOA

Estimating the effects of fiscal consolidations: A synthetic control approach

Tomás Manuel Lança Martins

Master in Economics

Supervisor:
PhD Ricardo Nuno Ferreira Paes Mamede, Associate professor,
ISCTE – IUL

September, 2021

Department of Economics

Department of Political Economy

Estimating the effects of fiscal consolidations: A synthetic control approach

Tomás Manuel Lança Martins

Master in Economics

Supervisor:
PhD Ricardo Nuno Ferreira Paes Mamede, Associate professor,
ISCTE – IUL

September, 2021

Acknowledgements

Although this thesis is an individual project, many people provided invaluable contributions to which I would like to express my gratitude.

To my family, especially to my parents for their unconditional support and encouragement and to my grandfather for his unconditional role on financing my academic path.

To my supervisor for his full availability and counsel throughout the last year.

To all teachers at ISCTE that played a role on enhancing my abilities and to all friends and colleagues that contributed in some way to the making of this dissertation.

Resumo

A presente tese aplica o método de controlo sintético como processo de estimação dos efeitos de consolidações orçamentais no PIB. A metodologia desenvolvida requer a identificação de episódios de tratamento (períodos em que se verifiquem consolidações orçamentais) e a identificação de episódios de controlo (quando a política orçamental foi relativamente neutra). A identificação destes episódios é feita usando as duas variáveis mais comuns na literatura recente. Como estas variáveis são muito diferentes e cobrem períodos e países distintos, a estimação é feita separadamente em duas bases de dados e apenas os resultados são comparados. A justificação para a implementação do método de controlo sintético foi a possibilidade de este conseguir fornecer estimações robustas para um único tratamento, com a intenção de estudar a heterogeneidade dos efeitos de consolidações orçamentais. Através da avaliação do método de estimação em placebos, foi possível concluir que esta metodologia gera demasiada variância, fazendo com que estimações individuais sejam inadequadas. No entanto, as estimações dos efeitos médios não são enviesadas e fornecem algumas conclusões úteis: consolidações orçamentais geram perdas significativas de PIB, consolidações baseadas em redução de despesa são menos contracionárias e consolidações implementadas a seguir a uma recessão geram maiores perdas. Estas conclusões são robustas aos dois métodos de identificação (quando podem ser testadas com ambos) e estão em linha com as conclusões obtidas na literatura recente.

Palavras-chave: Consolidação orçamental, Política orçamental, Método de controlo sintético.

JEL: E62, H60

Abstract

This thesis applies the synthetic control method to estimate the effects of fiscal consolidations on GDP. The framework developed requires the identification of suitable multi-year treatment episodes as well as periods that can serve as controls (where fiscal policy was relatively neutral). The identification of treatment and control episodes is done through the two most used measures of fiscal policy in the recent literature. Since the two measures are very different and have little overlap in terms of countries and period covered, the estimation is done in two different datasets and only the main results are compared. The rationale behind the implementation of the synthetic control method was to explore the possibility that it could provide robust estimations for the effects of individual fiscal consolidations. However, through the evaluation of the estimation method on placebos, it seems that this methodology generates too much variance, making individual estimations unreliable. Nevertheless, the average estimated effects are unbiased and can provide some useful insights: fiscal consolidations tend to generate significant losses in GDP, consolidations based on expenditure reduction are less contractionary than those based on tax increases and consolidations implemented following a recession generate higher costs. These conclusions are robust to both identification methods (when it could be tested with both) and are also in line with recent studies that use more common estimation methods.

Keywords: Fiscal consolidation, Fiscal policy, Synthetic control method.

JEL: E62, H60

Index

Introduction

Every year, in every nation, a government budget detailing the yearly policies and their expected outcome in public finances is drafted. Economies may thrive or wither by the decisions inscribed in this document, and the welfare of every citizen is indissolubly dependent on it. As such, understanding the aggregated effects of fiscal policy seems an important undertaking.

This thesis studies fiscal consolidations – that is, episodes of fiscal policy in which one or successive government budgets intend to lower the public deficit. Since the duty and incentive of any government is to maximize the welfare of its citizens, lowering the net spending of the state is generally an undesirable decision as the direct effect of such measures is to reduce the contemporaneous disposable income of the domestic private sector. However, there are some common reasons that lead governments to pursue fiscal consolidations: it may be in the best interest of the country to rein in public finances to avoid long-run imbalances, the economy may be booming too much and at risk of generating inefficient bubbles or strong inflation, the price of issuing new government bonds may climb too high or simply there may be no willing lenders. These episodes seem particularly important to investigate since in the cases where consolidations are implemented somewhat forcefully, public deficits may be drastically reduced without much concern for the effect on the economy.

Following the sovereign debt crisis in Europe, many countries implemented ambitious consolidation programs whose contractionary effects were severely underestimated (Fatás & Summers, 2018). How could policy makers botch the planning of such critical interventions to this degree may partly be explained by a lack of insight on the effects of fiscal policy in the scientific community. With the COVID-19 pandemic, governments all over the world massively increased their deficits and public debt levels have been climbing to record highs, making fiscal consolidations likely to be necessary in the near future. Having reasonable estimations of the effects of fiscal policy and how and when to best implement these consolidations is crucial to prevent needless suffering. Hopefully, policy makers will be mindful of the progress on research over the last decade and mistakes of the past can be avoided. This thesis hopes to be a small contribution to this endeavor.

The major contribution of this thesis is the appliance of the synthetic control method (SCM) to estimate effects of fiscal consolidations. Unlike most studies where fiscal consolidations are interpreted as yearly exogenous shocks, the framework developed in this thesis requires the identification of multi-year consolidation episodes. A counterfactual scenario of no fiscal consolidation is then estimated by constructing a synthetic country from a donor pool of non-treated countries. The synthetic countries are constructed in a way that maximizes the similarity of pre-treatment growth paths and pre-treatment economic characteristics that are likely to influence post-treatment growth. Due to the high variability of GDP and the relatively low number of controls for most treatment episodes, estimations of individual treatment effects are likely to be inadequate. Therefore, the analysis focuses only on estimating average treatment effects on the treated (ATT). Heterogeneity is then explored by estimating the ATT across sub-groups.

The rest of this thesis is divided in four chapters. The first chapter presents a synthesis of the recent literature on the empirical study of fiscal consolidations. In the second chapter, the methodology followed is discussed. The third chapter presents and discusses the results of the ATT estimations and explores possible sources of heterogeneity. The last chapter concludes.

CHAPTER 1

Literature Review

The literature on the macroeconomic effects of fiscal policy is vast and consensus has been difficult since studies often reached different conclusions. However, since the global financial crisis and the widespread fiscal consolidations that followed, a growing number of researchers dedicated their time to the topic. The inflow of research and debate that ensued has led to very significant improvements on the measurement of these effects over the last decade.

Usually, the literature that focuses on fiscal consolidations defines them as episodes in which government budget balances improve due to active fiscal policies – discretionary fiscal policy. The challenges begin right here: an improvement of the budget balance might be caused by a discretionary decision of policy makers, but it can also be a consequence of improving economic conditions. When the economy grows, tax income increases and government spending decreases (for example, in unemployment benefits), leading to an improvement of the budget balance, even though fiscal policy remained unchanged.

To deal with this issue the most common variable used by earlier studies was the cyclically adjusted primary balance (CAPB). The CAPB is essentially an estimation of the budget balance that would be verified if GDP was permanently at its potential (or if the economy was at full employment), purged of interest payments on public debt. Guajardo et al. (2014) raise concerns against the use of this indicator since it may still include changes due to factors other than discretionary fiscal policy, like a boom in the stock market or one-of accounting operations. A further issue is that discretionary fiscal policy may be a response to developments affecting the economy (for example, governments might reduce consumption due to a risk of overheating in a boom). According to this argument, fiscal policy may itself be endogenous to the most common outcome variables, violating exogeneity assumptions required by most common estimation methods.

To surpass these issues, Pescatori et al. (2011) construct a database of fiscal consolidation measures that are narratively identified to be primarily motivated by the intention to reduce the budget deficit. This approach was pioneered by Romer and Romer (2010), who studied the effects of US tax changes identified to be exogenous to the business cycle. The database constructed by Pescatori et al. (2011) include data on 17 OECD economies during the 1978-2009 period. This database and subsequent extensions became henceforth one of the most used by researchers to identify discretionary fiscal consolidations and minimize endogeneity issues (for example, Alesina et al., 2015, 2017, 2018; Banerjee and Zampolli, 2019; Dell'Erba et al., 2018; Guajardo et al., 2014). However, Jordà and Taylor (2016) test the exogeneity of the narratively identified variable and find that it can still be predicted by past output growth and other likely confounding variables, raising concerns for methods that require the exogeneity assumption.

On the other hand, recent studies have also started to use the structural primary balance (STPB) to identify discretionary fiscal policy changes (for example, Fatás and Summers, 2018; Fragetta and Tamborini, 2019). This indicator is an improvement of the CAPB, where the budget balance is also controlled for variations in the financial sector and asset prices and the effects of one-off expenditure or revenue items. The STPB addresses the main identification problems of the CAPB evidenced by Guajardo et al. (2014), but is still very likely to suffer from endogeneity issues, since changes in discretionary fiscal policy motivated by the business cycle are not excluded. Recent studies that utilize the STPB or the narrative identification approach estimate bigger contractionary effects for fiscal consolidations than the ones based on the CAPB.

A common way to tackle the endogeneity of fiscal policy is to use vector autoregressions (VAR). VAR allow to account for the possible causal effect past economic variables have on current fiscal variables and to simulate impulse response functions of exogenous shocks to fiscal variables on economic variables (for example, Auerbach and Gorodnichenko, 2012; Guajardo et al., 2014). Other studies that use the narrative approach, and consequently have less concerns of endogeneity, use moving averages or local projections in their estimations (for example, Alesina et al., 2015, 2017; Dell'Erba et al., 2018).

Even though the simpler approach would be to estimate the average expected response of output to a fiscal consolidation, most studies try to account for non-linearities in this response. It has been theorized, for example, that fiscal multipliers may be very different if a fiscal consolidation is pursued on "normal" times or when the economy is in a recession (Delong et al., 2012). Estimating an average response based on aggregate empirical evidence of both cases would say little to policymakers about the expected outcome of concrete fiscal consolidations and would not provide useful insight as to when and how to best implement them. Recent studies therefore focus on estimating heterogeneous effects across different characteristics of fiscal consolidations, like the composition (for example, if it is tax-based or expenditure-based), timing (sharp versus drawn-out fiscal consolidations) or country characteristics at the time the consolidation is implemented (for example, accounting for the cyclical position of the economy, public debt, trade openness or the response of monetary policy).

In what concerns the composition, the main conclusion that emerges from the literature is that fiscal consolidations that focus on reducing expenditure have less contractionary effects than ones focusing on increasing revenue. For example, Alesina and Ardagna (2010) study the difference in outcomes when consolidations are based on reducing spending or increasing taxation. They examine the effects of large changes in fiscal stance (defined as an improvement or deterioration of the CAPB bigger than 1.5% of GDP) and find that fiscal adjustments based on spending cuts and little tax increases are more likely to reduce deficits and debt over GDP ratios and are less likely to create recessions than those based on tax increases. Although the approaches used in that paper have become highly outdated, subsequent studies that study the heterogeneous effects of composition (Alesina et al., 2015, 2017, 2018, 2019; Attinasi and Klemm, 2016; Attinasi and Metelli, 2017; Dell'Erba et al., 2018; Fragetta and Tamborini, 2019) corroborate the finding that adjustments based on spending cuts are less contractionary than ones based on tax increases.

Nonetheless, even if the short-run effect of spending-based consolidations is less contractionary, the choice of cutting spending can have other adverse consequences. Bardaka et al. (2021) find a long-run negative relationship between austerity and total factor productivity, especially for spending-based austerity. They theorize that choosing to consolidate by cutting spending might undermine public expenditure on infrastructure projects, technology, innovation, education, and health necessary to build physical, human, and social capital that enhances productive capacity.

When looking at the heterogeneity of fiscal consolidations based on country characteristics, the focus of the research has been on how the effects of fiscal consolidations might differentiate if implemented on a boom versus a recession. Jordà and Taylor (2016) estimate the average treatment effect (ATE) of fiscal consolidations for when the economy is in a boom or in a slump. They discuss the validity of the CAPB and the narrative approach as instruments for evaluating fiscal policy and ultimately utilize an augmented inverse propensity score weighting to tackle endogeneity issues. They find that fiscal consolidations implemented in a recession are much more contractionary and significant than when implemented during a boom (respectively, 3.5% and 1.8% less of GDP over 5 years for each 1% of GDP of fiscal consolidation). Alesina et al. (2018) and Auerbach and Gorodnichenko (2012) utilize the VAR methodology to explore the same heterogeneity and reach similar conclusions.

Some papers explored the relationship between fiscal consolidations following the financial crisis and official error forecasts of international institutions (Blanchard and Leigh, 2013; Fatás and Summers, 2018; Gechert et al., 2019). They found that fiscal multipliers were grossly underestimated and that fiscal consolidations were correlated with downward revisions of both GDP and potential GDP over the following years. They provide evidence that austerity in recessions causes additional damage on short-run economic conditions and that these effects persist over the long-run, permanently lowering future output (an effect known as hysteresis) and leading to self-defeating fiscal consolidations (that do not have the desired effect of lowering debt/GDP levels). Dell'Erba et al. (2018) apply a different approach (local projections) and use the narratively identified dataset (that covers a completely different period), confirming the findings that fiscal consolidations implemented in recessions lead to hysteresis effects in the medium-term.

One of the most discussed sources of heterogeneity in the recent literature is the constraining effect that the zero-lower bound (ZLB) may have on the responsiveness of monetary policy. The reasoning is that during normal times, when a government implements a fiscal consolidation program that has contractionary effects on the economy (for example, to reduce the long-term public debt path), the central bank might then reduce interest rates, cushioning the contractionary effects of the consolidation program. The intervention of the central bank guarantees that output does not fall much below potential and supports the success of the program. However, if the interest rate is already at or close to zero, the central bank might not have enough room to avoid a contraction in the economy. This effect may also be linked with the position of the business cycle – if an external shock drives the economy into a recession, the central bank will probably use monetary policy to cushion this shock. If a fiscal consolidation is started during the slump, the government is effectively adding to the adverse effects of the economic shock and the central bank is likely to be unable to respond. Although there is little empirical study on this subject, as the ZLB is a recent phenomenon, Amendola et al. (2020) study how the effective lower bound (ELB) conditions fiscal multipliers in the euro area. ELB is a concept that accounts for the fact that even if the interest rate is already at zero, the central bank can still exert its influence through other measures like quantitative easing. They estimate multipliers of fiscal policy in normal times between 0.3 and 1.4, while at the ELB multipliers are estimated to be between 1.6 and 2.9.

Following the line of reasoning that monetary policy is crucial in cushioning contractionary effects of fiscal consolidations, the hypothesis that country-level fiscal policy effects are magnified in the eurozone is a natural concern. Alesina et al. (2018) find significant differences in fiscal consolidations on countries that have sovereign monetary policy versus countries in a monetary union. They conclude that the existence of a sovereign domestic central bank dampens the contractionary effects of consolidations implemented during recessions. Terzi (2020) studies how the eurozone conditions the cost of fiscal consolidations. Through propensity score methods he matches the size of consolidations and country characteristics of eurozone countries that implemented big fiscal consolidation programs following the financial crisis to past similar episodes, constructing counterfactuals where the main difference is the membership of a monetary union. The analysis shows that fiscal consolidations in the eurozone countries were much costlier, with an additional average cumulated loss over 5 years of 11% of GDP compared to the counterfactual.

Recent papers have also started to study how other factors might influence the outcomes of fiscal consolidations. El-Shagi and Schweinitz (2021) explore the heterogenous effects that fiscal consolidations may have conditional on the state of government finances. They find that austerity is less contractionary when implemented in countries with fragile public finances. Andrés et al. (2020) model fiscal consolidations in the context of private deleveraging and conclude that small and gradual consolidations are conductive towards an earlier reactivation of the economy. Gechert et al. (2019) find that fiscal consolidations have reduced contractionary effects on small open economies. Banerjee and Zampolli (2019) estimate that contractionary effects are reduced in countries where public debt is high or facing a current account deficit, while consolidations in countries with weak private credit growth exhibit larger contractionary effects.

The present dissertation builds on the existent literature by introducing the SCM to the study of fiscal consolidations. The intent is to verify the robustness of the main sources of heterogeneity discovered by the literature though a new methodology and explore additional possible sources.

CHAPTER 2

Methodology

2.1 Applying the SCM to the study of fiscal consolidations

The SCM was first developed in Abadie and Gardeazabal (2003) and Abadie et al. (2010) to estimate the effect of an intervention on an aggregate outcome variable of a single treated unit. The SCM was inspired on comparative case studies, where the path of some outcome variable of an intervened aggregate unit is compared to the outcome path on similar aggregate units that were not intervened. Its main innovation is to propose a data driven procedure to create a synthetic control from a donor pool of comparable aggregate units that did not undergo the studied intervention. The synthetic control should be chosen in a way that minimizes the mismatch of pre-intervention outcome values and that minimizes the difference between covariates that could be important at explaining the outcome variable. As the synthetic control is a weighted average of units in the donor pool, it provides a better comparison unit than any other real unaffected unit. For a detailed discussion on the methodology and recent applications in the synthetic control literature, see Abadie (2020).

My main motivation for applying the SCM to the study of fiscal consolidations is the conviction that it may be a better method for identifying heterogeneity and the factors behind it. Contrary to most estimation methods, the SCM focuses on trying to provide a robust estimate of the causal effects of one particular treatment. If this goal is achieved for many episodes of fiscal consolidations, heterogeneity can be studied by comparing the estimated effects of each consolidation episode and investigating how particular characteristics influence those effects (for example, if the country is in a recession at the time of the consolidation, or if the consolidation is mostly based on expenditure reduction). However, to provide a good basis for generalization, it is crucial that the number of fiscal consolidations for which effects are estimated is as big as possible. Therefore, the goal is to create synthetic controls in a way that can be generalized and in which the validity of estimations obtained can be properly evaluated.

The criteria I chose to evaluate the validity of the estimation method, is to compare the average treatment on the treated (ATT) and the average root mean squared prediction error

(RMSPE) of the estimation on placebos. The best method is the one for which the estimated ATT of the placebos is closer to 0 and the post-treatment RMSPE is lowest. If the ATT of placebos is close to 0, then the ATT estimated on the treatments should be unbiased. Minimizing the post-treatment RMSPE on placebos ensures that the difference in outcomes on the treated episodes is caused mostly by the presence of the treatment and minimizes the variance of the estimation. Having a distribution of estimated effects on placebos also allows to identify the treatments that have a good pre-treatment fit and still have significant estimated effects, without having to subjectively discard episodes. The robustness of individual estimations is done following the proposals discussed in Abadie (2020).

After empirically selecting the specifications that provide the most accurate estimations, heterogeneity can be studied by comparing the ATT of meaningful subsamples. Due to the choice of evaluating fiscal consolidation episodes instead of yearly fiscal shocks, comparing estimated effects across subsamples or with previous works in the literature provides an additional challenge. I propose two ways of normalizing estimated effects: 1) estimate an impulse response function for a fiscal consolidation, with size of 1% of previous year GDP, at time t; 2) estimate the average wealth lost for each percentual point of debt (measured as % of GDP) avoided in the first n years after the treatment start.

To estimate the impulse response function (IRF) one needs to assume that the ATT at time t is an appropriate estimation of the effect of a consolidation with the average size of consolidations implemented at time t and that consolidation effects are linear. Under these assumptions, the estimated effect at time $n > 0$ due to a consolidation of 1% of GDP at time $t = 0$ is given by:

$$E_n = \frac{(Y*_n - Y_n) - \sum_{i=0}^{n-1} E_i C_{n-i}}{C_0} \tag{1}$$

While at time $n = 0$ the estimated effect is simply:

$$E_0 = \frac{(Y*_0 - Y_0)}{C_0} \tag{2}$$

Where $Y*$ is the average synthetic control output, Y is the average treated country output and C is the average yearly treatment.

The estimated wealth lost for each percentual point of GDP of public debt avoided up to year *n,* for a consolidation implemented at time *t = 0,* is given by:

$$W_n = \frac{\sum_{i=0}^{n}(Y*_i - Y_i)(1+u)^{n-i}}{\sum_{i=0}^{n} C_i(1+r)^{n-i}} \qquad (3)$$

Where r is the interest rate paid on public debt and u can be interpreted as the marginal product of capital or as a discount factor for utility derived from consumption. For simplification, I assume that r = u, although it should be considered that the lost wealth will be overstated (understated) if r is higher (lower) than u.

2.2 Identifying fiscal consolidation episodes

As discussed in the literature review, there are two main approaches to measure fiscal consolidations: assessing quantitative shifts in fiscal variables or narratively identifying measures that were taken with the intent of reducing public debt or budget deficits.

When using the quantitative shifts approach, the main concern is that the definition of what qualifies as a fiscal consolidation is largely arbitrary, and there is no consensus in the literature on the size of the cut-of-point, nor on the amount of years that should be considered.[1] Another usual concern is the possible endogeneity of the fiscal consolidations identified this way, as some of them may be motivated by developments affecting the economic outlook (Guajardo et al., 2014). Since every country will be compared to a counterfactual that mimics the pre-treatment economic conditions of the treated country, endogeneity should be appropriately addressed. For example, if a government implemented a fiscal consolidation to reduce the risk of overheating in a boom, the counterfactual selected would also be booming, meaning that the estimated effects should not be biased.

Even when using the narrative approach, some authors have raised concerns with the predictability of fiscal consolidations using past realizations of GDP or other economic variables. For example, Hernández de Cos and Moral-Benito (2016) find that even narratively identified spending based fiscal consolidations can be predicted by past realizations of GDP and that estimated multipliers may be biased. As discussed above, these concerns are appropriately addressed by the SCM framework. Nonetheless, there are other

[1] See Kleis and Moessinger (2016) for an overview of quantitative definitions of fiscal consolidations used in the literature.

problems with the narrative dataset that seem more concerning: 1) The number of countries with available data is small and as a result it is impossible to create an adequate synthetic control for some of the years where most of the countries were implementing fiscal consolidations; 2) The dataset only provides information on the discretionary fiscal policy measures that were identified to be motivated by the desire to reduce public debt or budget deficits. As such, the control group may be formed by countries that have implemented fiscal consolidations for different reasons or that implemented expansionary fiscal policies. Ideally, the control group would be formed by countries that had a neutral fiscal policy stance over the comparison period. This lack of information may increase the variance and bias of estimations.

Since both approaches are far from perfect but offer significant relative advantages, I identify two separate sets of treatments and controls, each based on a different approach and on a different sample. Effects of fiscal consolidations are estimated and analyzed in a similar manner for each set, providing an important robustness check for any conclusions that may be found.

In the first set, treatment and control episodes are identified using information on the yearly total consolidation from the databased created in Alesina et al. (2015). The database covers the period from 1978 to 2014 for 16 OECD countries. The criteria used to select treatment episodes is that a new treatment episode starts any time a positive yearly fiscal consolidation is identified, preceded by two years of no fiscal consolidation. This way, consecutive or one-year apart fiscal consolidations are evaluated as if being part of the same episode. As discussed in Alesina et al. (2018), governments often plan multi-year fiscal consolidations and adjust those plans over the following years. If fiscal consolidations are interpreted as a yearly time-series, that time-series is autocorrelated. Estimating treatment effects every year there is a consolidation would likely result in overstated effects.

The optimal control group would be countries that never experienced the treatment over both the pre-treatment matching period and the post-treatment evaluation period. As fiscal consolidations are a relatively common occurrence, a less-optimal criteria for selecting control episodes is used: a country can serve as a control of a consolidation implemented at year t if, during the period from t-2 to t+4, no yearly fiscal consolidation was verified. This criterium ensures that there are at least 5 years of valid comparison on outcomes and that the outcome of controls is not affected by recent fiscal consolidations, while maintaining a sufficient number of donors to construct synthetic controls from.

The second set is constructed from data on the structural primary balance (STPB) from the IMF WEO database (October 2020). Any country can be selected as a treatment or control as long as it has available data. Treatment episodes are identified when the cumulated variation of STPB is higher than 3% of GDP over a period of 5 years and there was a fiscal consolidation of more than 1% of GDP in the first year. If year t is selected as the start of a treatment for country j, the years from t+1 to t+4 cannot be selected as the start of new treatments for country j.

A country is selected as an appropriate control for a treatment starting at t if, during the periods from t-2 to t+4 and t to t+4, the cumulative change of STPB was between -1.5% and 1.5% of GDP. These criteria are chosen to ensure that there is always a sizeable treatment that can be measured and that the number of treatments and controls is substantial. The fairly long period considered in treatment selection ensures that the identified fiscal consolidations have a permanent nature and are not quickly reversed.

2.3 Data sources

As previously discussed, the main variables used to measure fiscal consolidations are the narratively identified total consolidation (Alesina et al., 2015) and the STPB. The outcome variable is an index of constant GDP with base 100 in the year before the start of a consolidation. Other macroeconomic data is taken from the World Bank database, the IMF WEO October 2020 and from Barro and Lee (2013). These additional variables were chosen on the basis that they are likely to influence post-treatment GDP growth and should be considered in the construction of synthetic controls, or because they may be an interesting source of heterogeneity to explore. Table A1 (in the annex) provides information on all variables collected. Tables A2 and A3 (in the annex) present the fiscal consolidation episodes that were identified according to the criteria discussed before.

CHAPTER 3

Results and Discussion

3.1 Estimation on the narrative dataset

This section presents the results from the estimation of average fiscal consolidation effects over the narrative dataset. The selection of country donors is always based on the minimization of pre-treatment standardized GDP paths, but multiple models were applied to test the robustness of estimations (Table 1).

Table 1 – Models applied in the narrative dataset

Model	Selection of variables	Custom V	Contemporaneous	List of variables
A	All variables that had at least one pre-treatment observation in every control were selected as possible control variables.	Yes	Yes	GDP; GDP_capita; Inflation; Unemployment; Young_dependency_ratio; Old_dependency_ratio; Investment; Population_growth; Secondary; Tertiary; Trade_openness.
B	All variables that had at least one pre-treatment observation in every control were selected as possible control variables.	No	Yes	GDP; GDP_capita; Inflation; Unemployment; Young_dependency_ratio; Old_dependency_ratio; Investment; Population_growth; Secondary; Tertiary; Trade_openness.
C	Only outcome variable was used.	No	Yes	GDP
D	All variables were selected as possible control variables.	Yes	Yes	REER, Young_dependency_ratio, Population_growth, Industry_share, Domestic_credit_to_private, Secondary
E	All budget related and all variables that had at least one pre-treatment observation in every control were selected as possible control variables.	Yes	Yes	Unemployment, Labor_participation_rate, Young_dependency_ratio, Old_dependency_ratio, Government_revenue, Secondary, Tertiary, Trade_openness
F	All variables that had at least one pre-treatment observation in every control were selected as possible control variables.	Yes	No	GDP; GDP_capita; Inflation; Unemployment; Young_dependency_ratio; Old_dependency_ratio; Investment; Population_growth; Secondary; Tertiary; Trade_openness.

Note: All variables used are the 5-year pre-treatment means, excluding model C where individual values for GDP from t-5 to t-2 were used. When the importance of each variable is fixed (custom V), the values of relative importance come from a LASSO model that predicts 5-year GDP growth applied to control episodes (when the treatment is known to be absent). Variables with coefficient 0 are dropped from the model and coefficients are rescaled to sum to 1. Further details on the estimated LASSO models can be found in the annexes.

Table 2 outlines the placebo statistics for the various models. The models that rely on selecting the importance of control variables through the minimization of pre-treatment differences in outcome (models B and C) provide a much better pre-treatment fit, but this does not seem to significantly improve the quality of the post-treatment period fit. On the other hand, the models that select the importance of variables using a LASSO estimation seem likely to be less biased (models A and E). Models D and F are dropped since model D allows for too few treatment episodes and model F seems biased and less precise (model F would only be viable if fixed time-effects were relatively unimportant – which does not seem to be the case).

Table 2 – Estimation on the narrative dataset placebos

Model	Bias	P-value	Pre-RMSPE	Post-RMSPE	Treatments	Controls
A	2.496	0.3506	2.448	4.576	23	108
B	4.144	0.1272	1.291	4.356	23	108
C	4.210	0.1230	1.233	4.257	25	108
D	-2.118	0.4827	2.126	2.903	7	44
E	0.268	0.9341	2.966	4.446	19	82
F	4.974	0.1263	2.114	5.467	50	132

Note: "Bias" is the estimated ATT on the placebos over a period of 6 years; "P-value" is the p-value of a t-test that tests if the estimated ATT is significantly different from 0.

Figure 1 provides an overview of the estimated effects of the average fiscal consolidation in the sample. On average, over the first six years following the start of a fiscal consolidation, GDP in treated countries uninterruptedly falls compared to the counterfactual and is significantly smaller than the counterfactual across all models after six years.[2]

[2] This occurs because the average treatment on the sample is positive over all years. This does not necessarily mean that the expected effects of a consolidation at time 0 would be increasingly worse over the first six years nor that GDP would still be below the counterfactual at the sixth year.

Figure 1 – Estimated ATT of fiscal consolidations (narrative dataset)

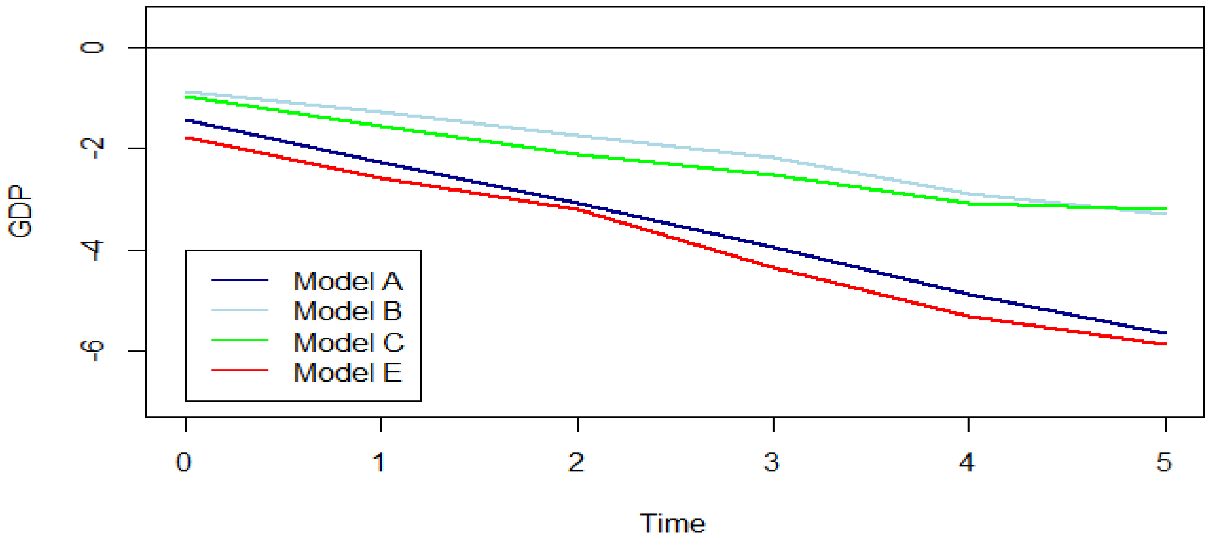


Figure 2 presents the estimations for the IRF of a fiscal consolidation of 1% of GDP at time 0, as discussed in the methodology. According to the estimations, the worse effects of a consolidation are felt in the same year it is implemented. While this would not be too surprising, there is no valid explanation for the drop in GDP verified across all models from the third to the fourth year nor for a consolidation having bigger negative effects in the fourth and fifth year than on the second and third. The most likely explanation is that the estimated effects on the first year are overstated. For example, this could be due to anticipatory effects of predicted or announced consolidations for following periods that start to have effects prior to their implementation. If the estimation of the contemporaneous effects at time 0 are overstated, the estimation for the effects on the following period will be understated and the IRF will be somewhat biased across all periods. Due to this problem, I find it prudent to not draw any conclusions for the evolution of the gap on GDP across time. Instead, I use the IRF as a way to derive the estimated cumulated cost of a consolidation implemented at time 0 over a period of six years.

Figure 2 – GDP gap due to a fiscal consolidation at time 0 of 1% of GDP (narrative dataset)

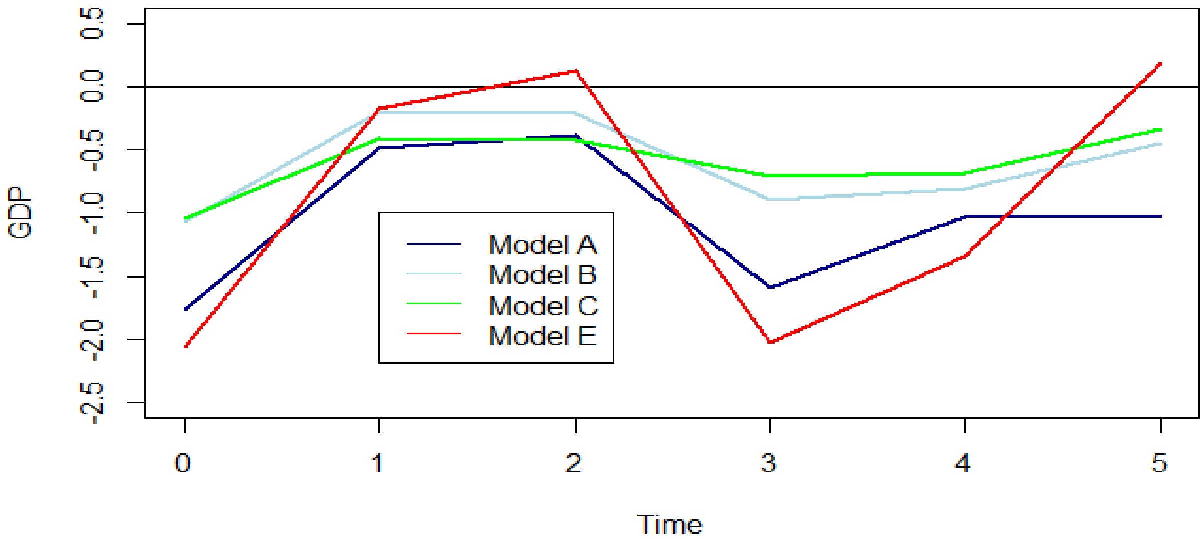


Table 3 presents significance statistics for the estimated effects and standardizes the results in a way that can be compared between samples, as discussed in the methodology. As expected, fiscal consolidations seem to lead to significant reductions in GDP (on average). A consolidation of 1% of GDP at time 0 leads to a cumulative loss estimated between -3.75% and -6.54% of pre-treatment GDP over the next six years. For each 1% of GDP in avoided debt there is an estimated loss between -0.61% and -1.06% of GDP.

Table 3 – Average effects of fiscal consolidations (narrative dataset)

Model	ATT	Cumulated IRF	Loss for 1pp of avoided debt	Bootstrap p-value	T-test p-value	Pre-RMSPE	Treatments
A	-21.273	-6.539	-1.059	0.0001	0.0001	2.702	23
B	-12.295	-3.790	-0.612	<0.0001	0.0171	1.428	23
C	-13.484	-3.747	-0.656	<0.0001	0.0005	1.265	25
E	-23.124	-5.728	-1.010	<0.0001	0.0012	3.324	19

Note: The ATT is the cumulated GDP gap between the treated and the counterfactual; The T-test p-value reports the significance of the estimated ATT; Bootstrap p-value is the ratio of drawn samples of placebos that have an estimated ATT more extreme than the true estimated ATT across 10000 random placebo samples (the samples have the same number of fake treatments as the number of treatments studied in each model).

3.2 Estimation on the STPB dataset

This section presents the results from the estimation of average fiscal consolidation effects over the STPB dataset. The selection of country donors is always based on the minimization of pre-treatment standardized GDP paths, but multiple models were applied (Table 4).

Table 4 – Models applied in the STPB dataset

Model	Selection of variables	Custom V	Contemporaneous	List of variables
A	All variables that had at least one pre-treatment observation in every control were selected as possible control variables.	Yes	Yes	GDP, Current_account, Labor_participation_rate, Young_dependency_ratio, Old_dependency_ratio, Investment, Population_growth, Budget_balance, Government_revenue, Tertiary
B	All variables that had at least one pre-treatment observation in every control were selected as possible control variables.	No	Yes	GDP, GDP_capita, Current_account, Labor_participation_rate, Young_dependency_ratio, Old_dependency_ratio, Investment, Population_growth, Industry_share, Budget_balance, Government_revenue, Tertiary
C	Only outcome variable was used.	No	Yes	GDP
D	All variables were selected as possible control variables.	Yes	Yes	GDP, Current_account, REER, Labor_participation_rate, Young_dependency_ratio, Investment, Industry_share, Domestic_credit_to_private, Government_revenue, Secondary, Tertiary

Note: All variables used are the 5-year pre-treatment means, excluding model C where individual values for GDP from t-5 to t-2 were used. When the importance of each variable is fixed (custom V), the values of relative importance come from a LASSO model that predicts 5-year GDP growth applied to control episodes (when the treatment is known to be absent). Variables with coefficient 0 are dropped from the model and coefficients are rescaled to sum to 1. Further details on the estimated LASSO models can be found in the annexes.

The estimation on the placebos seems to not be significantly biased when control variables are included. Similarly to the narrative sample, models where the importance of control variables is defined through LASSO (A and D) seem to be less biased, although the pre-treatment fit is poorer. Overall, all models generate higher post-treatment RMSPE than in the narrative sample, even though pre-treatment RMSPE are similar. A possible explanation is that defining control episodes through the STPB leads to the selection of control episodes where the change in fiscal policy is small but not completely neutral, increasing the variance of the estimations.

Table 5 – Estimation on the STPB dataset placebos

Model	Bias	P-value	Pre-RMSPE	Post-RMSPE	Treatments	Controls
A	2.283	0.5849	2.138	7.321	48	155
B	5.220	0.2327	1.207	7.494	47	155
C	6.684	0.1332	1.036	7.664	55	155
D	3.882	0.5592	2.639	7.765	26	71

Note: "Bias" is the estimated ATT on the placebos over a period of 6 years; "P-value" is the p-value of a t-test that tests if the estimated ATT is significantly different from 0.

Figure 3 provides an overview of the estimated effects of the average fiscal consolidation in the STPB sample. Compared to the narrative sample, it seems that the fall on GDP is sharper and, on average, GDP in treated countries begins to reapproach the counterfactual after the fifth year. This is likely due to the identification of the start of consolidation episodes in the STPB sample having as a condition that the consolidation verified in year 0 is bigger than 1% of GDP. On the narrative sample, many consolidations are smaller than 1% of GDP at time 0, while similar small consolidations are not included in the STPB sample.

Figure 3 – Estimated ATT of fiscal consolidations (STPB sample)

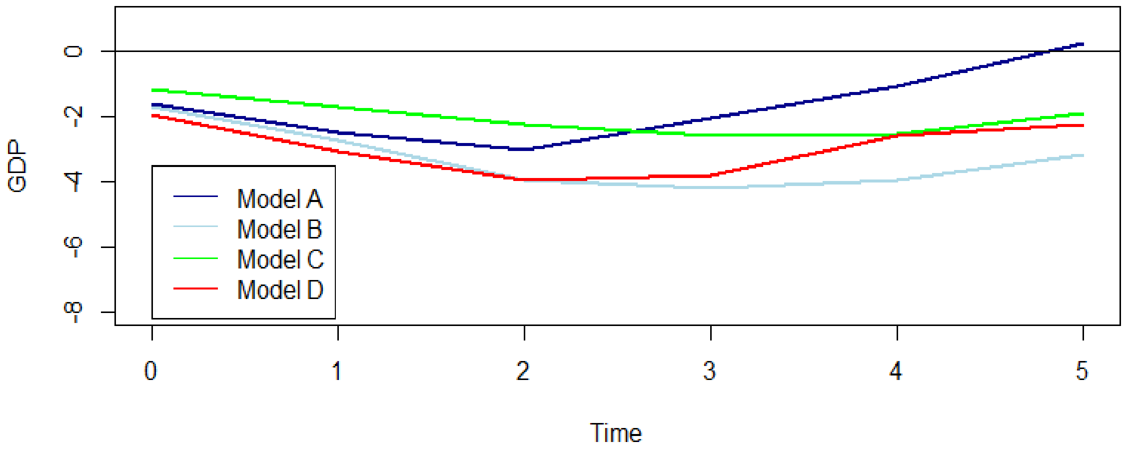
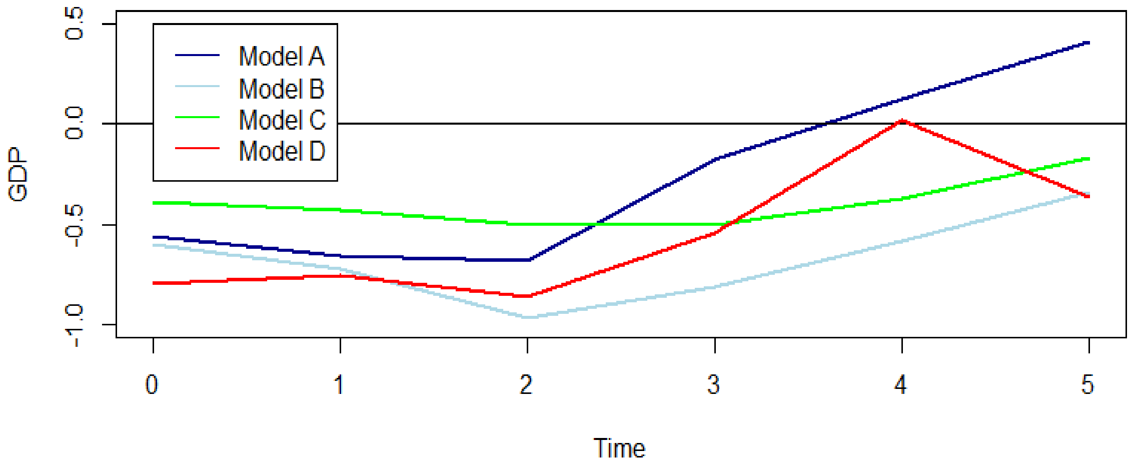


Figure 4 presents the estimations for the IRF of a fiscal consolidation of 1% of GDP at time 0, as discussed in the methodology. Contrary to the IRFs estimated for the narrative sample, the biggest GDP gap is verified in the third period. Having double the observations and requiring that consolidations start with a sharp adjustment explains why the estimated IRFs are less volatile and why the estimations of contemporaneous effects are less overstated. Overall, it seems that a fiscal consolidation creates a GDP loss that increases up to the third period. GDP in treated countries starts to reapproximate to the counterfactual over the following periods and, in the sixth period, the estimated average gap in GDP due to a consolidation is relatively close to zero. While the significance of these estimations cannot be tested, fiscal consolidations seem to have fading effects, although GDP seems to remain below the counterfactual even after six years.

Figure 4 – GDP gap due to a fiscal consolidation at time 0 of 1% of GDP (STPB sample)



Results for the estimation of the effects of fiscal consolidations in the STPB sample are synthesized in Table 6. Overall, significant costs are estimated across all models. Although the estimated ATT for model A and C are not sufficiently different from 0 to be significant according to the t-test, they are significantly below the estimated ATT on the placebos. A consolidation of 1% of GDP at time 0 leads to a cumulative loss estimated to be between -1.55% and -4.02% of pre-treatment GDP over the next six years. For each 1% of GDP in avoided debt there is an estimated loss between -0.35% and -0.70% of GDP. The estimated effects are smaller than the ones obtained in the narrative sample, even though there is some overlap. Since the samples are widely different in the period and countries covered, it cannot be known if the difference on the estimated effects is due to the identification method or due to heterogeneity.

Table 6 - Average effects of fiscal consolidations (narrative dataset)

Model	ATT	Cumulated IRF	Loss for 1pp of avoided debt	Bootstrap p-value	T-test p-value	Pre-RMSPE	Treatments
A	-9.931	-1.545	-0.350	0.0234	0.1321	3.635	48
B	-19.661	-4.024	-0.699	<0.0001	0.0498	2.344	47
C	-12.143	-2.370	-0.413	0.0013	0.1749	2.265	55
D	-17.589	-3.303	-0.609	<0.0001	0.0462	3.218	26

Note: The ATT is the cumulated GDP gap between the treated and the counterfactual; The T-test p-value reports the significance of the estimated ATT; Bootstrap p-value is the ratio of drawn samples of placebos that have an estimated ATT more extreme than the true estimated ATT across 10000 random placebo samples (the samples have the same number of fake treatments as the number of treatments studied in each model).

3.3 Heterogeneity

To evaluate heterogeneity across different consolidations the models A and B from the previous sections were selected. These models do not seem to be significantly biased and provide a robustness check on the way the synthetic control is built. Additionally, these models allow the inclusion of as many consolidations as possible.

3.3.1 Composition

The most studied source of heterogeneity has been the difference between consolidations based on tax increases and consolidations based on spending cuts. However, assessing differences between these two types can only be done with the narrative dataset, where information about the nature of measures implemented exists.

The sample was divided in three groups: the 25% of consolidations most based on tax increases, the 25% of consolidations most based on spending cuts and the remaining 50% that serve as a control group of relatively balanced consolidations. To classify the nature of each consolidation, all measures up to five years since the start of the treatment are considered.

The relatively balanced consolidations that form the "control" group correspond to consolidations that have between 44% and 69% of the value of all the measures implemented classified as spending cuts. Balanced consolidations are estimated to be costlier when compared to the estimations over the whole sample for each correspondent model. It seems that consolidations mostly based on spending cuts (>69%) lessen the average estimated costs associated with fiscal consolidations.

Spending based consolidations have significantly smaller costs according to model A, and according to model B these consolidations even have positive effects on GDP on average (although very close to zero). A consolidation of 1% of GDP at time 0 is estimated to lead to a cumulative difference of -0.94% or 2.7% compared to the counterfactual, over the next six years. For each 1% of GDP in avoided debt there is an estimated difference of -0.46% or 0.07% of GDP.

Consolidations mostly based on tax increases (where less than 44% of measures are spending cuts) have estimated effects bigger than the control group, although not sufficiently different to be significant according to a 95% bootstrap confidence interval. The estimated loss due to a consolidation of 1% of GDP is expected to be bigger than 8% or 11% of GDP over six years. Further details on the results are presented in table 7.

Table 7 – Estimated effects of fiscal consolidations grouped by composition of measures

Model	Measure	25% Tax based	25% Spending based	Control	95% Confidence Interval
A (Narrative)	IRF	-11.226	-0.943*	-7.850	-12.031 | -3.273
	Wealth	-1.517	-0.464*	-1.179	-1.701 | -0.795
B (Narrative)	IRF	-8.122	2.722*	-5.601	-8.402 | -2.236
	Wealth	-0.753	0.066*	-0.832	-1.123 | -0.519

Note: "Control" group are the relatively balanced 50% of the distribution of treatment episodes (ordered in terms of treatment composition); "Wealth" is the estimated average loss of GDP due to the avoidance of 1% of GDP in debt; Confidence interval is built through bootstrap: the measures of IRF and Wealth are calculated for a random sample of half of the control group 1000 times; the values presented are the ones for the ordered 25^{th} and 975^{th} sample.

The results obtained regarding heterogeneity due to the composition of fiscal consolidations are in line with the existing literature. As discussed in the literature review chapter, most studies find that spending-based consolidations are much less recessionary, with estimated effects close to zero. Explaining the reasons behind this heterogeneity goes beyond the scope of this thesis, but a discussion on possible explanations and on the recent literature regarding this topic can be found in Alesina et al. (2019).

3.3.2 Growth

To evaluate how fiscal consolidations may have differing impacts due to the state of the economy, the simplest approach is to look at the growth in the year preceding the start of the consolidation. Although the position of the economy may change during the implementation of consolidation programs, these changes should be, to some degree, affected by the treatment. Therefore, the year before the treatment is the best exogenous observation to assess the cycle of the economy at the time of implementation (and is the period when programs are first drafted and announced, although anticipatory effects may already be present).

The procedure followed is similar to the one in section 3.3.1. Consolidation episodes are grouped into three samples based on GDP growth before the treatment: the 25% that registered higher growth, the 25% with lower growth and the control group.

The estimations using the narrative sample seem somewhat inconclusive regarding heterogeneity in treatment growth. According to the IRF measure, consolidations preceded by higher growth do not seem to have significantly different effects compared to the control group, while for the lowest growth group the cumulated IRF is only significantly lower in model B. When considering the Wealth measure[3], the costs of consolidations seem significantly lower in the highest growth group, while the lowest growth group has similar values compared to the control. This evidence suggests that costlier consolidations are correlated to lower pre-treatment growth, although the significance of such conclusion is questionable.

When using the STPB sample, the conclusions regarding heterogeneity due to growth seem more robust. On the highest growth group there are no significant differences associated with the effects of fiscal consolidations when compared to the control group. On the other hand, all measures in both models seem to suggest that consolidations following recessions (the cut-off point for the lowest growth group is very close to zero) are costlier than in the control group. The estimated cost following a 1% GDP consolidation in countries that experienced a recession in the previous year is 4.62% or 6.96% of GDP over six years. In the control group the same cost seems to not be significant, with estimations pointing to a value of 0.18% or 2.41%. Further details on the results are presented in table 8.

[3] See note below Table 7.

Table 8 – Estimated effects of fiscal consolidations grouped by pre-treatment growth

Model	Measure	25% Lowest	25% Highest	Control	95% Confidence Interval
A (Narrative)	IRF	-11.893	-6.337	-7.504	-13.923 | -4.521
	Wealth	-0.921	-0.743*	-1.281	-2.279 | -0.787
B (Narrative)	IRF	-10.651*	-3.375	-3.147	-5.835 | -1.297
	Wealth	-0.848	0.037*	-0.602	-1.073 | -0.315
A (STPB)	IRF	-4.615*	-0.184	-0.181	-3.723 | 4.692
	Wealth	-0.925*	-0.016	-0.179	-0.697 | 0.537
B (STPB)	IRF	-6.955*	-2.654	-2.410	-6.883 | 2.585
	Wealth	-1.207*	-0.380	-0.479	-1.166 | 0.288

Note: "Control" group are the middle 50% of the distribution of treatment episodes (in terms of pre-treatment growth); "Wealth" is the estimated average loss of GDP due to the avoidance of 1% of GDP in debt; The confidence interval is built through bootstrap: the measures of IRF and Wealth are calculated for a random sample of half of the control group 1000 times; the values presented are the ones for the ordered 25^{th} and 975^{th} sample.
* - estimated effect is significantly different from the control group at the 95% CI

3.3.3 Other sources

A similar procedure to the ones applied in the previous sections was conducted to study heterogeneity regarding other pre-treatment variables. However, results were mostly inconclusive and not robust across samples. The sample was divided across different levels of pre-treatment public debt, unemployment, investment, trade openness, domestic credit to private sector or by the sharpness of implementation of the fiscal consolidation, but the results were contradictory across samples, or no significant conclusion could be derived. Further details on the results for these estimations can be found in the appendix.

CHAPTER 4

Conclusion

The goal of this thesis was to innovate the empirical study of fiscal consolidations by introducing a novel method of estimating their effects. The main issues identified by the literature in estimating the effects of fiscal consolidations are the possible endogeneity of the measures implemented, and the likely heterogeneity regarding their effects. I identified the SCM as a promising method of research in this area as it ideally would tackle both issues. Since the estimation is obtained by building a counterfactual mimicking the pre-treatment status of the treated country, concerns regarding endogeneity are lessened as the incentives to pursue fiscal consolidation regarding the status of the economy should also be present in the counterfactual. In addition, as the SCM focuses on generating a valid estimation for a single treatment, heterogeneity could, in principle, be studied by modelling the individual estimated effects dependent on differences in treatment or country characteristics.

In practice, however, it was not possible to analyze heterogeneity based on individual estimated effects, given the large variance regarding the estimations. To assess the variance (and bias) of the estimation method, I applied it to the placebos and found that the post-treatment RMSQPE of the placebos is only slightly smaller than the RMSQPE of the estimation on the treatments. This means that it would be discreditable to assume that the individual estimated effects are reliable estimations for each individual treatment. Since any estimated effects has to be divided by the treatment, the issue is exacerbated in countries that pursued small consolidations, as most of the estimated effects are likely to be a product of chance rather than of the treatment. This makes it impossible to model heterogeneity based on individual estimated effects, as heterogeneity in the estimation is very much a product of randomness or hidden variables, and small consolidations have extreme estimated effects that are clearly fallacious. As such, I explored heterogeneity by estimating the ATT across sub-groups.

Despite the large variance, the estimations tend to be seemingly unbiased, and the estimated ATT emerge as clearly significant in most models. It can be concluded that fiscal consolidations tend to lead to significant costs in terms of GDP loss. A permanent positive shock of 1% of GDP in the budget balance is estimated to lead to a reduction of GDP between 1.55% and 6.54% of GDP over 6 years (between 0.26% and 1.09% per year). Dividing the sample into subgroups, it is possible to conclude that consolidations based on spending reduction tend to be much less contractionary than tax-based ones. Although this hypothesis can only be tested in the narrative sample, the other variables for which heterogeneity was analyzed for seem to evidence less meaningful differences across subgroups and no other variable is strongly correlated with the composition of fiscal consolidations. It seems that the nature of measures implemented matter in terms of the costs of fiscal consolidations. Regarding pre-treatment growth, the results seem to suggest that fiscal consolidations tend to be costlier when implemented following a recession. Other possible sources of heterogeneity were explored but results were contradictory across samples or simply not significant. The conclusions found in this thesis regarding heterogeneity due to the nature of policies and to the economic cycle are already common in the recent literature, as exposed in the literature review. The implication for policymaking is that fiscal consolidations should be avoided at times of recession and should be focused on spending reduction, if the loss of GDP is an important concern.

The main limitation I encountered when applying the SCM was the lack of comparison units for most years of the estimation. The database on narratively identified fiscal policies only covers sixteen countries, making it impossible to build a contemporaneous counterfactual for many consolidation episodes. Fifty treatment episodes were identified but only twenty-five had at least three non-treated countries to which to compare them too (which is already a very small number, making the synthetic controls unlikely to be very precise at mimicking the treated country). The possibility of building the synthetic control with non-contemporaneous donors was tried but it led to an increase in bias and variance (as can be seen by model F, in section 3.1). The STPB variable offered a possible solution, but it was only available starting on the 2000s for most countries. Also, the STPB is probably still a bad variable to study fiscal policy as it relies on dubious estimations of potential GDP and does not provide information on the nature of the measures. An expansion of the Alesina et al. (2015) and Pescatori et al. (2011) databases to additional countries and to additional fiscal policy measures, other than consolidation measures motivated by improving budget deficits, would certainly be a major improvement for any studies on fiscal policy.

Further research should be done on studying how the nature of policies tends to affect the cost of consolidations (by dividing them further than the simple “tax increases” and “spending reductions”), as it seems to be a major source of heterogeneity. A similar approach to the one followed in this thesis could also be applied to the study of fiscal expansions, to test the symmetry of conclusions regarding the size of multipliers and heterogeneity.

Bibliographical References

Abadie, A. (2020). Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects. *Journal of Economic Literature*.

Abadie, A., Diamond, A., & Hainmueller, A. J. (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. https://doi.org/10.1198/jasa.2009.ap08746

Abadie, A., & Gardeazabal, J. (2003). The economic costs of conflict: A case study of the Basque country. *American Economic Review*, *93*(1), 113–132. https://doi.org/10.1257/000282803321455188

Alesina, A., & Ardagna, S. (2010). Large Changes in fiscal policy: Taxes versus spending. *Tax Policy and the Economy*, *24*, 35–68. https://doi.org/10.1086/649828

Alesina, A., Azzalini, G., Favero, C., Giavazzi, F., & Miano, A. (2018). Is it the "how" or the "when" that Matters in Fiscal Adjustments? *IMF Economic Review*, *66*(1), 144–188. https://doi.org/10.1057/s41308-017-0047-z

Alesina, A., Barbiero, O., Favero, C., Giavazzi, F., & Paradisi, M. (2017). The Effects of Fiscal Consolidations: Theory and Evidence. In *NBER WORKING PAPER SERIES*. https://doi.org/10.1016/j.jinteco.2017.11.007

Alesina, A., Favero, C., & Giavazzi, F. (2015). The output effect of fiscal consolidation plans. *Journal of International Economics*, *96*(S1), S19–S42. https://doi.org/10.1016/j.jinteco.2014.11.003

Alesina, A., Favero, C., & Giavazzi, F. (2019). Effects of Austerity: Expenditure- and Tax-based Approaches. *Journal of Economic Perspectives*, *33*(2), 141–162.

Amendola, A., Di Serio, M., Fragetta, M., & Melina, G. (2020). The euro-area government spending multiplier at the effective lower bound. *European Economic Review*, *127*, 103480. https://doi.org/10.1016/j.euroecorev.2020.103480

Andrés, J., Arce, Ó., Thaler, D., & Thomas, C. (2020). When fiscal consolidation meets private deleveraging. *Review of Economic Dynamics*, *37*, 214–233. https://doi.org/10.1016/j.red.2020.02.002

Attinasi, M. G., & Klemm, A. (2016). The growth impact of discretionary fiscal policy measures. *Journal of Macroeconomics*, *49*, 265–279. https://doi.org/10.1016/j.jmacro.2016.08.004

Attinasi, M. G., & Metelli, L. (2017). Is fiscal consolidation self-defeating? A panel-VAR analysis for the Euro area countries. *Journal of International Money and Finance*, *74*, 147–164. https://doi.org/10.1016/j.jimonfin.2017.03.009

Auerbach, A. J., & Gorodnichenko, Y. (2012). Measuring the output responses to fiscal policy. *American Economic Journal: Economic Policy*, *4*(2), 1–27. https://doi.org/10.1257/pol.5.3.320

Banerjee, R., & Zampolli, F. (2019). What drives the short-run costs of fiscal consolidation? Evidence from OECD countries. *Economic Modelling*, *82*(October 2018), 420–436. https://doi.org/10.1016/j.econmod.2019.01.023

Bardaka, I., Bournakis, I., & Kaplanoglou, G. (2021). Total factor productivity (TFP) and fiscal consolidation: How harmful is austerity? *Economic Modelling*, *94*(February), 908–922. https://doi.org/10.1016/j.econmod.2020.02.032

Barro, R. J., & Lee, J. W. (2013). A new data set of educational attainment in the world, 1950-2010. *Journal of Development Economics*, *104*, 184–198. https://doi.org/10.1016/j.jdeveco.2012.10.001

Blanchard, O. J., & Leigh, D. (2013). Growth forecast errors and fiscal multipliers. *American Economic Review*, *103*(3), 117–120. https://doi.org/10.1257/aer.103.3.117

Dell'Erba, S., Koloskova, K., & Poplawski-Ribeiro, M. (2018). Medium-term fiscal multipliers during protracted economic contractions. *Journal of Macroeconomics*, *56*, 35–52. https://doi.org/10.1016/j.jmacro.2017.09.004

Delong, J. B., Summers, L. H., Feldstein, M., & Ramey, V. A. (2012). Fiscal Policy in a Depressed Economy [with Comments and Discussion]. *Brookings Institution Press*, *Brookings*, 233–297.

El-Shagi, M., & Schweinitz, G. von. (2021). Fiscal policy and fiscal fragility: Empirical evidence from the OECD. *Journal of International Money and Finance*. https://doi.org/10.1016/j.jimonfin.2020.102292

Fatás, A., & Summers, L. H. (2018). The permanent effects of fiscal consolidations. *Journal of International Economics*, *112*(April 2017), 238–250. https://doi.org/10.1016/j.jinteco.2017.11.007

Fragetta, M., & Tamborini, R. (2019). It's not austerity. Or is it? Assessing the effect of austerity on growth in the European Union, 2010-15. *International Review of Economics and Finance*, *62*(March), 196–212. https://doi.org/10.1016/j.iref.2019.03.013

Gechert, S., Horn, G., & Paetz, C. (2019). Long-term Effects of Fiscal Stimulus and Austerity in Europe. *Oxford Bulletin of Economics and Statistics*, *81*(3), 647–666. https://doi.org/10.1111/obes.12287

Guajardo, J., Leigh, D., & Pescatori, A. (2014). Expansionary austerity? International evidence. *Journal of the European Economic Association*, *12*(4), 949–968. https://doi.org/10.1111/jeea.12083

Hernández de Cos, P., & Moral-Benito, E. (2016). On the predictability of narrative fiscal adjustments. *Economics Letters*, *143*, 69–72. https://doi.org/10.1016/j.econlet.2016.03.024

Jordà, Ò., & Taylor, A. M. (2016). The Time for Austerity: Estimating the Average Treatment Effect of Fiscal Policy. *Economic Journal*, *126*(590), 219–255. https://doi.org/10.1111/ecoj.12332

Kleis, M., & Moessinger, M.-D. (2016). The Long-Run Effect of Fiscal Consolidation on Economic Growth: Evidence from Quantitative Case Studies. *SSRN Electronic Journal*, *16*. https://doi.org/10.2139/ssrn.2801590

Pescatori, A., Leigh, D., Guajardo, J., & Devries, P. (2011). A New Action-Based Dataset of Fiscal Consolidation. *IMF Working Papers*, *11*(128), 1. https://doi.org/10.5089/9781455264407.001

Romer, C. D., & Romer, D. H. (2010). The macroeconomic effects of tax changes: Estimates based on a new measure of fiscal shocks. *American Economic Review*, *100*(3), 763–801. https://doi.org/10.1257/aer.100.3.763

Terzi, A. (2020). Macroeconomic adjustment in the euro area. *European Economic Review*, *128*. https://doi.org/10.1016/j.euroecorev.2020.103516

Annexes

Table A1 – Data sources

Variables	Data Source
GDP (constant LCU)	World Bank
GDP per capita (constant 2010 US$)	World Bank
Inflation, consumer prices (annual %)	World Bank
Current account balance (% of GDP)	World Bank
Real effective exchange rate index (2010 = 100)	World Bank
Unemployment (% of total labor force)	World Bank
Labor force participation rate (% of total population ages 15+)	World Bank
Age dependency ratio, young (% of working-age population)	World Bank
Age dependency ratio, old (% of working-age population)	World Bank
Population growth (annual %)	World Bank
Gross fixed capital formation (% of GDP)	World Bank
Industry, value added (% of GDP)	World Bank
Domestic credit to private sector (% of GDP)	World Bank
Trade openness ((Exports + Imports) in % of GDP)	World Bank
General government revenue (% of GDP)	IMF WEO 2020
General government gross debt (% of GDP)	IMF WEO 2020
General government net lending/borrowing (% of GDP)	IMF WEO 2020
General government primary net lending/borrowing (% of GDP)	IMF WEO 2020
Structural balance (% of potential GDP)	IMF WEO 2020
Total fiscal consolidation (narratively identified)	Alesina et al. (2015)
Percentage of secondary schooling attained (population over 25 years)	Barro and Lee (2013)
Percentage of tertiary schooling attained (population over 25 years)	Barro and Lee (2013)

Table A2 – Treatment episodes identified (narrative sample)

Countries	Treatment start
Australia	1985
Austria	1980
Austria	2001
Belgium	1982
Belgium	1990
Belgium	2010
Canada	2010
Germany	1982
Germany	1991
Germany	2003
Spain	1989
Spain	1992
Finland	1992

Countries	Treatment start
France	1987
France	1991
United Kingdom	2010
Ireland	1982
Italy	1991
Italy	2004
Italy	2010
Japan	2003
Portugal	2000
Portugal	2005
Portugal	2010
United States	1985

Table A3 – Treatment episodes identified (STPB sample)

Countries	Treatment start
Argentina	2002
Barbados	2014
Belarus	2012
Bosnia and Herzegovina	2001
Bosnia and Herzegovina	2010
Bulgaria	2012
Costa Rica	2003
Croatia	2007
Croatia	2014
Cyprus	2013
Czech Republic	2010
Denmark	2001
Egypt	2014
El Salvador	2001
France	2011
Germany	2011
Greece	2009
Grenada	2003
Grenada	2014
Guyana	2007
Hungary	2007
Iceland	2010
Ireland	2011
Israel	2003
Israel	2010
Italy	1992
Japan	2014
Jordan	2010

Countries	Treatment start
Latvia	2011
Lebanon	2007
Lithuania	2009
Malaysia	2001
Malta	2004
Malta	2009
Mexico	1998
Morocco	2003
Morocco	2013
Netherlands	2012
New Zealand	2012
Panama	2005
Philippines	2011
Poland	2011
Portugal	2011
Romania	2010
Russia	2000
Russia	2010
Slovak Republic	1998
Slovak Republic	2011
Spain	2010
Suriname	2003
Thailand	2003
Ukraine	2005
United Kingdom	2010
United States	2011
Uruguay	2002

Figure A1 – LASSO estimation to generate the V matrix for model A and F (narrative sample)

GDP	1.8836350
GDP_capita	-2.6056126
Inflation	-3.0360116
Unemployment	4.2762509
Labor_participation_rate	.
Young_dependency_ratio	5.3718156
Old_dependency_ratio	-5.0884696
Investment	0.8368940
Population_growth	-0.9692256
Secondary	2.4439582
Tertiary	-1.6919998
Trade_openness	2.2739300

	RMSE	Rsquare
Training sample performance	5.932208	0.7103263
Test sample performance	8.073903	0.7197132

V matrix (variables have the same order as in the coefficient table, excluding those with estimated coefficient = 0)

0.06180350 0.08549214 0.09961386 0.14030706 0.17625337 0.16695657 0.02745913 0.03180103 0.08018814 0.05551581 0.07460938

Figure A2 – LASSO estimation to generate the V matrix for model D (narrative sample)

GDP	.
GDP_capita	.
Inflation	.
Current_account	.
REER	-2.0550286
Unemployment	.
Labor_participation_rate	.
Young_dependency_ratio	1.7237084
Old_dependency_ratio	.
Investment	.
Population_growth	0.4878613
Industry_share	1.5520264
Domestic_credit_to_private	-0.2282339
Public_debt	.
Budget_balance	.
Budget_primary_balance	.
Government_revenue	.
Secondary	0.6433916
Tertiary	.
Trade_openness	.

	RMSE	Rsquare
Training sample performance	3.953966	0.573427
Test sample performance	6.542371	0.1627398

V matrix (variables have the same order as in the coefficient table, excluding those with estimated coefficient = 0)

0.30716767 0.25764483 0.07292123 0.23198332 0.03411440 0.09616854

Figure A3 – LASSO estimation to generate the V matrix for model E (narrative sample)

```
GDP                          .
GDP_capita                   .
Inflation                    .
Unemployment                 4.0220230
Labor_participation_rate    -0.7022225
Young_dependency_ratio       5.1764264
Old_dependency_ratio        -2.6980460
Investment                   .
Population_growth            .
Budget_balance               .
Budget_primary_balance       .
Government_revenue          -1.0353502
Secondary                    1.5193714
Tertiary                    -1.1053985
Trade_openness               2.1839563
```

Training sample performance

```
    RMSE   Rsquare
6.130752 0.7292251
```

Test sample performance

```
    RMSE   Rsquare
6.648848 0.2880553
```

V matrix (variables have the same order as in the coefficient table, excluding those with estimated coefficient = 0)

```
0.21808100 0.03807571 0.28067474 0.14629269 0.05613847 0.08238293 0.05993661 0.11841786
```

Figure A4 – LASSO estimation to generate the V matrix for model A (STPB sample)

```
GDP                         -5.3858980
GDP_capita                   .
Inflation                    .
Current_account              1.4005807
Unemployment                 .
Labor_participation_rate     1.8540091
Young_dependency_ratio       2.4513750
Old_dependency_ratio        -3.0706503
Investment                   6.1684474
Population_growth            0.2484537
Industry_share               .
Public_debt                  .
Budget_balance              -0.6371356
Budget_primary_balance       .
Government_revenue          -0.3726240
Secondary                    .
Tertiary                    -1.3305304
Trade_openness               .
```

Training sample performance

```
    RMSE   Rsquare
11.45944 0.6162143
```

Test sample performance

```
    RMSE   Rsquare
9.287164 0.5376968
```

V matrix (variables have the same order as in the coefficient table, excluding those with estimated coefficient = 0)

```
0.23498986 0.06110815 0.08089149 0.10695491 0.13397425 0.26913294 0.01084018 0.02779859 0.01625780 0.05805182
```

Figure A5 – LASSO estimation to generate the V matrix for model D (STPB sample)

```
GDP                         -5.0787584
GDP_capita                   .
Inflation                    .
Current_account              2.1351828
REER                        -3.2883831
Unemployment                 .
Labor_participation_rate     1.1652796
Young_dependency_ratio       3.2344416
Old_dependency_ratio         .
Investment                   6.2818379
Population_growth            .
Industry_share               1.7755421
Domestic_credit_to_private  -1.0286977
Public_debt                  .
Budget_balance               .
Budget_primary_balance       .
Government_revenue          -2.3579866
Secondary                   -1.8737562
Tertiary                    -0.4974919
Trade_openness               .
```

	RMSE	Rsquare
Training sample performance	9.256905	0.7490666
Test sample performance	12.07852	0.5303412

V matrix (variables have the same order as in the coefficient table, excluding those with estimated coefficient = 0)

0.17685326 0.07435165 0.11450855 0.04057753 0.11263019 0.21874707 0.06182819 0.03582146 0.08211015 0.06524821 0.01732373

Table A4 – Estimated effects of fiscal consolidations grouped by pre-treatment public-debt

Model	Measure	25% Lowest	25% Highest	Control	95% Confidence Interval
A (Narrative)	IRF	-2.622	-7.181	-7.641	-15.747 | 0.585
	Wealth	-0.876	-1.149	-1.258	-2.108 | -0.787
B (Narrative)	IRF	2.881*	-4.450	-5.672	-11.486 | 0.788
	Wealth	0.217*	-0.856	-0.922	-1.419 | -0.365
A (STPB)	IRF	-5.571*	-0.095	-0.405	-4.026 | 3.366
	Wealth	-0.941*	-0.144	-0.165	-0.782 | 0.458
B (STPB)	IRF	-14.127*	0.419	-1.768	-7.579 | 4.335
	Wealth	-2.207*	-0.063	-0.347	-1.279 | 0.658

Table A5 – Estimated effects of fiscal consolidations grouped by pre-treatment trade-openness

Model	Measure	25% Lowest	25% Highest	Control	95% Confidence Interval
A (Narrative)	IRF	-3.138	-8.475	-8.144	- 15.797 | 19.070
	Wealth	-0.443*	-1.246	-1.268	-2.174 | -0.749
B (Narrative)	IRF	1.367	-6.286	-5.959	-11.106 | 18.791
	Wealth	0.346*	-0.925	-0.925	-1.391 | -0.481
A (STPB)	IRF	-2.893	0.420	-2.466	-6.268 | 2.045
	Wealth	-0.607	-0.038	-0.454	-1.053 | 0.279
B (STPB)	IRF	-5.090	-4.244	-4.627	-10.543 | 1.759
	Wealth	-0.906	-0.741	-0.748	-1.639 | 0.213

Table A6 – Estimated effects of fiscal consolidations grouped by pre-treatment investment

Model	Measure	25% Lowest	25% Highest	Control	95% Confidence Interval
A (Narrative)	IRF	-11.398*	-6.750	-4.471	-7.362 | -2.448
	Wealth	-1.719*	-1.008	-0.649	-1.019 | -0.329
B (Narrative)	IRF	-7.541*	-1.506	-2.635	-5.939 | -0.091
	Wealth	-1.199*	-0.337	-0.353	-0.832 | 0.092
A (STPB)	IRF	1.388	-6.637*	0.153	-3.861 | 4.387
	Wealth	-0.003	-1.162*	-0.016	-0.667 | 0.657
B (STPB)	IRF	3.260*	-9.223	-5.990	-13.236 | 1.069
	Wealth	0.380*	-1.524	-0.954	-2.077 | 0.138

Table A7 – Estimated effects of fiscal consolidations grouped by pre-treatment private credit

Model	Measure	25% Lowest	25% Highest	Control	95% Confidence Interval
A (Narrative)	IRF	-5.497	15.053	-7.538	-106.434|73.289
	Wealth	-0.805	-1.322	-1.698	-2.46 | -0.74
B (Narrative)	IRF	-1.858	11.563	-3.902	-85.370 | 60.064
	Wealth	-0.340	-1.029	-0.954	-1.440 | -0.265
A (STPB)	IRF	-1.519	-5.033	-1.794	-5.209 | 2.784
	Wealth	-0.351	-1.036*	-0.341	-0.878 | 0.365
B (STPB)	IRF	-0.290	-7.231	-2.950	-7.462 | 3.741
	Wealth	-0.171	-1.320*	-0.482	-1.188 | 0.547

Table A8 – Estimated effects of fiscal consolidations grouped by pre-treatment unemployment

Model	Measure	25% Lowest	25% Highest	Control	95% Confidence Interval
A (Narrative)	IRF	-10.856	-5.010	-4.964	- 15.907 | 4.556
	Wealth	-1.565	-0.858	-1.018	-1.872 | -0.303
B (Narrative)	IRF	-6.163	-2.751	-2.501	-11.688 | 5.722
	Wealth	-0.806	-0.399	-0.735	-1.363 | -0.210
A (STPB)	IRF	-0.626*	2.647*	-4.039	-6.428 | -1.372
	Wealth	-0.207*	0.294*	-0.712	-1.072 | -0.280
B (STPB)	IRF	-3.829	3.334*	-7.983	-12.639 | -3.302
	Wealth	-0.671	0.429*	-1.280	-1.983 | -0.590