The Flat Tax and Tax Revenues – A Synthetic Control Approach

Aleksandar KOSULIEV
"Angel Kanchev" University of Ruse, Bulgaria
akosuliev@uni-ruse.bg

Elizar STANEV
"Angel Kanchev" University of Ruse, Bulgaria
eastanev@uni-ruse.bg

Abstract. *In the 1990s and the 2000s a number of countries with economies in transition introduced the flat tax – a proportional tax on personal income, replacing their progressive systems of taxation. Eventually, some of the countries have reverted back to these systems, while others have kept the flat tax. The arguments in favour of the latter suggest that it stimulates economic growth and boosts tax revenues because of its simplicity, transparency, and lower impetus for tax avoidance. While analysis of the effects of the flat tax are often ideologically charged, we try to empirically test how it has affected tax revenues in selected countries from Central and Eastern Europe by applying the synthetic control method. This is a statistical method for causal inference in observational studies, which is used for the evaluation of treatment effects or policy interventions on a single unit over time. The synthetic control serves as a counterfactual to estimate what the development of the unit would have been without the intervention and compare it with the actual development. To our knowledge, this is the first time when this method has been used to study this subject matter and while the results are not conclusive, they suggest that the adoption of the flat tax has a negative effect on revenues and contribute to the debate about the direction of the development of the tax system and of the implementation of tax reforms.*

Keywords: taxation, flat tax, tax revenues, synthetic control, tax reform.

Introduction

With deepening globalization, various taxation systems across the globe have traditionally been the subject of numerous academic studies, as their analysis provides key insights into international capital flow, investment initiatives, domestic consumption and income disparity. As a system that influences almost all aspects of modern economies, measuring the effects of particular taxation policy changes is unavoidably contemporary, due to the high levels of uncertainty and complexity of economic relations, incentivized or limited by taxation interference.

In the pursuit of stimulating economic activity and lowering tax evasion, some governments around the world have introduced and implemented a proportional tax rate (also called a "flat tax"). Flat taxation is often used in both a specific and a broader sense. In the purest theoretical sense, it encompasses a single rate levied across a whole stratum of society or economic activities (e.g. on personal income, as a corporate tax, etc.), but some scholars and institutions also employ it as a catch-all term for taxation systems that include two-step (rarely three-step) bifurcated proportional rates, according to a certain tax base threshold (although some may argue that the latter should rather be classified as a progressive type of taxation). The commonality between the two seems to center around the simplicity and relatively low number of tax brackets (or lack thereof), imposed by the financial authorities of the particular state.

The main challenge of measuring the effects of a particular taxation reform is the single timeline of events, meaning that for a certain set of economic variables, present for a set amount

DOI: 10.2478/picbe-2025-0127
© 2025 A. Kosuliev; E. Stanev, published by Sciendo.
This work is licensed under the Creative Commons Attribution 4.0 License.

of time for a given state, there exists a single and unique combination of said variables with the governmentally-imposed taxation practices. This limits the ability of researchers to empirically observe alternatives, or stated simply – limits our capabilities of finding out “what could have been”. The problem, of course, persists across the board in many aspects of economic research. In order to address this problem, we approach the main question posed within this paper – whether the transition from a progressive tax system to a flat tax rate has a positive effect on tax revenues. For this purpose we employ the synthetic control method.

Literature review

In the 1990s and the 2000s a lot of countries, mostly transition economies from the former socialist block, experimented with the introduction of a flat tax rate on personal income. Some of them reverted back to progressive taxation, others kept the rates flat. This storm of policy changes in tax systems spurred a large volume of literature, dedicated to the effects of flat tax adoption. A part of it is dedicated to the consequences for tax revenues and we provide a brief and selective overview, focusing on methodology and conclusions.

The first group of studies either prioritizes theoretical arguments or the empirical part is mostly descriptive and is stripped of the application of a rigorous methodology for causal inference. An example of this is Nikolova (2016), who reports that the adoption of the flat tax in Bulgaria in 2008 lead to a nominal increase in the revenues from the income tax in national currency. This alone however is not enough to conclude that the flat tax was a success in the context of revenue generation. Keen et al. (2006) compare tax revenues before and after the introduction of the flat taxes (in percent of GDP) for eight Central and Eastern European (CEE) countries. For all but three the share went down and the unweighted average of the changes for all of them was lower. Bartha (2014) reports a drop in the personal income tax revenue in national currency in nominal terms and (accounting for inflation) even more in real terms.

A second group of studies employs micro-level data and applies either causal inference to assess the effects of past events or simulation techniques to predict various scenarios of change in the tax system. Gorodnichenko et al (2008) analyze the Russian 2001 flat rate income tax introduction utilizing data from a household panel for 1998-2004. They use difference-in-differences and regression discontinuity approaches and reach the conclusion that the tax reform stimulated better voluntary tax compliance, inferred from the observation that households with a reduction in the marginal tax rates experienced growth in their income by 10% higher than that in consumption. A study, based on household data from the EU-SILC database about Romania, does a micro-simulation of a hypothetical progressive income taxation schedule to replace the flat rate tax (Popescu et al., 2019). The authors use EUROMOD – a tax-benefit model for the European Union and run simulations for two progressive taxation scenarios. They report a slight negative effect upon the total income tax revenues (as a share of GDP), but point out that it may be the result of under representation of higher income households in the database.

A third group of studies conducts country level analysis of flat tax rate reforms. An example of this is Wheaton (2021), who runs a difference-in-differences analysis on a panel of twenty flat tax adopters. With regards to tax revenues, he concludes that potential decreases did not prevail because of increased labour supply and investment, stimulated by the reform. Another possible explanation is increased tax compliance.

In summary, the results are mixed, but the more rigorous econometric studies are conducive to the view that flat tax rates have a positive effect (or at least don’t affect negatively) the revenues from the personal income tax.

Methodology

We examine the effect of flat tax adoption on the revenues from individual and household taxes, expressed as a share of GDP. For this purpose we use the synthetic control method.

Experimental studies often are either not possible, prohibitively costly, or ethically unacceptable when assessing the effect of policy interventions at the level of units that include a large number of people and could potentially affect their welfare. When policymakers introduce changes in the institutional or regulatory environment, this opens an opportunity for researchers to study the phenomenon but not under the ideal conditions of a controlled experiment. Standard econometric tools such us regression analysis may not be suited for the purpose of impact evaluation and the quasi-experimental nature of many interventions has given rise to new instruments and methods to evaluate casual effects (for an overview see Gertler et al., 2016; Huntington-Klein, 2022).

The synthetic control method allows for a quantitative approach towards comparative case studies. Regression analysis is applicable in cases when there is a large number of observations, which allows for the identification of statistically significant effects, subject to the validity of a set of assumptions. Regressions on panel data allow for both cross and time dimensions and while potentially capturing the dynamics of the events they require many observations in multiple periods. A special case of a fixed effects panel regression model with interaction effects is the difference-in-differences model. It is applicable when there are two periods – a pre-treatment and a post-treatment, with some of the units exposed to the intervention in the second period, while others are not. For the researcher, real world policy interventions often don't offer the luxury of multiple (treated) units, which is where the synthetic control method comes in useful.

Changes in policy that affect a single (or only a few) aggregate unit(s) (countries, districts, organizations etc.) can't be assessed on their own as they lack a proper counterfactual. If the unit is treated, the same unit simply doesn't exist post-treatment in its non treated state. Instead of pure speculation on what would have happened, the synthetic control method provides a quantifiable approach to examine the effect from the intervention. If the treated unit has a perfect match with an untreated unit, comparisons will be straightforward. While some untreated units might be very much similar to the treated, usually no actually existing units will be completely identical to it. However such an identical unit could be artificially created for the purposes of the study, the so called synthetic control. The synthetic control is a composite of non treated units that approximates as much as possible the trend in the outcome variable of the treated unit in the pre-treatment period.

The assumption is that after the intervention the synthetic control would have continued the original trend, whereas the treated unit would deviate from it (if the intervention has an effect). As per Abadie et al. (2011), Abadie et al. (2014) and Abadie et al. (2021) the effect from the intervention α_{it} for unit i in time t is:

$$\alpha_{it} = Y_{it}^{I} - Y_{it}^{N} \tag{1}$$

where Y_{it}^{I} is the observed outcome for unit i at time t for the unit exposed to the intervention, and Y_{it}^{N} would be the outcome for unit i at time t in the absence of the intervention. Let's assume that there are $i = 1, \dots, J + 1$ units (where unit 1 is treated and the rest are not), and $t = 1, \dots, T$ periods (with T_0 pre-intervention periods, where $1 \leq T_0 \leq T$). As Y_{1t}^{I} is observed, the estimation of α_{1t} requires just the estimation of Y_{1t}^{N} . The latter is not observed and must be approximated by the synthetic control. The authors show that the estimator of α_{1t} could be given by:

$$\widehat{\alpha_{1t}} = Y_{1t} - \sum_{j=2}^{J+1} w_j^{*} Y_{jt} \tag{2}$$

where $\sum_{j=2}^{J+1} w_j^{*} Y_{jt}$ is the sum of the weighted values of the outcome variable for non treated units (i.e., outcome variable values for the synthetic control). The question is then, how are the weights (w_j^{*}) for each country estimated. The weights are determined during the pre-intervention period. There is a $(J \times 1)$ vector of weights $= (w_2, \dots, w_{j+1})$, with each weight being equal to or greater than 0 and all weights summing to 1. Each value of W is a weighted average of non treated units and each of these values can represent a synthetic control. If we assume that X_1 is a $(k \times 1)$ vector of pre-intervention characteristics for the treated unit, and X_0 is a $(k \times J)$ matrix of the values of these characteristics for the non treated units that form the donor pool (k being the number of predictors for the treated unit, the outcome variable can also be included), we choose the value of W that minimizes the difference between the pre-intervention characteristics of the treated unit and a synthetic control. Or, to put it more succinctly, we choose the value of W that minimizes $X_1 - X_0W$.

$$W^* = arg\,\min_{W} \|X_1 - X_0W\| \tag{3}$$

The results of the application of the synthetic control method are easy to understand as their presentation is graphical and straightforward. The outcome variables are plotted for the treated unit and for the synthetic control. If the two curves overlap or deviate slightly from each other in the pre-intervention period, the synthetic control is most likely a good fit. Given that, if the two curves diverge in the period after the intervention, this is an indicator of the intervention's effect.

There are additional steps to ensure the inferential character of the results and their credibility. The first such step is the generation of placebos, which can be either in-time or unit placebos. With in-time placebos we take the treated unit and change the intervention point to a period when the intervention actually didn't happen. When studying the economic costs of the German unification in 1990, Abadie et al. (2014) create an in-time placebo as if the unification was in 1975, and find no difference in the trends of the outcome variable (real GDP per capita for actual and synthetic West Germany). The unit placebo test removes the treated unit from the pool and calculates synthetic controls for each of the remaining units vis-a-vis the others, which are used for the generation of the synthetic control. In the German unification example, a placebo is calculated for Norway, assuming they underwent a similar event in 1990, another placebo is calculated for Austria, then Italy, and so on (in total 16 non-treated OECD countries, included in the pool). Ideally, the paths of the outcome variables will overlap for the treated unit and its synthetic control before the intervention and they will take different paths after that. For the placebos, they will overlap in both periods. In most cases, this idealized situation will not be realized and the graphical representation will not suffice to decide how the treated unit effect compares to the placebos.

We need an objectively quantifiable way to evaluate the intervention effect and for that purpose we calculate the mean square prediction error (MSPE) (or sometimes the root MSPE) in the pre-intervention period and in the post-intervention period. The MSPE measures the lack of fit between the path of the actual outcome for the unit and the path of its synthetic control. For the treated unit in the pre-intervention period it is represented as follows:

$$MSPE = \frac{1}{T_0}\sum_{t=1}^{T_0}\left(Y_{1t} - \sum_{j=2}^{J+1} w_j^* Y_{jt}\right)^2 \tag{4}$$

The smaller the MSPE, the better the fit, which is what we want to get in the pre-intervention period. But after the intervention, in case it is effective, we would expect bigger values for MSPE for the treated unit (but not for the placebos). Then the ratios of post- and pre-intervention MSPEs are calculated for each unit ($ratio = \frac{postMSPE}{preMSPE}$). The bigger the ratio for the treated unit, the better. As there can be no cutoff value to say how big is good enough (after all, it depends on the existence of an intervention effect and its size, and the effect might or might not be there), the ratio for the treated unit is compared to the ratios of the placebos. Their ratios should be small and not much different from 1. But due to chance, quality and availability of the data, choice of predictors and units in the donor pool, the placebos ratios could deviate from 1. If an intervention effect exists, the treated unit should have the highest or one of the highest ratios in the pool. This is the case in the German unification example (Abadie et al., 2014), as the RSMPE for West Germany is much higher than for all the other 16 countries from the donor pool. If a random country is picked from the sample, the probability of obtaining a ratio as high as that of West Germany would be 1/17 ≃ 0.06. The p -value is calculated as follows:

$$p = \frac{1+\sum_{j=2}^{J+1} \mathbb{1}(condition)\left(\frac{MSPE_j^{post}}{MSPE_j^{pre}} \geq \frac{MSPE_1^{post}}{MSPE_1^{pre}}\right)}{J+1} \tag{5}$$

where $MSPE_j^{post}$ is the post-treatment mean squared prediction error (MSPE) for unit j, $MSPE_j^{pre}$ is the pre-treatment MSPE for unit j, $j = 1$ the treated unit, J is the number of units in the donor pool, and $J + 1$ is the total number of units, $\mathbb{1}(condition)$ is the indicator function, which equals 1 if the condition inside holds and 0 otherwise.

There are a number of countries from Central and Eastern Europe (CEE) with flat tax on individual and household incomes (see Table 1 of Krekó et al. (2023) for the years of introduction). But we are constrained by data availability, idiosyncratic shocks, and the staggered adoption of the flat tax regarding the choice of countries to which to apply the synthetic control method. First, available data for the variables in interest go back to the mid-1990s at best. The pre-intervention period can't be very long and this lowers the fit of the synthetic control. Second, some countries, e.g. Bulgaria, have periods of economic trouble and instability in the pre-treatment period, which affect the outcome variable (in this case "revenues from taxes on individual and household income as a share of GDP") and we have to skip these years. And third, the staggered adoption of the flat tax in CEE excludes some countries as potential donors to the pool (e.g. Slovakia did it in 2004 and can't be used in the donor pool for Romania and Bulgaria, where it happened in 2005 and 2008 respectively). Thus the donor pool is dominated by countries from Southern Europe or Western Europe. They lack the similarity of the characteristics of other CEE countries, which potentially decreases the fit of the synthetic control to the treated unit (also a CEE country).

With this in mind, we apply the synthetic control method to the introduction of the flat tax in Slovakia (in 2004), Romania (in 2005), Bulgaria (2008), Czech Republic (2008), and Hungary (2011). The inclusion of Hungary is a little controversial, because the transition from progressive taxation took them two years, starting in 2011 (see Krekó et al. (2023)). We also test the intervention in reverse, when Slovakia went back to progressive taxation in 2013.

We follow the steps outlined above (except for the in-time placebos, because our time series are short), using software for the calculations of the equations (*R* (version 4.4.1) and *R Studio* for data processing, the *tidysynth* package (version 0.2.0) for the synthetic control method). In the choice of predictor variables for the generation of the synthetic control, we have tried to identify those, which could potentially influence the outcome variable (taxes on individual or household income as % of GDP). We have applied an OLS regression (R^2=0,30) to validate our choice, even though the OLS is a very rough estimator and not very appropriate in this case, so the results serve only as a guidance. GDP growth is the only indicator with no statistical significance in the OLS model, but we consider it too important to omit as a predictor in the synthetic control generation. As expected, a higher share of the compensation of employees in GDP is associated with higher revenues from income taxes (the regression coefficients is 0,20). In line with the expectations are the inversely related unemployment rate and informality of the economy index (with coefficients respectively -0,12 and -0,14). Trade openness and inflation are also inversely related, but their coefficients are small (-0,02 and -0,06). We have included a dummy variable for the flat tax and it is statistically significant, indicating that its presence is associated with 1,5 percentage points lower revenues from individual and household income taxes (as a share of GDP). The data and its sources are described in the Table 1 in the Appendix.

While we keep the predictors the same across all cases, the number of countries in the donor pool vary with each case. In general, we try to include as many countries as possible for which data is available and the time series go back in time far enough, in order to calculate the weights and achieve a better fit for the synthetic control in the pre-treatment period. The donor pool can be inferred for each case from Figure 3 in the Appendix.

Results and discussions

The paths for the outcome variable (taxes on individual or household income as % of GDP) for the treated units and their synthetic controls for each case are shown on Figure 1. The graphs, related to the subsequent diagnostics, are located in the Appendix – Figure 2, which shows the placebo tests, and Figure 3 with the MSPE ratios. It is obvious that the synthetic control fits poorly in the case of Slovakia (2004), Romania (2004), and Bulgaria (2008) adopting the flat tax. Adhikari & Alm (2016) achieve much better fit of the synthetic controls for Slovakia and Romania in their study of the effect of the flat tax adoption on GDP growth, but their outcome variable and predictors are different and the time series in the pre-treatment period go all the way to the early 1990s. Back to our own study, in the case of both Slovakia and Romania, apart from the poor fit we observe low MSPE ratios (as shown on Figure 3) and high p-values (respectively 0,39 and 0,41). Bulgaria is different in this respect. It has the highest MSPE ratio of all countries in the pool and the lowest possible p-value given the pool's size (0,06). Yet, we would be jumping to conclusions if we declare that the introduction of the flat tax is associated with lower share of tax revenues (coming from Figure 1, where the path of the synthetic control is above the path of actual Bulgaria). The divergence with the synthetic control starts five years before the tax reform. The weight of just one country (Greece) in the construction of the control is almost 100%, which is another reason to have reservations about the result.

Hungary and the Czech Republic show a significantly better fit of the synthetic control. Probably the reason for this is the length of the pre-intervention period (almost twice as much). For both countries the path of the outcome variable for the synthetic control is above the treated unit after the adoption of the flat tax. As Figure 2 shows, the difference is bigger than that for most placebos (in the case of Hungary – for all). Hungary has the highest MSPE ratio in its pool, while Portugal and Sweden exceed the MSPE ratio of the Czech Republic in the Czech pool. The ratios are reflected in the higher p-value for the latter (0,18), whereas for Hungary it is 0,06. While it doesn't go below the generally accepted level of 0,05, it is very close to it and the lowest possible in a pool of 17 units. Hungary also has a balanced pool of donors, with Greece, Slovenia, Belgium and the Netherlands contributing with double digit weights in the synthetic control. For the Czech Republic the contribution of Greece is more than 87%.

Finally, we want to examine the case of going back from the flat tax to progressive taxation. The Slovak Republic offers this opportunity with its tax reform in 2013. The pre-intervention period is not very long (2005-2013) and the number of countries in the pool is small, but the path of the synthetic control matches the treated country relatively well. However the match goes beyond the year of intervention, and the p-value is 0,57. Unlike the case of flat tax adoption in Slovakia, Romania, and Bulgaria, the low MSPE ratio and high p-value is not indicative of a poor fit. The likeliest explanation is that the transition from flat to progressive taxation in Slovakia simply didn't have an effect on the tax revenues.

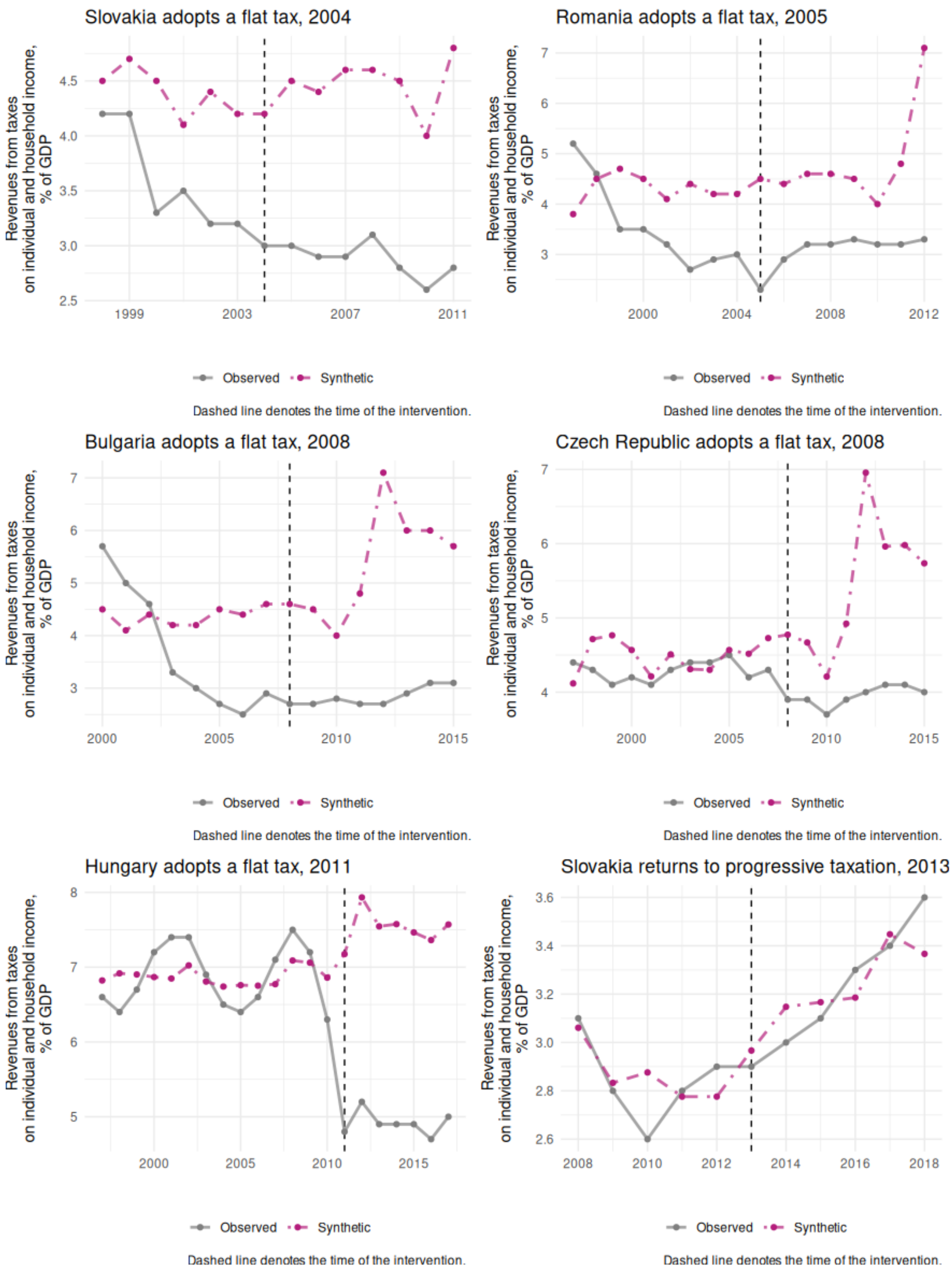

Figure 1. Observed paths and synthetic controls

Source: Authors' own research results.

Conclusion

To our knowledge, this is the first time the synthetic control approach has been used to study the effect of the introduction of the flat tax on the revenues from individual and household income. Our results are inconclusive due to certain data limitations, but they tend towards the view that the introduction of the flat tax leads to lower revenues. This is in contrast to the results of some previous studies that employ other methods of analysis. We should emphasize that tax reforms, while falling into broad categories such as flat tax or progressive schedules, can differ substantially in details. The introduction of a flat tax may replace a progressive tax system, but the new effective tax rate could be lower or higher than before, depending on the rate level chosen, thus taking revenue changes in one direction or another. Various tax deductions and exemptions could further complicate the analysis. A possible direction of research is to expand the current study, accounting for effective tax rates before and after the tax reform.

References

Adhikari, B., & Alm, J. (2016). Evaluating the Economic Effects of Flat Tax Reforms Using Synthetic Control Methods. In *Southern Economic Journal* (Vol. 83, Issue 2, pp. 437–463). Wiley. https://doi.org/10.1002/soej.12152

Nikolova, D. (2016). Flat Income Tax and Tax Revenue. In Hadjinedelcheva, K. (Eds.), In *Flat Tax in Bulgaria: History, Introduction, Results* (pp. 37-43). Institute for Market Economics

Bartha, Z. (2014). Mid-term Effects of the Flat Rate Personal Income Tax in Hungary. *Humanities and Social Sciences*, vol. 19, 21 (3/2014), pp. 25-42, https://doi.org/10.7862/rz.2014.hss.34

Gertler, P., Martinez, S., Rawlings, L. B., Premand, P., & Vermeersch, C. M. J. (2016). Impact Evaluation in Practice: Second Edition. Inter-American Development Bank. https://doi.org/10.18235/0006529

Gorodnichenko, Y., Martinez-Vazquez, J., & Peter, K. S. (2008). Myth and Reality of Flat Tax Reform: Micro Estimates of Tax Evasion Response and Welfare Effects in Russia. *National Bureau of Economic Research*. https://doi.org/10.3386/w13719

Huntington-Klein, N. (2022). The effect. An Introduction to Research Design and Causality. Philadelphia, PA: Chapman & Hall/CRC.

Keen, M., Kim, Y., Varsano, R. (2006). The “Flat Tax(es)”: Principles and Evidence, IMF Working Paper WP/06/218.

Krekó, J., Erős, H., Greskovics, B., Hajnal, Á., & Scharle, Á. (2023). The redistributive effect of the Hungarian flat tax and family allowance system. In *Acta Oeconomica* (Vol. 73, Issue 4, pp. 481–503). Akademiai Kiado Zrt. https://doi.org/10.1556/032.2023.00011

Popescu, M. E., Militaru, E., Stanila, L., Vasilescu, M. D., & Cristescu, A. (2019). Flat-Rate versus Progressive Taxation? An Impact Evaluation Study for the Case of Romania. In *Sustainability* (Vol. 11, Issue 22, p. 6405). MDPI AG. https://doi.org/10.3390/su11226405

Wheaton, B. (2021). Essays on Political Economy and Macroeconomics. Doctoral dissertation, *Harvard University Graduate School of Arts and Sciences*. Available at https://nrs.harvard.edu/URN-3:HUL.INSTREPOS:37368194

Appendix

Table 1. Description of predictor variables in the construction of the synthetic controls

Variable name	Measurement unit	Source	Database
Taxes on individual or household income including holding gains; General government	Percentage of gross domestic product (GDP)	Eurostat	Main national accounts tax aggregates https://doi.org/10.2908/GOV_10A_TAXAG
GDP growth rate	Percentage change over previous year. Own calculations using "Chain linked volumes (2010), million euro"	Eurostat	Gross domestic product (GDP) and main components (output, expenditure and income) https://doi.org/10.2908/NAMA_10_GDP
Inflation rate	Price index (implicit deflator), percentage change on previous period, national currency	Eurostat	Gross domestic product (GDP) and main components (output, expenditure and income) https://doi.org/10.2908/NAMA_10_GDP
Trade openness (Exports + Imports of goods and services)	Percentage of gross domestic product (GDP)	Eurostat	Gross domestic product (GDP) and main components (output, expenditure and income) https://doi.org/10.2908/NAMA_10_GDP
Total unemployment	Percentage of unemployed in the population from 15 to 74 years	Eurostat	Unemployment rates by sex, age and citizenship (%) https://doi.org/10.2908/LFSA_URGAN
Compensation of employees	Percentage of gross domestic product (GDP)	Eurostat	Gross domestic product (GDP) and main components (output, expenditure and income) https://doi.org/10.2908/NAMA_10_GDP
DGE estimates of informal output	Percentage of official GDP	The World Bank	Informal Economy Database Elgin, C., M. A. Kose, F. Ohnsorge, and S. Yu. 2021. "Understanding Informality." CERP Discussion Paper 16497, Centre for Economic Policy Research, London.

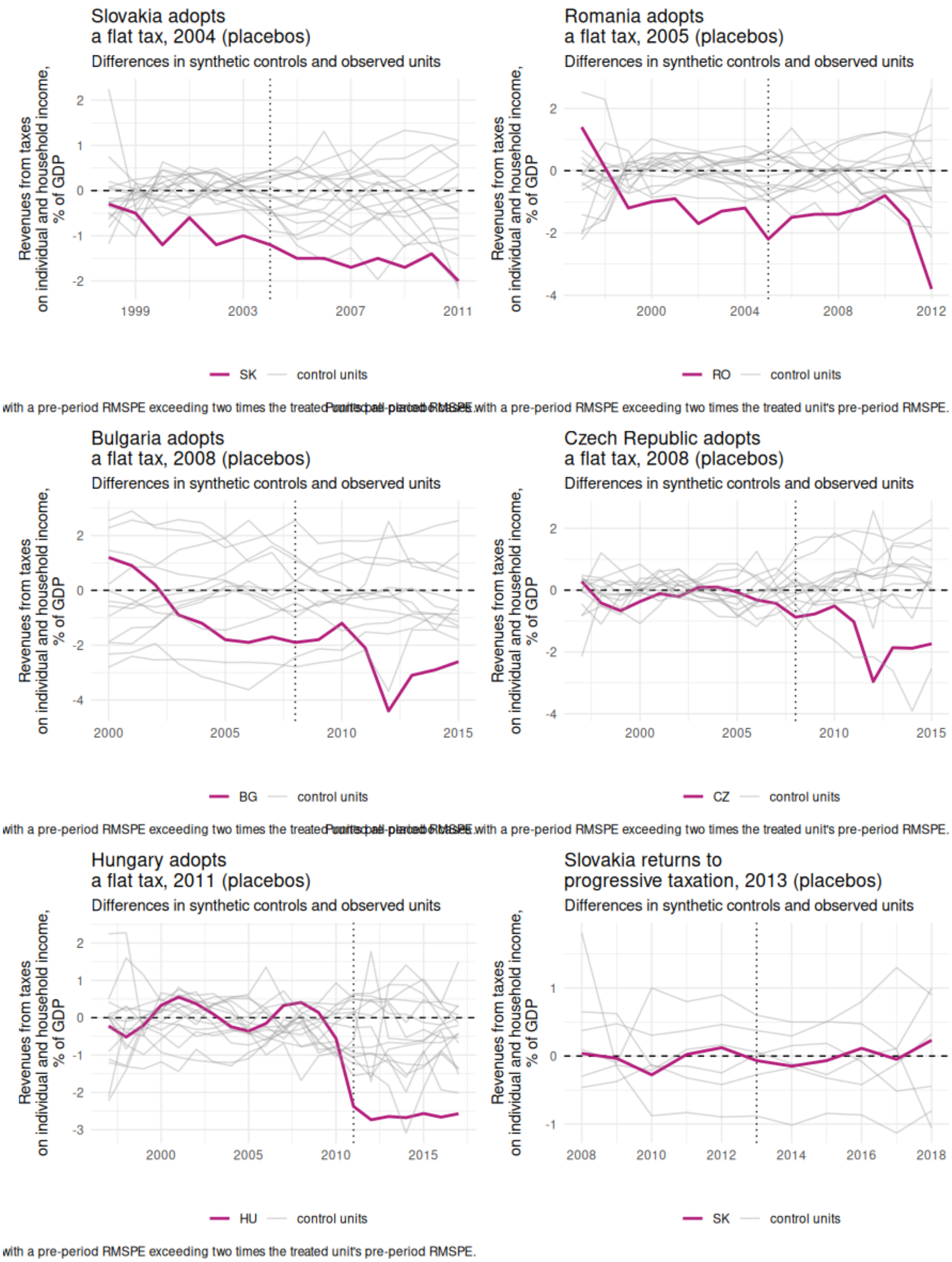

Figure 2. Treated units and their placebos

Source: Authors' own research results.

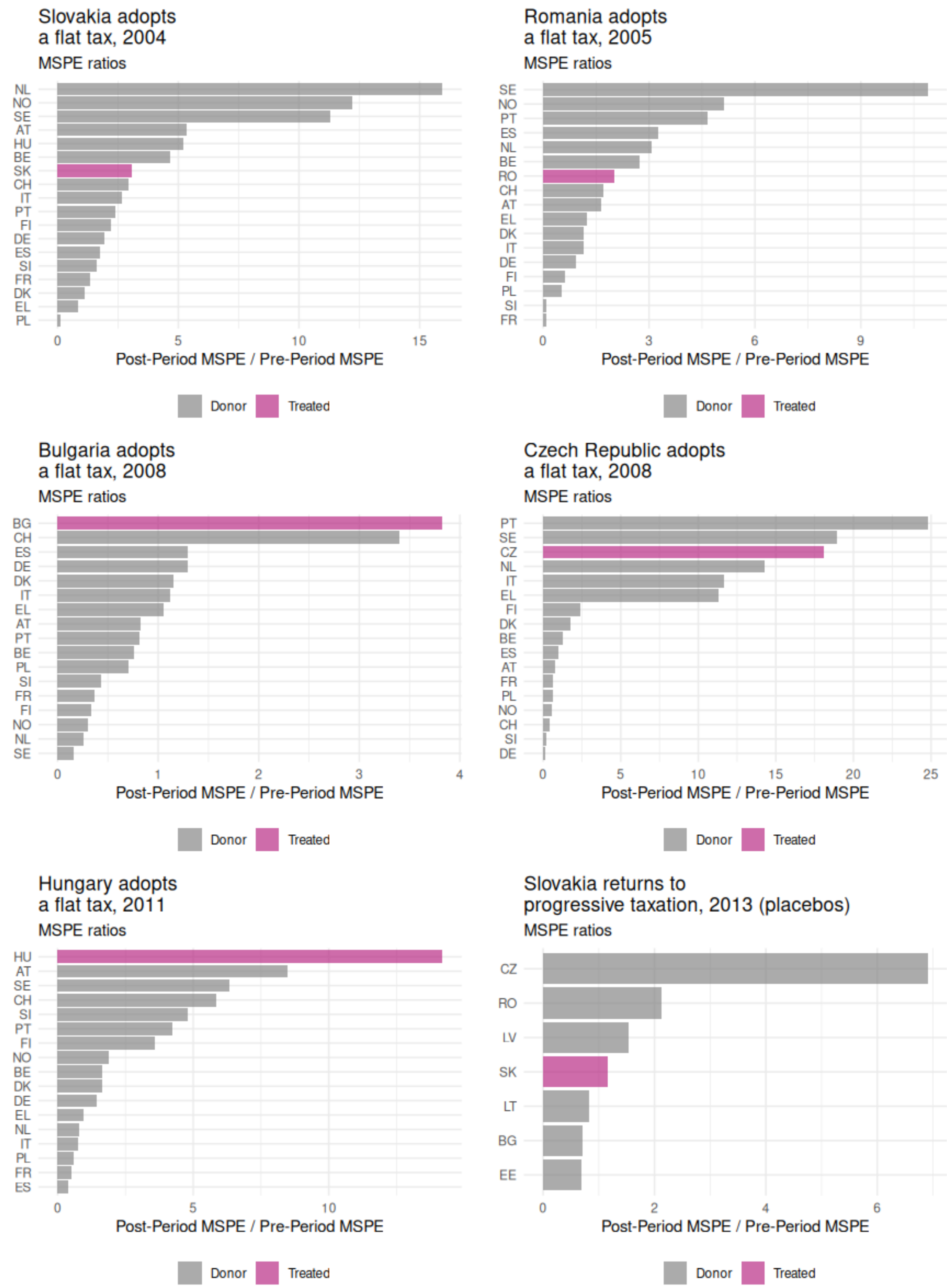

Figure 3. MSPE ratios

Source: Authors' own research results.

DOI: 10.2478/picbe-2025-0127, pp. 1655-1665, ISSN 2558-9652 |
Proceedings of the 19th International Conference on Business Excellence 2025