On Analysis of Economic Impact of the 2014 Russian invasion of Ukraine

Miray Ozcan, Viktor Tsvil, Yashvardhan Sharma

Minerva University

CS130: Statistical Modeling: Prediction and Causal Inference

Prof. A. Diamond

April 21, 2023

Table of Contents

On Analysis of Economic Impact of the 2014 Russian invasion of Ukraine

Executive Summary

In this project, we summarize, replicate, and extend parts of the paper "The Economic Costs of Hybrid Wars: The Case of Ukraine" by Bluszcz and Valente (2020). In this paper, by extending the findings of the original paper, we intend to provide more robust conclusions about the economic impact of the Donbas war on Ukraine's GDPpc by three primary efforts: 1) examining the long-term effects of the 2014 Russian invasion on Ukraine's economy through extending the post-treatment period to 2021, 2) using more rigorous criteria for the donor pool, and 3) adjusting the weights of predictors to match them better.

Section 1: Introduction

Bluszcz and Valente (2020) explores the following causal inference question: "How did the 2014 Russian invasion of Ukraine (Donbas war) affect (1) the Ukrainian economy as a whole and (2) war-affected regions, as measured by GDP per capita, in years 2013-2017?". The Donbas war is an armed conflict between Ukraine and pro-Russian Russia-backed separatists in the Luhansk and Donetsk regions in Eastern Ukraine. The region used to be the core of Ukrainian exports and production, accounting for 25% of Ukraine's exports (Bluszcz and Valente 2020); therefore, the conflict is expected to have a strong negative impact on Ukraine's economy (measured by GDPpc[1]).

The researchers employ the synthetic control method (SCM) in their research (Bluszcz and Valente 2020). They create the donor pool of Post-Soviet (excluding Russia) and Eastern

[1] GDPpc - GDP per capita

Bloc countries that did not experience any shocks in the pre-war period before 2013. Finally, they construct the synthetic control unit by matching donor pool countries and Ukraine on 6 variables (Human Development Index, political regime, inflation, gross fixed capital formation as % of GDP, and the dependence on trade with Russia). Bluszcz and Valente (2020) account for previous shocks in Ukraine (the revolution of 2004 and gas crisis in 2009) to separate the effects of past shocks and conclude that Ukraine's loss in GDP caused by the Russian invasion averaged $1310 (13.9%) of GDPpc. They also conduct (1) in-place placebo tests for all countries from the donor pool to show that the obtained results are the most extreme in the case of Ukraine as well as (2) leave-one-out robustness checks. All replication data is available at the link (Bluszcz & Valente, 2020).

Finally, the researchers perform SCM on Luhansk and Donetsk regions directly affected by the Donbas war. They find that the average GDPpc in the Luhansk and Donetsk region decreased by $3355 (52%) and $4294 (42%) respectively, caused by this war.

Section 2: Replication

The link to the code can be found in Appendix A.

Replication Method Explanation - Synthetic Control

The Synthetic Control Method (SCM) is a statistical technique that helps researchers understand the effects of some intervention when conducting a traditional RCT experiment is not an option, potentially due to the nature of the treatment or the inconvenience of randomly assigning units to comparison groups. To discuss causal effects in the research context, it is essential to identify the alternative outcome if the treatment wasn't applied to a particular unit. By generating a "synthetic" version of the unit that received the intervention, SCM addresses the fundamental problem of causal inference, the inability to observe two outcomes simultaneously. To create the synthetic control unit, researchers use data from other units (e.g., individuals, cities, or countries) that did not receive the intervention but have similar characteristics to the unit that received the intervention. This group of comparison units is called the "donor" pool. The weighted sum of the donor pool units can create a "synthetic control" that closely resembles the unit of interest during the pre-intervention period. After the intervention, any differences between the actual outcome of the treated unit and the synthetic one can be attributed to the intervention, helping researchers make causal inferences.

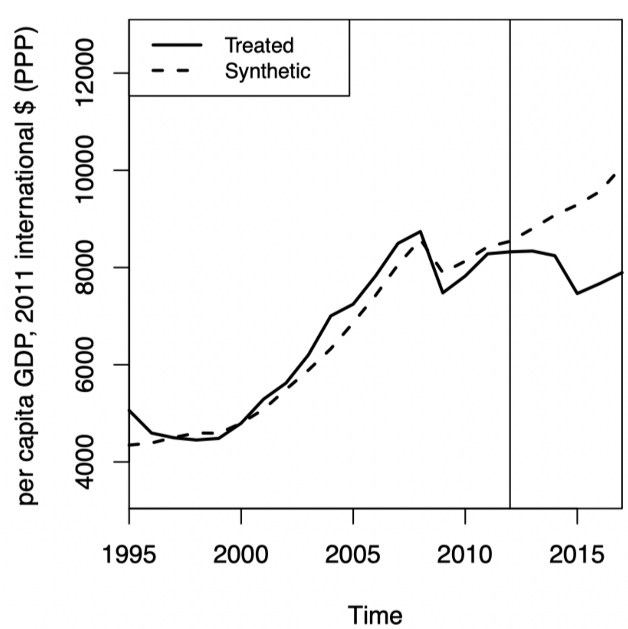
Bluszcz & Valente (2020) created a synthetic control for Ukraine by combining data from similar countries unaffected by the war, primarily from the former Soviet Union and Eastern Bloc, excluding Russia. They calculated optimal weights for each country in the donor pool to resemble Ukraine regarding GDPpc and relevant pre-war factors. This synthetic Ukraine allowed them to estimate the causal effect of the Donbas war on Ukraine's GDP by comparing the actual and synthetic GDPpc after the war, 2013-2017.

Three key assumptions were made to ensure the validity of the findings: 1) the outcome of all countries follows a linear model, 2) optimal weights for control countries should range between 0 and 1, and 3) countries in the donor pool must be unaffected by the intervention or any other shocks specific to that country. Therefore, the synthetic control group was carefully selected to avoid bias.

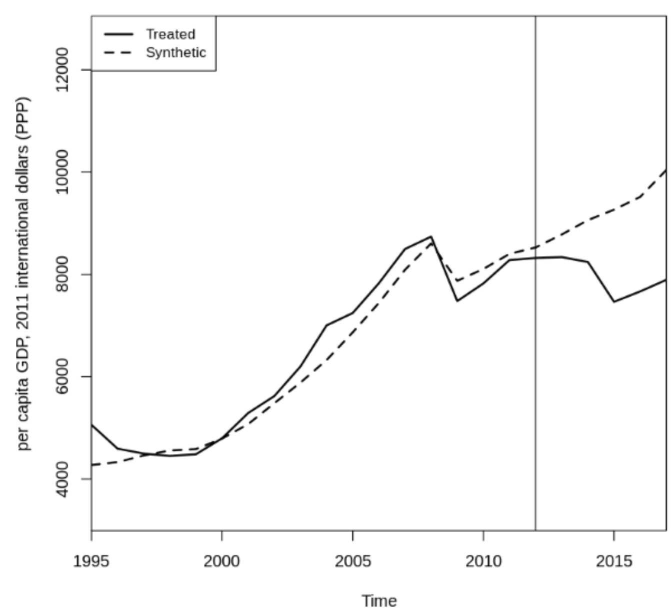
SCM has several benefits in this context. It can estimate causal effects even when randomized experiments are impossible, account for unobserved confounding factors while constructing the synthetic control, and precisely estimate sparse weights for control units. This allows for assessing their contribution to the counterfactual and potential bias directions. These advantages make the SCM particularly valuable for examining the impact of complex interventions, such as wars, on economic outcomes like GDPpc.

Replication Figure

Fig. 1, comparing the original and replication of Figure 2 from the original paper, displays the trends of per capita GDP of Ukraine and its synthetic counterpart. Like the original figure, this replication clearly shows that both Ukraine and its synthetic counterpart follow a very similar path until 2012 and deviate considerably afterwards. The replicated figure reports the average treatment effect of -1411$, as compared to -1438$[2] in the original paper, showing the quality of the replication. The difference can be attributed to random calculation effects. Additionally, we noticed that not all variables were well-balanced, which was a discrepancy that persisted in the original paper. We cover this issue in the Extension section.



Original



Replication

[2] Note that this number, as well as Fig. 2, reports the unadjusted average treatment effect, as compared to the lower one that adjusted for the gas disputes with Russia in 2009. Since the focus is on replicating Fig. 2 from Bluszcz and Valente (2020), we discuss and analyze related findings.

Figure 1. Original (left) and replication (right) of Figure 2 from Bluszcz & Valente (2020). The figures match each other closely. The average treatment effect differs by 27$, which is insignificant in the current context.

Section 3: Extension

This section contains three extensions of the original paper. First, it extends the post-treatment period from 2013-2017 to 2013-2021 to track the causal effect of war over time. Second, it takes a more conservative approach to form the donor pool. Third, it (1) adjusts the weights of different predictor variables to achieve a better balance in the original paper and our first extension and (2) explains why such adjustments are not needed in Part 2.

Part 1: Extending the Post-Treatment Period

Bluszcz and Valente (2020) perform their analyses during the post-treatment period of 2013-2017. We extended the period to 2013-2021 as the effects of war can have a long-term impact on economic outcomes, and the longer post-treatment period would provide more insights into the long-term effects of the conflict. As part of this extension, we used GDP Per Capita PPP (constant 2017 international USD) instead of the data with the 2011 base rate in the original paper, as the 2011 data was no longer available for all countries. The data was collected from the World Bank (n.d.). As seen in Fig. 2, there is no significant difference after changing the base rates. Similar to the original plot, both follow a very similar path until 2012 and deviate considerably from the controls, especially in the extended post-treatment plot.

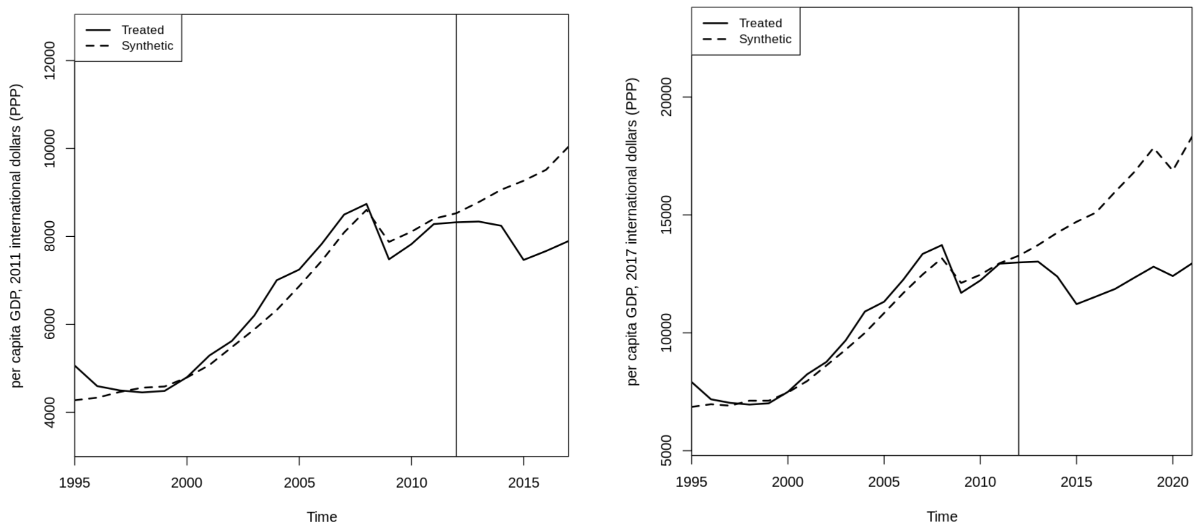


Figure 2. Trends in GDP per capita: Ukraine vs. synthetic Ukraine with post-treatment period of 2013-2017 using 2011 base rate (left) and extended post-treatment period of 2013-2021 using 2017 base rate (right).

The mean post-treatment difference (computed as the average difference between the GDPpc(2017) of the observed and synthetic unit) in the original period of 2013-17 was $-3091.5, whereas for the extended period (2013-21) it is $-4031.38. Clearly, the difference in GDP per capita between Ukraine and the synthetic control unit has increased in this extended period. This suggests that the impact of the 2014 Russian invasion of Ukraine on the country's economy has been longer-lasting than previously estimated. This could be due to several reasons, such as the disruption in the supply chains, uncertainty for businesses in the conflict-affected areas leading to decline in foreign investment, the displacement of people, and reduced tourism in Ukraine.

It is also noteworthy that the slight dip in 2020 in the extended plot might be due to the impact of COVID-19 on all countries' GDPs; however, this has not been taken as a confounding variable since it affected both the treated and synthetic units.

Part 2: Rigorous Criteria for Donor Pool

The original paper mentioned removing those Post-Soviet and Eastern Bloc countries from the donor pool that had some significant shocks during the pre-treatment period. However, the “significant” part was not thoroughly defined. To our understanding, only Georgia, Russia, Uzbekistan, and Turkmenistan were removed from the donor pool, while multiple other countries experienced wars, civil unrest, and other shocks that were likely to have long-term impacts on the economies of those countries.

We hypothesize that removing from the donor pool countries that had 6+ months-long conflicts from 1990 to 2012 can lead to higher estimates of the causal effects of the 2014 Russian invasion of Ukraine. We use a longer conflict consideration window (from 1990) despite the pre-treatment period starting in 1995, as economic impacts of such conflicts tend to last for years. For that purpose, we exclude the following additional countries from the donor pool:

1. Armenia, due to the Nagorno-Karabakh War (Encyclopaedia Britannica, 2023).
2. Azerbaijan, due to the Nagorno-Karabakh War (Encyclopaedia Britannica, 2023).
3. Kazakhstan, due to the Tajikistan Civil War (PERI, n.d.).
4. Kyrgyz Republic, due to the Tajikistan Civil War (PERI, n.d.) and the Kyrgyz Revolution of 2010 (Nichol, 2010).
5. Tajikistan, due to the Tajikistan Civil War (PERI, n.d.) and the Tajikistan Insurgency (Crisis Group Asia, 2011).

The weights used for creating synthetic control for Ukraine using this new donor pool are in Table 1. The average estimated causal effect using this donor pool is -4140$, compared to -4031$ (GDPpc) found using the original donor pool. This supports the idea that the excluded countries might have resulted in an underestimation of the causal effect of interest.

Country	Weight	Country	Weight
Belarus	0	Lithuania	0
Bulgaria	0.196	Poland	0.011
Czech	0	Romania	0
Estonia	0	Slovakia	0
Hungary	0.001	Slovenia	0
Latvia	0.165	Moldova	0.626

Table 1. The weights of the donor pool countries for the synthetic control.

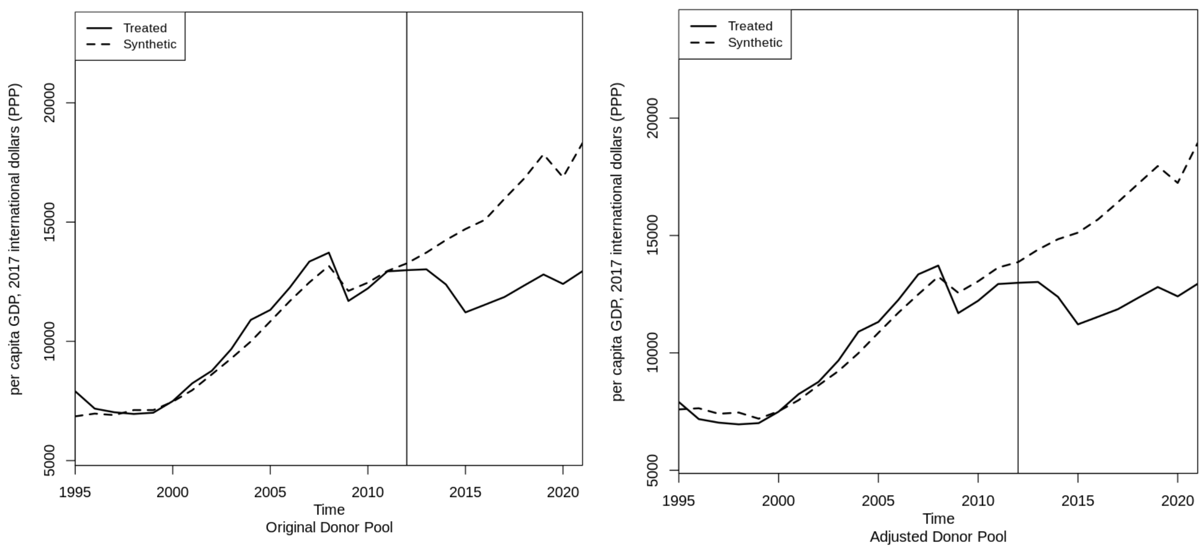


Figure 3. Trends in GDP per capita. Left panel: using the original donor pool as per Bluszcz and Valente (2020). Right panel: excluding 5 additional countries due to their prolonged shocks between 1990 and 2012.

Covariate	Weight	Real	Synthetic
Inflation	0.004	36.44	23.04
GFCF	0.032	20.91	22.89
TradeDep	0.001	0.23	0.19
HDI	0.001	0.70	0.68
Polity	0.001	6.50	6.83
GDPpc(2000)	0.577	4797.38	4797.03
GDPpc(2012)	0.385	8322.17	8538.05

Table 2. Outcome predictor means and weights in Bluszcz and Valente (2020)

Covariate	Weight	Real	Synthetic
Inflation	0.367	36.44	36.855
GFCF	0	20.912	22.33
TradeDep	0	0.23	0.277
HDI	0.007	0.701	0.682
Polity	0.001	6.50	6.953
GDPpc(2000)	0.61	7496.813	7514.168
GDPpc(2012)	0.015	12985.087	13866.29

Table 3. Outcome predictor means and weights (after excluding five countries in the donor pool)

To observe the effect of leaving out a total of five countries from the donor pool on the pre-treatment variables' balance, we compare the "outcome predictor means and weights" tables in Bluszcz & Valente (2020) and our extension.

As we can see in Table 2, the predictor means are closely matched for real and synthetic units, representing a better balance in the pre-treatment period. However, the real and synthetic means for the "Inflation" predictor are still quite off (i.e., real: 36.44 and synthetic: 23.04). On the other hand, the Synth algorithm we run after excluding five countries from the donor pool results in strictly closer real and synthetic means for this predictor (i.e., real: 36.44 and synthetic: 36.85), as observed in Table 2.

In SCM, when the predictor means are closer for the synthetic control unit and the real unit, the constructed synthetic control unit is a good match for the real unit in the pre-treatment period. In other words, the synthetic control unit has similar characteristics to the treated unit before the treatment occurred. This closeness in the pre-treatment period is crucial in the SCM, as it helps establish that the synthetic control unit is a valid counterfactual for the treated unit, allowing for a more reliable estimation of the treatment effect.

When we holistically consider the real and synthetic predictor means in Tables 2 and 3, we observe that our extension provides close values for other predictors as well. Therefore, we infer that this method provides a better balance on the outcome predictors than Bluszcz & Valente's (2020).

Part 3: Achieving Better Balance on Predictors

As discussed in Part 2, the “Inflation” predictor isn’t well-balanced in Bluszcz & Valente (2020). Consequently, when we extend only the post-treatment period in the synthetic control analysis from 2013-2017 to 2013-2021 in Part 1, we still get an imbalance for this variable, potentially leading us to make incorrect conclusions about the causal effect.

To account for this imbalance in the original paper and Part 1, we suggested manually manipulating the weights of predictors by decreasing the weights of the most influential predictors and increasing the weights of less influential ones. We also added a special predictor for the “GDPpc” in 2008 after observing pre-treatment period discrepancies between the trends of the real and synthetic units around the year 2008. These additions provided a better balance on the outcome predictors compared to Bluszcz & Valente (2020), as shown in Tables 2 and 4.

Covariate	**Weight**	**Real**	**Synthetic**
Inflation	0.031	36.44	32.523
GFCF	0.005	20.912	22.89
TradeDep	0.265	0.23	0.23
HDI	0	0.701	0.676
Polity	0.024	6.50	6.239
GDPpc(2000)	0.075	7496.813	7761.04
GDPpc(2012)	0.239	12985.087	13465.43
GDPpc(2018)	0.360	13719.271	13327.168

Table 4. Outcome predictor means and weights (added custom weights and a special predictor)

Referring back to our extension in Part 2, we realize that excluding five countries from the donor pool, which had conflict shocks in the pre-treatment period, naturally provided optimal weights and better balance on predictors without demanding to adjust covariate weights manually or adding special predictors. Since excluding those particular five countries from the donor pool accounts for potential biases and confounds due to pre-treatment period shocks and provides a good balance on the predictors, the causal effect measured between 2013-2021 will likely provide reliable results alternative to Bluszcz & Valente (2020).

Conclusion

In this paper, we brought extension proposals to Bluszcz & Valente (2020) in three parts to provide reliable average treatment effect estimation methods for the economic impact of the Donbas war on Ukraine's GDPpc that account for potential biases and confounds. The first part of our extension focused on updating Bluszcz & Valente's (2020) analysis with newer data; while they measured the average treatment effect for the post-treatment period 2013-17 with the 2011$ base rate, we extended it to 2013-21 with the 2017$ base rate. The average treatment effect for 2013-17 with 2017$ was -3091.5$, and for 2013-21 with 2017$ was -4031.38$. Based on these results, the difference in GDP per capita between Ukraine and the synthetic control unit has increased in the extended post-treatment period. Later, the second part of our extension focused on improving the rigorous criteria for the donor pool by leaving out five donor countries as we identified the conflicts they experienced might lead to biased average treatment estimates. We observed that leaving out these particular countries resulted in a better outcome balance on the "Inflation" predictor; real: 36.44 and synthetic: 23.04 in Bluszcz & Valente (2020), and real: 36.44 and synthetic: 36.85 in our extension. We find the treatment effect -4140 USD$2017, compared to -4031 USD$2017 (GDPpc), which indicates that Bluszcz & Valente's (2020) findings might have underestimated the causal effect due to a less restrictive donor pool. Our last extension focused on improving the predictor balance in Bluszcz & Valente (2020) by manually changing the predictor weights and adding special predictors and pointed out the lack of necessity for such modifications after utilizing our second extension. Finally, we believe the extension methods proposed in this paper can provide reliable average treatment estimate alternatives building on Bluszcz & Valente's (2020) work.

Individual Contributions

This paper was written by Miray Özcan, Yashvardhan Sharma, and Viktor Tsvil. Overall, there was thorough collaboration of all team members on each section of the paper. Introduction was mostly written by Viktor; Replication section was written by Miray; Extension section was written collaboratively by all team members. Additionally, each section was iteratively reviewed and commented on by each team member. The replication part was mostly coded by Yashvardhan, and the extension part resulted from group collaboration.

Use of AI Tools

Grammarly, an AI-based text editing tool was used to edit the paper for errors in grammar, mechanics, punctuation, word use, and spelling. Additionally, ChatGPT+ was used as a paraphrasing tool to make certain sentences more clear and concise for the intended audience.

Word count: 2082 words.

References

Bluszcz, J., & Valente, M. (2020, July 1). The Economic Costs of Hybrid Wars: The Case of Ukraine. *Defence and Peace Economics*. https://doi.org/10.1080/10242694.2020.1791616

Bluszcz, J., & Valente, M. (2020, July 1). *Replication data for: The economic costs of hybrid wars: The case of Ukraine*. Mendeley Data. Retrieved April 11, 2023, from https://data.mendeley.com/datasets/9bwk589xz6/1

Crisis Group Asia. (2011, May 24). *Tajikistan: The Changing Insurgent Threats*. Refworld. Retrieved April 20, 2023, from https://www.refworld.org/pdfid/4dddf5622.pdf

Encyclopaedia Britannica. (2023, April 11). *Nagorno-Karabakh | History, Map, & Conflict | Britannica*. Encyclopedia Britannica. Retrieved April 20, 2023, from https://www.britannica.com/place/Nagorno-Karabakh

Nichol, J. (2010, June 15). *The April 2010 Coup in Kyrgyzstan and its Aftermath: Context and Implications for U.S. Interests*. Congressional Research Service. Retrieved April 20, 2023, from https://sgp.fas.org/crs/row/R41178.pdf

PERI. (n.d.). *Modern Conflicts: Modern Profile*. Political Economy Research Institute - University of Massachusetts Amherst. Retrieved April 20, 2023, from http://peri.umass.edu/fileadmin/pdf/Tajikistan.pdf

World Bank Open Data. (n.d.). World Bank Open Data. World Bank Open Data | Data. Retrieved April 13, 2023, from https://data.worldbank.org/

Appendix A

The code to replicate our findings can be found at the following link:

https://drive.google.com/drive/folders/1MJfuSINd1vk-pb18GFRzgZlrxbg84XHK?usp=sharing