A replication of Jones & Marinescu (2022)

Etienne Bacher
Mario Herrera-Rodriguez
Diego Marino Fages
Felix Stips

Abstract: Jones and Marinescu (2022) study the employment effects of a universal cash transfer in Alaska. Using a synthetic control method, they find that the transfer had no negative effects on employment. We reproduce the results using their replication package and investigate if the results hold when using a different software to run the analysis. We also use different estimation techniques and perform sensitivity checks to assess robustness of the results. We find some differences in the size and significance of the average treatment effects on labor force participation and hours worked when we use a different software (R) and various extensions of the synthetic control method. We also find smaller coefficients on part-time employment when including more covariates. However, these differences do not contradict the main conclusion of the paper.

Keywords: synthetic control; labour market; universal cash transfer; replication.

1. Introduction

This paper replicates the paper by Jones and Marinescu (2022), henceforth JM, who study the effects of a universal cash transfer introduced in Alaska in 1982. The authors explore whether this policy led to a negative employment response. The main dataset is individual-level data from the Current Population Surveys (CPS), provided either by IPUMS or MORG, which the authors aggregate to the state level.

To analyze the effects of this policy, they use the synthetic control method. This method consists in the creation of a synthetic state using a weighted combination of several untreated states. This synthetic state is built so that it matches the trend of the outcome of the treated unit, Alaska, during the pre-treatment period. If the pre-treatment fit is good, then the synthetic state can be considered as a credible counterfactual for Alaska in the post-treatment period, and the treatment effect can be estimated as the difference between the two trends in the post-treatment period. The authors used data from 1977 to 2014, although data for one outcome (the number of hours worked in the previous week) was only available since 1979. The main result of the paper is that the authors do not find a negative employment response to the cash transfer, but an increase in part-time work.

We conduct a series of exercises to study the reproducibility and robustness of the conclusions of JM. First, we check whether the replication package allows us to reproduce the main results of the paper. We also reproduce key parts of the synthetic control estimation in R to check whether the results are sensitive to the choice of software. Next, we repeat the empirical estimation using different covariates. We also vary the post-treatment period to assess how it affects the magnitude of the treatment effect. Moreover, we use two different estimation methods derived from the synthetic control framework to compare their results with those in the paper. Finally, we separate the placebos used in the paper between time and unit placebos.

2. Reproducibility

As a first step, we checked whether the Stata code could run as-is and whether its outputs correspond to the results reported in the paper. We found three typos in the code, which were easy to fix and did not compromise the reproducibility of the paper. Two input datasets required to recreate the main datasets were missing from the replication package. The data was provided by the authors upon request. With these changes, we were able to reproduce the results of the paper.

Table 1: Replication of Table 1 of the original paper using the R package `tidysynth`.

	Alaska	Employment rate	Labor-force participation	Part-time rate	Hours worked last week
IPUMS data					
employed	0.639	0.639			
activelf	0.712		0.707		
parttime	0.103			0.104	
age1	0.108	0.102	0.098	0.099	
age2	0.154	0.135	0.125	0.124	
age3	0.691	0.642	0.672	0.672	
female	0.503	0.51	0.504	0.506	
ind1	0.361	0.361	0.337	0.359	
ind2	0.097	0.099	0.087	0.092	
ind3	0.035	0.063	0.064	0.034	
ind4	0.191	0.185	0.18	0.18	
ind5	0.078	0.119	0.161	0.152	
educ1	0.229	0.242	0.263	0.281	
educ2	0.396	0.388	0.42	0.397	
MORG data					
hourslw	37.98				37.93
age1	0.074				0.078
age2	0.155				0.139
age3	0.759				0.746
female	0.435				0.411
ind1	0.148				0.176
ind2	0.051				0.129
ind3	0.292				0.278
ind4	0.123				0.134
ind5	0.385				0.283
educ1	0.11				0.191
educ2	0.387				0.384

Note: Column 2 shows the average value of the covariates in the pre-treatment period for Alaska (treated unit). Columns 3-6 show these values for the synthetic control.

Next, we tried to reproduce the main findings using R (R Core Team, 2022). Two packages were used for this: `Synth` (v. 1.1-6, Abadie et al., 2011) and `tidysynth` (v. 0.2.0, Dunford, 2023). The objective was to reproduce the results from the balance checks in Table 1 and the main results in Table 2 of the original paper. Table 1 shows the summary statistics for Alaska and the synthetic control obtained with R. The first panel gives the averages for Alaska, which are identical to those of the original paper. The next rows show the summary statistics for the placebo states used for each of the four outcomes. We can observe minor differences in the averages for employment, labor force participation, and part-time employment as compared to the averages in the paper. These differences are due to the fact that the estimation in R assigns slightly different weights. However, the differences are rather minimal, occurring at the second or third decimal place. The differences for the hours worked last week outcome are slightly larger, indicating larger differences in the matching for that outcome.

Table 2 shows the results from Table 2 of the original paper and, below, from our replication using R. For the first two outcomes, employment rate and part-time rate, we find similar results: the coefficients are very close, the statistical significance is the same as in the original paper, and the pre-period RMSE are almost identical. However, we find different results for the effect of the unconditional transfer on the participation in the labor force and on the number of hours worked in the previous week. For labor force participation, we find an average post-treatment effect twice larger than JM (0.027 in our replication against 0.012), and most importantly we find that this effect is statistically significant at the 10% level. On the other hand, for the number of hours worked in the previous week, we find a smaller and statistically insignificant effect (-0.442 against -0.796). The results for `tidysynth` and `Synth` are very similar.

Table 2: Replication of Table 2 of the original paper using the R package `tidysynth`.

	Employment rate	Part-time rate	Labor-force participation	Hours worked last week
Original				
Average effect	0.001	0.018	0.012	-0.796
p-value	0.942	0.020	0.331	0.084
RMSE	0.005	0.003	0.013	0.394
Replication with R				
Average effect	0.004	0.017	0.027	-0.442
p-value	0.593	0.008	0.052	0.142
RMSE	0.004	0.001	0.009	0.058

The difference in p-values between Stata and R probably results from differences in the implementation of the placebo computations. `Tidysynth` computes the weights of each variable only once and then applies these weights to all synthetic controls with placebo states, while the Stata code runs the synthetic control separately for each placebo state. The different permutations may explain the difference in p-values, but they cannot explain the differences in coefficients. These latter differences are most likely due to the fact that R and Stata use slightly different control groups, as indicated also by the differences in Table 1.

3. Replication

We tested the robustness of the main findings (Table 2), by performing several robustness checks:

1. Including other covariates trying six easy-to-implement specifications (Table 3):

Panel B replicates the results from Table 2 of the original paper. This controls for education composition, age structure, share of female, and industry composition of the workforce (for definitions see p. 323 of the original paper). In Panel A, we remove industry controls from the set of covariates. In the remaining panels, we subsequently add other covariates that were mentioned in the paper. These include GDP per capita in Panel C, oil revenues as a share of GDP in Panel D, net migration rates in Panel E, and government expenditure per capita in Panel F. Details on the exact definition of the variables can be found in the original paper, which we follow exactly by using the original variables from the replication package.

For the employment rate in column (I), varying the covariates results in a slightly larger positive employment coefficient. The specification from the paper (Panel B) has the best pre-period fit and we did not include p-values for our analysis. Importantly, a slightly positive employment coefficient would not contradict the main finding of the paper that the employment response to cash transfer is not negative. Thus, the finding holds in this analysis. In column (II) we look at the part-time employment rate and find that adding more covariates decreases the size of the coefficient slightly without deteriorating the pre-period fit. For example, when including GDP per capita, oil revenues as share of GDP and net migration as matching variables in Panel E, we find an average effect of 0.011, that is, one third smaller of that in the paper, while achieving a smaller root mean squared error in the pre-treatment period. However, without p-values, we cannot conclude whether the coefficient would still have been significant. In column (III) we look at labor-force participation and find that the coefficient remains virtually unchanged when including further covariates. The results for hours worked in column (IV) are also robust to changing the covariates. The coefficient remains in the same ballpark but the pre-treatment fit diminish even more.

Taken together, the results from Table 3 indicate that the main results of the paper are quite robust to a change of covariates. As a caveat to this statement, it should be noted that it was not possible to collect additional covariates from the original data within the time frame of this project and the analysis is thus restricted to covariates mentioned in the paper.

2. Varying the post-treatment sample period length:

Perhaps the main weakness of the paper is the use of a short pre-treatment and long post-treatment period, which may lead to spurious pre-treatment fit (Abadie, 2021). While this issue is of conceptual nature and not within the realm of this replication, it does highlight another researcher degree of freedom: when to stop the post-treatment period. The authors used the maximum number of years available. Here, we assess if the results are robust to the use of a shorter post-treatment period. To test this, we re-estimated the main specification from Table 2 of the original paper varying the number of years post-treatment included in the estimation sample. The results of this exercise are provided in the Appendix in Figures 1 - 4. The Y-Axis gives the average effect (equivalent to $\hat{\alpha}_1$ in the original table) and the X-Axis gives the number of years after 1982 which are included in the sample. Again, we omitted inference for time limitations.

Figure 1 provides the results for employment and shows that the effect decreases strongly with time up until ten years. While we don't know whether the coefficient would be significant, the short-run coefficient of 0.03 seems rather large. The mirror image is shown in Figure 3 for part-time employment, which initially is just below zero, but starts to increase over time. The increase here is stronger in the first

Table 3: Replication of Table 2 of the original paper using different covariates

	(I) Employment rate	(II) Part-time rate	(III) Labor-force participation	(IV) Hours worked last week
Panel A: age, education, and gender groups				
Average effect	0.015	0.020	0.023	-0.866
RMSE	0.009	0.003	0.011	0.656
Panel B: age, education, gender, and industry groups (original)				
Average effect	0.001	0.018	0.012	-0.796
RMSE	0.005	0.003	0.013	0.394
Panel C: age, education, gender, industry groups and gdp pc				
Average effect	0.025	0.009	0.018	-0.824
RMSE	0.006	0.003	0.014	0.538
Panel D: age, education, gender, industry groups, gdp pc and oil gdp				
Average effect	0.025	0.010	0.014	-0.899
RMSE	0.008	0.003	0.014	0.613
Panel E: age, education, gender, industry groups, gdp pc, oil gdp, and net migration				
Average effect	0.026	0.011	0.014	-0.866
RMSE	0.008	0.002	0.014	0.638
Panel F: age, education, gender, industry groups, gdp pc, oil gdp, net migration, and govt. expenditure				
Average effect	0.023	0.003	0.015	-0.863
RMSE	0.008	0.006	0.014	0.631

Notes: The table replicates the results of Table 2 of Jones and Marinescu (2022). P-values were not calculated.

post-treatment years and flattens later. While these differences may or may not be statistically significant, they are, in any case, not qualitatively significant since they do not alter the paper's main result of no negative employment response. If anything, a shorter post-treatment period would have lead the authors to find stronger demand effects for these outcomes. The results for labor force activity in Figure 2 have a two-humped shape with maximums around 5 and 15 years after the treatment period length. They remain in the same range and are nowhere near a negative coefficient. Hours worked during the last week is the only outcome where the length seems to matter for the sign of the coefficient. The curve in Figure 4 is initially positive and crosses the origin only if we include at least ten years post-treatment.

In sum, the choice of post-treatment length matters qualitatively for the results on hours worked (though we don't know if statistically), quantitatively for employment and part-time employment (though not qualitatively in changing the message of the paper), and not at all for labor force activity.

3. Using augmented synthetic control and synthetic DiD estimation (Table 4):

To check the robustness of the results to the estimation procedure used, we implement two recent extensions of the canonical synthetic control method. First, following Ben-Michael et al. (2021), we estimate an "augmented" synthetic control model that corrects the treatment effects for inexact pre-treatment fit. Additionally, we use the method proposed by Arkhangelsky et al. (2021), which performs a synthetic differences-in-differences estimation. In essence, this introduces time and unit fixed effects as well as unit and time weights in the spirit of the synthetic control framework.

Table 4 shows the replication of the treatment effects of the paper, as well as the parameters using the two alternative methods. We omitted placebo estimation for the synthetic control methods due to time constraints. The bias-corrected synthetic control leads to effects closer to zero for the part-time rate. On the other hand, it leads to larger coefficients for employment rate, labor force participation, and the number of hours worked. The results from the synthetic differences-in-differences are also quite different to the replication results. The effect on labor force participation turns slightly negative and is significant at the 10 percent confidence level, and the effect of hours worked decreases in size and loses its significance.

Overall, the results from this exercise exhibit significant variability. The differences are strongest in the case of the synthetic differences-in-differences estimator. However, while showing a surprising amount of variability, the results do not contradict the main message of the paper as the employment effect remains null and even the effect on labor force participation is rather small.

Table 4: Replication of Table 2 of the original paper using other estimation methods

	Employment rate	Part time rate	Labor force part.	Hours worked
ATE (Replication)	0.001	0.018	0.012	-0.796
ATE (Bias corrected)	0.015	0	0.018	-1.331
ATE (Synth. DiD)	0.002	0.002	-0.008	-0.08
SE (Synth. DiD)	0.004	0.003	0.004	0.07

Notes: The table replicates the results of Table 2 of Jones and Marinescu (2022). It adds two other econometric methods: Line 2 show the estimates using the bias correction proposed by Ben-Michael et al. (2021). Lines 5 and 6 implement the synthetic control differences-in-differences estimator described by Arkhangelsky et al. (2021).

4. Using either time or state placebos:

Panel B of Figures 2 and 3 in the original article builds confidence bands around the synthetic control estimate using all possible combinations of years and states as placebos showing wide and dense confidence intervals. We reproduce these figures in the Appendix in Figures 5 - 8.

Although having all possible combinations of year and state placebos can help in computing the p-values, it might obscure a clear visual comparison if we want to understand the rank relative to specific kinds of placebos. Therefore, we replicated the results in Stata distinguishing between state and year placebos. Figures 9 - 12 show the results using states and Figures 13 - 16. The results are similar with the exception of a short lived but positive effect on employment.

4. Conclusion

We explore the reproducibility and replicability of the results of Jones and Marinescu (2022). Regarding reproducibility, the outputs of their Stata code match the results of their paper. We could also reproduce some of their results using R, but found some small differences in effects. The positive effect on participation in the labor force became marginally significant and the negative effect on hours worked became insignificant. This difference highlights the importance of cross-software replicability.

We also assess the robustness of the findings to changes in covariates, post-treatment period length and estimator choice. Using more covariates used to fit the synthetic control group reduces the effect size on part-time employment. Changing the post-treatment period can lead to positive employment effects, smaller effects on part-time employment, and non-negative effects on hours worked. However, we did not assess the statistical significance of these differences. Using the augmented synthetic control method by Ben-Michael et al. (2021) we observe smaller effects on part-time employment. Using the synthetic differences-in-differences estimator by Arkhangelsky et al. (2021) also gives a very small negative effect on labor force participation, but turns the effect on hours worked insignificant. Lastly, using only state or year placebos may show a positive employment effect in the initial years after the cash transfer came into place.

Overall, we consider the replication to be successful. The main result of no negative employment response holds throughout all replication attempts. Although the results for other outcomes are less robust, we did not find a consistently different set of results that would lead us to reject the claims of the paper. The excellent replication package and clear documentation helped our work considerably.

References

Abadie, A. (2021). Using synthetic controls: Feasibility, data requirements, and methodological aspects. *Journal of Economic Literature*, *59*(2), 391–425.

Abadie, A., Diamond, A., & Hainmueller, J. (2011). Synth: An R package for synthetic control methods in comparative case studies. *Journal of Statistical Software*, *42*(13), 1–17. https://www.jstatsoft.org/v42/i13/

Abadie, A., & L'Hour, J. (2021). A penalized synthetic control estimator for disaggregated data. *Journal of the American Statistical Association*, *116*(536), 1817–1834. https://doi.org/10.1080/01621459.2021.1971535

Arkhangelsky, D., Athey, S., Hirshberg, D. A., Imbens, G. W., & Wager, S. (2021). Synthetic difference-in-differences. *American Economic Review*, *111*(12), 4088–4118. https://doi.org/10.1257/aer.20190159

Ben-Michael, E., Feller, A., & Rothstein, J. (2021). *The augmented synthetic control method* (Working Paper No. 28885). National Bureau of Economic Research. https://doi.org/10.3386/w28885

Dunford, E. (2023). *Tidysynth: A tidy implementation of the synthetic control method* [R package version 0.2.0].

Flood, S., King, M., Ruggles, S., & Warren, J. R. (2015). Integrated public use microdata series, current population survey: Version 4.0.[machine-readable database]. *Minneapolis: University of Minnesota*, *1*.

Jones, D., & Marinescu, I. (2022). The labor market impacts of universal and permanent cash transfers: Evidence from the alaska permanent fund. *American Economic Journal: Economic Policy*, *14*(2), 315–40. https://doi.org/10.1257/pol.20190299

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/

Appendix

Figure 1. ***Treatment period length - Employment***

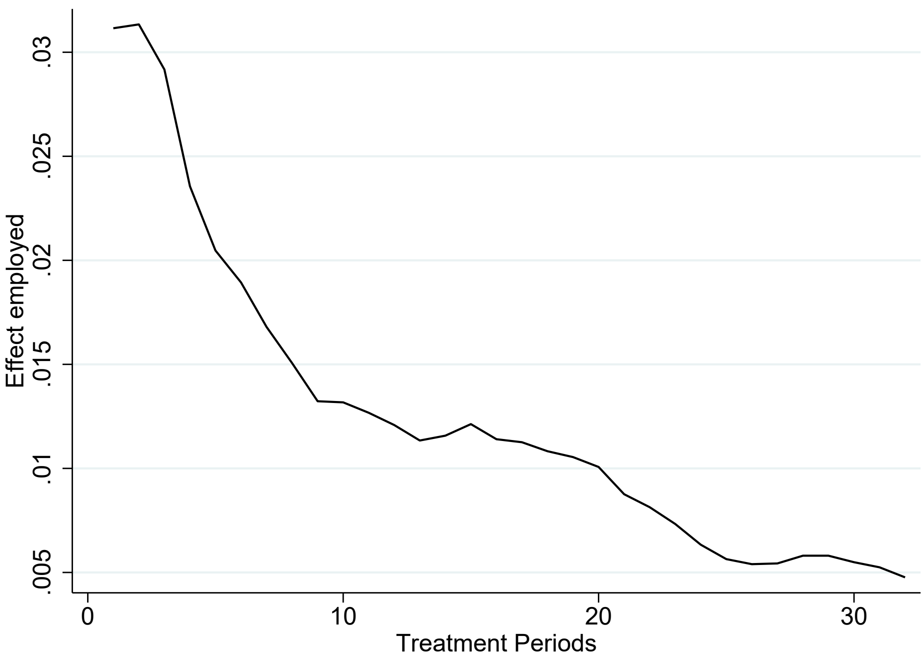

Figure 2. ***Treatment period length - Labor force activity***

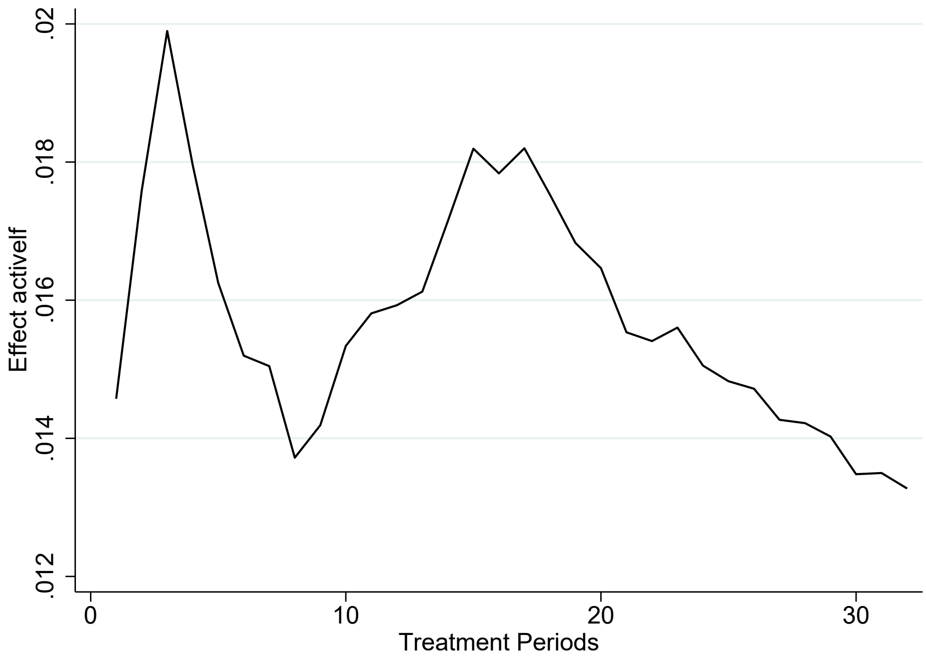

Figure 3. ***Treatment period length - Part-time employment***

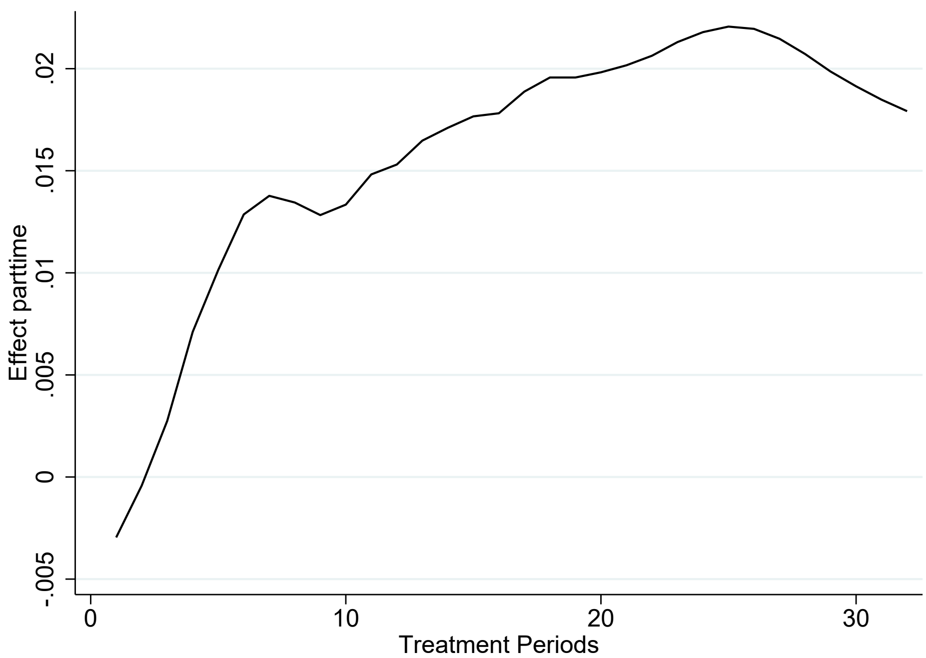

Figure 4. ***Treatment period length - Hours worked***

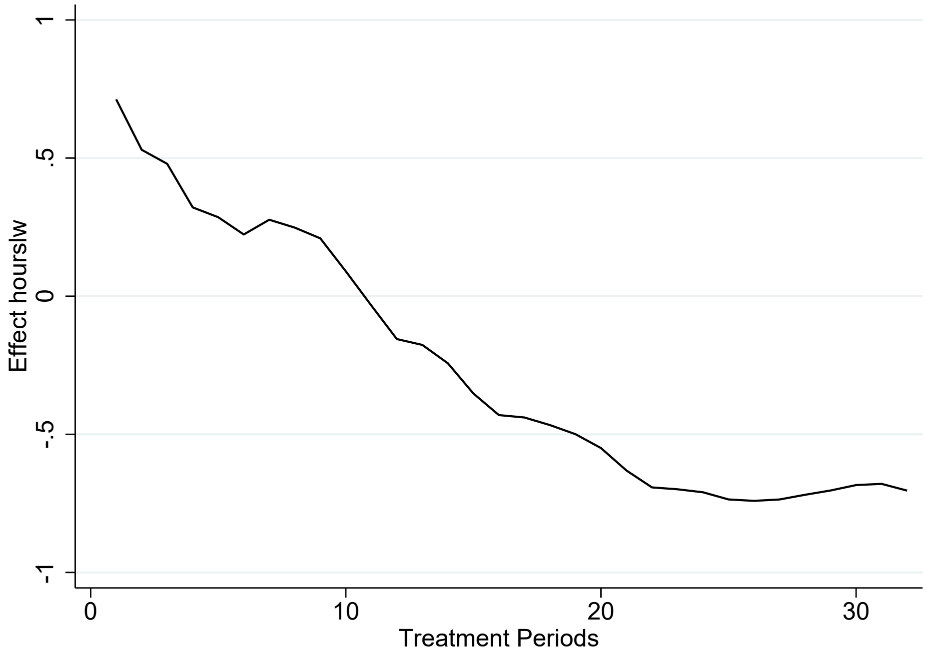

Figure 5. ***Placebo states and years - Employment***

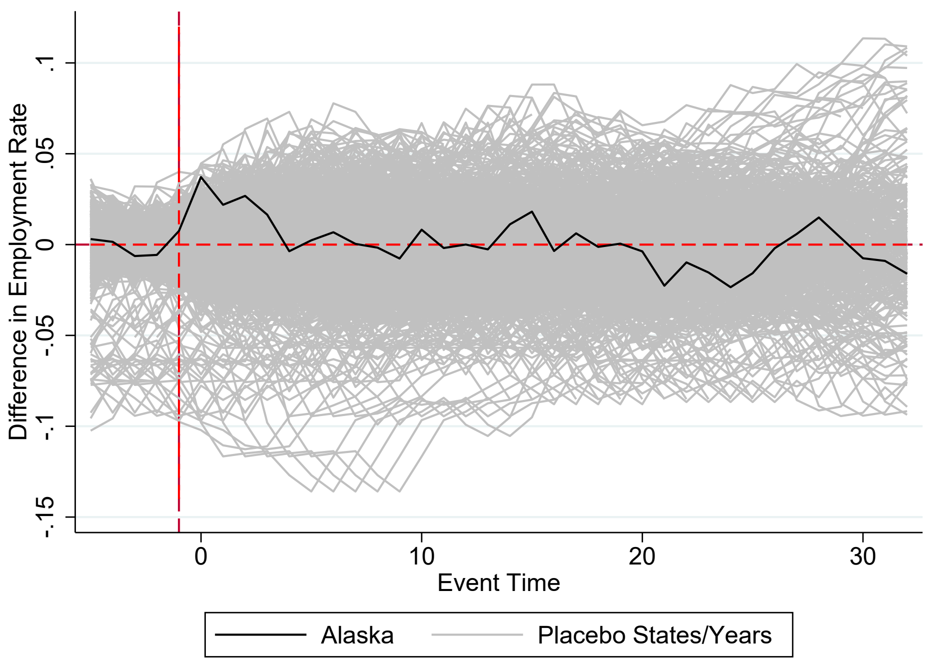

Figure 6. ***Placebo states and years - Labor force activity***

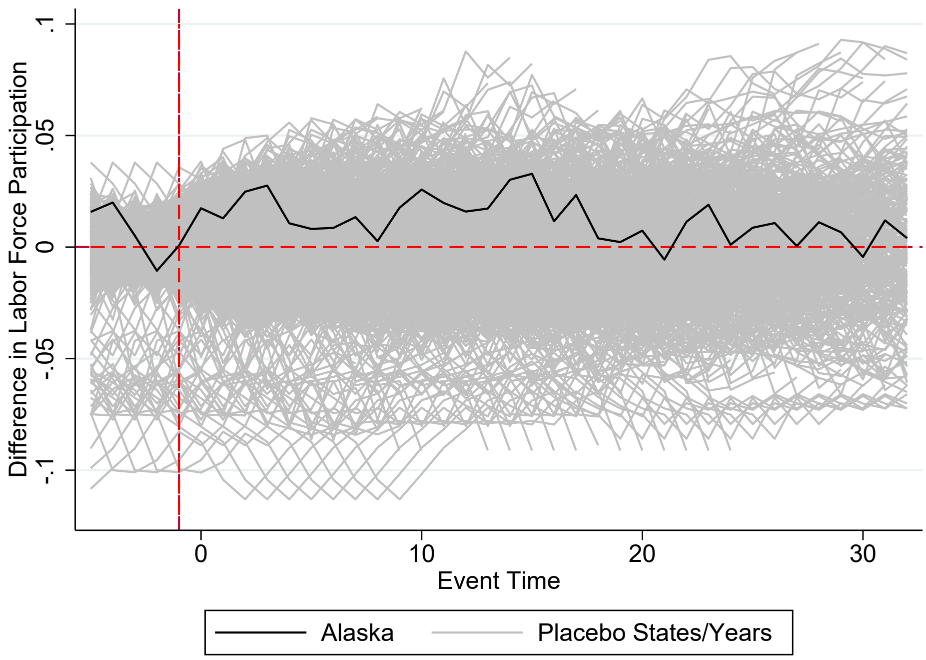

Figure 7. ***Placebo states and years - Part-time employment***

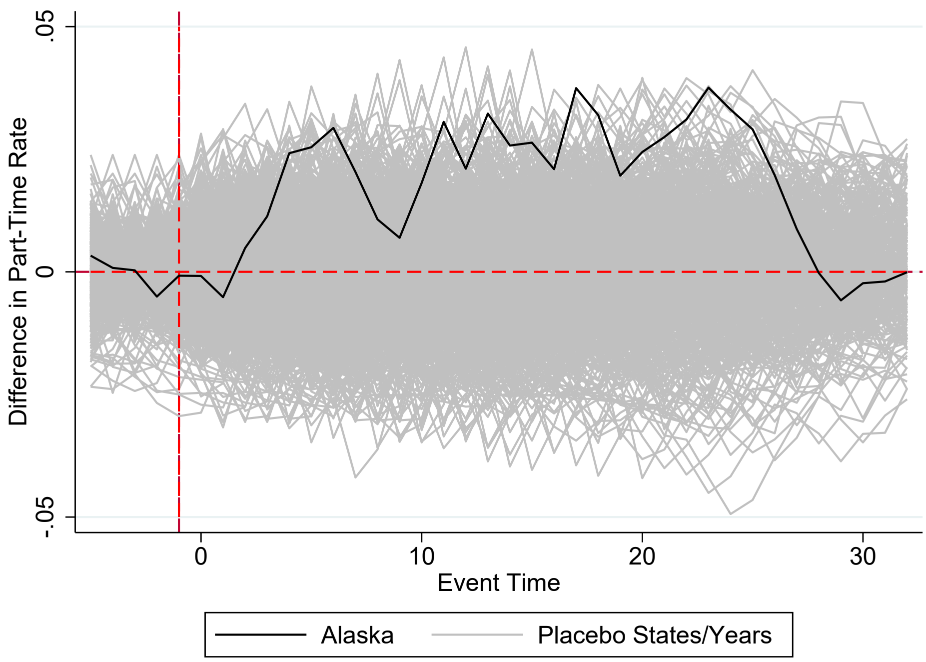

Figure 8. ***Placebo states and years - Hours worked***

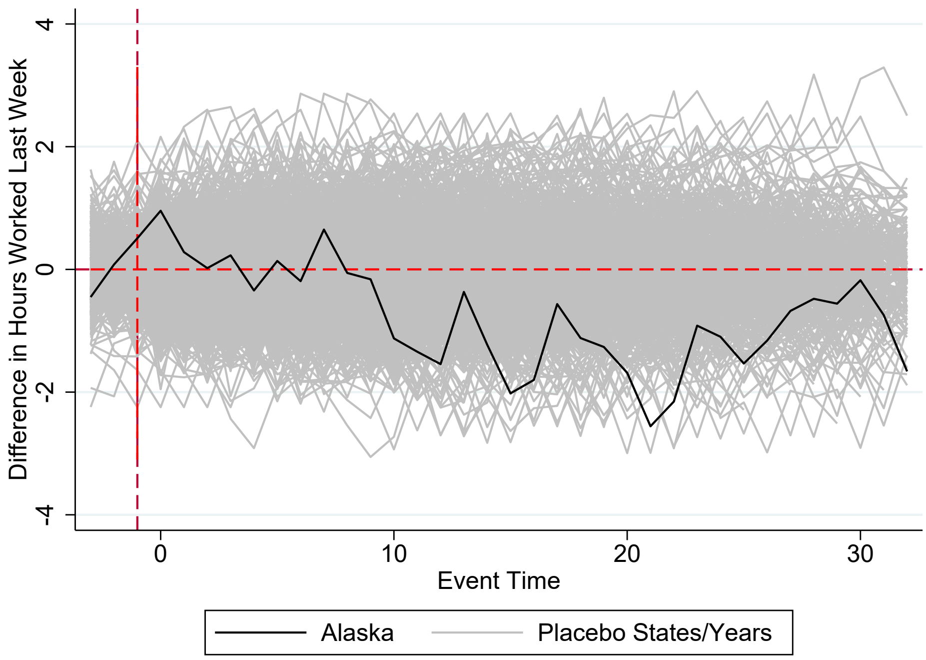

Figure 9. ***Placebo states - Employment***

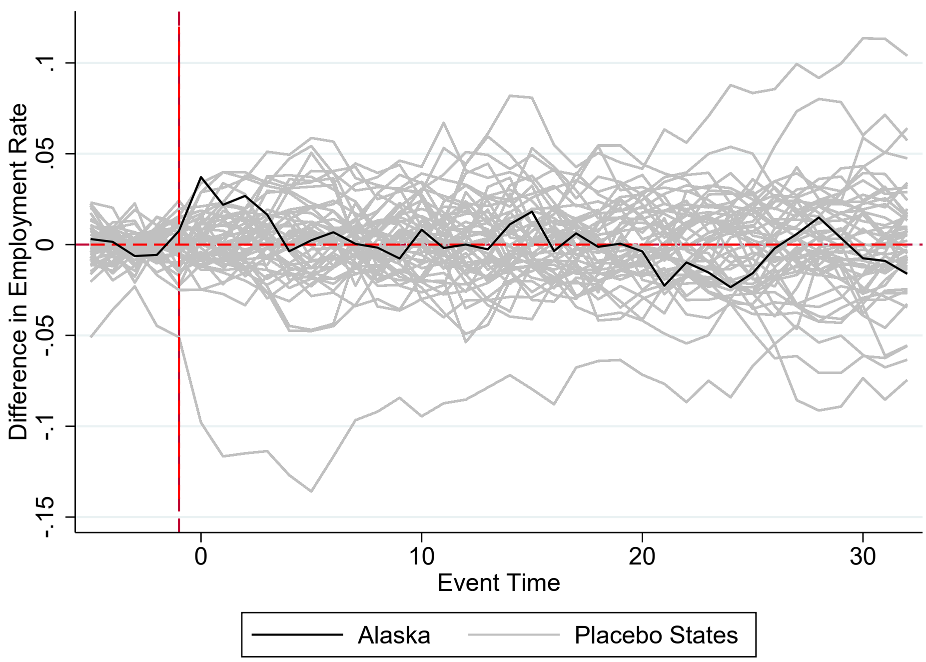

Figure 10. ***Placebo states - Part-time employment***

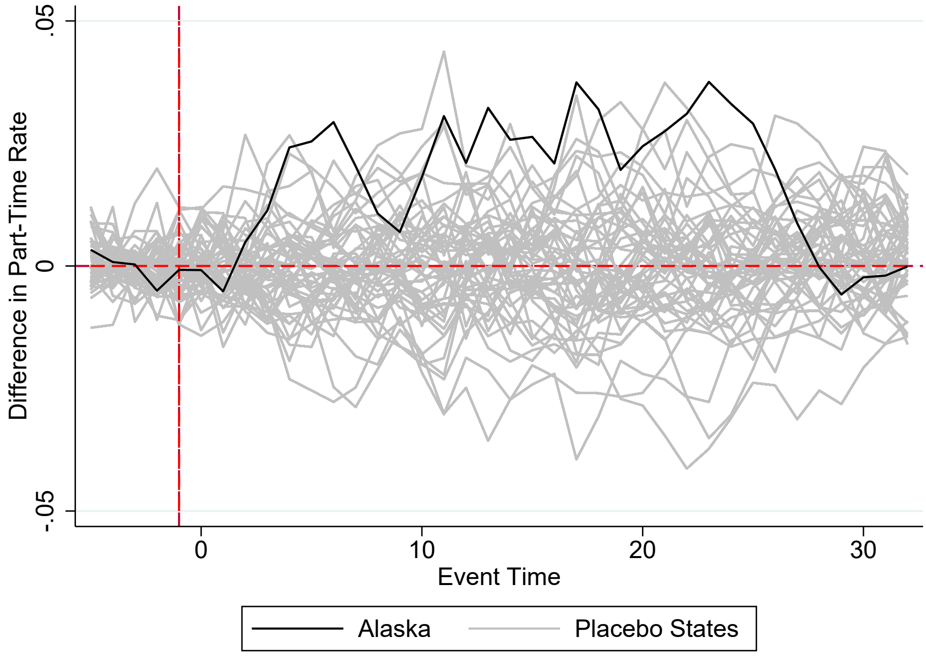

Figure 11. ***Placebo states - Labor force activity***

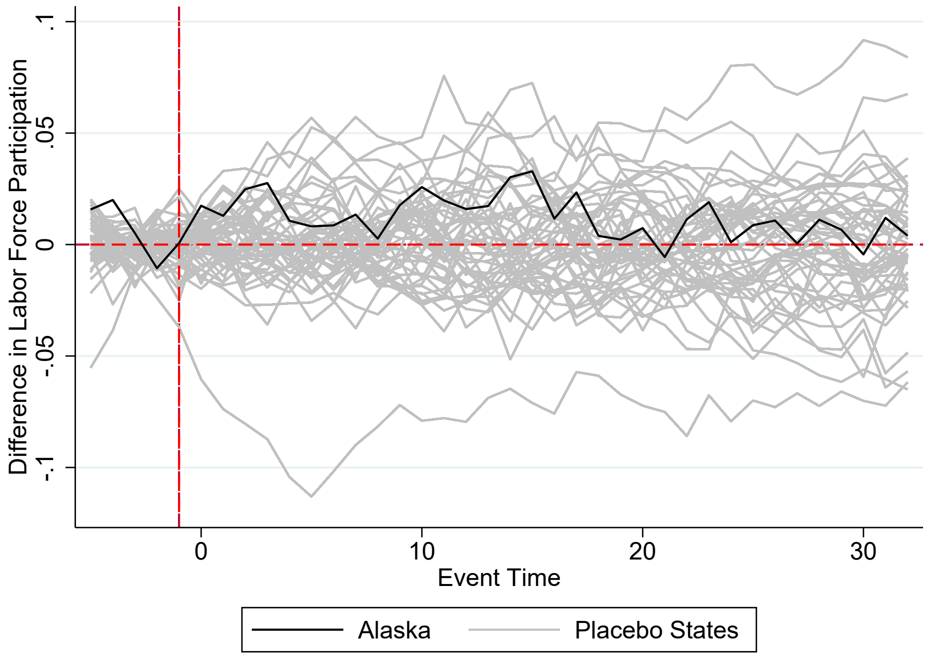

Figure 12. ***Placebo states - Hours worked***

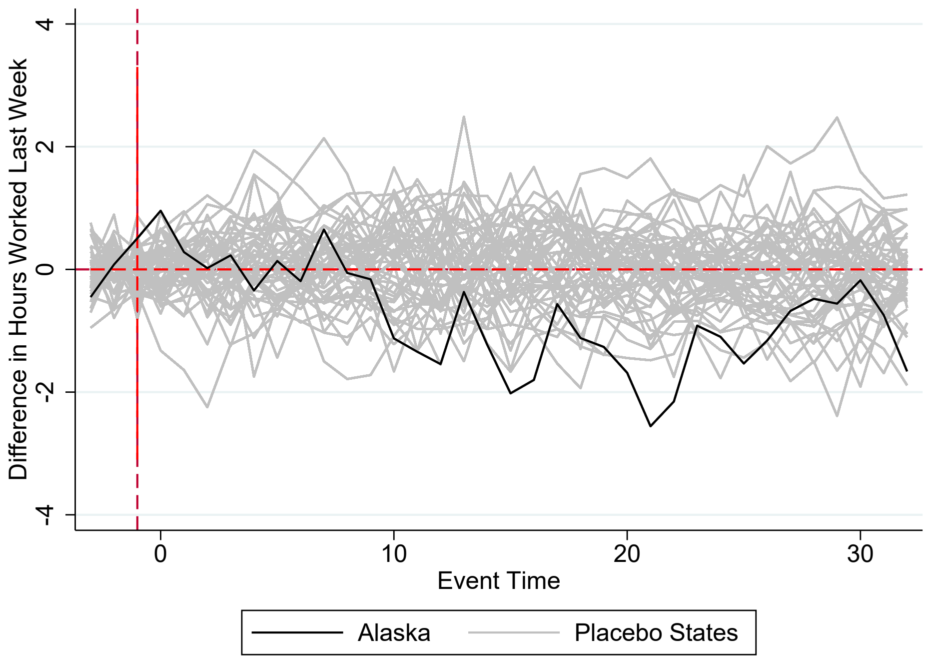

Figure 13. ***Placebo years - Employment***

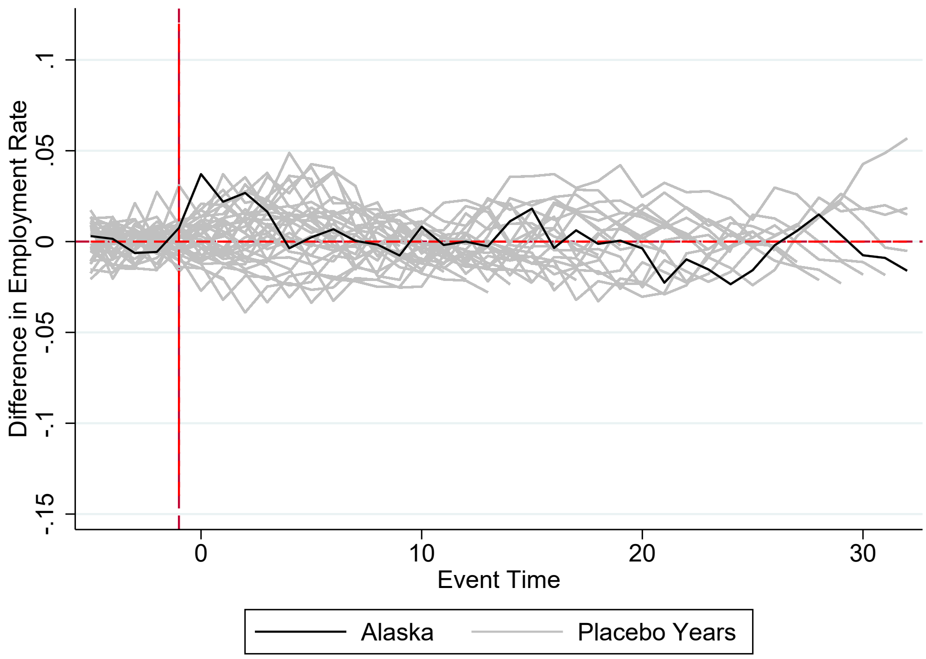

Figure 14. ***Placebo years - Labor force activity***

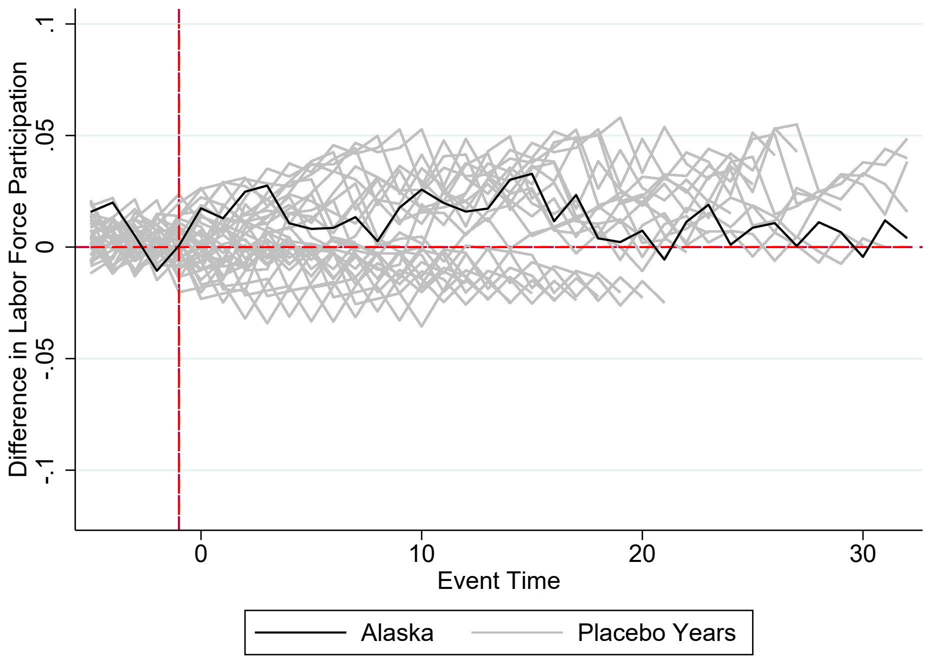

Figure 15. ***Placebo years - Part-time employment***

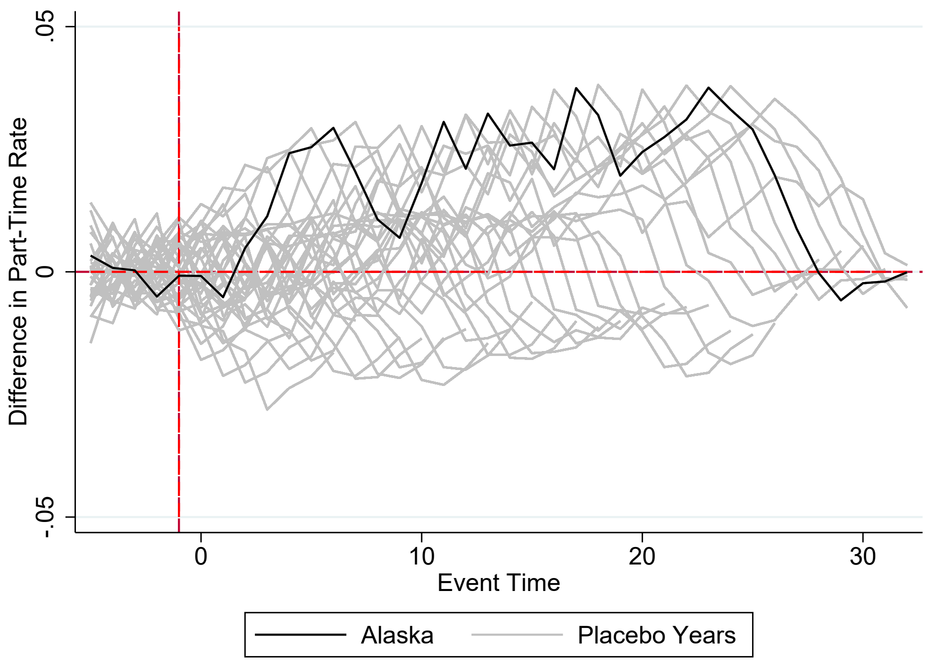

Figure 16. ***Placebo years - Hours worked***