To: Daniel de Kadt and Stephen B. Wittels

From: Alma Gashi, Anosha Rahim and Haitham Alhad Hyder

Date: December 20, 2019

Re: Synthetic controls by Google Implementation contribute to the Causal Inference for "Democratization and Economic Output in Sub-Saharan Africa"

Executive Summary

The original paper performs synthetic control to conclude that democratization had a positive treatment effect on the economic output of Mali. We replicate the results using two methods: synthetic control method ('Synth' Package), and the CausalImpact package by Google. The analysis is further extended by performing placebo tests in-space and in-time to test the reliability of observed treatment effect. Both packages produce similar results, but the CausalImpact package is better at achieving a closer pre-treatment match. In addition, while the placebo in-time did not show an effect, the in-space placebo tests show negative effect in all weighted control units (except Burkina Faso). This could be due to shocks to the units that affected their GDP, such as civil wars reported in several Sub-Saharan nations between 1975-2008, spillover effects arising from trade or diplomatic relations between the neighboring countries, or another underlying cause that affected the GDP of these countries.[1]

Overall, we recommend using the CausalImpact package because it achieves a much better pre-treatment match than synthetic control, and unlike synthetic control, it does not assume that control units are independent of each other. It is highly recommended that countries in the control pool must be selected with care. In particular, avoid countries with potential spillover effects such as allies, countries that have experienced shocks such as wars, or countries experiencing an altogether effect on GDP from another treatment.[2]

Introduction

This memo uses two effective methods currently present for assessing the causal impact on a case study to assess Mali's democratization and economic output between 1975-2008. The causal question is as follows:

What is the effect of democratization that occurred in 1991 on Mali's economy (GDP)?

We use 19 neighboring Sub-Saharan countries that did not experience democratization to form a synthetic control that approximates the counterfactual for Mali. By doing so, we measure the causal

---

[1] #audience: An appropriate level of jargon is used to address the authors, and concepts such as SUTVA, causal inference, and placebo in-space are contextualized to Mali. By tailoring the language and the content to their interests, we have ensured that they will be engaged while reading the decision memo.

[2] #composition: Instead of making the paper too technical by using complex jargon and dry details, we have used a more engaging diction that makes it easy for people to follow the concepts. This is particularly done by using similar terms and repeating them, which may be necessary to retain the attention of the authors of this paper, who may be particularly busy people.

impact of the 1991 events. Similarly, we use a different package that approximates Mali's counterfactual using a Bayesian structural time-series model coined by *Google*. We finally compare the results we get from both methods to see if there is a true causal impact in democratizing a country. The key finding from the paper is that democratization increased Mali's economic output after 1991, but our findings in the weighted-controls suggest that the case may be different.

Methods

Data pre-processing

The synth function fails to run when the values of a certain column all have the same value in all the controls. Therefore, to solve this error, we had to remove the following predictors, *"imfadj", "imfadj_lag1", "imfadj_lag2"*, that were present in the paper's implementation but not in our's to allow for better placebo tests.

In our implementation of the CausalImpact function, we have to pass in predictors that do not contain any NA value. Therefore, we filter for columns that do not contain any NA value before, passing the data to the CausalImpact function.

Moreover, we also shifted the time range from 1975 to 1980 for all the graphs, to avoid a spike (an extreme outlier) in Mali's data that occurs in 1978.

The Synthetic Control Method

The 'Synth' package in R utilizes weighted units of neighboring countries to synthesize a control unit for Mali. In this case, 19 autocratic countries near Mali are used to create a sufficiently matched pre-treatment plot. This is done by using matching to achieve a balance on the covariates of Mali and the synthetic control for the time period before 1991 i.e. before democratization. This implies that if Mali's economic output is the same as the economic output of the control unit before treatment (democratization), we can be fairly confident that the difference between Mali and the control unit's economic output is reflective of the impact of democratization only.

In order to further test the reliability of the results, we conduct placebo tests in-space and in-time. This means that we trick the code into thinking that treatment occurred either in a different (autocratic) country or a different year. In placebo tests, we expect to see no effect, and if we happen to observe any effect, it is an indication of the unreliability of our results, which indicates that the treatment was not the one creating the observed effect.

Causal-Impact Method

The synthetic control method creates a counterfactual for the treatment, and finds the difference between the two; this has a couple of disadvantages, First, it assumes that units do not affect each other, which is not the case in our data (Bertrand, Dufflo, Mullainathan, 2002). Moreover, it means that the difference between the treat and control is constant which is rarely true in time-series data.

In order to measure the quality of a synthetic control we check how close it approximates the treated unit's pre-treatment characteristics. The 'Synth' package provides weights for control units lying between 0 and 1 (of which their sum equals 1). However, the new method we are suggesting doesn't look at other characteristics, rather at how well the synthetic control approximates the history of our outcome variable (log of GDP) before treatment.

To solve these problems we need to use a time-series model, such as the CausalImpact package (A Bayesian Structural Time-series model). This allowed us to create a more flexible model that takes into account local trends and seasonal differences in economic output in the pre-treatment period for Mali.

Results

Replication of Mali Treatment Effect

There is a strong indication that democratization had a positive impact on the economic output of Mali after political reforms began in 1991, as seen in Figure 1 below. Both methods have created a synthetic control unit that matches on covariates before 1991. This is a crucial factor in determining whether there is a treatment effect because the synthetic control must be as similar in the pre-treatment period to Mali as possible. As a result, we can be fairly confident that economic output of Mali increased after democratization began.

While both methods show sufficient matching in the pre-treatment period for Mali, Figure 2 seems to have a better and closer balance than Figure 1. It seems that while the synthetic control method is sufficiently effective, Bayesian structural time series method is slightly better in obtaining this matched balance.

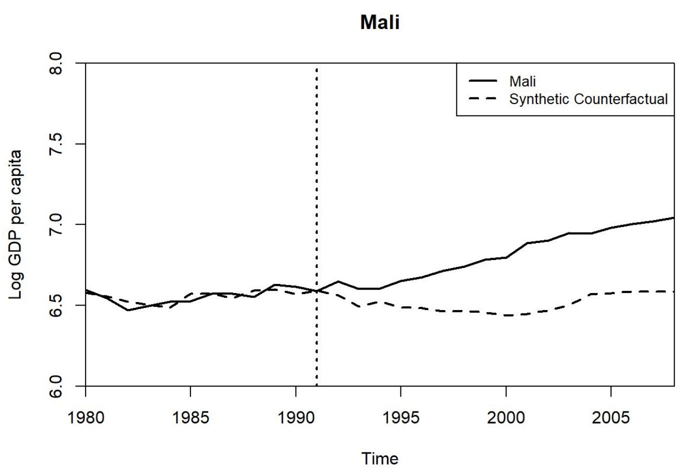

Figure 1. Replication of Mali from the original paper.

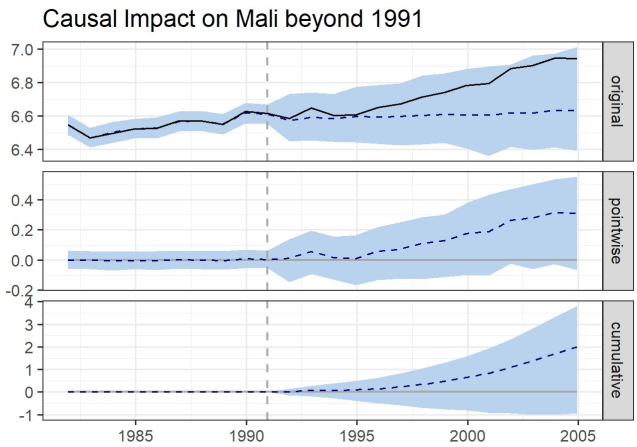

Figure 2. Replication of Mali using a new package (CausalImpact).[3]

Placebo extensions

When changing the treatment for Mali, from 1991 to 1986, there is no observed treatment effect until 1990 - when the actual pre-treatment period ended (Fig. 3). From this, we can conclude that the treatment effect is only apparent when treatment happened, which contributes to causal inference; in other words, it leads us to think the effect only appears when there is a treatment.

Furthermore, strengthening the causal inference, the in-space placebo shows the lack of treatment effect in Burkina Faso, which is one of the weighted control countries, with 16% weight (Fig. 3). This suggests that, although Burkina Faso resembles Mali, there is a treatment effect in Mali but not in Burkina Faso. If there are no confounders, then Mali must have experienced some treatment that *caused* this effect, which Burkina Faso did not experience - in this case, democratization. We also considered higher-weighted controls such as Togo, Burundi, Chad, and Nigeria (**see Appendix D**); they all show negative effect which means that we should doubt the treatment effect on Mali discussed earlier, as well. The reason why we see negative effects in these countries could be either from violation of *SUTVA* (Stable Unit Treatment Value Assumption) or from a confounded treatment that affected all the weighted-control units. In that case, the choice of weighted-controls has room for improvement. Instead of these control units, we suggest to choose as control units, countries that:

1. do not have spillover effects,
2. are not experiencing shocks in the outcome variable (GDP) during the time that Mali is experiencing the treatment,
3. are not experiencing an altogether effect in GDP from another common treatment.

[3] #studyreplication: Replicated the Figure 2 from the paper with two approaches (Synth Package and CausalImpact) to increase the reliability of the initial study. This is partial replication, as we choose to replicate only one element of the whole study.

This finding indicates that the original paper rejects the null hypothesis (that there was no treatment effect in Mali), without taking into consideration the negative effects exposed in in-space placebo[4]. Instead, we suggest that we fail to reject the null hypothesis, after the negative effects in other countries during the treatment time in Mali.[5]

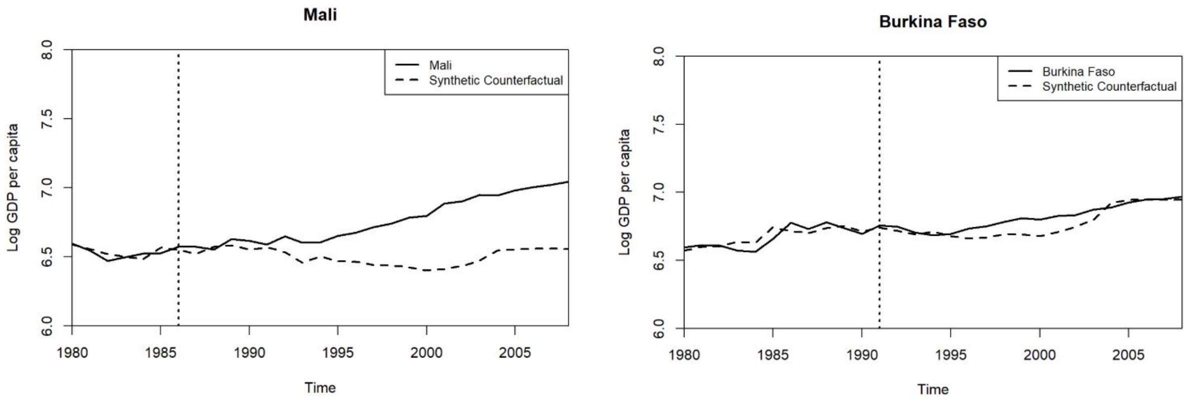

Figure 3. Mali's in-time placebo (treatment shift from 1991 to 1986) on the left, and in-space placebo (Burkina Faso switched with Mali) on the right.

Google Extension

Similar results are seen from Figure 4 from Causal Impact Library placebos. The addition we have in the cumulative plots shows the amount of treatment experienced in total in the post-period as compared to the predicted counterfactual (control). Mali has a positive cumulative effect while Burkina Faso has almost zero of such effect. The other countries that were used to create Mali's synthetic control show a negative effect in the log gdp (gross domestic per capita) (**see Appendix D**). This extension helps us confirm the conclusions we have from the 'Synth' package plots.

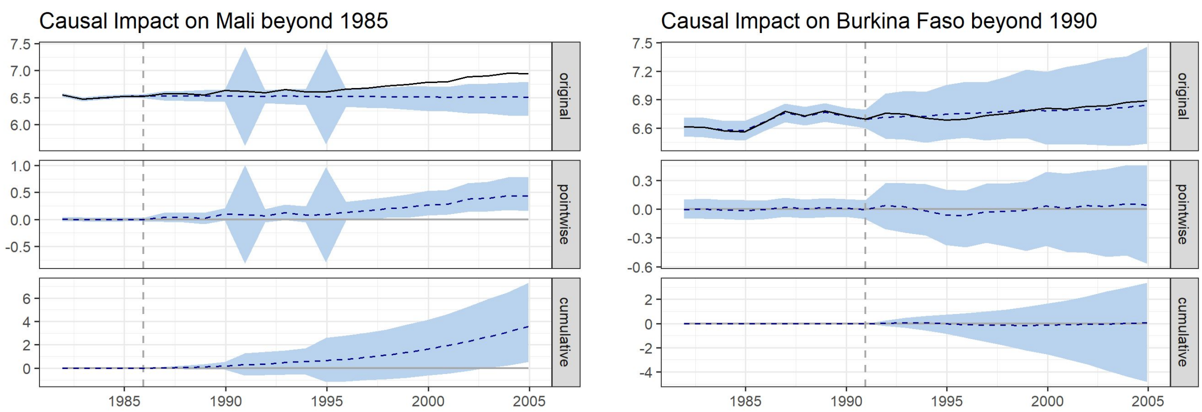

Figure 4. An in-time (left) and in-space (right) placebo test for Mali using CausalImpact Package.

[4] #controlgroups: Suggested improvements for the control group by explaining how the existent one weakens the causal inference from the Synthetic Control results, due to negative effects exposed in placebos.
[5] #induction: Throughout the paper we have questioned the causal inference (which is a type of induction) drawn by the original paper, adding our arguments about how the causal inference is not as robust as suggested.

Conclusion and Recommendations for Future Analysis

In conclusion, while both the synthetic control and causal impact method shows similar results, pre-treatment matching is far better in causal impact method. Moreover, we find that besides Burkina Faso, all other weighted control countries see a negative effect in placebo tests (**see Appendix D**), indicating that their GDP might have been affected by either the treatment through spillover effect or some other underlying cause. Considering these findings, it is recommended that countries should be carefully selected to be in the control pool such that their GDP is not impacted by any other major factors such as wars and relations with other countries. Moreover, we highly recommend using the causal impact method for this analysis, because it shows an almost perfect pretreatment match (Figure 2 and 5), and is known to take local trends and seasonal fluctuations of GDP into account as well.[6]

[6] #casestudy: We performed an in-depth analysis of the results of this case study about Mali, and found out that measuring GDP as a result of democratization requires picking the right control units in order to avoid creating a control unit whose GDP is affected by some other underlying cause. This is a non-apparent but important consideration, which can give this paper considerable quantitative rigor in making qualitative conclusions about Mali.

References

Bertrand, M., Duflo, E., & Mullainathan, S. (2003). How Much Should We Trust Differences-In-Differences Estimates?

Brodersen, K. H., Gallusser, F., Koehler, J., Remy, N., & Scott, S. L. (2015). INFERRING CAUSAL IMPACT USING BAYESIAN STRUCTURAL TIME-SERIES MODELS. *The Annals of Applied Statistics, 9*, 247-274.

De Kadt, D., & Wittels, S. B. (2016). Democratization and Economic Output in Sub-Saharan Africa. *The European Political Science Association, 7*(1), 63-84.

Larsen, K. (2016, January 13). *Making Causal Impact Analysis Easy*. Retrieved from MultiThreaded: https://multithreaded.stitchfix.com/blog/2016/01/13/market-watch/

Appendix A: The code

The notebook with the code used for the replication and extensions:

https://github.com/Inventrohyder/CS112-Final-19/blob/master/ProjectNotebook.pdf

Data, R-file, notebooks:

https://github.com/Inventrohyder/CS112-Final-19

Appendix B: Replication from Synth Package

```
# Replication function
replicate <- function(
  unitID,
  fullname,
  begin,
  end,
  tr2,
  final,
  low,
  high
){

  data <- afripanel[afripanel$WBCode==unitID | afripanel$cont_dem_ind==1,]

  controls <- unique(data$WBCode[data$WBCode!=unitID&data$WBCode!="ETH"&data$WBCode!="SDN"])

  prep <- dataprep(
    foo=data,
    predictors=c(
      "lngdpmadlag",
      "lngdpmadlag2",
      "lngdpmadlag3",
      "lngdpmadlag4",
      "lnpop",
      "ki",
      "openk",
      "civwar",
      "civwarend",
      "pwt_xrate",
      "pwt_xrate_lag1",
      "pwt_xrate_lag2",
      "pwt_xrate_lag3",
      "eximdiff",
      "eximdiff_lag1",
      "eximdiff_lag2"
    ),
    dependent="lngdpmad",
    unit.variable="wbcode2",
    time.variable="year",
    treatment.identifier=unitID,
    controls.identifier=controls,
    time.predictors.prior=c(begin:end),
    time.optimize.ssr=c(begin:tr2),
    time.plot=c(begin:final),
    unit.names.variable="WBCode"
  )

  out <- synth(prep)

  path.plot(synth.res=out, dataprep.res=prep,
            Ylab="Log GDP per capita", Legend=c(fullname, "Synthetic Counterfactual"), tr.intake=tr2,
            Ylim=c(low,high) , Main=fullname
  )

}

## Figure 2 Replication
replicate("MLI", "Mali", 1980, 1990, 1991, 2008, 6, 8)
```

Appendix C: [Extension] In-time placebo

```
## Figure 2 Replication in-time placebo
replicate("MLI", "Mali", 1980, 1985, 1986, 2008, 6, 8)
```

Appendix D: [Extension] In-space placebos

```
## Figure 2 Replication in-space placebo
replicate("BFA", "Burkina Faso", 1980, 1990, 1991, 2008, 6, 8)

## Figure 2 Replication in-space placebo
replicate("TCD", "Chad", 1980, 1990, 1991, 2008, 6, 8)

## Figure 2 Replication in-space placebo
replicate("NGA", "Nigeria", 1980, 1990, 1991, 2008, 6, 8)

## Figure 2 Replication in-space placebo
replicate("BDI", "Burundi", 1980, 1990, 1991, 2008, 6, 8)

## Figure 2 Replication in-space placebo
replicate("TGO", "Togo", 1980, 1990, 1991, 2008, 6, 8)
```

Appendix E: [Replication+Extension] CausalImpact Package

```
# Replication function from Google Extension
show_impact_n <- function(
  Country,
  begin,
  end,
  treatYear
){
  data <- afripanel[which(afripanel$Country == Country), ]
  predictors=c(
      "lngdpmadlag",
      "lngdpmadlag2",
      "lngdpmadlag3",
      "lngdpmadlag4",
      "lnpop",
      "ki",
      "openk",
      "civwar",
      "civwarend",
      "pwt_xrate",
      "pwt_xrate_lag1",
      "pwt_xrate_lag2",
      "pwt_xrate_lag3",
      "eximdiff",
      "eximdiff_lag1",
      "eximdiff_lag2",
      "wbank",
      "wbank_lag1",
      "wbank_lag2"
  )

  outcome <- 'lngdpmad'
  time.points <- as.Date(as.character(data$year), "%Y")

  data <- data[, c(outcome, predictors)]
  data<-data[!is.na(data[outcome]),]
```

```
  data <- data %>% select_if(not_any_na)


  data <- zoo(data, time.points)
  data <- data[index(data) > as.Date(begin, '%Y') & index(data) < as.Date(end, '%Y')]

  nextYear <- as.Date(as.character(as.numeric(treatYear) + 1), "%Y")
  treatYear <- as.Date(treatYear, "%Y")

  start_date <- start(data)
  end_date <- end(data)

  pre.period <- as.Date(c(start_date, treatYear))
  post.period <- as.Date(c(nextYear, end_date))

  impact <- CausalImpact(data,
                         pre.period,
                         post.period,
                         model.args = list(
                           niter = 1000,
                           nseasons = 52)
                         )
  return(impact)
}

# Graphs:
impact <- show_impact_n('Mali', '1980', '2005', '1990')
plot(impact) + ggtitle ("Causal Impact on Mali beyond 1991")

impact <- show_impact_n('Mali', '1980', '2005', '1985')
plot(impact) + ggtitle("Causal Impact on Mali beyond 1985")

impact <- show_impact_n('Burkina Faso', '1980', '2005', '1990')
plot(impact) + ggtitle("Causal Impact on Burkina Faso beyond 1990")
```