I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14228

The Impact of Mass Antigen Testing for COVID-19 on the Prevalence of the Disease

Martin Kahanec
Lukáš Lafférs
Bernhard Schmidpeter

MARCH 2021

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14228

The Impact of Mass Antigen Testing for COVID-19 on the Prevalence of the Disease

Martin Kahanec
Central European University, CELSI and UEBA

Lukáš Lafférs
Matej Bel University

Bernhard Schmidpeter
JKU Linz and IZA

MARCH 2021

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

The Impact of Mass Antigen Testing for COVID-19 on the Prevalence of the Disease

More than 100 million people have been infected and 2.5 million people have died of COVID-19 globally as of February 2021. Mass antigen testing could help to mitigate the pandemic and allow the economy to re-open. We investigate the effects of mass antigen testing on the pandemic, using data from a uniquely designed nation-wide testing implemented in Slovakia in Autumn 2020. After the first round, only districts above an ex-ante unknown prevalence threshold were re-tested. Comparing districts above and below the threshold using a quasi-experimental design, we find that repeated mass antigen testing reduced infections by about 25-30% and decreased R_0 by 0.3 two weeks after the re-testing; the effects on incidence peaked around that time and all effects gradually dissipated afterward. These results suggest that mass testing could be an effective tool in curbing the spread of COVID-19, but for lasting effects regular retesting would be necessary.

JEL Classification: D04, I18, J22

Keywords: COVID-19, COVID-19 policies, antigen testing, mass testing, non-pharmaceutical interventions

Corresponding author:
Martin Kahanec
CELSI
Zvolenska 29
821 09 Bratislava
Slovakia

E-mail: kahanecm@spp.ceu.edu

1 Introduction

More than one year after the first documented cases of COVID-19 in Wuhan, China in 2019, most countries are still struggling to contain this highly contagious disease. According to data provided by Johns Hopkins University, more than 100 million people around the world have been infected and almost 2.5 million people have died of COVID-19 as of February 15, 2021.[1] To protect their most vulnerable citizens and to slow down the spread of the disease, many governments have imposed strict policy measures such as requiring social distancing, stay-at-home orders, as well as local and nation-wide lock-downs. While there is evidence that some of these policies have been successful in at least slowing the numbers of infections (e.g. Chernozhukov, Kasahara and Schrimpf, 2020), they also have both directly and indirectly affected labor supply and demand, investment, consumption, expectations and other economic variables, taking a heavy toll on economies. World's GDP is projected to fall by more than 4% in 2020 (IMF, 2020). The projected decline in GDP is even more pronounced in advanced economies. Economic distress caused by the pandemic policy measures has also affected broader aspects of peoples' lives and well-being (e.g. Arenas-Arroyo, Fernandez-Kranz and Nollenberger, 2021; Brodeur, Clark, Fleche and Powdthavee, 2021).

Facing such detrimental effects on both the economy and society, and with the prospect of widely accessible vaccination still distant – especially for low-income countries, policy makers have been looking for alternative ways of containing the pandemic. Mass testing for COVID-19 as a potential tool of containing the pandemic has received particular attention.[2] Regional and local mass antigen testing has been carried out in several countries such as the UK, China, South Korea, Austria, Luxembourg, and Slovakia. Evidence on whether and how mass testing can work to curb the spread of COVID-19 is scant, however. There is hardly any evidence on how effective mass testing has been can be in comparison to alternative strategies. Informing policy makers on the question whether mass testing can be an effective and cost-efficient way of re-opening the economy is hence an urgent call.

Proponents of mass testing maintain that it indeed is a cost-efficient policy for identifying and quarantining potentially infectious individuals. This would in turn help to reduce the number of cases and the velocity of the spread of the diseases (e.g. Pavelka, Van-Zandvoort, Abbott, Sherratt, Majdan, COVID, Analỳz, Jarčuška, Krajčí, Flasche et al., 2021). If mass testing could accomplish this, policy makers would be able to avoid costly social-distancing policies, keep schools and the economy open, and avert the far-reaching social and psychological costs.

[1] The data was taken from https://coronavirus.jhu.edu/.

[2] Many countries build on a mass Covid vaccination strategy to re-open the economy in the future. Delivery shortages as well as uncertainty about the share of the public which is willing to get finally vaccinated do not make this a viable short-term strategy, however.

On the other hand, opponents of mass testing argue that it can create a false sense of security and may lead individuals to behave less carefully; see the discussion in Mahase (2020). Cheaper rapid antigen (Ag) tests often used in mass testing events, generally have lower sensitivity and specificity compared with the more expensive polymerase chain reaction (PCR) tests, likely leading to a higher rates of false negatives and false positives. This could undercut the credibility of such testing and anti-Covid-19 measures in general. In contrast to the original intention of mass testing, larger numbers of false positives would confine a large share of workers wrongly in quarantine putting an unjustified pressure on the economy (Pettengill and McAdam, 2020). Negative test certificates would likely reduce peoples' caution in social contacts and risky behaviors; false negatives would unknowingly continue spreading the disease. Furthermore, even if mass testing initially worked, there remains the important question on how long the potential benefits can be expected to last.

In this study, we evaluate the impacts of mass antigen testing on the spread of the COVID-19 pandemic. Exploiting a unique quasi-experimental setting whereby districts in Slovakia above an ex-ante unknown threshold of positive test incidence had to repeat the mass testing event, we are able to identify the impact of repeated mass antigen testing on the spread of the disease.

We find that, compared to the districts which were only tested once, in the districts which were tested repeatedly (those above the threshold), the number of infections fell on average by up to 30% and the reproduction number (R_0) decreased by 0.3 two weeks after the second round of testing. Our results hold also when we only consider districts marginally above and below the thresholds in our analysis, albeit our estimates become less precisely estimated due to the smaller sample size. Exploring the dynamics behind these effects, our results indicate a maximum reduction in Covid-19 incidence around 15 days after the second round of testing and a reversal to a statistical zero effect afterward. By the end of our observation period circa three weeks after the second round, we do not find any effects of re-testing on R_0 either.

Our results have important implications for policy makers. First, they show that repeated mass testing can be an effective tool for bringing the spread of the disease under control. The immediate benefits can be quite substantial, in our analysis a second round of mass testing decreased new infections by about one quarter to one third and reduced the reproduction number by 0.3. Second, our results also highlight the necessity to conduct mass testing on a regular basis if a sustained mitigation of the pandemic were to be achieved. Whether such strategy is feasible, cost-effective, let alone whether it would be the best alternative, are different, which we do not address in this study.

We make several contributions to the small but rapidly growing literature on the effects of mass antigen testing on the Covid-19 pandemic (Atkeson, Droste, Mina and

Stock, 2020; Baqaee, Farhi, Mina and Stock, 2020; Mina, Parker and Larremore, 2020; Mina, Miller, Quigley, Prentiss, McKinnon and Comer, 2020; Pavelka et al., 2021; Platiel, Zheng and Walensky, 2020; Callaway and Li, 2020).[3] To the best of our knowledge, this is the first paper systematically evaluating and testing the potential benefits of repeated mass antigen testing using a unique quasi-experimental design.[4] The unique setting in which districts were required to conduct a second round of mass testing if the positive rate was above an a priori unknown and arbitrary threshold enables us to employ a quasi-experimental empirical strategy that enhances the external validity of our results, as it relies on arguably weaker identification assumptions than model-based evaluation methods (e.g. Atkeson et al., 2020). As the dynamic patterns of the pandemic we estimate are at least in part driven by behavioral responses after being tested, our study also contributes to the literature which investigates individual behaviors as important factors in explaining the spread of COVID-19 (e.g. Pettengill and McAdam (2020)).

Closest related to our study is the work by Pavelka et al. (2021) exploring the impact of mass antigen testing in Slovakia by comparing the spread of infections across municipalities and in different rounds of antigen testing, complementing spatial statistical methods with a microsimulation model. They find that the decrease in prevalence compared to a scenario of unmitigated growth cannot be fully explained by non-pharmaceutical interventions implemented before the mass antigen testing. They interpret this finding as evidence of an impact of antigen testing (and the ensuing isolation and quarantine of positively tested individuals) on the spread of the disease. In another study using the data from mass antigen testing in Slovakia, Boďová and Kollár (2020) study the spatial patterns of the epidemic in Slovakia. They conclude that the mitigating effect of repeated antigen testing increased with the measured prevalence of the disease in the earlier round of testing.

The paper proceeds by discussing the mass antigen testing that took place in Slovakia in the autumn 2020. Our empirical approach as well as the data used are outlined in Section 3; we report the results in Section 4. Section 5 concludes. We present the results of several robustness checks in Appendix B.

2 Mass Testing in Slovakia

In late 2020, Slovakia became the first country in the world that introduced nationwide mass rapid antigen testing intended to detect new COVID-19 cases early and halt

[3]The literature on the economic consequences of COVID-19 is reviewed by Brodeur, Gray, Islam and Bhuiyan (2020).

[4]Similar ideas are explored in a blog post by Šuster (2021), which compares COVID-19 trends in the districts that were tested and those that were not tested in the second round of mass antigen testing in Slovakia.

the spread of the disease. With a total population of 5.45 million people, residents aged between 10 and 65 years and older adults in employment, or about 80% of the population, were eligible for voluntary tests.

Before conducting mass testing, Slovakia implemented several containment measures to control infections, such as partial schools closings and restrictions on indoor gastronomy and leisure activities. During the week prior to the first wave of mass testing, authorities asked citizens to limit their movement. To ensure compliance with the measures, police conducted random checks. The Slovakia's government also organized preliminary pilot testing from October 23 to 25, 2020, in four districts: Bardejov, Dolný Kubín, Námestovo and Tvrdošín.

The first wave of nation-wide testing took place from October 31 to November 1, 2020. Around twenty thousand healthcare professionals and forty thousand army personnel and volunteers helped to test residents of all the country's 79 districts. Even though the participation in mass testing was voluntary, those who did not participate in mass testing were also obliged to quarantine for ten days. Citizens with positive test results, members of their households, and their self-traced contacts, had to quarantine for ten days. Citizens with a negative test result had to present their test certificates to their employers to be able to be physically present at work. Authorities also conducted random inspections in public places and individuals had to present a certified negative test results when asked for. Fines up to 1659 Eur applied for citizens violating these regulations. These strict measures led to a high participation rate of around 85% of the eligible population (Pavelka et al., 2021).

Shortly after the first testing round, on November 2, 2020 the government announced that in all districts with a prevalence rate of 0.7% or higher a second round of mass testing will be conducted on November 7 and 8, 2020. The threshold of 0.7% was ex-ante unknown and was chosen arbitrarily. The second round of testing was conducted in the 45 districts. As it was the case during the first round, participation was voluntary but the same restrictions as in the first round were imposed on non-participating citizens and those tested positive.[5]

The testing scheme leads to three different types of districts:

1. Districts with non-pharmaceutical measures and two waves of testing (1st and 2nd)
2. Districts with non-pharmaceutical measures and the first wave of testing only
3. Districts with non-pharmaceutical measures participating in the pilot scheme

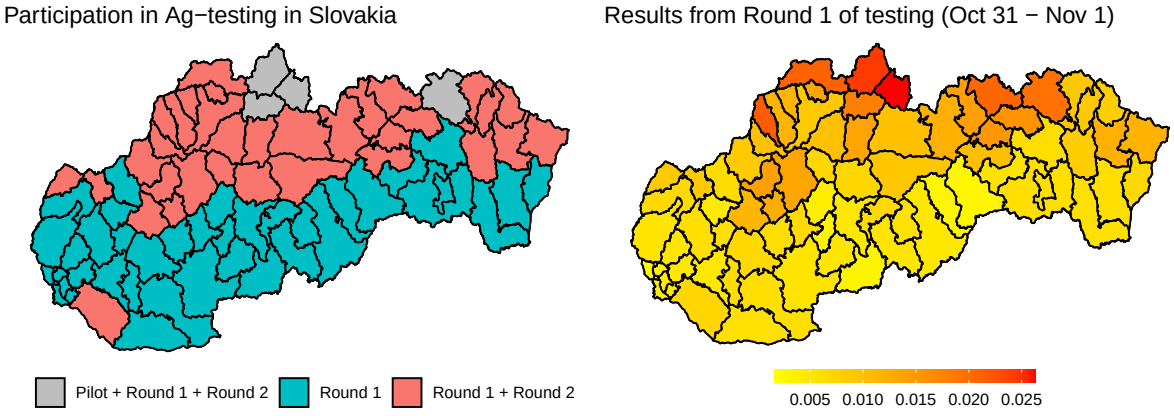
[5]As a result of the mass testing efforts, 50,466 tests turned positive, with the proportion of positives in all tested varying from 3.91% during the pilot to 1.01% in the first wave of mass testing and 0.62% in the second wave (Pavelka et al., 2021).

In our analysis, we only consider districts belonging to Group 1 and Group 2. Districts in Group 1 constitute our treatment group and districts belonging to Group 2 are our control group. Due to their unique setting, we do not consider the four districts participating in the pilot scheme in our analysis (Group 3).

In our study we exploit that whether a given district would be subjected to the second round of testing had been unknown prior to the announcement and whether a district happened to fall above or below the stipulated threshold was "as good as random" in a sufficiently narrow neighborhood of the threshold. This enables us to estimate the average treatment effect of the second round of mass antigen testing in the vicinity of the 0.7% threshold.

In Figure 1, we provide an overview over the location of the different districts in our sample and the share of positively tested persons in the first round. As one can see on the left hand side of the Figure, most of the districts subjected to two rounds of testing are located in the north of Slovakia. While this would suggest a clear spatial spread of Covid, the conclusion is not that clear when looking at the share of positive tests during the first round by district, presented on the right hand side. There is some substantial variation both within and between our treatment and control group. In some cases, the difference in two adjacent districts where one is assigned to a second round of testing while the other district is not can be as little as 0.02 percentage points (districts Nitra and Zvolen). This supports the view that the applied threshold of 0.7% is indeed quasi randomly chosen.[6] We provide additional information about the epidemiological situation before the mass testing in Appendix A.

Figure 1: Participation of different districts in mass Ag-testing and results from the first round.



[6]The data is available under `https://github.com/Institut-Zdravotnych-Analyz/covid19-data`

3 Data and Empirical Approach

3.1 Data

We make use of data provided the Institute of Health Analyses (IHA), an analytical unit of the Ministry of Health of the Slovak Republic. For all 79 districts in Slovakia, the IHA collects data on the daily number of infections within a district. In addition, it also collects information on the number of conducted and positive PCR and rapid antigen tests.[7] We also obtain the information on the share of positive tests by district for the pilot, first, and second round of mass testing from the IHA.[8]

From the IHA data we construct two measures which reflect the spread of Covid-19. Our first measure is the 7-day rolling average of infections on the district level. While one could also use the daily number of new infections, the 7-day rolling average is less noisy and more robust to intra-week variation in testing intensity.

Our second measure is the reproduction number R_0. This measure reflects how many additional people one infected person will infect and can be thought of as capturing the velocity of the spread. We calculate the R_0 at a time T as $\left(\sum_{\tau=T-1}^{T-7} y_\tau\right) / \left(\sum_{\tau=T-6}^{T-12} y_\tau\right)$, where τ is a day and y the numbers of new infections. This measure was also used in epidemic nowcasting in Germany (Hamouda et al., 2020).

3.2 Empirical Approach

To measure impact of repeated mass testing on the prevalence and velocity of COVID-19, we consider a simple Difference-in-Difference model

$$y_{it} = \beta_0 + \beta_1(above_i \cdot t) + \beta_2 t + \beta_3 above_i + \epsilon_{it}, \tag{1}$$

where y_{it} is the outcome measured in district i at time t. Variable y_{it} represents the two variables we use to depict the spread of Covid-19: the 7-day rolling average of new infections and R_0. In our main analysis we consider two time periods: the pre-period for which $t = 0$ (Nov 8, 2020) and the post-period for which $t = 1$ (Nov 22, 2020).

[7] The IHA reports PCR test results only for 72 districts, because districts of the two largest cities - Bratislava I, II, III, IV, V and Košice I, II, III, IV are merged together into Bratislava district and Košice district, respectively. We therefore also merged the Ag test results for these districts and worked 72 districts only. After removing the four districts that were included in the pilot testing, we are left with 68 districts.

[8] The IHA also provides data on hospitalization. There are some problems with this information, however. For example, individuals will be admitted in the nearest hospital which may or may not be in a different district. Therefore, the number of hospitalizations will not reflect the proceeding on the district level. Moreover there are some districts that do not have a hospital. Given these large drawbacks we decided not to use this information in our analysis.

The scalar *above* is a binary indicator which takes a value of one if a district had to be re-tested in the second round. As discussed above, this is the case if during the first round of testing the infection rate was equal to or above 0.7%. Our coefficient of interests is β_1 which we interpret as the mean effect of repeated mass testing on our outcome variables in the vincinity of the 0.7% threshold.

One concern with estimating Equation (1) using the full sample might be that the full sample includes districts relatively far from the assignment threshold and thus districts with likely fundamentally different Covid-19 prevalence trends. As a robustness check, we therefore also estimate a version of Equation (1) only considering districts with first-round prevalence rates only marginally below and above the threshold.[9] In fact, to understand the role of the band-width around the 0.7% threshold, we provide estimates based on samples defined using the whole range of meaningful band-widths around the threshold. These estimates also help us to gauge how important selection on unobserved district characteristics might be in our setting.

Another important assumption is the one about the lag that we apply to measure the effects of the second round of testing. The baseline model is based on the assumption that it takes several days until any symptoms develop, the individual registers and obtains a date for testing, and the results are reported in the statistics. In line with the literature reviewed above, we assume this to take 14 days. However, in our analysis we test the sensitivity of our results to this assumption by reporting the estimated effects for the whole range of possible lags up to three weeks after the second round. This approach also enables us to shed light on the duration of the effects of mass antigen testing.

In the Appendix, we also provide multiple robustness and placebo checks. We estimate our model based on Equation (1) using other, arbitrary threshold levels (0.5% and 1.2%, rather than the true threshold of 0.7%) and arbitrary date of the second round of testing (November 1 instead of the true dates of November 7 and 8). As it turns out, we do not find similar any significant effects in these placebo tests, lending support to our identification strategy. Similarly, our results, also reported in the Appendix, appear to be robust with respect to measuring the pandemic by using only the PCR test statistics (instead of PCR+antigen testing) as well as the inclusion of mobility trends in Equation (1).

[9] Due to the small sample size, employing a non- or semi-parametric regression discontinuity design is not feasible in our setting.

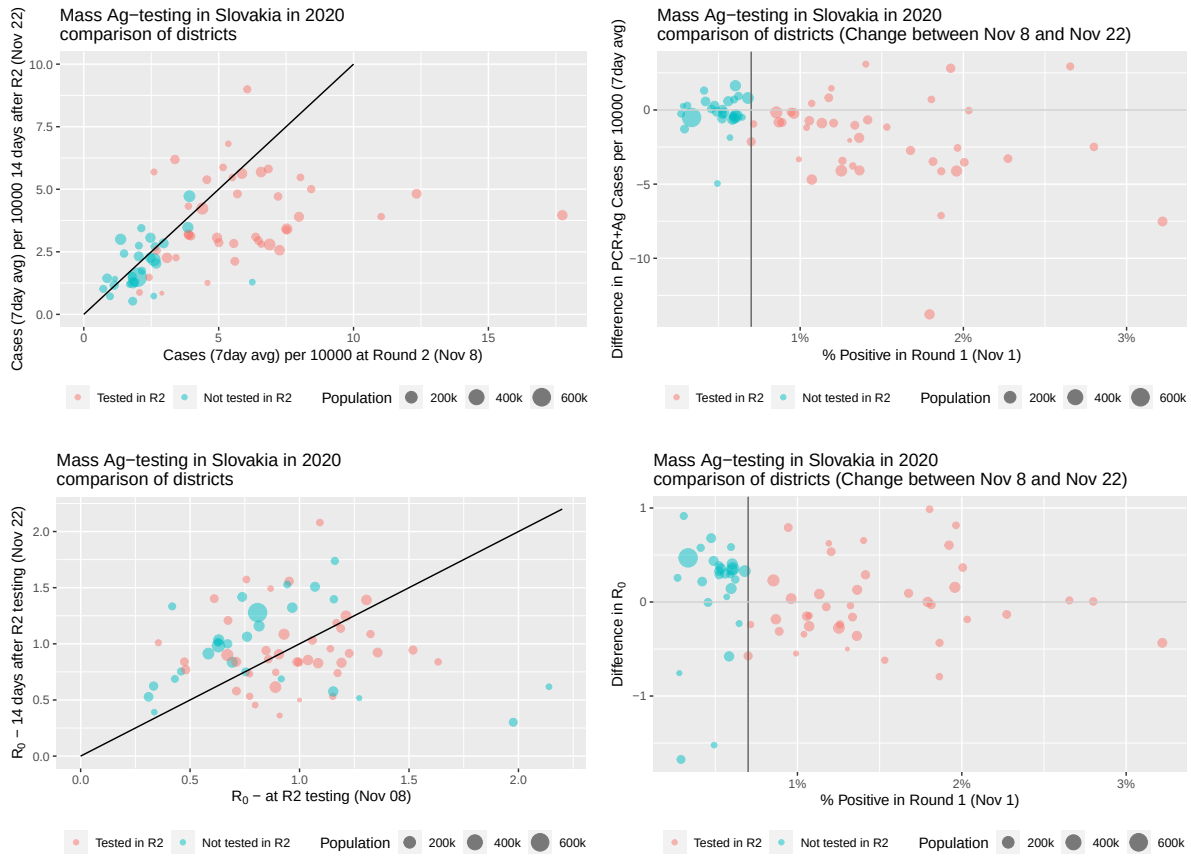
4 Results

4.1 Descriptive Evidence

In this section we provide some first, descriptive evidence about the prevalence of Covid-19 in the treated and non-treated districts.

In Figure 2 we explore the association between the prevalence of Covid-19 measured in from the first round of mass testing with the difference betweenchange in normalized positive cases (PCR or antigen) between November 8 (second round) and November 22 (14 days after the second round) measured in by 68 districts. [10] We observe that re-tested districts experienced larger drops in infections than those exempt from the second round (top-left panel), which in turn experienced an increase in the reproduction number R0 compared to the re-tested ones (bottom-left panel). The top- and bottom-right panels, comparing prevalence and the reproduction number measured at the time of the second round of testing and two weeks later, show similar patterns.

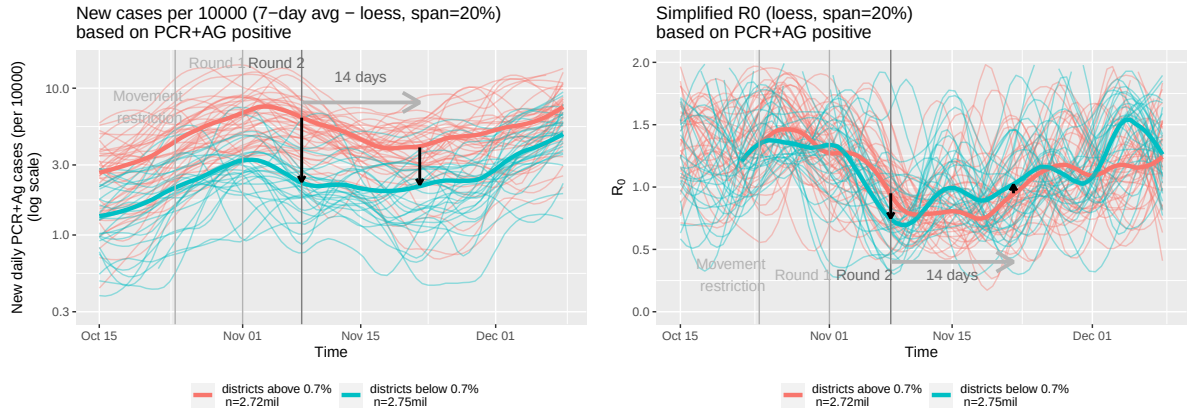
Figure 2: Association between the results from the first round of mass testing and the R_0 and various metrics two weeks after the Round 2.



[10] Recall that we removed four districts that participated in the pilot testing one week before the first round. The epidemic situation in these four districts was far worse than in the rest of the country.

In Figure 3 we report the trends in prevalence and the reproduction number for the full sample of 68 districts, with the thick lines representing the averages for the treated and non-treated districts. Consistently with the patterns observed in Figure 2, we observe that differences in both the number of cases and simplified R_0 between tested and non tested districts shrank between Rounds 1 and 2 (Figure 3). The treated districts improved more than the non-treated ones in terms of the infection rate, whereas the reporduction number increased in the non-treated districts more than in the treated ones.

Figure 3: Evolution of infections and R_0 in the districts below and above the threshold 0.7% for the full sample.



4.2 Baseline specification

The previous section provided descriptive evidence that suggests that mass antigen testing might reduce the spread of Covid-19. In this section, we present estimates from our Difference-in-Difference specification; as proposed in Equation (1), estimated for four outcome variables: the 7-day average number of cases per 10 000 citizens, the logarithm thereof, R_0, and $\log R_0$.

Table 1 presents the estimation results based on the overall sample. Each regression is weighted by the district population size. Looking at Column (1), we see that the second wave of mass antigen testing reduced the 7-day average in infections measured 14 days after the repeated testing by approximately 2.3 cases per 100.000 inhabitants. This constitutes a quite sizable reduction of 36% (Column 2).

Columns (3) and (4) report the impact on R_0 and $\log R_0$ respectively. Our estimates suggest that the second round of testing decreased simplified R_0 by approximately 0.28 more in the treated districts than the non-treated ones, corresponding to a reduction by 31%. Loosely speaking, these results imply that repeated mass testing reduces the number of people to whom ten infected persons pass on the infection two weeks after the second round by 2.8.

While each national setting is in many ways unique and therefore comparable only to some extent, it is worth comparing our results to estimates for other non-pharmaceutical interventions in the literature. For example, Chernozhukov et al. (2020) estimate that face masks reduced the weekly growth rate in the number of cases by around 10% in the U.S. They also find that stay-at-home-orders reduced the number of new cases by between 6% to 63%. Putting these findings into perspective, they would imply that repeated mass testing has the potential to be more effective than mandatory mask wearing and may even be as effective as stay-at-home orders, at least in the short run. On the other hand, our estimates are considerably smaller than those estimated by Pavelka et al. (2021). They estimate that the decrease in prevalence following the mass testing within one week was up to 70% after adjusting for geographical clustering, epidemiological situation at the time of the first round and an expected growth rate in infection prevalence.

Table 1: Effect of Repeated Mass Testing on Covid Infections - Full Sample

	Dependent variable:			
	Cases	log Cases	R_0	log R_0
	(1)	(2)	(3)	(4)
Tested in R2 × Time	−2.252***	−0.358**	−0.279***	−0.314***
	(0.632)	(0.144)	(0.100)	(0.112)
Tested in R2	4.047***	1.005***	0.186***	0.235***
	(0.447)	(0.102)	(0.071)	(0.079)
Time	−0.125	−0.092	0.271***	0.306***
	(0.446)	(0.101)	(0.071)	(0.079)
(Intercept)	2.243***	0.739***	0.787***	−0.302***
	(0.315)	(0.072)	(0.050)	(0.056)
Observations	136	136	136	136
R^2	0.463	0.535	0.105	0.114

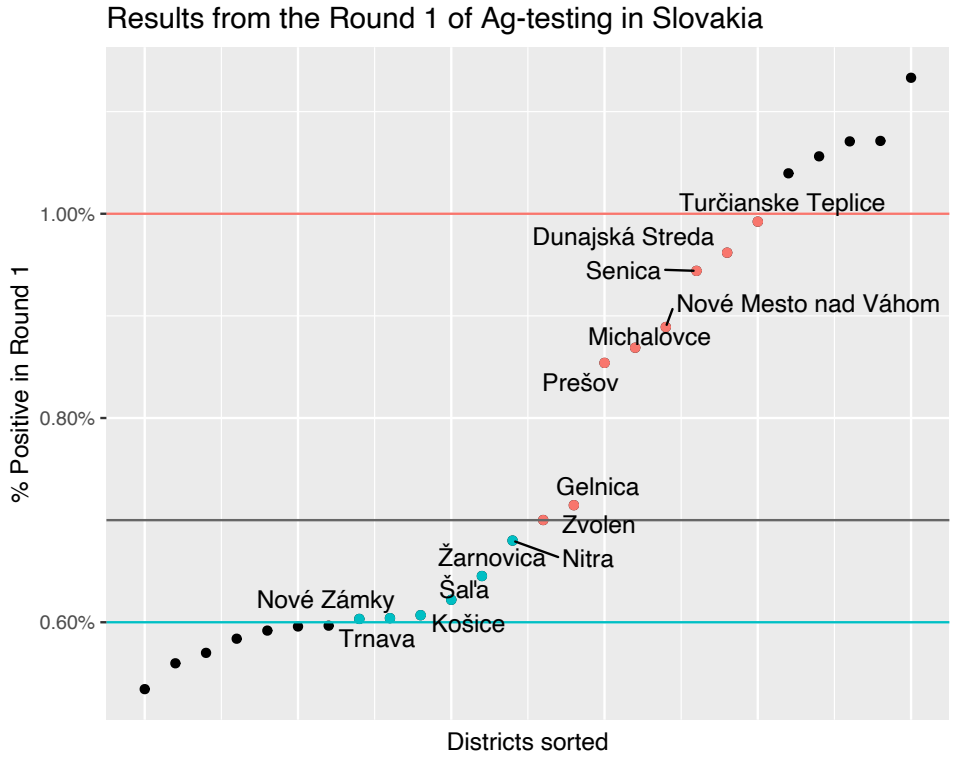
Note: Districts weighted by their population size. *p<0.1; **p<0.05; ***p<0.01

One concern with our results reported in the previous section might be that we also included districts far off from the 0.7% threshold in the comparison. That could be an issue, as the epidemiological situation and other observed and unobserved characteristics in those districts might have been substantively different from those nearer the threshold and hence their inclusion in the "above" and "below" groups might not be "as good as random". In this section, we therefore assess the robustness of our results considering only districts close to the 0.7% threshold. There are obvious costs of the inclusion of fewer districts in the analysis: the confidence intervals widen as the statistical precision of our estimates drops.

We restrict the sample to a comparably sized "above" and "below" groups of districts that are closer to the threshold level of 0.7%. The choice of the districts is presented in Figure 4. The "above" the threshold group of districts that were re-tested in the second round include districts where the prevalence as measured in round 1 was between 0.7%

and 1% with a cumulative population size of around 650 000; the "below" group consists of districts with a prevalence in round 1 in the range between 0.6% and 0.7% with the population of approximately 750 000.

Figure 4: Choice of districts in our restricted sample.



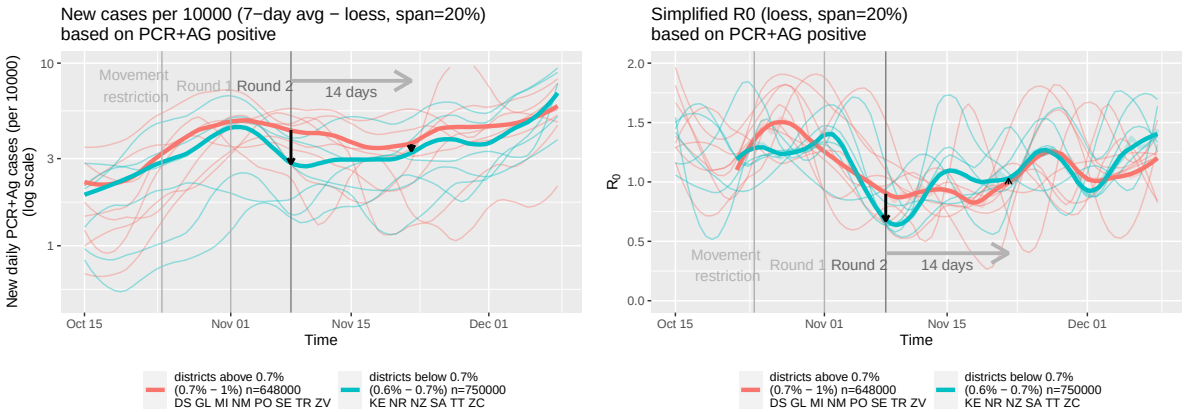
The results for Equation (1) estimated on this restricted sample are presented in Table 2. We find quantitatively very similar effects of the second round of testing on the number of cases and R_0. For example, we estimate now a drop in R_0 by around 39% which is even slightly large than in our full sample; see Column (4). The estimated effects are now, however, only marginally significant. This is not surprising given the large drop in our sample size.

Table 2: Effect of Repeated Mass Testing on Covid Infections - Restricted Sample

	Dependent variable:			
	Cases	log Cases	R_0	log R_0
	(1)	(2)	(3)	(4)
Tested in R2 × Time	−1.007	−0.350	−0.332*	−0.385*
	(0.904)	(0.293)	(0.173)	(0.189)
Tested in R2	1.451**	0.463**	0.285**	0.315**
	(0.639)	(0.208)	(0.123)	(0.133)
Time	0.328	0.136	0.328**	0.395***
	(0.616)	(0.200)	(0.118)	(0.128)
(Intercept)	2.769***	0.941***	0.699***	−0.378***
	(0.435)	(0.141)	(0.083)	(0.091)
Observations	28	28	28	28
R^2	0.193	0.181	0.286	0.318

Note: Districts weighted by their population size. *p<0.1; **p<0.05; ***p<0.01

Figure 5: Evolution of infections and R_0 in the districts below and above the threshold 0.7% for the restricted sample.



4.3 Size of the reference groups

In order to explore the sensitivity of the regression results to the choice of the reference groups, we estimate a series of regressions for different sizes of the "below" and "above" groups.[11] Figure 6 presents the results. On the horizontal axis we plot the maximal size of both the "below" and "above" groups. The black line denotes the respective regression coefficients and the gray area depicts the 90% (darker) and 95% (brighter) confidence intervals from the model based on Equation (1).[12] On the left hand size of each these four graphs, we plot the results estimated on relatively small reference groups, consisting from only a few districts. The statistical uncertainty of these estimates is rather high. On the right hand side we plot the results based on the full sample. The grey vertical dashed line shows the results for the size of the larger of the "above" and "below" equal to 750 000, which is close to the choice of infection rate bandwidth made in the restricted sample presented in the previous section.[13]

Several patterns become apparent. The downward slope of the curve representing the coefficient $\hat{\beta}_1$ on the effect of the second round of testing on the 7-day average of new cases in the upper left pane suggests a regression to mean effect less present on the log scale (upper right pane). On the other hand, the regression coefficients for R_0, appears rather stable for different population sizes of the reference groups (bottom panes).

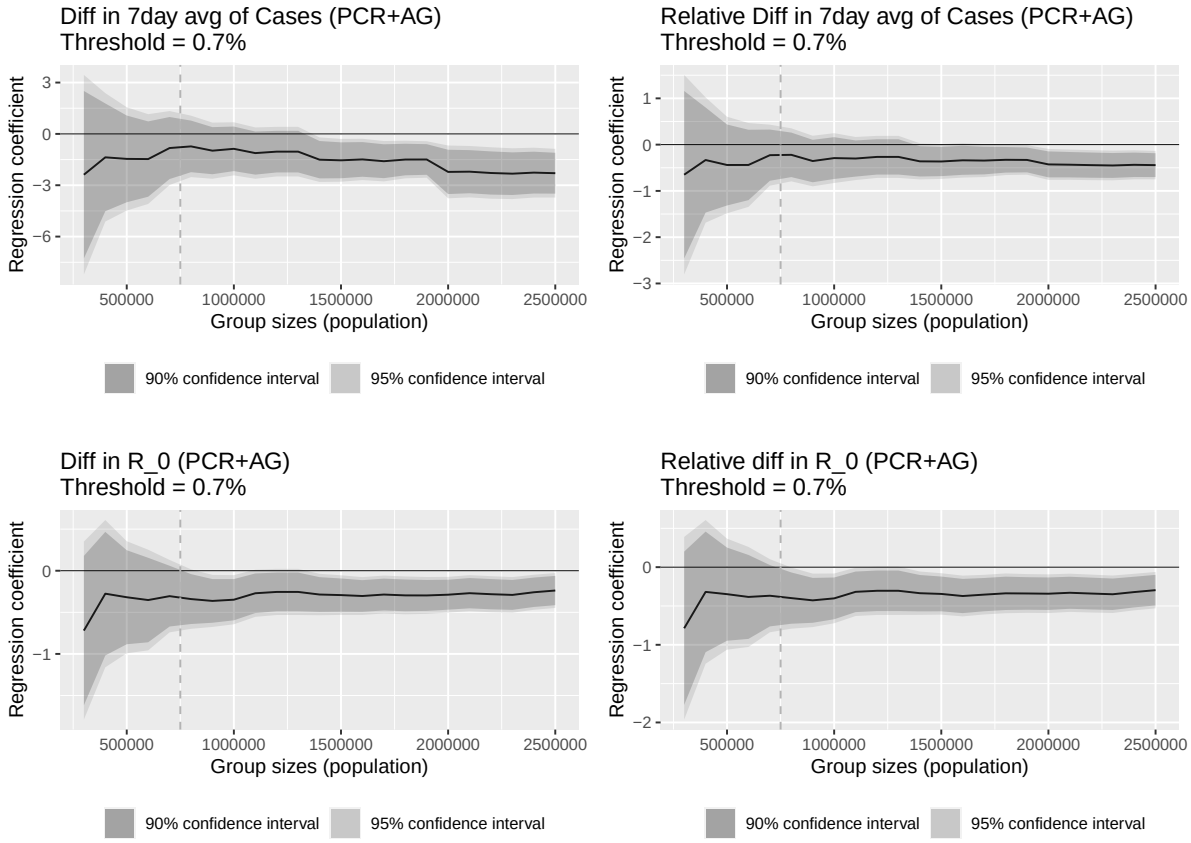
Overall, the reduction due to the second round of testing in the reported infection cases is about 30% and the reduction in R_0 is about 0.3 two weeks after the re-testing. Statistical uncertainty of these estimates is sizeable and 95% confidence intervals for many specifications include zero. However, we observe that the estimates become statistically significant as certain threshold sizes of "above" and "below" groups are reached: circa 1.3 million for cases and 0.8 million for R_0.

[11] We did not make use of symmetric bandwidths around the 0.7% thresholds as the sizes of the groups would be very dissimilar and thus complicating the statistical comparison.

[12] More precisely, we sequentially include non-treated districts in the "below" group (and similarly, treated districts in the "above" group) until the cumulative population in these group crosses a specified threshold.

[13] These results are not completely the same, as the range of percentages that correspond to this line would not consist of round numbers, such as 0.6% and 1% in the restricted sample.

Figure 6: Estimated regression coefficient $\hat{\beta}_1$ with confidence intervals based on Equation (1) as a function of the size of the groups below and above the threshold.



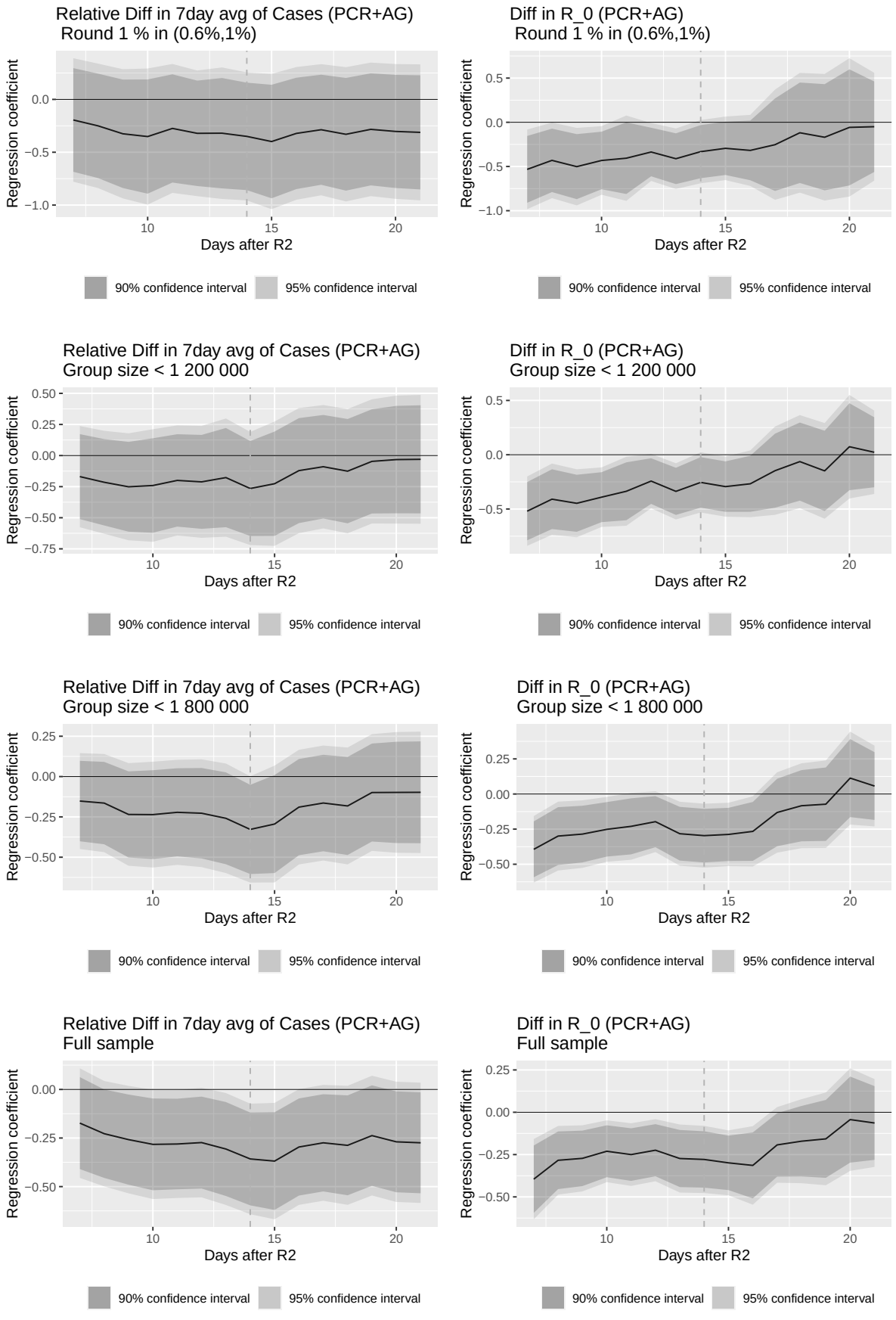
4.4 The Timing of the Effects

In light of the results reported in the previous section, it is interesting to see if mass antigen testing had a long-lasting impact on infections and velocity of the disease. Our initial evidence pointed towards a convergence in infections between our treatment and control group some time after the second round. More generally, in this section we study the timing of the effects of mass testing on Covid-19 during the three weeks following the second round of mass antigen testing.

Figure 7 visualizes the regression coefficients of interest for log cases and R_0 as a function of the number of days after Round 2 for the following specifications:

(1) Districts with Round 1 results between $(0.6\%, 1\%)$ as described in Subsection 4.2
(2) Districts with the "above" and "below" group sizes each less than 1 200 000
(3) Districts with the "above" and "below" group sizes each less than 1 800 000
(4) The full sample

Figure 7: Estimated regression coefficient $\hat{\beta}_1$ with confidence intervals based on equation (1) as a function of time at which the outcome was measured. Results are shown for different subsamples. Dashed vertical line denotes two weeks.



For the specification described in Subsection 4.2 (subsample with round-one results between $(0.6\%, 1\%)$) the effect on log cases is estimated at -30% after 14 days, but with wide confidence intervals. The curve for R_0 shows a gradual decrease of the absolute value of the effect in time (convergence to zero).

In the second specification with group sizes limited up to 1 200 000, we observe that the effect on the 7-day average of cases attains its maximum (in absolute value) two weeks after the second round and gradually fades out thereafter. The effect on R_0 attains its maximum right after the second round and gradually converges to zero over the three weeks following the re-testing. Very similar patterns emerge in the specification with the group sizes that have population size limited up to 1 800 000, as well as in the full sample.These results for large "below" and "above" groups may be subject to a stronger regression-to-mean effect and we should be cautious when interpreting them.

Overall, we consider the second and third specifications as the most salient to triangulate the true effect. Based on our results, Our best conjecture is that the effect of the second round, conditional on the first round, was a 25-30% reduction in prevalence and reduction in R_0 by approximately 0.3. After three weeks the effects diminished, with point estimates indicating a reduction in the 7-day average of new cases at less than 10% and a zero effect on R_0; however, the coefficients after three weeks are estimated with a large statistical uncertainty and cannot be distinguished from zero.

4.5 Robustness and falsification tests

In order to scrutinize the robustness of our results we conduct several falsification tests, explore an alternative measure of incidence of Covid-19, and look into how behavioral mobility responses might have affected our findings. First, we study whether we find any effects of 'placebo' treatments that are known not to have real effects on the spread of Covid-19 by construction. To that effect, we construct three placebo tests: we redefine the treatment by replacing the true threshold of 0.7% with false values of 1.2% and 0.5% and we change the true date of the treatment (November 7-8) to a false date one week earlier (November 1). The results reported in Appendix B.1-B.3, respectively, show essentially no effects in any of these placebo specifications, suggesting that our findings pass the falsification tests.

To further test the robustness of our results, we note that PCR and antigen testing only provide proxy measures of the true prevalence and incidence of the pandemic. In the analysis above we use the results from both types of testing to minimize the measurement problems possibly caused by people substituting one type of testing for the other. However, to verify whether our results hold when we use only PCR tests as our measure of the spread of Covid-19, we re-estimate our models based on PCR data only. The results

of such analysis reported in Appendix B.4 show that the results are quantitatively and qualitatively similar to those presented above for the aggregate of PCR and antigen tests.

Finally, we study whether and how behavioral mobility responses to antigen testing affect our results. Specifically, we include the 7-day rolling average of Google workplace mobility (GoogleLLC, 2021) in our regression models.[14] Again, all the key results remain qualitatively and quantitatively the same as those presented above (See Appendix B.5). Interestingly, mobility trends come out as insignificant determinants of our measure of the spread of Covid-19, suggesting that either workplace mobility does not matter for the spread of Covid-19 or the elasticity of behavioral responses in workplace mobility to testing is too low to result in different effects in treated and non-treated districts. Given that we observe very similar patterns in workplace mobility in treated and non-treated districts, we conjecture that the latter hypothesis might be the case.

5 Conclusion

Repeated mass testing has been widely discussed by policy makers as a possible instrument that can be deployed to mitigate the spread of Covid-19 while also be able to keep the economy open. Despite this attention to the topic, there is surprisingly little evidence if and how repeated mass testing affects the pandemic. We exploit a unique quasi-experimental setting of mass antigen testing in Slovakia which randomly assigned some districts to a second rounds of mass testing to examine the effect of repeated mass testings on the spread of Covid-19.

Our results suggest that 14 days after the second round of testing new infections decreased by between 25% to 30%, and R_0 decreased by 0.3. Investigating the patterns of these effects over time, we find that the impact on new infections gradually faded out after circa two to three weeks. Several falsification and robustness checks confirm the validity of our estimates.

Our findings have important policy implications. Based on our results, repeated mass antigen testing could be a viable tool for bringing the spread of the disease under control. The immediate benefits can be quite substantial, in our setting the second round of mass antigen testing decreased new infections by around one third; however the effect dissipate after 2-3 weeks. Our results thus highlight the necessity to conduct mass testing on a regular basis and quite frequently, if sustainable effects are to be achieved. To the best of our knowledge, no country offers such a testing regime, however.

We also note that before any decision about mass antigen testing is made, a proper cost-benefit analysis must evaluate not only the potential benefits of mass testing, but

[14] All the other mobility trends reported by Google at the required level of granularity have too many missing values to enable meaningful testing.

also its costs. These include the direct costs of testing, as well as any indirect costs on the society. Only a proper cost-benefit analysis of mass antigen testing vis-a-vis alternative instruments will enable policy makers to decide, whether mass antigen testing offers best value for the last resource spent on it.

While we think that our study makes an important contribution to the discussion about the potential benefits of mass testing, one should also recognize its limitations. We note that we estimate the effect of the second round only, but conditional on that the first round occurred one week before. Given that there were many more infections isolated during the first round of testing than in the second round, we conjecture that the effect of the first round may have been somewhat larger than the effects that we estimate for the second round. It is also important to note that during the week after each round of the mass nation-wide antigen testing, citizens who were tested positive, their contacts, as well as those not tested were required to self-quarantine. The second round of testing might have had additional effects through different behavioral responses of citizens, policy makers, law enforcement agencies, testing sites and health workers, et cetera in treated and non-treated districts. We do not disentangle these effects and our estimates have to be interpreted as a joint effect of Round 2 antigen-testing together with all the policy restrictions and behavioral responses associated with this round of testing.

There might be different explanations why the effect of antigen-testing started to slowly fade away in time. We conjecture that the principal mechanism was that the direct mitigating effects of quarantining of positive cases and their contacts (as well as those who chose not to get tested), which dominated on the first few days after the testing, were overridden by the adverse effects of the behavioral responses of those, who tested negatively and, as a result, gained a false sense of security. In addition, the initial mitigating effect on peoples' behavior in districts denoted a "high-risk" might have gradually dissipated and weakening compliance with (and expiration of) quarantine measures could potentially explain part of the effect reduction. Limited data availability makes it impossible to test any ot these explanations directly, however. It should also be noted that our work does not evaluate the pecuniary and non-pecuniary costs or any other direct and indirect effects of mass testing on other outcomes, such as labor supply and demand. All these are open questions which should be answered in further work.

References

Arenas-Arroyo, E., Fernandez-Kranz, D. and Nollenberger, N. (2021), 'Intimate Partner Violence and Forced Coexistence and Economic Stress: Evidence from the COVID-19 Pandemic', *Journal of Public Economics* **194**, 104350.

Atkeson, A., Droste, M. C., Mina, M. and Stock, J. H. (2020), 'Economic benefits of covid-19 screening tests', *National Bureau of Economic Research WP 28031* .

Baqaee, D., Farhi, E., Mina, M. J. and Stock, J. H. (2020), 'Reopening scenarios', *National Bureau of Economic Research WP 27244* .

Bod'ová, K. and Kollár, R. (2020), 'Characteristic spatial scales of sars-cov-2 pandemics: lessons from mass rapid antigen testing in Slovakia', *medRxiv* .

Brodeur, A., Clark, A. E., Fleche, S. and Powdthavee, N. (2021), 'COVID-19, Lockdowns and Well-being: Evidence from Google Trends', *Journal of Public Economics* **193**, 104346.

Brodeur, A., Gray, D., Islam, A. and Bhuiyan, S. J. (2020), 'A Literature Review of the Economics of COVID-19', *IZA Discussion Paper No 13411* .

Callaway, B. and Li, T. (2020), 'Understanding the effects of tennessee's open covid-19 testing policy: Bounding policy effects with nonrandomly missing data', *arXiv preprint arXiv:2005.09605* .

Chernozhukov, V., Kasahara, H. and Schrimpf, P. (2020), 'Causal impact of masks, policies, behavior on early covid-19 pandemic in the us', *Journal of Econometrics* **220**(1), 23–62.

GoogleLLC (2021), 'Google covid-19 community mobility reports', https://www.google.com/covid19/mobility/. Accessed: 2021-01-22.

Hamouda, O. et al. (2020), 'Schätzung der aktuellen entwicklung der sars-cov-2-epidemie in deutschland–nowcasting'.

IMF (2020), 'World Economic Outlook - A Difficult Ascent', *Report* .

Mahase, E. (2020), 'Covid-19: Mass testing is inaccurate and gives false sense of security, minister admits', *BMJ* **371**, m4916.

Mina, M. J., Parker, R. and Larremore, D. B. (2020), 'Rethinking covid-19 test sensitivity—a strategy for containment', *New England Journal of Medicine* **383**(22), e120.

Mina, M., Miller, S., Quigley, M., Prentiss, T., McKinnon, J. E. and Comer, S. (2020), 'Analyzing conflicting results in rapid point-of-care covid-19 testing', *mimeo* .

Pavelka, M., Van-Zandvoort, K., Abbott, S., Sherratt, K., Majdan, M., COVID, C., Analỳz, I. Z., Jarčuška, P., Krajčí, M., Flasche, S. et al. (2021), 'The impact of population-wide rapid antigen testing on sars-cov-2 prevalence in slovakia', *Science* .

Pettengill, M. A. and McAdam, A. J. (2020), 'Can We Test Our Way Out of the COVID-19 Pandemic', *Journal of Clinical Microbiology* **58**(11).

Platiel, A. D., Zheng, A. and Walensky, R. P. (2020), 'Assessement of SARS-CoV2 Screening Strategies to Permit the Safe Reopening of College Campuses in the United States', *JAMA Network Open* (7), e2016818.

Šuster, M. (2021), 'Ide plošákom sploštit' krivka?', https://martinsuster.blog.sme.sk/c/553710/ide-plosakom-splostit-krivka.html. Accessed: 2021-01-25.

A Additional Information on PCR and antigen testing

In this section, we give a brief overview over the prevalence of PCR and antigen tests over time. We also provide an overview over the number of positive tests as well as numbers of hospitalizations over time.

As it is apparent from Figure A.1, the number of positive PCR tests went down after the mass testing. At the same time, we see an increase in antigen positive tests. Around mid October, antigen testing sites were introduced at various places in Slovakia, where it was possible to get tested for free. Availability of these free antigen tests increased over time as demonstrated in the right pane of Figure A.1. With the wider availability, it is likely that a larger share of the population has switched to antigen tests, rather than the more expensive PCR tests. Given that there was a large variation in the antigen testing capacities on the district level, we included antigen tests into our analysis.

Figure A.1: Evolution of positive tests and the number of administered tests.

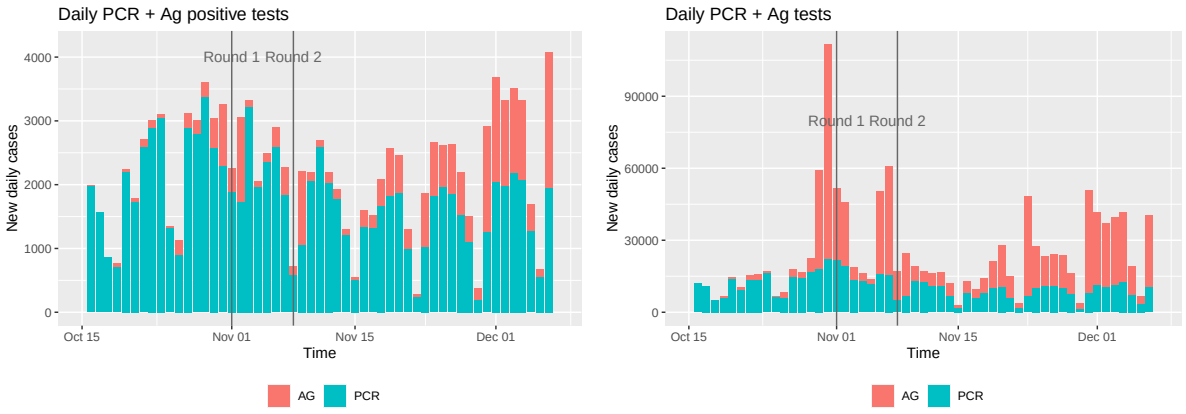


Figure A.2 shows that positivity of PCR tests went up in the week after the round 2 of testing and the went down the week after. Positivity of the antigen tests appeared to be somewhat more stable with a spike approximately 10 days after the second round of testing.

Hospitalizations (Figure A.3) decreased around the weekend of the second round of antigen testing. The right pane shows the simplified R_0 of hospital admissions.[15] From these figures alone it is not possible to disentangle the potential effect of the antigen testing as several other policy measures were in place, such as closed schools and movement restrictions. However, we see some improvement in the R_0, which fell below 1 for approximately two weeks. Data on hospitalizations are independent of the testing capacities

[15]Simplified R_0 evaluated at time T is equal to $\left(\sum_{\tau=T-1}^{T-7} y_\tau\right) / \left(\sum_{\tau=T-6}^{T-12} y_\tau\right)$, where τ is a day. This measure was used in epidemic nowcasting in Germany (Hamouda et al., 2020).

and therefore contain a lot of information about the epidemic situation although with a time lag.[16]

Figure A.2: Evolution of percent positive for PCR and Ag-tests.

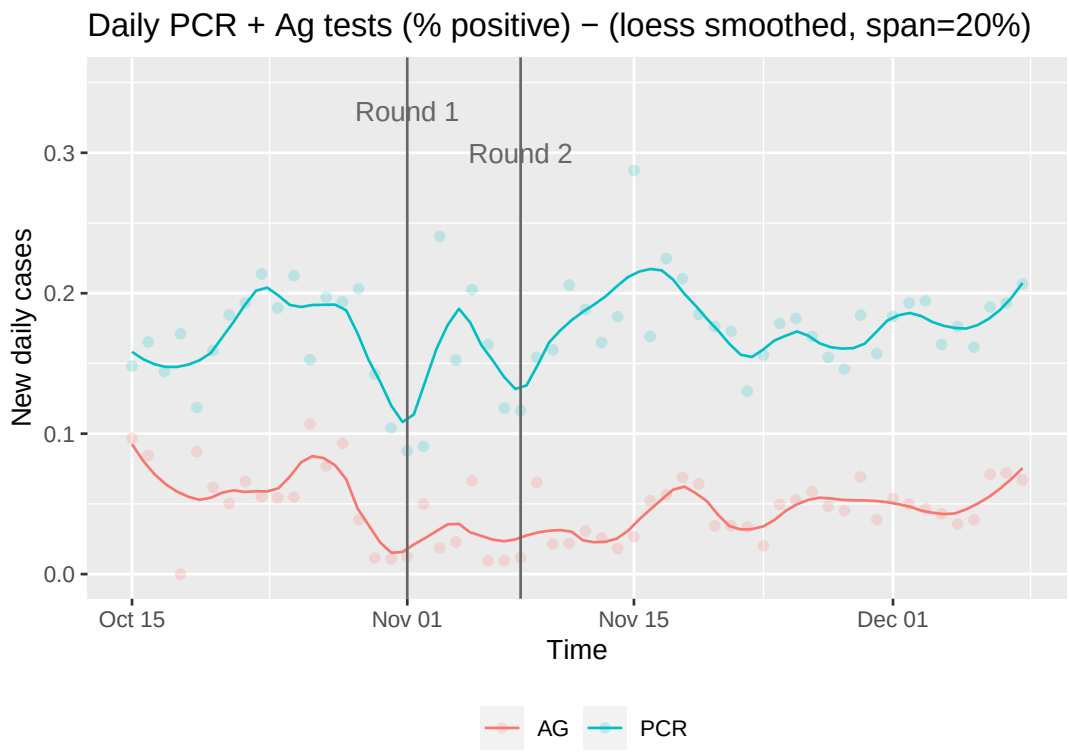
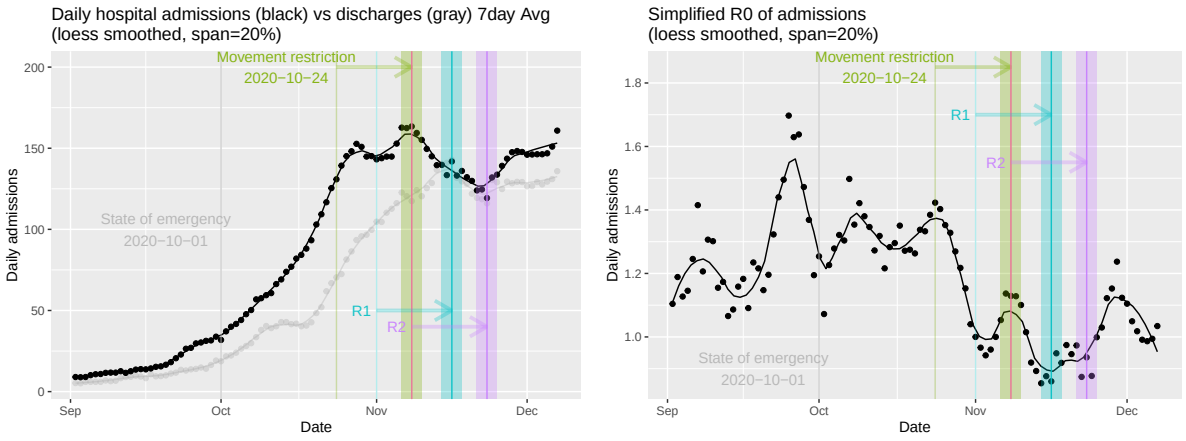


Figure A.3: Hospital admissions, discharges and R_0. Arrows point 15 days ahead.



[16]There is no reliable data on the numbers of hospitalization on the district level in Slovakia. Not all districts have hospital ane the are many spillovers from the neighborhood districts.

B Sensitivity and robustness

In order to investigate whether our results might be driven by some spurious correlations, we conducted similar analysis as under our main specifications presented above but with different arbitrary placebo thresholds for defining the "above" and "below" groups: 1.2% and 0.5%. The regression results, figures and sensitivity to the size of references groups are presented in the Appendices B.1 and B.2.

For the placebo specification with threshold 1.2% we see a negative but non-significant regression coefficient of interest for all the reference group sizes that did not include any non-tested district. All the other coefficients, both for thresholds 1.2% and 0.5%, are not significant and very close to zero. The curves for coefficients for R_0 are non-significant and closely match the horizontal axes, supporting the causal interpretations of the results with the true threshold 0.7%.

We also tried a false date at which the districts were treated. Instead of the true date Nov 8, we set Nov 1 as the date of placebo Round 2 instead, and then we measured outcomes 14 days after (on Nov 15). Figure B.7 in Appendix B.3 shows that the estimated effects for this placebo test are non-significant and close to zero, apart for the specification with very large group sizes for the logarithm of R_0. This essentially states that the velocity of the disease in districts with worse epidemiological situations improved more than in the districts with better situations. This is intuitive as we have seen this pattern in most of the other specifications and can be attributed to the regression to mean effect.

Furthermore, Appendix B.4 presents the results based on PCR tests only.[17] These results are not qualitatively different from our main analysis.

We included a mobility measure as regressor in our analysis, the results from this analysis are presented in Appendix B.5. Inclusion of the mobility into our regressions leads to somewhat stronger but less precisely estimated effects, due to missing data on mobility in several districts.

[17] We thank Martin Šuster for this suggestion.

B.1 Placebo 1 - different threshold (1.2%)

Figure B.1: Choice of districts in the placebo with threshold 1.2%.

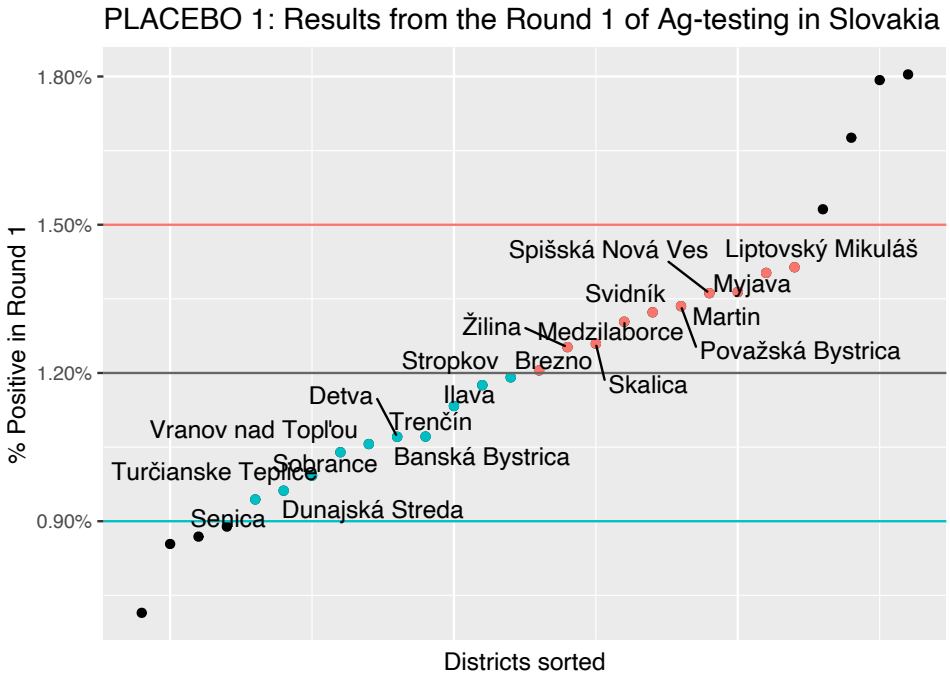


Figure B.2: Evolution of infections and R_0 in the districts below and above the Placebo threshold 1.2%.

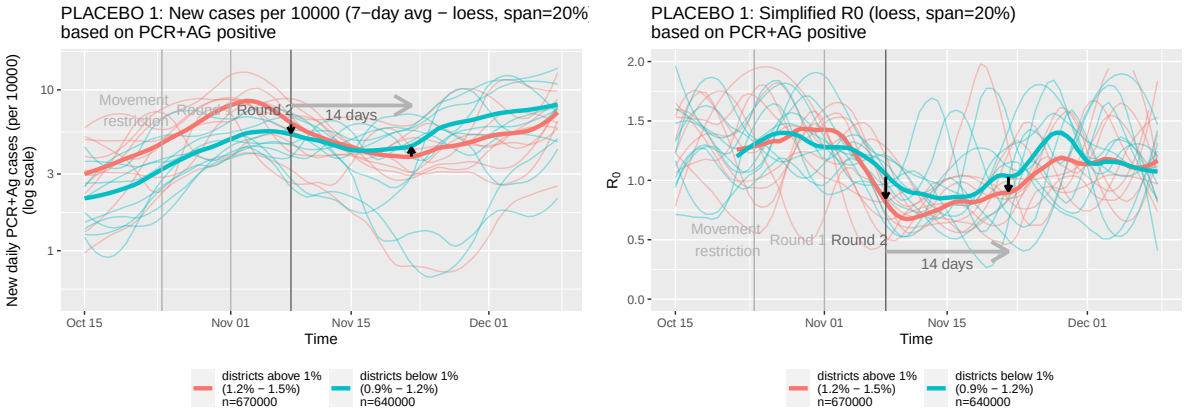
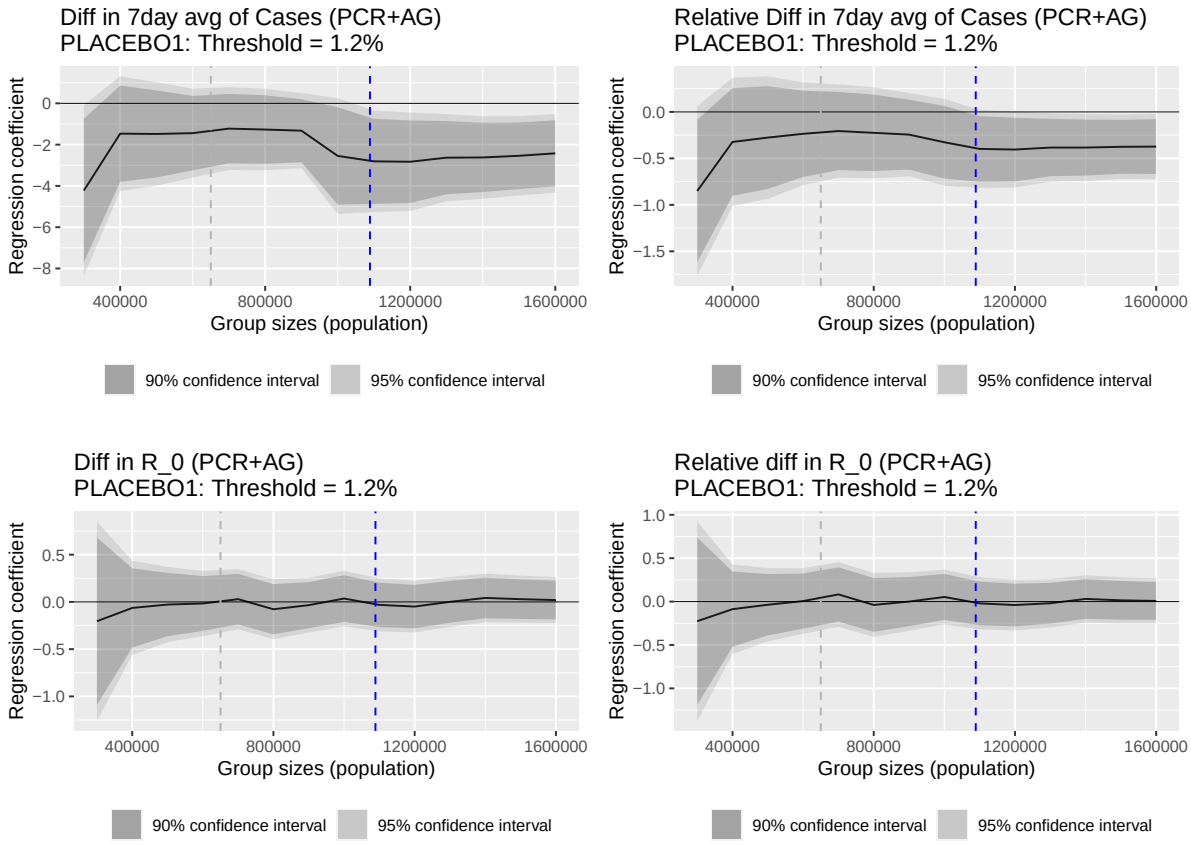


Table B.1: Placebo 1: Regression results based on eq (1)

	Dependent variable:			
	Cases	log Cases	R_0	log R_0
	(1)	(2)	(3)	(4)
Tested in R2 × Time	−1.325	−0.229	−0.052	−0.005
	(0.989)	(0.246)	(0.160)	(0.186)
Tested in R2	0.889	0.169	−0.210*	−0.246*
	(0.700)	(0.174)	(0.113)	(0.131)
Time	−1.101	−0.279	0.016	−0.011
	(0.707)	(0.176)	(0.114)	(0.133)
(Intercept)	5.280***	1.611***	1.070***	0.046
	(0.500)	(0.125)	(0.081)	(0.094)
Observations	40	40	40	40
R^2	0.293	0.241	0.197	0.167

Note: Districts weighted by their population size. *p<0.1; **p<0.05; ***p<0.01

Figure B.3: Regression coefficient as a function of maximal size of the groups below and above the threshold. Vertical dash line stands for the size of the placebo specification groups and vertical blue line depicts the size of the below group that does not include any district that was not tested in R2. The results to the left of this blue line are all tested in R2.



B.2 Placebo 2 - different threshold (0.5%)

Figure B.4: Choice of districts in the placebo with threshold 0.5%.

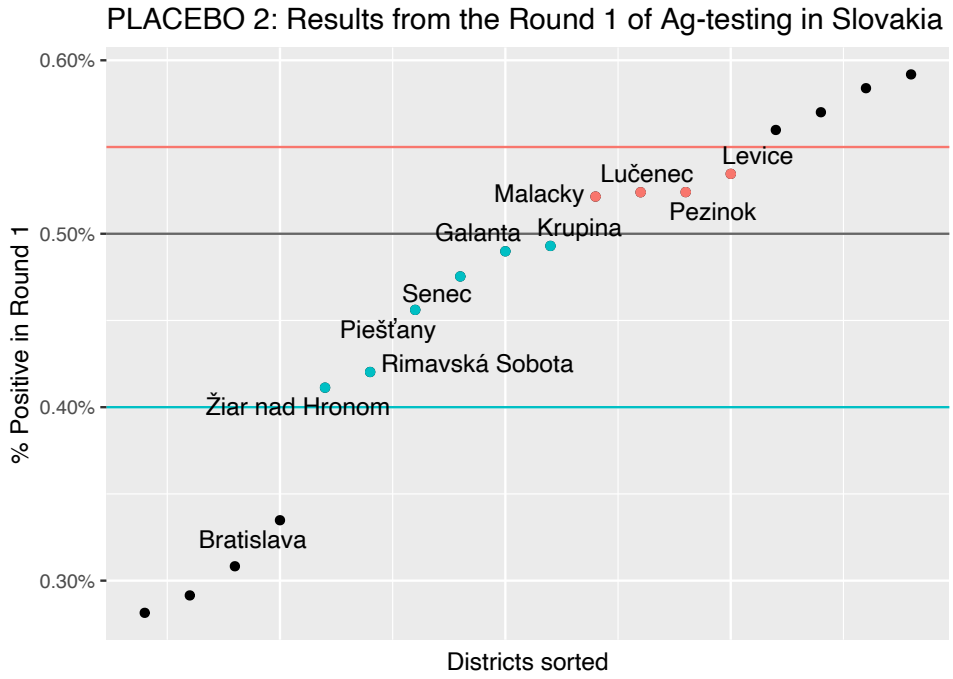


Figure B.5: Evolution of infections and R_0 in the districts below and above the Placebo threshold 0.5%.

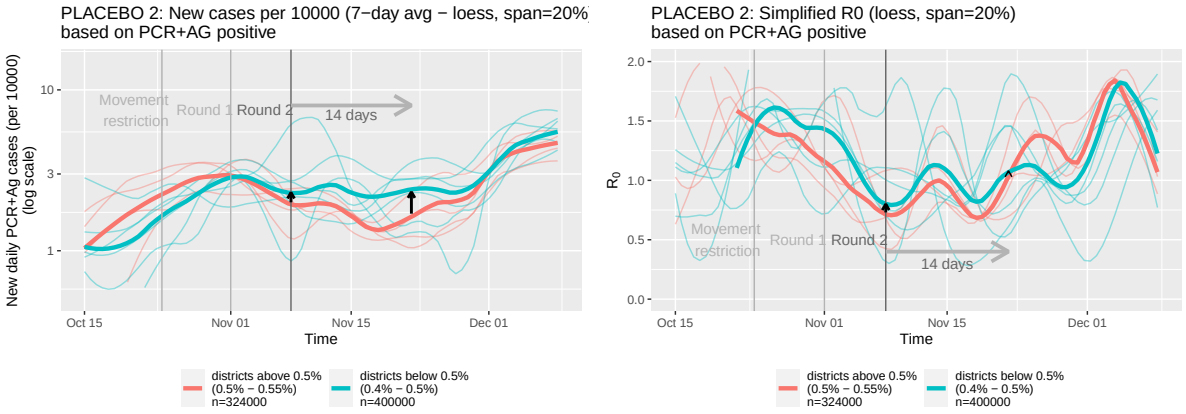
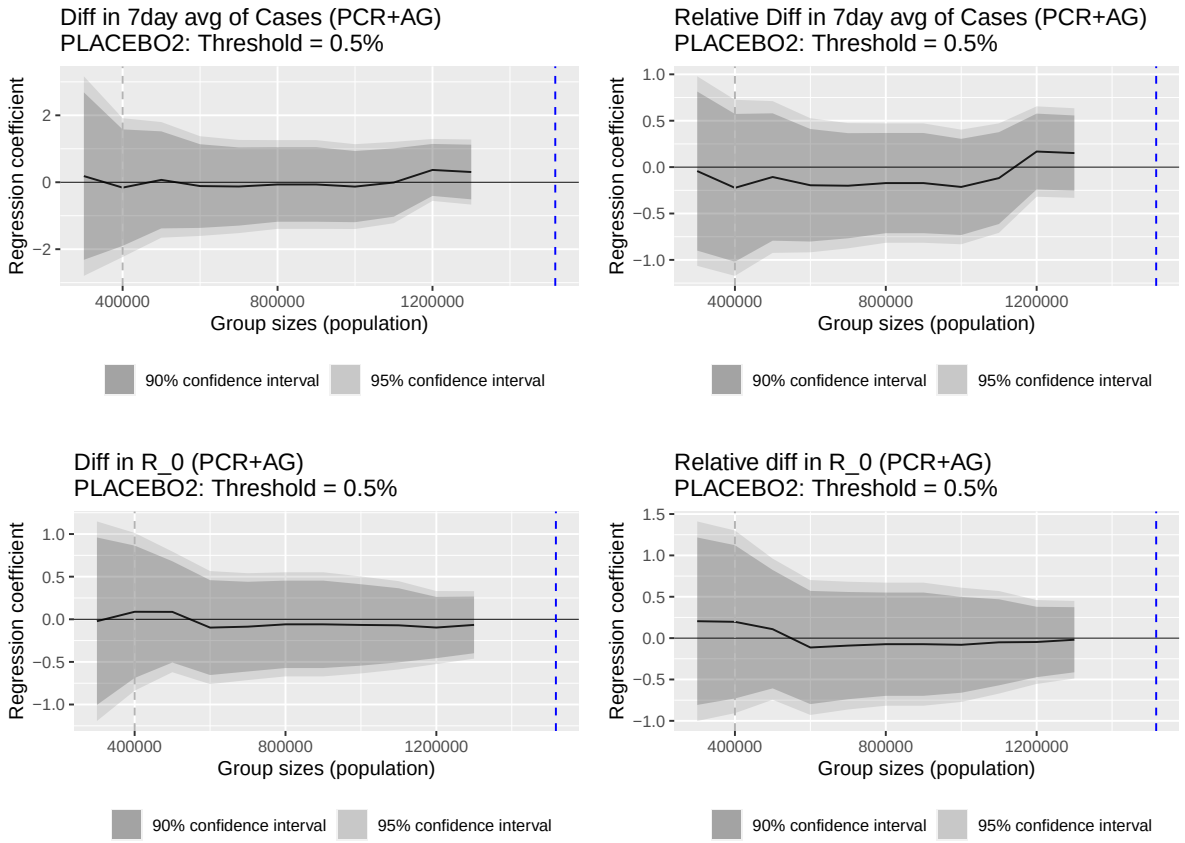


Table B.2: Placebo 2: Regression results based on eq (1)

	Dependent variable:			
	Cases	log Cases	R_0	log R_0
	(1)	(2)	(3)	(4)
Tested in R2 × Time	−0.298	−0.242	0.057	0.121
	(0.795)	(0.372)	(0.370)	(0.436)
Tested in R2	−0.454	−0.128	−0.169	−0.163
	(0.562)	(0.263)	(0.261)	(0.308)
Time	0.046	0.100	0.284	0.316
	(0.532)	(0.249)	(0.247)	(0.291)
(Intercept)	2.347***	0.727***	0.856***	−0.283
	(0.376)	(0.176)	(0.175)	(0.206)
Observations	20	20	20	20
R^2	0.135	0.122	0.177	0.168

Note: Districts weighted by their population size. *p<0.1; **p<0.05; ***p<0.01

Figure B.6: Regression coefficient as a function of maximal size of the groups below and above the threshold. Vertical dash line stands for the size of the placebo specification groups and vertical blue line depicts the size of the above group that does not include any district that was tested in R2. The results to the left of this blue line are all based on districts not tested in R2.



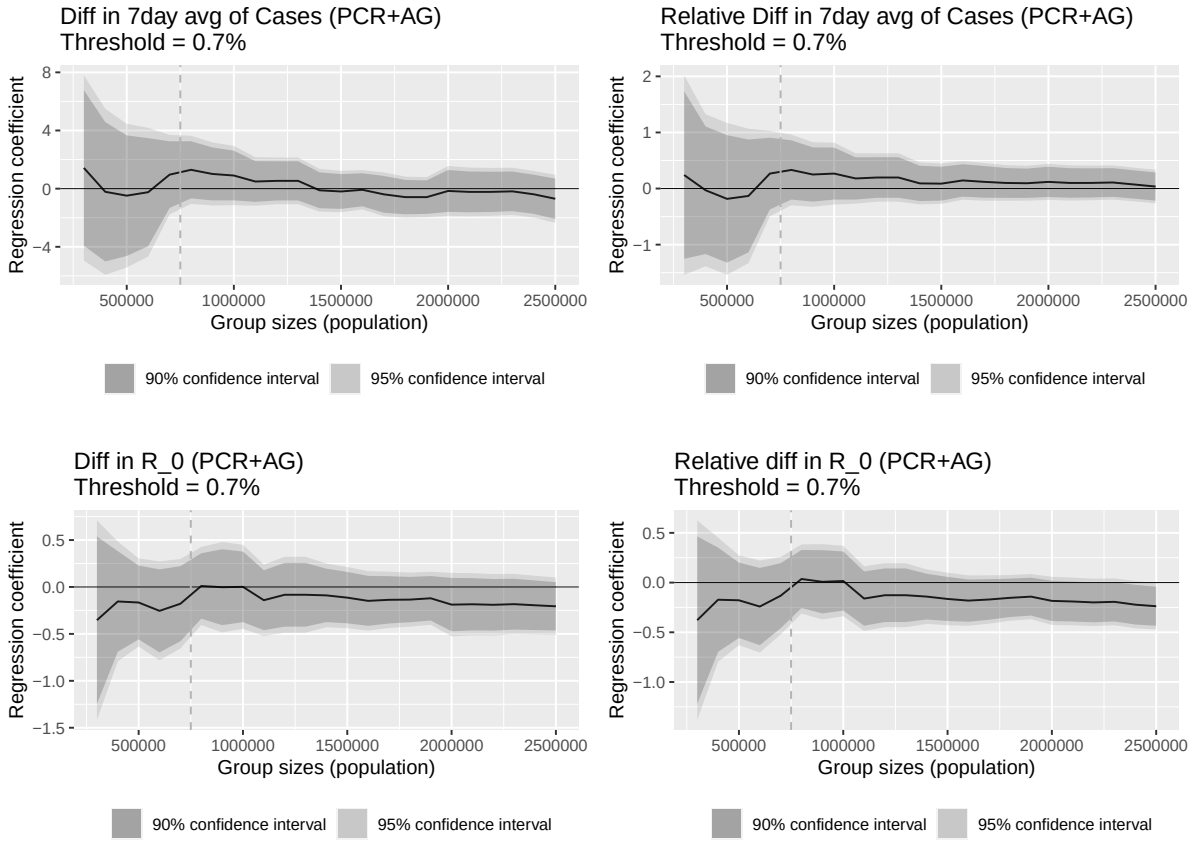
B.3 Placebo 3 - different treatment date (Nov 1)

Table B.3: Restricted Sample: Regression results based on eq (1) - different (incorrect) date for Round 2

	Dependent variable:			
	Cases	log Cases	R0	log R0
	(1)	(2)	(3)	(4)
Tested in R2 × Time	0.470	0.145	−0.109	−0.082
	(1.141)	(0.312)	(0.258)	(0.201)
Tested in R2	0.350	0.124	−0.138	−0.119
	(0.807)	(0.221)	(0.182)	(0.142)
Time	−1.572*	−0.402*	−0.337*	−0.291**
	(0.777)	(0.213)	(0.176)	(0.137)
(Intercept)	4.448***	1.393***	1.494***	0.385***
	(0.550)	(0.150)	(0.124)	(0.097)
Observations	28	28	28	28
R^2	0.223	0.211	0.324	0.360

Note: Districts weighted by their population size. *p<0.1; **p<0.05; ***p<0.01

Figure B.7: Estimated regression coefficient $\hat{\beta}_1$ with confidence intervals based on equation (1) as a function of the size of the groups below and above the threshold. Round 2 date was (incorrectly) set for Nov 1.



B.4 Results based on PCR tests only

Around the time of autumn mass-testing in 2020, Slovakia was on it's path of building permanent antigen testing sites, where citizens could get tested for free. The availability of these sites were increasing in time as documented in the right panel of Figure A.1. We included the antigen tests in our main analysis because of the variability in the antigen testing capacities on the district level. PCR and antigen tests however differ in terms of their properties (sensitivity and specificity) and logistics associated with their administration - in contrast to PCR tests, a citizen could walk in the antigen testing site and learn the test result within minutes. Given the different time lags we explore the results if we would base our calculations on the PCR tests only in Appendix B.4.

Tables and and Figure B.8 present results for the restricted sample described in Subsection 4.2. While the effect on the cases is statistically insignificant, the effect on R_0 is more pronounced at more than 0.4 Similar to our main specification, we investigated the sensitivity of these results to the size of the "below" or "above" groups. These results are presented in Figure B.9 and display similar patterns compared to the case with PCR and antigen tests combined.

Figure B.8: Evolution of infections and R_0 in the districts below and above the Placebo threshold 0.7% for the Restricted sample described in Subsection 4.2.

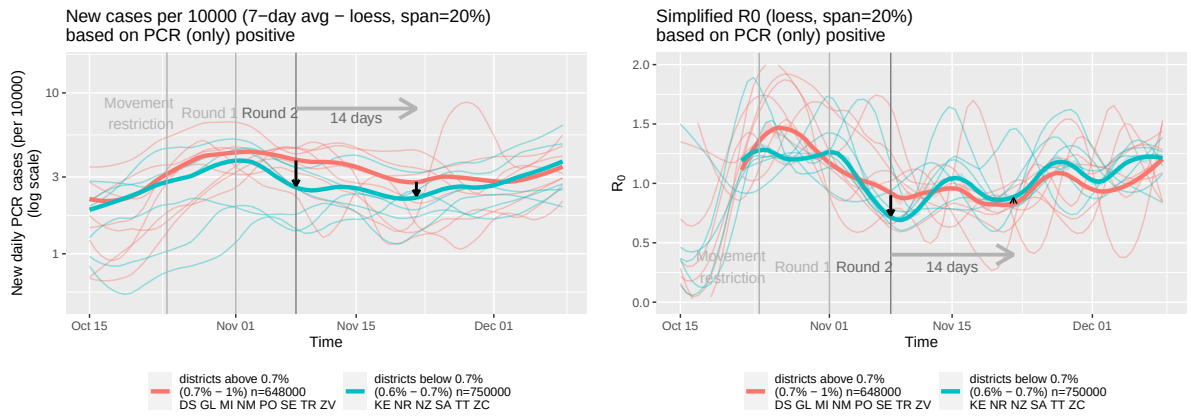
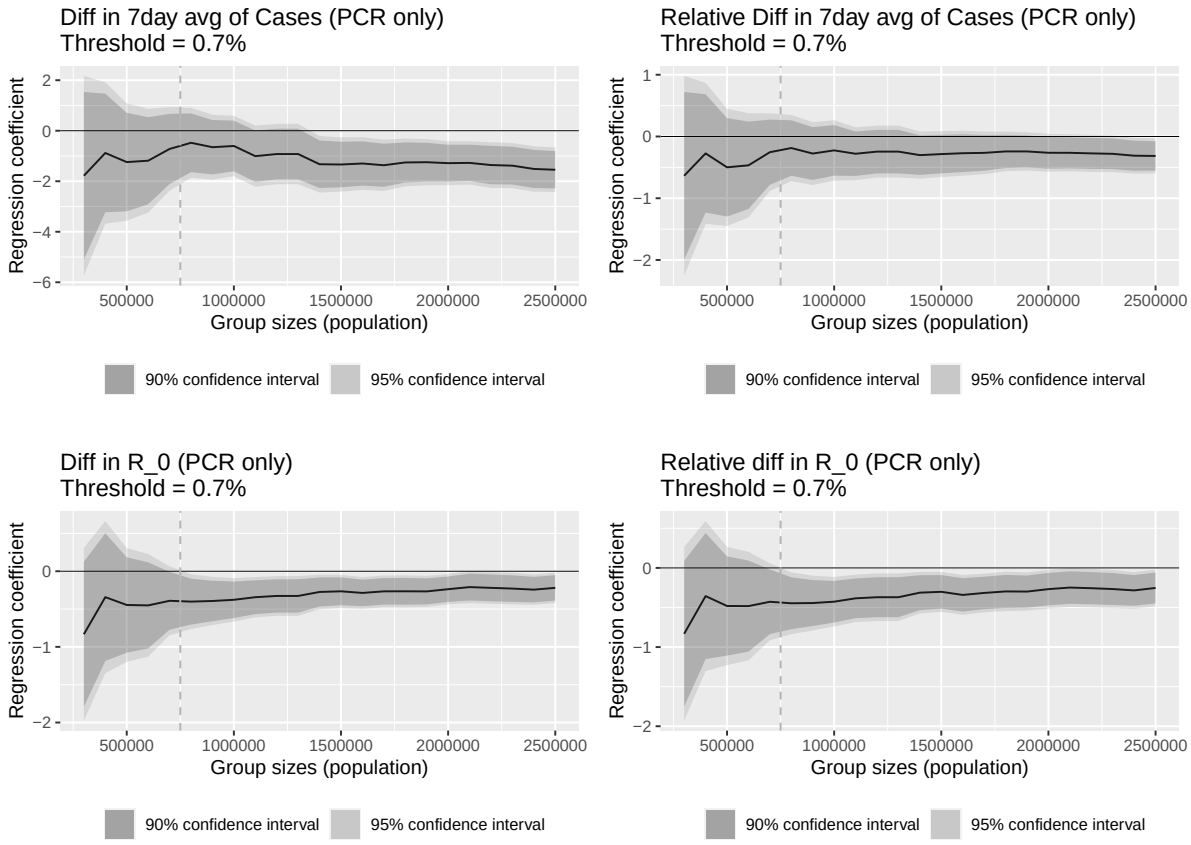


Table B.4: Restricted Sample: Regression results based on eq (1) for PCR cases only

	Dependent variable:			
	Cases	log Cases	R0	log R0
	(1)	(2)	(3)	(4)
Tested in R2 × Time	−0.741	−0.301	−0.387**	−0.420**
	(0.706)	(0.277)	(0.179)	(0.191)
Tested in R2	1.222**	0.431**	0.290**	0.286**
	(0.499)	(0.196)	(0.126)	(0.135)
Time	−0.299	−0.098	0.152	0.189
	(0.481)	(0.189)	(0.122)	(0.130)
(Intercept)	2.558***	0.876***	0.735***	−0.321***
	(0.340)	(0.134)	(0.086)	(0.092)
Observations	28	28	28	28
R^2	0.299	0.255	0.199	0.185

Note: Districts weighted by their population size. *p<0.1; **p<0.05; ***p<0.01

Figure B.9: Estimated regression coefficient $\hat{\beta}_1$ with confidence intervals based on equation (1) as a function of the size of the groups below and above the threshold. PCR tests only were used in calculations.



B.5 Results with workplace mobility included as a regressor

In order to explore if different mobility patterns in the regions can predict our outcomes of interest we employed the following model specification

$$y_{it} = \beta_0 + \beta_1(above_i \cdot t) + \beta_2 t + \beta_3 above_i + \beta_4 mob_i + \epsilon_{it}, \tag{2}$$

where mob_i stands for 7day rolling average of Google workplace mobility measure at $t = 0$ (Nov 8 2020) and $t = 1$ (Nov 22 2020) (GoogleLLC, 2021).

We present results for both the full sample and for the restricted sample, because the sample size is lower in both cases as for some districts this measure is missing. Apart from the *Workplace* mobility, the rate of missing data on a district level was about 50% for all the other types of Google mobility measures (*Retail and Recreation*, *Grocery and Pharmacy*, *Parks*, *Transit Stations*, *Residential*).

We estimate qualitatively similar results and while mobility is insignificant. This is not surprising as we observe little differences in the mobility patterns in the different districts as presented on Figure B.11.

Table B.5: Full Sample: Regression results based on eq (1) with Mobility

	Dependent variable:			
	Cases	log Cases	R0	log R0
	(1)	(2)	(3)	(4)
Tested in R2 × Time	−2.537***	−0.415***	−0.310***	−0.329***
	(0.719)	(0.154)	(0.108)	(0.121)
Tested in R2	4.316***	1.029***	0.202**	0.250***
	(0.513)	(0.110)	(0.077)	(0.087)
Time	−0.075	−0.038	0.300***	0.324***
	(0.516)	(0.111)	(0.078)	(0.087)
Mobility	−0.003	0.010	0.0002	−0.003
	(0.052)	(0.011)	(0.008)	(0.009)
(Intercept)	2.152	1.025***	0.786***	−0.375
	(1.442)	(0.310)	(0.217)	(0.243)
Observations	104	104	104	104
R^2	0.496	0.583	0.145	0.149

Note: Districts weighted by their population size. *p<0.1; **p<0.05; ***p<0.01

Table B.6: Restricted Sample: Regression results based on eq (1) with Mobility

	Dependent variable:			
	Cases	log Cases	R0	log R0
	(1)	(2)	(3)	(4)
Tested in R2 × Time	−0.926	−0.319	−0.319	−0.360*
	(1.021)	(0.320)	(0.193)	(0.201)
Tested in R2	1.408*	0.457*	0.286*	0.320**
	(0.748)	(0.234)	(0.142)	(0.147)
Time	0.490	0.177	0.379**	0.445***
	(0.745)	(0.234)	(0.141)	(0.146)
Mobility	0.053	0.009	0.012	0.010
	(0.115)	(0.036)	(0.022)	(0.023)
(Intercept)	4.146	1.194	1.001*	−0.132
	(2.970)	(0.931)	(0.562)	(0.583)
Observations	22	22	22	22
R^2	0.231	0.223	0.364	0.422

Note: Districts weighted by their population size. *p<0.1; **p<0.05; ***p<0.01

Figure B.10: Estimated regression coefficient $\hat{\beta}_1$ with confidence intervals based on equation (2) as a function of the size of the groups below and above the threshold.

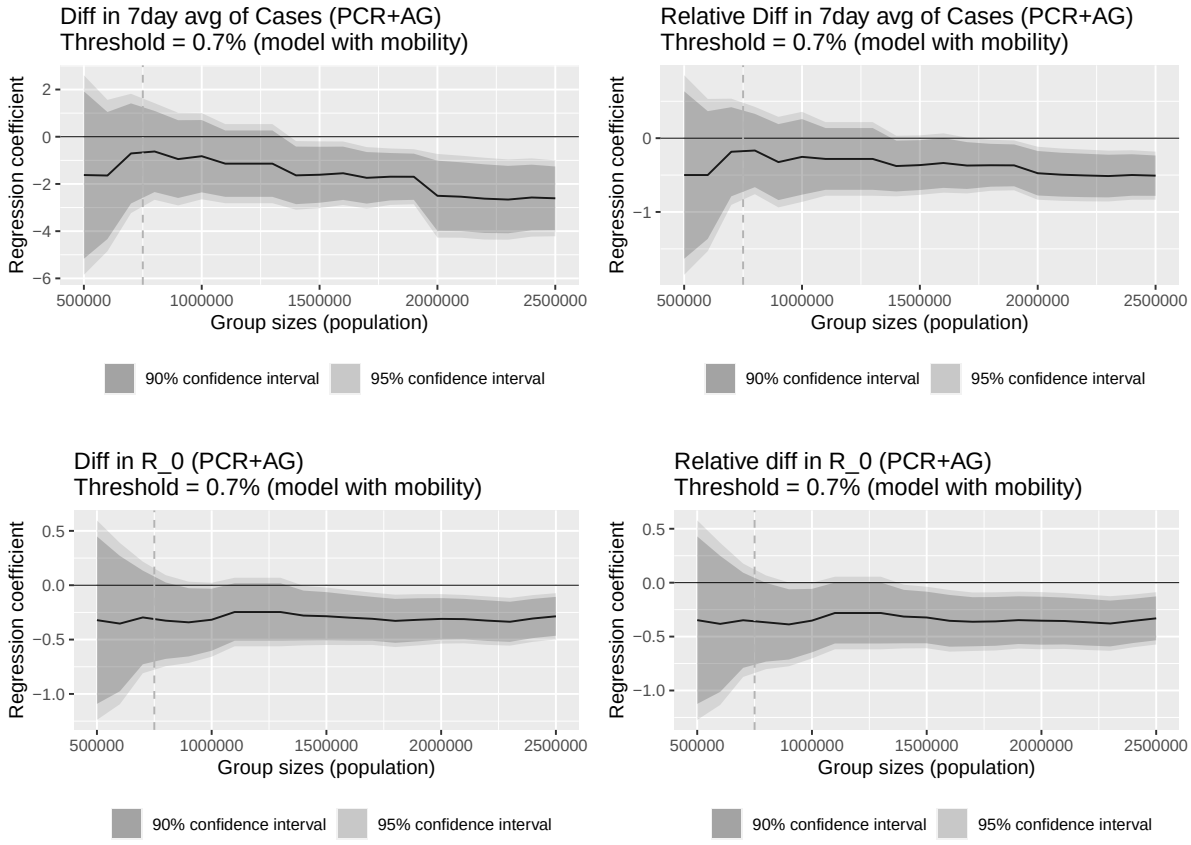


Figure B.11: Google mobility data - Transport.