Center for Research On the Wisconsin Economy
University of Wisconsin- Madison

The Impact of Statewide Stay-at-Home Orders: Estimating the Heterogeneous Effects Using GPS Data from Mobile Devices

Junjie Guo

Center for Research on the Wisconsin Economy, UW-Madison

April 29, 2020

Abstract

In response to the spread of COVID-19, most states in the U.S. have ordered non-essential businesses to close and residents to not leave home unless necessary. This paper estimates the impact of these stay-at-home orders on the mobility of Americans using GPS data from mobile devices. Recognizing the heterogeneity in both mobility and the orders across states, I use the synthetic control method at the state level to estimate the impact of individual orders instead of pooling them together for an average effect. I find the impact does vary significantly across orders/states. For example, the estimates suggest that the orders in Michigan and Wisconsin increased the fraction of their residents at home all day by about 5.5 and 4 percentage points, respectively, while the corresponding estimate for Ohio is small and insignificant. In addition to the effectiveness of the orders in limiting the spread of COVID-19, these estimates are also informative of the responsiveness of Americans when the orders would be lifted.

1 Introduction

In response to the spread of COVID-19, most states in the U.S. have ordered non-essential businesses to close and residents to not leave home unless necessary. Although public health experts have argued that such dramatic measures are necessary, and there is suggestive evidence that social distancing is working in slowing down the spread of the virus, the negative impact of locking down the economy is also significant, as suggested by the record number of initial unemployment insurance applications. Partly because of this negative impact, some state governors have resisted the call to issue a stay-at-home order, and states with such an order in place are under pressure to lift the ban and reopen the economy.[1]

The impacts of stay-at-home orders on both the spread of COVID-19 and the economy depend on how effective they are in limiting the mobility of Americans, which is a question without an obvious answer. Partly because these orders are not strictly enforced, not everyone is taking them seriously, as suggested by many media reports about people ignoring the orders and leaving home unnecessarily.[2] Moreover, as these orders were issued when COVID-19 was already widespread and were preceded by other social distancing guidelines, it's likely that many Americans had already chosen to stay home as much as possible by the time these orders were issued, leaving little room for them to have a large impact. Finally, even if we observe a change in mobility after an order went into effect, it's possible that part of the change is a voluntary response to the continuing spread of COVID-19 instead of a direct impact of the order.

This paper estimates the impact of statewide stay-at-home orders on the mobility of Americans using GPS data from mobile devices (almost all of which are cellphones). Specifically, I use a daily measure of the percentage of mobile devices without a GPS ping observed from outside their home location as a proxy for the fraction of Americans at home all day. I will refer to this measure as the home rate, and study how it responds to stay-at-home orders. Instead of pooling the orders together for an average effect, I take each order as a different treatment and estimate its impact separately. This is important because, as shown later in the paper, both mobility and the orders are quite heterogenous across states.

[1] "Dr. Anthony Fauci, director of the National Institute of Allergy and Infectious Diseases, suggested Thursday that the federal government should impose a nationwide stay-at-home order to help prevent the spread of the coronavirus in the United States." https://www.politico.com/news/2020/04/03/fauci-endorses-national-stay-at-home-order-162794

"Fauci says coronavirus hospitalizations are dropping because social distancing is working". https://www.cnn.com/2020/04/09/health/us-coronavirus-thursday/index.html

"U.S. jobless claims reach 26 million since coronavirus hit, wiping out all gains since 2008 recession". https://www.nbcnews.com/business/business-news/u-s-jobless-claims-reach-26-million-coronavirus-hit-wiping-n1190296

"Governors of a half-dozen states worry about the economic fallout of forcing businesses to close". https://www.wired.com/story/why-some-states-resisting-social-distancing-measures/

"Coronavirus: US faced with protests amid pressure to reopen". https://www.bbc.com/news/world-us-canada-52348288

[2] "Crowds ignore social distancing rules to watch USNS Comfort". https://nypost.com/2020/03/30/crowds-ignore-social-distancing-rules-to-watch-usns-comfort/

Empirically, I use the synthetic control method to estimate the impact of each order at the state level. Specifically, for each state (referred to as the treated state) with a stay-at-home order, I focus on the average impact in the first seven days since the order went into effect. This allows me to use the states without an order by the seventh day as potential controls. The synthetic control is then a weighted average of the potential controls, where the weights are chosen such that the synthetic control mimics the behavior of the treated state before the order was announced. Under the assumption that the outcome variable in the treated state would evolve as it does in the synthetic control in the absence of the order, the impact of the order is estimated by the differences in the outcome variable between the treated state and the synthetic control after the order went into effect.

For example, Wisconsin's Safer-at-Home order was issued on March 24 and went into effect on March 25 after Governor Evers announced his plan to do so on March 23. To estimate its impact on the home rate in the first seven days (March 25-31), I use the 15 states without an order announced by March 31 as potential controls, and select the weights such that the resulting synthetic control matches Wisconsin as close as possible before March 23 in terms of both the home rate and some other predictors including the number of confirmed cases of COVID-19. The differences in the home rate between Wisconsin and the synthetic control in the period of March 25-31 are interpreted as the impact of the order in the first seven days. On average, I find the difference is about 4 percentage points. That is, Wisconsin's Safer-at-Home order is associated with an increase in the fraction of Wisconsinites (or those with cellphones) at home all day by about 4 percentage points. To put this number into context, the home rate in Wisconsin increased by 20 percentage points from its pre-pandemic level of 19 percent on Tuesday March 10 to 39 percent on Tuesday March 31. So the Safer-at-Home order contributed about 20 percent to this increase.

After showing the estimate is statistically significant, I provide some evidence suggesting it reflects a causal impact of the Safer-at-Home order. Specifically, I run a placebo test by back-dating the order's effective date to March 16, one week before it was announced on March 23. After using the data before March 16 to construct a new synthetic control, I find no significant difference between Wisconsin and the new synthetic control in the period of March 16-22, suggesting the (new) synthetic control does a good job of approximating Wisconsin's behavior in the absence of the Safer-at-Home order.

The home rate is obviously an extreme measure of (im)mobility. To provide additional evidence, I use another daily measure likely at the other extreme: the percentage of mobile devices that spent 3 or more hours at a location other than their home during the period of 8 am - 6 pm in local time. As these devices are most likely to be associated with work-related behaviors, I will refer to this measure as the work rate. Applying the synthetic control method to this new outcome variable, I find Wisconsin's Safer-at-Home order is associated with a reduction in the work rate by about 2 percentage points.

Applying the same method to stay-at-home orders in other states, I find significant heterogeneity in their impacts. For example, the estimates suggest the order in Michigan raises the home rate and reduces the work rate by about 5.5 and 2.8 percentage points, respectively, while both estimates for Ohio are small economically and insignificant statistically. This is striking given that the orders in the two neighboring states went into effect on the same day March 24 after they were issued within a short span of less than 24 hours. The results, however, are supported by a direct comparison between the two states.

Specifically, I find there was essentially no systematic difference in the home rate between the two states before the orders were issued. Then there was an almost discrete jump in the difference in the few days when the orders were announced and subsequently went into effect. The size of the jump is about 5 percentage points in favor of Michigan, very close to the difference in the estimated impacts between the two states. Since the orders went into effect on March 24, the difference in the home rate between the two states has been fluctuating around 5 percentage points without any increasing or decreasing trend. This is in sharp contrast to the difference in the confirmed cases of COVID-19 between the two states, which has been increasing continuously since March 18, about a week before the orders went into effect. This contrast suggests that the jump is most likely a result of the heterogenous impacts of the two orders. This conclusion is also supported by evidence from weekly initial unemployment insurance claims, the number of which was smaller in Michigan than in Ohio in the weeks before the orders in the two states were issued but has been much larger in Michigan than in Ohio since the orders were issued. This suggests that the order in Michigan is indeed more restrictive, which could account for its large impact on mobility.

In short, this paper finds significant heterogeneity in the impact of stay-at-home orders across states. This could arise from either the differences in the orders themselves, like which activities are exempt and which are not, or the differences in enforcement and compliance, or both. A detailed investigation of the causes, while beyond the scope of this paper, should be an important direction for future work. Among other things, the investigation could be informative of how to make the right policy to achieve the desired impact.

In addition to the effectiveness of the orders in limiting the spread of COVID-19, the estimates in this paper are also informative of the responsiveness of Americans when the orders are lifted. For example, relative to Ohio where the estimated impact of the order is small and insignificant, residents in states like Michigan and Wisconsin where the estimated impact is much larger would probably be more responsive and engage in more economic and social activities outside home when the order is lifted. Whether and when such increased activities are desirable should be a key factor in determining when and how to reopen the economy.

The next section discusses the relevant literature. The three sections after that present the data, the empirical strategy, and the results, respectively. The final section concludes the paper by discussing some of the implications.

2 Related Literature

Originally proposed in Abadie and Gardeazabal (2003) and Abadie et al. (2010) with the aim to estimate the effects of aggregate interventions, the synthetic control method has been described by Athey and Imbens (2017) as "arguably the most important innovation in the policy evaluation literature in the last 15 years". Abadie (forthcoming) provides a detailed discussion of the method and the related literature.

GPS and related data are being used by more and more studies. For example, the SafeGraph data used in this paper has been used by Chen and Rohla (2018) and Athey et al. (2019) to measure the effects of political polarization on the length of Thanksgiving dinners and to estimate a novel measure of racial segregation, respectively. These studies also demonstrate that the data is well balanced across U.S. demographics and space. Due to its wide, balanced and timely coverage, and the fact that it's now freely available to the research community, the SafeGraph data is being used by numerous researchers addressing various questions related to the COVID-19 pandemic. For example, Allcott et al. (2020) use the data to study partisan differences in social distancing, Benzell et al. (2020) use the data to study how to ration social contact during the pandemic by shutting down some locations while leaving others open, Farboodi et al. (2020) use the data to study the internal and external effects of social distancing in a pandemic, and Williams (2020) uses the data to measure the impact of the pandemic on the Wisconsin economy.

The mobility impact of stay-at-home orders is also addressed by Engle et al. (2020) and Painter and Qiu (2020), with the later using the same SafeGraph data as I do in this paper. Both papers pool the orders together for an average effect, which essentially is identified by comparing states with a stay-at-home order with those without. The underlying assumption is that, in the absence of a stay-at-home order, states with and without a stay-at-home order are comparable with each other, at least in terms of changes in mobility if a fixed effect is included. In the next section, I provide some evidence against this assumption by showing that (1) states experienced different growth in mobility before the nation's first stay-at-home order was issued in California, which is true even conditional on changes in confirmed cases of COVID-19, and (2) the different growth is correlated with the timing of when a state issued a stay-at-home order. Consequently, I use the synthetic control method to address the heterogeneity and estimate the impact for individual orders.

3 Data

Table 1 lists the states with and without a stay-at-home order. For states with an order, it also lists the date when the order was announced and, in parentheses, the date when the order went

Table 1: Statewide Stay-at-Home Orders

Announcement date	State abbreviation (Effective date)
3/19	CA (3/20)
3/20	IL (3/22), NY (3/23), CT (3/24)
3/21	NJ (3/22)
3/22	DE (3/24), LA (3/24), OH (3/24)
3/23	MI (3/24), NM (3/24), OR (3/24), WA (3/24)
	HI (3/25), IN (3/25), MA (3/25), WV (3/25), WI (3/25)
3/24	VT (3/26)
3/25	CO (3/26), ID (3/26)
3/26	KY (3/27), MN (3/28), MT (3/28), NH (3/28)
3/27	AK (3/29), NC (3/31)
3/28	RI (3/29), KS (3/30)
3/30	MD (3/31), VA (3/31), AZ (4/1), DC (4/1), TN (4/1)
3/31	NV (4/1), ME (4/2), TX (4/2)
4/1	PA (4/2), FL (4/3), GA (4/4), MS (4/4)
4/3	AL (4/5), MO (4/6)
4/6	SC (4/8)
No order	AR, IA, ND, NE, OK, SD, UT, WY

into effect. The table is based on the information collected by the New York Times,[3] with two modifications. First, the effective date is the first full day when an order became effective. Specifically, if an order went into effect after 8 am on a particular day, the next day is recorded as the effective date. Secondly, the announcement date is actually the date when an order was first reported. For most states, it's the same as the date when the order was signed. However, for states like Wisconsin where the order was announced and reported before it was signed, I use the announcement date instead. The name of the order varies across states. To avoid confusion, I use the name stay-at-home to discuss a generic order, and use the specific name when discussing the order in a particular state, like the Safer-at-Home order in Wisconsin.

To measure the mobility of Americans, I use the Social Distancing Metrics (SDM) from SafeGraph.[4] The data is aggregated from GPS pings of tens of millions of mobile devices, almost all of which are cellphones. For each mobile device, a common nighttime location is determined over a 6 week period to a Geohash-7 granularity (~153m x ~153m). For ease of reference, this common nighttime location is referred to as the device's "home". I use four variables available daily at the census blog group (CBG) level: (1) Device count: the number of mobile devices homed in each census block group; (2) Completely home device count: Out of the device count, the number of devices which did not leave the Geohash-7 in which their home

[3]See Which States and Cities Have Told Residents to Stay at Home. By Sarah Mervosh, Denise Lu and Vanessa Swales. Updated April 7, 2020. https://www.nytimes.com/interactive/2020/us/coronavirus-stay-at-home-order.html

[4]Attribution: SafeGraph, a data company that aggregates anonymized location data from numerous applications in order to provide insights about physical places. To enhance privacy, SafeGraph excludes census block group information if fewer than five devices visited an establishment in a month from a given census block group. Documentation for the SDM data can be found at https://docs.safegraph.com/docs/social-distancing-metrics.

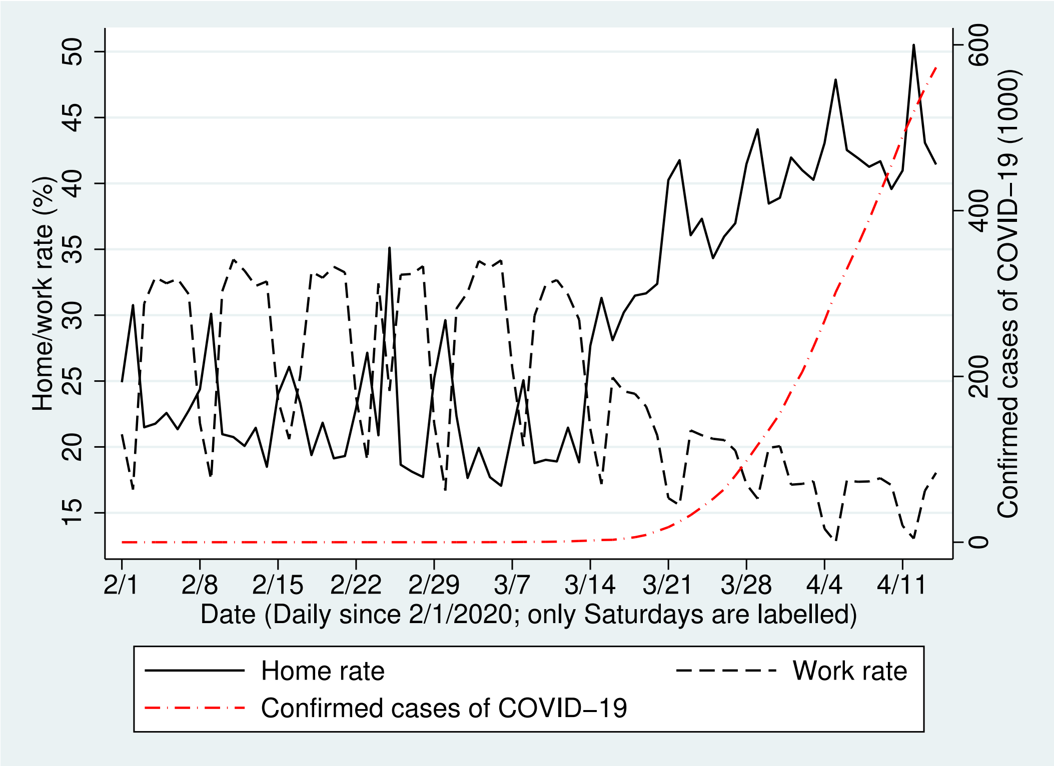

Figure 1: Home Rate, Work Rate and COVID-19 in the U.S.

is located; (3) Part-time work behavior devices: Out of the device count, the number of devices that spent one period of between 3 and 6 hours at one location other than their Geohash-7 home during the period of 8 am - 6 pm in local time; and (4) Full-time work behavior devices: Out of the device count, the number of devices that spent greater than 6 hours at a location other than their home Geohash-7 during the period of 8 am - 6 pm in local time. I aggregate all four numbers to the state level, and use the percentage of devices that are completely home as a proxy of a state's population at home all day. Referred to as the home rate hereafter, this is the key outcome variable of interest in this paper. For additional evidence, I also use the fraction of devices that exhibit (either part-time or full-time) work behavior as a proxy for a state's population at work, and refer to it as the work rate hereafter.

Figure 1 plots the home rate, the work rate and the (cumulative) confirmed cases of COVID-19 for the U.S. from February 1 to April 14, the seventh day since the last stay-at-home order went into effect in South Carolina. The confirmed cases of COVID-19 is obtained from the data collected by the Center for Systems Science and Engineering at John Hopkins University.[5]

The figure reveals several messages. First, as expected, there are significant and systematic within-week fluctuations in both the home and the work rate, with more people away from work and at home during the weekend. Aside from these fluctuations, the two rates were relatively stable in February and the first two weeks of March, with the home rate at around 20 percent and the work rate slightly below 35 percent during weekdays. The latter is much lower than

[5]https://github.com/CSSEGISandData/COVID-19

the employment-population ratio, which was 61.1 percent in February before dropping to 60 percent in March.[6] This is largely because, by requiring a mobile device to spend 3 or more hours in a location other than home, the work rate misses a large chunk of the labor force that are either constantly on the move or don't have to move at all. For example, bus, taxi and truck drivers are unlikely to be counted towards the work rate, neither are computer programmers and customer service representatives working exclusively from home. So we shouldn't pay much attention to its level, and I am only using its change associated with stay-at-home orders for additional evidence.

Secondly, starting from around March 14, the home rate increased and the work rate decreased rapidly for over two weeks. Both seem to have stabilized by the end of March, with the home rate now above 40 percent and the work rate around 17 percent. So the home rate has doubled and the work rate halved from their respective levels before mid-March. In addition to the direct impact of COVID-19, where the number of confirmed cases was relatively flat and small before mid-March but has been increasing exponentially since then, the structural changes in the home and the work rate around mid-March are also likely to be precipitated by the White House's declaration of COVID-19 as a national emergency on March 13 and the subsequent social distancing guidelines on March 16. Importantly for this paper, by the time California issued the nation's first statewide stay-at-home order on March 19, COVID-19 was already widespread, and both the home and the work rate had changed a lot from their pre-pandemic levels. This may have left little room for stay-at-home orders to have a large impact. Moreover, the continuing and dramatic increase in the confirmed cases of COVID-19 since March 19 suggests that not all of the changes in the home and the work rate since then could be attributed to the stay-at-home orders.

One way to estimate the impacts of stay-at-home orders is to compare states with an order with those without. However, a simple comparison is unlikely to work because the two groups of states are potentially very different. For example, figure 2 plots the growth in the home rate and the confirmed cases of COVID-19 for each state and the District of Columbia from Wednesday March 4 to Wednesday March 18, the day before the nation's first statewide stay-at-home order was issued in California. Clearly, states already had different growth in the home rate before stay-at-home orders were issued, which is true even conditional on the growth in the confirmed cases of COVID-19. Specifically, the growth in the home rate during the two weeks was around or above 15 percentage points in each of the first five states with a stay-at-home order (CA, IL, NY, CT and NJ), which is way above the growth in most other states. It's thus unlikely that a comparison between the two groups of states would produce unbiased estimates of the impact of stay-at-home orders.

[6] https://www.bls.gov/charts/employment-situation/employment-population-ratio.htm

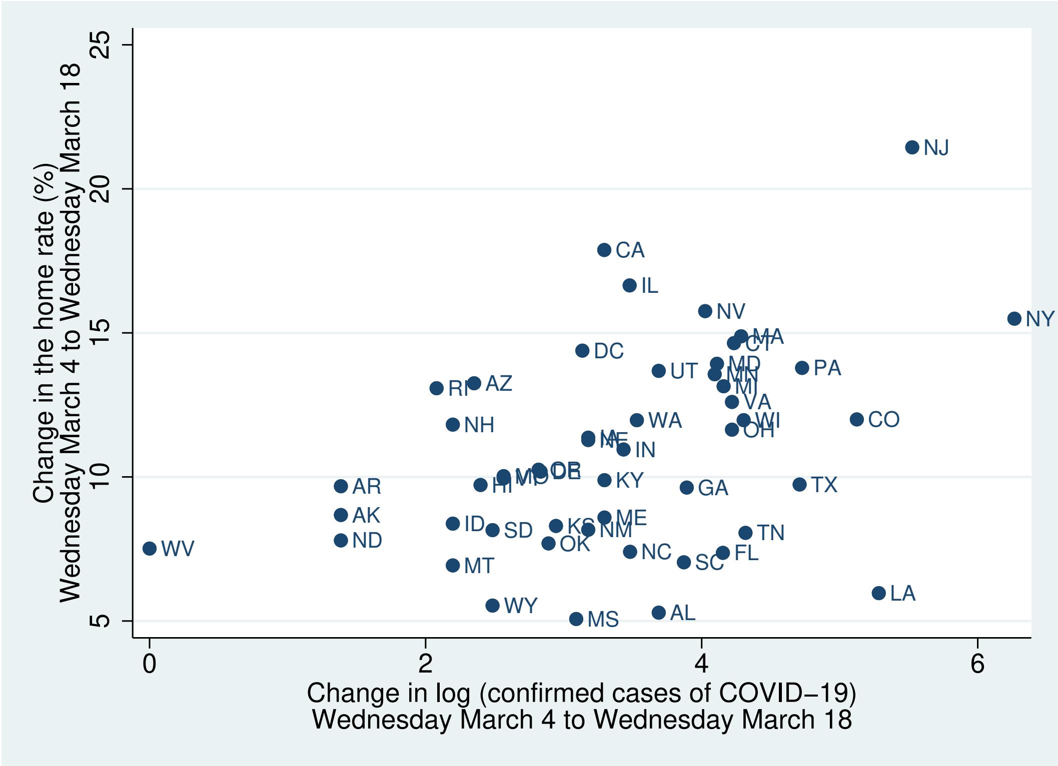

Figure 2: Growth in COVID-19 and the Home Rate Across States

4 Empirical Strategy

While states with and without a stay-at-home order may not be comparable either in general or individually, a weighted average of some of the states without a stay-at-home order may be comparable with another state in the absence of a stay-at-home order. This is the idea behind the synthetic control method, the empirical strategy used in this paper.

4.1 The Estimator

Formally, assume a state $s = 1$ issued a stay-at-home order on day d_a that went into effect on day d_1, and our goal is to estimate its impact in the first seven days from d_1 to d_7. We will also refer to this state as the treated state. For potential controls, we use states without a stay-at-home order announced by day d_7. Assume there are S of them, indexed by $s = 2, 3, \ldots, S+1$. Let y_{sd} be an outcome of interest in state s on day d. The average impact of the order in state $s = 1$ in the first seven days is estimated with the following synthetic control estimator

$$E_y = \frac{1}{7} \sum_{d=d_1}^{d_7} (y_{1d} - y_{0d}) \tag{1}$$

with

$$y_{0d} = \sum_{s=2}^{S+1} w_s y_{sd} \tag{2}$$

where w_s is the weight of state $s \geq 2$ subject to $w_s \in [0,1]$ and $\sum_{s=2}^{S+1} w_s = 1$. Let $X_s = \{x_{s1}, x_{s,2}, ..., x_{sK}\}$ be a set of K predictors of the outcome variable in state s, which, importantly, could include the values of the outcome variable before the order, i.e., y_{sd} for $d < d_a$. The weights $\{w_s\}$ are chosen to minimize

$$\sum_{k=1}^{K} v_k \left(x_{1k} - \sum_{s=2}^{S+1} w_s x_{sk} \right)^2 \tag{3}$$

where $\{v_1, v_2, ..., v_K\}$ is a vector of non-negative constants that in turn could be chosen by minimizing the root mean squared prediction error (RMSPE) before the order was announced

$$RMSPE_b = \left[\frac{1}{d_a - 1} \sum_{d=1}^{d_a - 1} (y_{1d} - y_{0d})^2 \right]^{\frac{1}{2}} \tag{4}$$

Mathematically, given a set of $\{v_1, v_2, ..., v_K\}$, we can use expression (3) to select the weights $\{w_s\}$ and use equation (2) to calculate y_{0d}. With y_{0d}, we can calculate $RMSPE_b$ given by expression (4). This process is repeated until $RMSPE_b$ is minimized, and the resulting y_{0d} is used to calculate the impact of the order E_y defined in expression (1).

Intuitively, the method constructs a synthetic control by weighting the potential controls properly such that the synthetic control and the treated state look similar before the order in the treated state was announced in terms of a set of predictors that may include the values of the outcome variable itself before the order was announced. Under the assumption that the outcome variable in the treated state would evolve as it does in the synthetic control in the absence of the order, the impact of the order is measured by the average differences in the outcome variable between the treated state and the synthetic control in the first seven days since the order went into effect.

Although no state used to construct the synthetic control had a stay-at-home order by the end of the sample d_7, it's possible that some of them took other measures to combat COVID-19 which affected their home rates between d_1 and d_7. In this sense, the estimator E_y is about the incremental impact of the stay-at-home order above and beyond these other measures, just as it is the incremental impact above and beyond other measures taken by the treated state before the order. The estimator E_y would be biased if the treated state implemented additional measures during the period of d_1 to d_7 that were more restrictive than the stay-at-home order. I am aware of no such additional measures, and stay-at-home orders are so far the most restrictive social distancing policy in the U.S.

Note that the data between the day of announcement d_a and the day before the order went into effect $d_1 - 1$ are not used to construct either the weights or the estimator. This allows the behavior of the outcome variable y during this period of time to be different from either before the order was announced or after the order went into effect or both, which could happen, for example, if consumers began stocking up for the lockdown during this period of time.

In practice, in addition to y_{sd} for $d < d_a$, the set of predictors X_s also includes the daily infection rate c_{sd} for $d < d_a$ defined as the number of confirmed cases of COVID-19 per 10,000 population, population density, the fraction of the population that are white, the fraction of the population that are female, the fraction of the population that are foreign born, the fraction of the population that are 29 years old or younger, the fraction of the population that are 65 years old or older, the fraction of the population without a high school diploma, the fraction of the population without any college degree, the fraction of the population in rural areas, and the share of the state's votes that went for Donald Trump in 2016.[7] These variables are included because they are potentially correlated with changes in the outcome variable (the home rate and the work rate). For example, age matters because older people are more vulnerable to COVID-19 and thus could be more vigilant by staying home, education matters because less educated workers are more likely to in occupations where the job cannot be done remotely from home, and political affiliation matters because it's correlated with individual beliefs about the threat of COVID-19.

Theoretically, we could either normalize $y_{sd} = 0$ for all s on a particular day d, or remove the day-to-day fluctuations within a week, or do both. In practice, I choose to do neither for two reasons. First, it's possible that changes in the outcome variable y depends on its level. For example, an increase in the home rate from 20 to 30 percent may not be the same as an increase from 30 to 40 percent, which would be ignored by a normalization. Secondly, it seems from figure 1 that the within-week fluctuations in both the home and the work rate may be weaker in April than in February. Part of this may be a direct impact of the stay-at-home orders, which is a case against removing the fluctuations. Alternatively, the estimator in expression (1) addresses the within-week fluctuations by averaging over the first seven days since a stay-at-home order went into effect. It should be noted, however, that the results reported below are robust at least qualitatively to both choices.

By restricting the weight for each state to be positive $w_s \in [0,1]$, the synthetic control estimator precludes extrapolation, and thus avoids using the potential controls that are completely dissimilar to the treated state. This is one of the advantages of the synthetic control method relative to conventional methods like regressions which do allow for extrapolation. This restriction means that we may not always find a good synthetic control for a treated state, which happens when a treated state is an extreme outlier along some dimensions compared with the potential controls. Flagged by a large $RMSPE_b$, this transparency is another advantage of the synthetic control method because we should caution against any estimators involving an outlier. In practice, I will focus on the 48 contiguous states because they are likely more comparable with each other, and I will only consider cases with a relatively small $RMSPE_b$. That is, I will not report the less credible estimates where the treated state is significantly different from the potential controls.

[7] Except for COVID-19, all other variables are downloaded from the MIT Election Lab at https://github.com/MEDSL/2018-elections-unoffical/blob/master/election-context-2018.md. The demographic variables there are calculated from the American Community Survey. The variables are originally available at the county level, and then aggregated to the state level.

4.2 Backdating

As a placebo test, I will backdate the effective date d_1 of a stay-at-home order to $d_1^b = d_a - 7$, and use the data before d_1^b to construct a new synthetic control $y_{0d}^b = \sum_{s=2}^{S+1} w_s^b y_{sd}$. This allows me to use the following backdated estimator to test whether the treated state $s = 1$ was significantly different in the seven days before the stay-at-home order was announced

$$E_y^b = \frac{1}{7} \sum_{d=d_a-7}^{d_a-1} \left(y_{1d} - y_{0d}^b \right) \tag{5}$$

If the estimate of E_y^b is significantly different from zero, it would be evidence against the validity of the estimator E_y in expression (1), because we should expect no significant difference between the treated state and the synthetic control before the order was announced. Otherwise, an insignificant estimate of E_y^b would make us more confident in E_y.

4.3 Inference

The statistical inference for a synthetic control estimator is different from that of traditional estimators. Fundamentally, this is because we have only one treated unit. Instead of a standard error, inference is based on the ratio of the RMSPEs

$$r = \frac{RMSPE_a}{RMSPE_b} \tag{6}$$

where $RMSPE_b$ is defined in expression (4), and $RMSPE_a$ is defined similarly for periods after the order went into effect

$$RMSPE_a = \left[\frac{1}{7} \sum_{d=d_1}^{d_7} \left(y_{1d} - y_{0d} \right)^2 \right]^{\frac{1}{2}} \tag{7}$$

Intuitively, if a stay-at-home order has no effect, we should expect $y_{1d} - y_{0d}$ to be close to zero for $d \geq d_1$, which, in turn, implies $RMSPE_a$ is small. On the other hand, a large $RMSPE_a$ is evidence for a significant effect. As the reliability of a synthetic control estimator is inversely related to $RMSPE_b$, the normalization is warranted.

The ratio r is calculated not only for the treated state $s = 1$ but also for each of the potential controls, which is done through permutation by switching the treatment status between $s = 1$ and each $s \in \{2, 3, ..., S+1\}$. Let r_s be the corresponding ratio when state s is assigned as the treated one, and let R be the rank (from the largest to the smallest) of r_1 in the set of $\{r_s\}_{s=1}^{S+1}$. Loosely speaking, we can think of $\frac{R}{S+1}$ as the p-value of the synthetic control estimator, and make inference by comparing it with pre-determined significance levels. One caveat is that this doesn't work well with small values of S. For example, if $S = 7$, the smallest "p-value" we can obtain is $\frac{R=1}{7+1} = 12.5\%$. To avoid falsely rejecting a significant estimate, for small values of S, we could consider an estimate to be significant if r_1 is dramatically larger than the rest of r_s for $s \geq 2$.

Expression (7) defines $RMSPE_a$ for a two-sided test. For one-sided tests, we can simply replace $(y_{1d} - y_{0d})$ in the expression with either its positive $(y_{1d} - y_{0d})_+$ or negative $(y_{1d} - y_{0d})_-$ part. In practice, I use one-sided tests for E_y. Specifically, I will test whether an estimate for the home rate is significantly positive, and whether an estimate for the work rate is significantly negative. I use two-sided tests for E_y^b to make sure the estimate is neither too positive nor too negative.

5 Results

I will first detail the estimates for Wisconsin, and then summarize the results for other states. I conclude this section by comparing Michigan with Ohio, two neighboring states with very different estimates.

5.1 Wisconsin

Figure 3 reports the daily home rate in Wisconsin and its synthetic control, with the level in the upper panel and the difference in the lower. The potential controls are the 15 states without a stay-at-home order by March 31, the seventh day since Wisconsin's Safer-at-Home order went into effect. The weights are 0.36 for Pennsylvania, 0.31 for Iowa, 0.28 for Wyoming, 0.02 for Georgia, 0.01 for Utah, and zero for the other ten states. The upper panel shows that the synthetic control mimics the behavior of Wisconsin very well until around the time when Wisconsin's Safer-at-Home order was announced on March 23. Since the order went into effect, the home rate in the synthetic control has been persistently below its level in Wisconsin.

The lower panel shows that, before March 23, the differences in the home rate between Wisconsin and its synthetic control were close to zero except for a spike on February 9 caused by a storm.[8] The differences in the few days before March 23 are positive and seem to be increasing. However, these differences are not statistically different from zero. Overall, the average difference before March 23 is a minimal 0.09 percent. The difference dropped on March 23 when Governor Evers announced his plan to issue a stay-at-home order on the next day.[9] This probably happened as Wisconsinites began to stock up for the lockdown. A similar announcement effect is observed in other states like Michigan and Ohio. The difference then increased significantly when the order was formally issued on March 24 and went into effect on March 25. The average in the first seven days since the order went into effect (March 25-

[8] According to the National Weather Service, a "fast moving, compact low pressure system brought snow to the area Sunday, February 9, 2020. The storm generated hazardous travel conditions across most of the state of Wisconsin. The heavy snow band with the storm extended across central and east-central Wisconsin, where a widespread snowfall of 6 to 10 inches was reported". https://www.weather.gov/grb/Feb09Snow

[9] https://www.wpr.org/governor-issue-stay-home-order-wisconsin

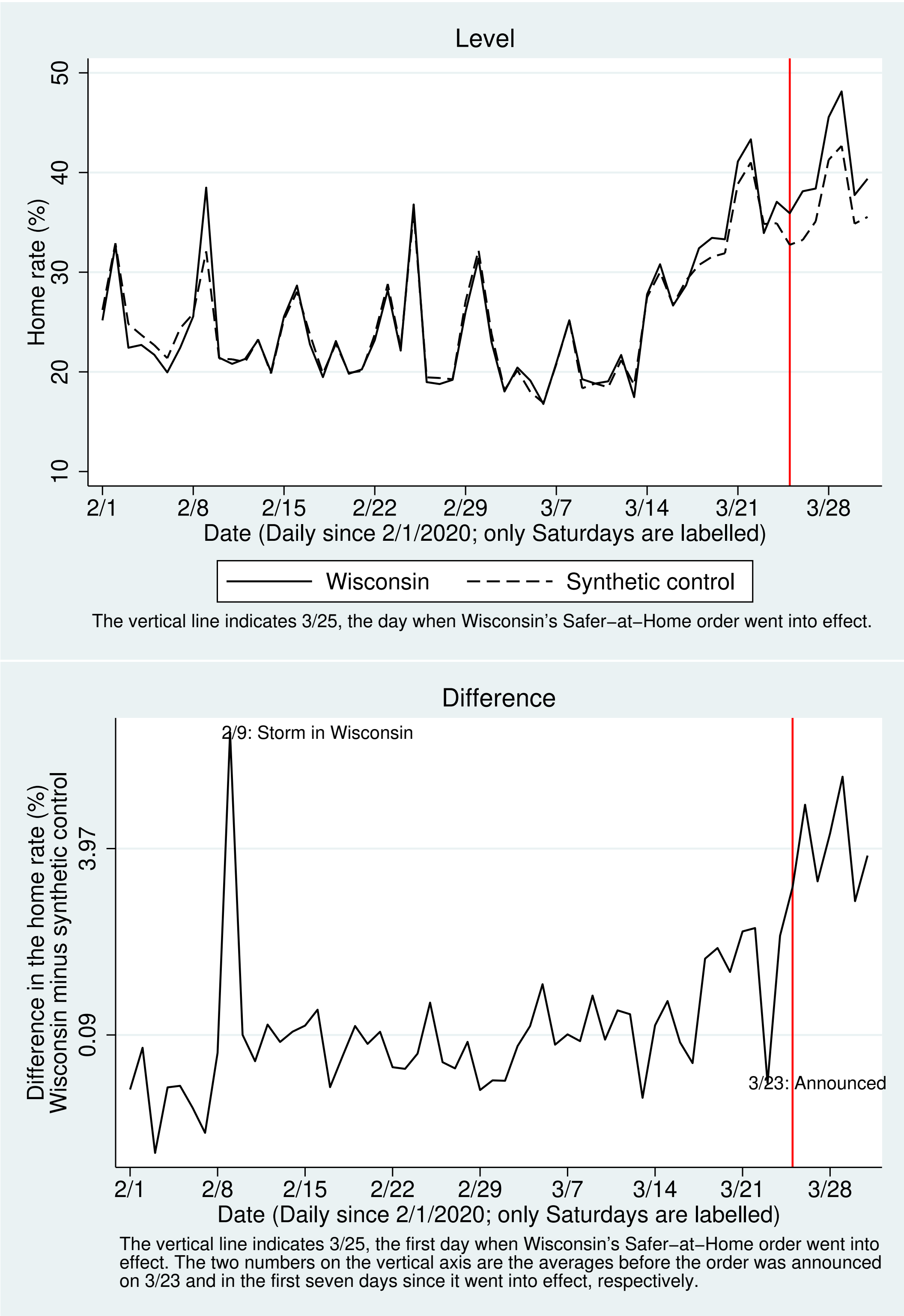

Figure 3: The Home Rate: Wisconsin vs Synthetic Control

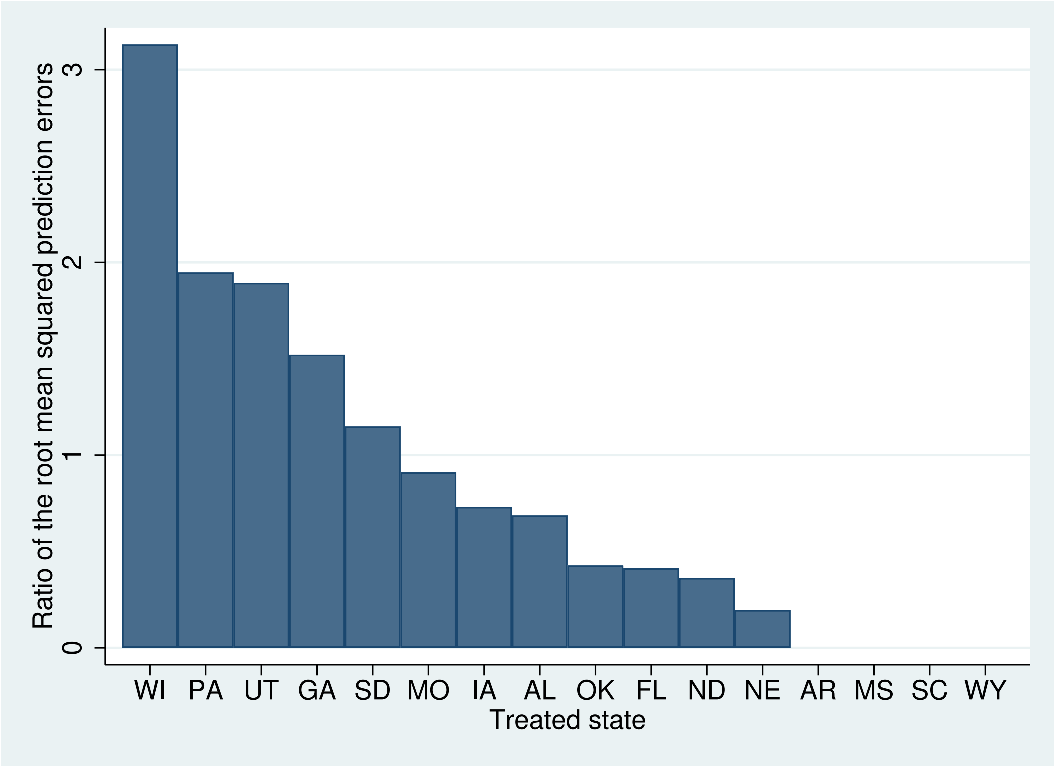

Figure 4: Ratio of the Root Mean Squared Prediction Errors

31), our synthetic control estimate E_y, is about 4 percent, suggesting that about 4 percent of Wisconsinites are staying at home all day because of the Safer-at-Home order. To put this number into context, the home rate in Wisconsin increased by 20 percentage points from its pre-pandemic level of 19 percent on Tuesday March 10 to 39 percent on Tuesday March 31. So the Safer-at-Home order contributed about 20 percent to this increase.

Figure 4 reports the ratio r of the RMSPEs when the treatment status is assigned to each of the 16 states. The ratios are calculated for a one-sided test of whether the estimate is significantly positive. Clearly, the ratio for Wisconsin is significantly larger than it is for the other 15 states, suggesting that the estimate is statistically significant.

Figure 5 plots the differences in the home rate between Wisconsin and a new synthetic control estimated using only data before March 16, the backdated effective date. The weights for this backdated synthetic control are 0.4 for Pennsylvania, 0.39 for Iowa, 0.21 for Wyoming, and zero for the other states. For comparison, the baseline estimate from figure 3 is also plotted. The two estimates are very close to each other, suggesting that the baseline estimate is not significantly affected by the data in the week before the order was announced. The average difference between Wisconsin and the backdated synthetic control in the seven days between March 16 and March 22, the backdated estimate E_y^b, is about 1 percent but not statistically significant, with Wisconsin ranking eighth among the sixteen states in terms of the ratio r of the RMSPEs for this backdated estimate. In comparison, the average difference between Wisconsin and the backdated synthetic control in the seven days between March 25 and March

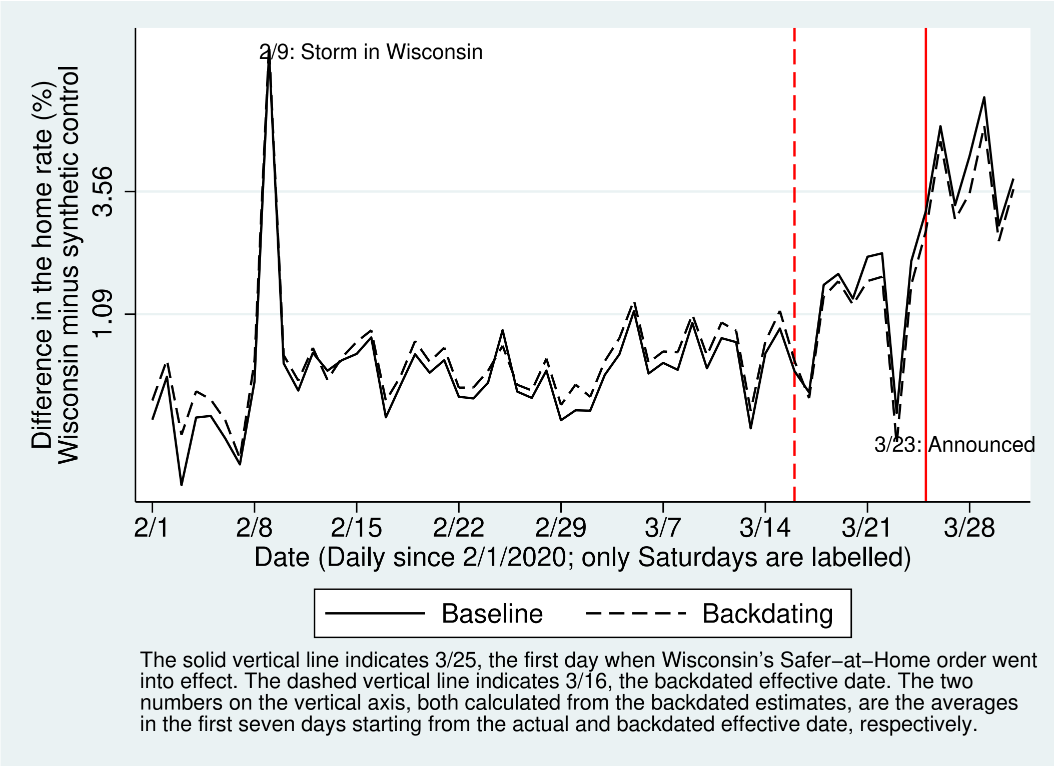

Figure 5: Difference in the Home Rate between Wisconsin and the Synthetic Control

31 is about 3.6 percent, very close to the baseline estimate of about 4 percent, and statistically significant, with Wisconsin ranking first among the sixteen states in terms of the ratio r of the RMSPEs for this estimate. Together, this backdating exercise suggests that the baseline estimate of 4 percent is the causal impact of Wisconsin's Safer-at-Home order on the home rate.

Finally, figure 6 reports the estimates when the work rate is used as the outcome variable instead of the home rate. The weights for the synthetic control in this case are 0.58 for Iowa, 0.2 for Pennsylvania, 0.15 for Wyoming, 0.08 for Georgia, and zero for other states. The average difference in the first seven days since the order went into effect is estimated to be about -2 percent and statistically significant - Wisconsin again ranks first among the 16 ratios of the RMSPEs for this estimate. This is further evidence that Wisconsinites are less mobile because of the Safer-at-Home order.

5.2 Other States

As discussed in the previous section, the synthetic control estimator would be significantly biased if the treated state is significantly different from the potential controls. In this case, no weighted average of the potential controls could mimic the behavior of the treated state, resulting in a large $RMSPE_b$ even for the best possible synthetic control. Consequently, we should restrict attention to cases with relatively small $RMSPE_b$. To be conservative, I only

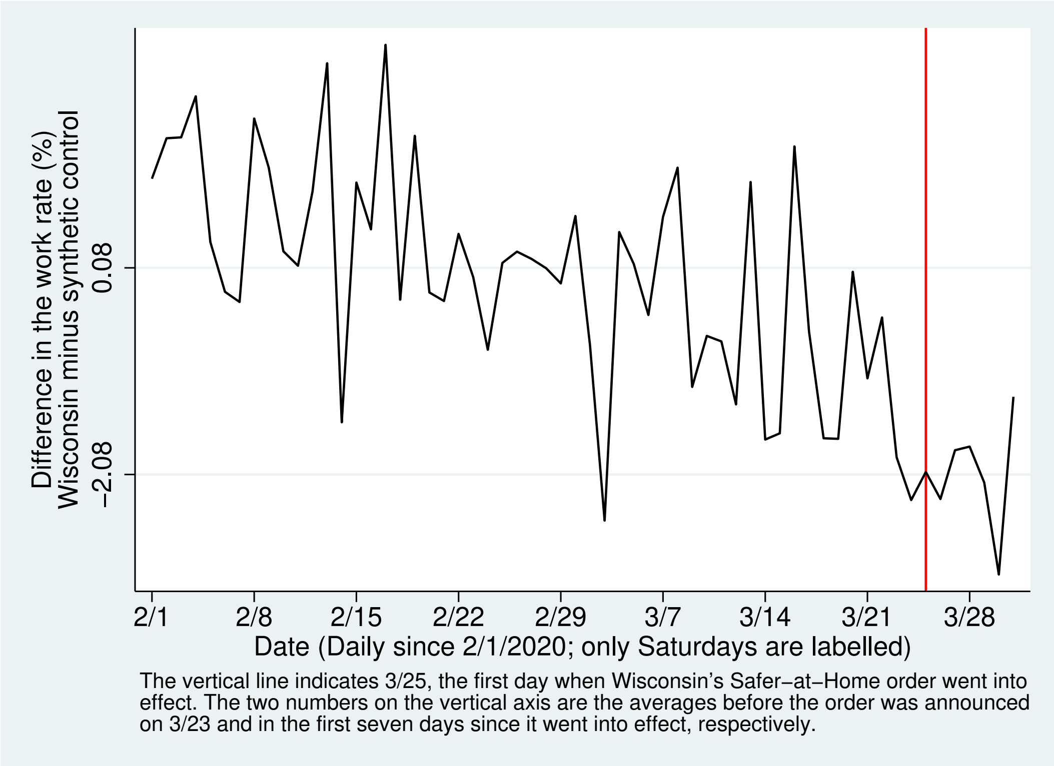

Figure 6: Difference in the Work Rate between Wisconsin and the Synthetic Control

consider treated states whose $RMSPE_b$ for the home rate is no larger than that of Wisconsin. This leaves me with seven states including Wisconsin. Table 2 reports the estimates for each of them, where the states are ordered by the estimated impact of their stay-at-home orders on the home rate. In addition to the home rate, results for the work rate are also reported for each of the seven states even though the $RMSPE_b$ for the work rate in Louisiana is slightly larger than the cutoff value set by the $RMSPE_b$ for the home rate in Wisconsin (1.5 versus 1.3).

Five out of the seven states are in the Midwest. This is not surprising because (1) coastal states like California, New Jersey and New York tend to be outliers as shown in figure 2, and (2) states in the South tend to be late in issuing a stay-at-home order, leaving them with fewer potential controls which makes it hard to find a good synthetic control. The seven states in the table, on the other hand, acted relatively quickly and announced their stay-at-home orders in a short span of four days (March 20-23) immediately after California did it first on March 19, leaving them with at least 15 states as potential controls (column 2).

The estimated impact on the home rate is essentially zero in Ohio, and it's small and insignificant in Delaware and Louisiana. The estimate in Indiana is about 2 percent and could be considered as marginally significant. At 4.1 percent, the estimate for Illinois is slightly larger than that of Wisconsin and statistically significant. Finally, Michigan has the largest estimate of about 5.5 percent. None of the backdated estimates are statistically significant, suggesting that the synthetic control estimator is capturing the causal effect of the stay-at-home orders.

Most of the estimates for the work rate are small and not statistically significant. Besides

Table 2: Estimates by State

	Controls		Baseline		Backdated	
Treated state	S	Top three by weight (%)	E_y	R	E_y^b	R
		The home rate				
Ohio	18	PA(65), MO(33), UT(3)	-0.69	8	-0.28	16
Delaware	18	PA(51), FL(37), GA(12)	1.36	5	-0.48	7
Louisiana	18	MS(33), AL(27), GA(22)	1.65	4	-0.04	15
Indiana	15	PA(41), AR(35), UT(12)	2.01	2	-0.12	12
Wisconsin	15	PA(36), IA(31), WY(28)	3.97	1	1.09	8
Illinois	22	MD(35), NV(18), GA(14)	4.10	2	1.00	8
Michigan	18	PA(49), ME(17), GA(12)	5.53	1	0.47	17
		The work rate				
Ohio	18	PA(77), MO(20), UT(2)	0.25	11	0.30	13
Delaware	18	PA(48), FL(35), GA(17)	-0.40	5	-0.25	7
Louisiana	18	MS(47), AL(24), GA(17)	-1.29	5	-0.20	18
Indiana	15	PA(52), AR(33), UT(11)	-1.51	3	-1.45	3
Wisconsin	15	IA(57), PA(20), WY(15)	-2.08	1	0.07	4
Illinois	22	MD(38), IA(35), TX(26)	-1.47	4	-1.40	5
Michigan	18	PA(50), ME(17), UT(11)	-2.80	2	-1.19	10

Notes: The seven states are ordered by the estimated impact on the home rate. The total number of potential controls is given by S in column 2, while column 3 lists the three states that contribute most to the synthetic control (with their weights in the parentheses). E_y is the synthetic control estimator defined in equation (1), and E_y^b is the backdated estimator defined in equation (5). For each estimator, R is the rank of the treated state's ratio of the RMSPEs. Loosely speaking, $\frac{R}{S+1}$ is the p-value of the corresponding estimate.

Wisconsin, Michigan is the only state with a marginally significant estimate of about -2.8 percent, which is about half the size of the estimated impact on the home rate as it is in Wisconsin.

5.3 Michigan versus Ohio

For further evidence on the heterogeneity of the impacts across states, I compare Michigan with Ohio, two neighboring states with very different estimates.

Less than 24 hours after Ohio issued a stay-at-home order on March 22, Michigan issued its own order on March 23, and both orders went into effect on March 24. The upper panel of figure 7 shows that, before the orders were issued, the two states look extremely like each other in terms of the home rate. A discrepancy emerged right around March 24. Although the home rate has been increasing in both states since March 24, there is no obvious trend in the difference between the two states.

The lower panel looks at the differences between the two states directly. Except for a couple of days caused by hazardous weather,[10] the daily differences in the home rate between the two states were very close to zero before the orders were issued. The difference then jumped almost immediately to about 5 percent after the orders went into effect on March 24, and has been fluctuating around that level since then without an obvious increasing or decreasing trend.

For comparison, the lower panel also plots the difference in the confirmed cases of COVID-19 between the two states. Neither state had a confirmed case before March 10. The difference then fluctuated closely around zero for about a week. Since March 18, the difference has been increasing continuously. This continuous increase is in sharp contrast to the almost discrete jump in the difference in the home rate around the time when the orders went into effect, which suggests that the latter is most likely a result of the heterogeneous impacts of the two orders. The same argument applies to the synthetic control estimates reported above.

Note that this simple comparison between the two states suggests that the impact of the order in Michigan is about 5 percentage points larger, which is very close to the difference in the synthetic control estimates reported above for the two states. This suggests that the synthetic control estimates are not significantly affected by states like Pennsylvania which features prominently in the synthetic control of many treated states.

Finally, table 3 provides additional evidence from weekly initial unemployment insurance claims. The three columns on the left show that Michigan had fewer weekly claims than Ohio in the three weeks before the orders in the two states were issued on March 22. Since then, Michigan has been having much more weekly claims than Ohio. On the other hand, the three columns on the right show that the weekly claims in the two states were roughly comparable during the same period in 2019. Although Michigan did have more weekly claims than Ohio in the first two weeks of April 2019, the differences are almost neglectable relative to those in

[10]The National Weather Service recorded a snow event in Ohio on February 7, 2020, and issued a Winter Weather Advisory of snowfall for Michigan on February 26, 2020. https://www.weather.gov/iln/20200207, and https://www.weather.gov/dtx/200226winteradvisory.

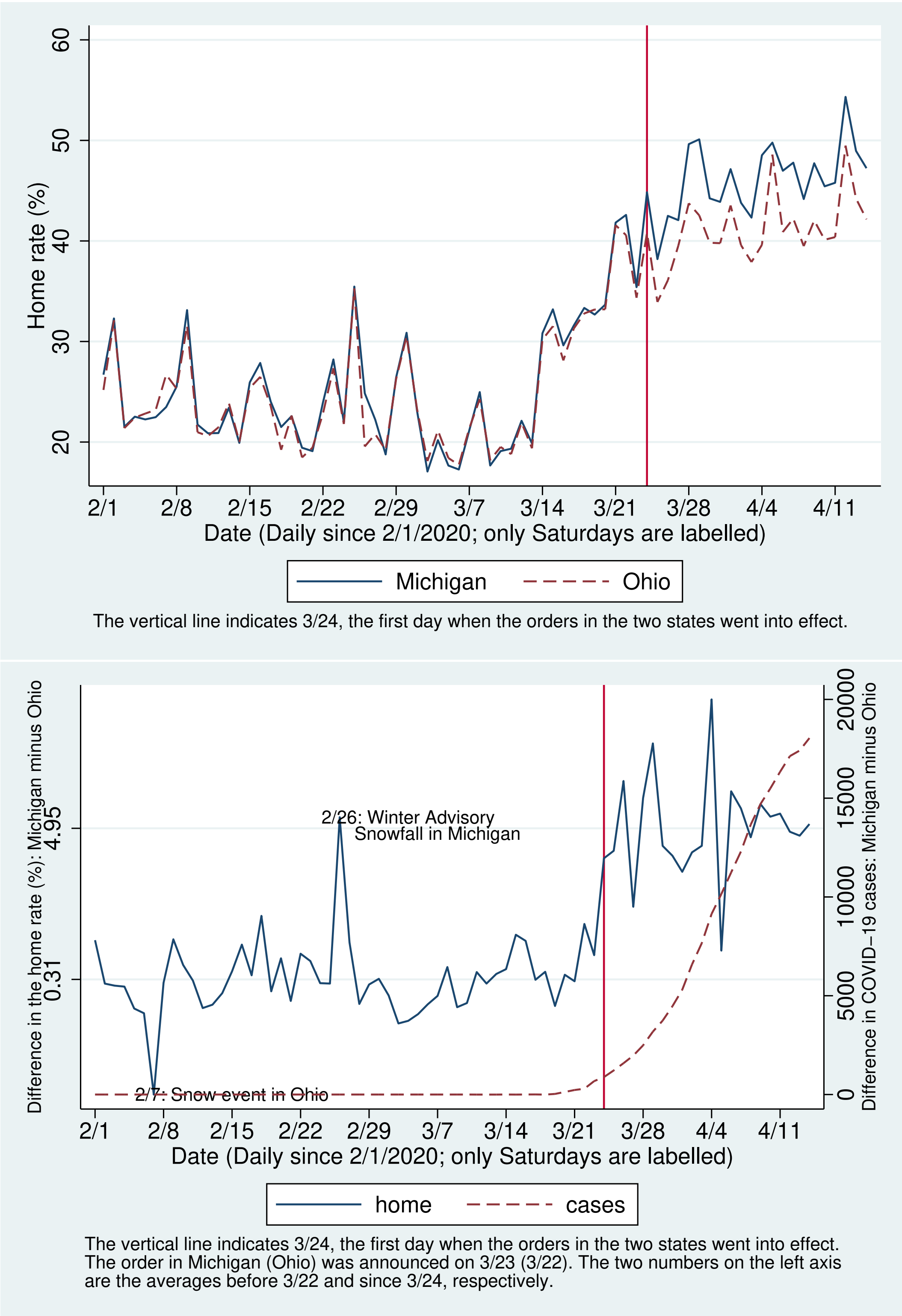

Figure 7: Difference in the Home Rate: Michigan versus Ohio

Table 3: Weekly Initial Unemployment Insurance Claims

Week ending	Michigan	Ohio	Week ending	Michigan	Ohio
March 7, 2020	5,150	6,545	March 9, 2019	5,927	6,569
March 14, 2020	5,338	7,046	March 16, 2019	5,156	6,449
March 21, 2020	128,006	196,309	March 23, 2019	4,887	5,981
March 28, 2020	304,335	274,288	March 30, 2019	4,737	6,197
April 4, 2020	388,554	226,191	April 6, 2019	6,682	6,011
April 11, 2020	222,207	159,317	April 13, 2019	5,579	5,235

April 2020. This is consistent the above estimates that Michigan's stay-at-home order has a much larger impact than Ohio's, and it suggests that at least part of the reason may be because the order in Michigan is more restrictive.

6 Conclusion

Using GPS data from mobile devices, this paper estimates the impact of statewide stay-at-home orders on mobility of Americans. I find significant heterogeneity in the impact across states. For example, the estimates suggest that the orders in Michigan and Wisconsin raise the fraction of their residents at home all day by about 5.5 and 4 percentage points, respectively, while the corresponding estimate for Ohio is small and insignificant.

Needless to say, the exact estimates depend on the data and measure. For example, we may find a significant impact of the order in Ohio by looking at different measures. It's also important to note that the estimates in this paper reflect the incremental impact of stay-at-home orders above and beyond other social distancing guidelines. The total effect of all social distancing guidelines could be much larger. Even larger is the total level of social distancing, as many Americans may be practicing it voluntarily.

The key finding is that there is significant heterogeneity in the impact of stay-at-home orders across states. This could arise from either the differences in the orders themselves, like which activities are exempt and which are not, or the differences in enforcement and compliance, or both. The evidence from weekly initial unemployment insurance claims in Michigan and Ohio suggests that the order in Michigan is more restrictive. A detailed investigation of the causes behind the heterogenous impact, while beyond the scope of this paper, should be an important direction for future work. Among other things, the investigation could be informative of how to make the right policy to achieve the desired impact.

While the estimates in this paper suggest that some stay-at-home orders are more restrictive than others, they don't imply that some orders are better than others. Like any actions, the optimal level of social distancing should balance its benefits with costs. Because states are different from each other, the cost-benefit analysis may result in different levels of optimal social distancing. For example, other things equal, states with more vulnerable people in terms of health should probably have more social distancing, while states with more economically vulnerable people should probably have less. More specifically, while the evidence suggests that the order in Michigan is more restrictive than that of Ohio, the difference between the two orders may be justified by the larger number of confirmed cases of COVID-19 in Michigan than in Ohio. More generally, it's possible that all orders are optimal, just as they could all be sub-optimal.

That being said, the significant differences across the estimates do provide some insights. For example, there probably aren't many pairs of states in the U.S. that are as comparable with each other as Michigan and Ohio. If the difference in the estimated impact of their orders is justified, we should probably expect more cross-state variation in social distancing.

Finally, the estimates in this paper could be useful as states are thinking about whether, when and how to modify their orders. For example, relative to Ohio where the estimated impact of the order is small and insignificant, residents in states like Michigan and Wisconsin where the estimated impact is much larger would probably be more responsive and engage in more economic and social activities outside home when the order is lifted. Whether and when such increased activities are desirable should be a key factor in determining when and how to reopen the economy.

References

Abadie, Alberto, "Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects," *Journal of Economic Literature*, forthcoming.

_ **, Alexis Diamond, and Jens Hainmueller**, "Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program," *Journal of the American statistical Association*, 2010, *105* (490), 493–505.

_ **and Javier Gardeazabal**, "The economic costs of conflict: A case study of the Basque Country," *American economic review*, 2003, *93* (1), 113–132.

Allcott, Hunt, Levi Boxell, Jacob Conway, Matthew Gentzkow, Michael Thaler, and David Y Yang, "Polarization and public health: Partisan differences in social distancing during the coronavirus pandemic," *NBER Working Paper*, 2020, (w26946).

Athey, Susan and Guido W Imbens, "The state of applied econometrics: Causality and policy evaluation," *Journal of Economic Perspectives*, 2017, *31* (2), 3–32.

_ **, Billy Ferguson, Matthew Gentzkow, and Tobias Schmidt**, "Experienced segregation," *Working paper*, 2019.

Benzell, Seth, Avinash Collis, and Christos Nicolaides, "Rationing social contact during the COVID-19 pandemic: Transmission risk and social benefits of US locations," *Working paper*, 2020.

Chen, M Keith and Ryne Rohla, "The effect of partisanship and political advertising on close family ties," *Science*, 2018, *360* (6392), 1020–1024.

Engle, Sam, John Stromme, and Anson Zhou, "Staying at home: Mobility effects of Covid-19," *Covid Economics*, April 2020, (4), 86–102.

Farboodi, Maryam, Gregor Jarosch, and Robert Shimer, "Internal and External Effects of Social Distancing in a Pandemic," *University of Chicago, Becker Friedman Institute for Economics Working Paper*, 2020, (2020-47).

Painter, Marcus and Tian Qiu, "Political beliefs affect compliance with covid-19 social distancing orders," *Covid Economics*, April 2020, *4*, 103–127.

Williams, Noah, "Measuring Wisconsin Economic Activity Using Foot Traffic Data," *Data brief, Center for Research on the Wisconsin Economy*, 2020.