J. F. Kennedy sq. 6
10000 Zagreb, Croatia
Tel +385(0)1 238 3333
www.efzg.hr/wps
wps@efzg.hr

EFZG WORKING PAPER SERIES
EFZG SERIJA ČLANAKA U NASTAJANJU
ISSN 1849-6857
UDC 33:65

No. 23-04

Robert J. Sonora and Martina Gottwald-Belinić

Does individual behavior converge to policy recommendations in times of pandemic?

Evidence from COVID-19 in US states

Does individual behavior converge to policy recommendations in times of pandemic? Evidence from COVID-19 in US states

Robert J. Sonora
email:robert.sonora@mso.umt.edu
Bureau of Business and Economic Research
University of Montana
Missoula, MT 59812 USA
and
Visiting Scholar
University of Zagreb, Faculty of Economics and Business
10000 Zagreb, Croatia

Martina Gottwald-Belinić
email: mbelinic@efzg.hr
Visiting Scholar
University of Zagreb, Faculty of Economics and Business
10000 Zagreb, Croatia

Abstract

The COVID-19 pandemic is an exceptional shock on human habitual behavior and provides a rare opportunity to analyze resilience in preferences. We use Google's mobility and policy stringency indices to investigate if policy maker and resident „preferences" align over the period. Differences in utility across the ten largest states in the United States should lead to idiosyncratic response on perceived cost of restrictions and associated risk attitudes in policy respond. We conduct structural break and rolling unit root tests on estimated residual. Our results suggest that individual behavior converges to the policy prescriptions within the time span up to 18 months.

Key words

COVID-19, individual risk behavior, structural break, unit root tests

JEL classification

C22, E70, H75, I12, I18

Does individual behavior converge to policy recommendations in times of pandemic? Evidence from COVID-19 in US states

Robert J. Sonora[a,b,*], Martina Gottwald-Belinić[b]

[a] *Bureau of Business and Economic Research, University of Montana, Missoula, MT 59812 USA*
[b] *Visiting Scholar, Faculty of Economics, University of Zagreb, Croatia*

Abstract

The COVID-19 pandemic is an exceptional shock on human habitual behavior and provides a rare opportunity to analyze resilience in preferences. We use Google's mobility and policy stringency indices to investigate if policy maker and resident "preferences" align over the period. Differences in utility across the ten largest states in the United States should lead to idiosyncratic response on perceived cost of restrictions and associated risk attitudes in policy respond. We conduct structural break and rolling unit root tests on estimated residual. Our results suggest that individual behavior converges to the policy prescriptions within the time span up to 18 months.

Keywords: COVID-19, individual risk behavior, structural break, unit root tests

JEL: C22, E70, H75, I12, I18

---

*Corresponding author

Email addresses: robert.sonora@mso.umt.edu (Robert J. Sonora), mbelinic@efzg.hr (Martina Gottwald-Belinić)

1. Introduction

> "The curve is shaped by public awareness. We're sort of lurching between crisis and complacency."
>
> Jennifer Nuzzo, epidemiologist at Johns Hopkins University, "'Lurching Between Crisis and Complacency': Was This Our Last Covid Surge?" *The New York Times* (Anthes [4])

By the end of 2022, according to the Johns Hopkins Coronavirus Center since the pandemic outbreak there have been over 101 million reported cases in the US with over 1.1 deaths. Furthermore, over 665 million vaccinations have been administered.[1]

According to Potter and Harries [25] the "determinants of policy effectiveness" for health policy models requires public administration system to include diverse social, cultural and economic motivators engaged in behavior. A state of pandemic represents a rare opportunity to empirically analyze individual behavior during a period of "lock-down" and "public fear" and behavioral respond to changes on policies intended to minimize infection rates and protect the health of the residents (Heuring [17]). According to the survey made among the citizens across the United States in May 2020, the reaction on widespread support to "stay-at-home" policies were actively accepted among the majority of the surveyed residents Czeisler et al. [12]. Individual reactions went even beyond the intended policy. People indicated that they would feel unsafe if restrictions were lifted. The anxiety level and concerns about the impact of disease on individual health as well as the health of their peers triggered risk averse behavior in population.

At its root, the decision to stay home during the COVID-19 crisis can be divided into two sub-behaviors, those influenced by the policy to restrict the movements and minimize the risk of spreading, and unobserved idiosyncratic individual choice. Using a subset of the largest ten US states, this paper examines how stay-home policies align with individual preferences. During the pandemic, individual states implemented mobility stringency policies to prevent the spread of COVID among the population. Googles phone data helps us to track mobility and identify patterns of behavioral changes: mobility should decrease as stringency increases and differences in preferences across the states should lead to idiosyncratic responses to policy

[1] See the Coronavirus Resource Center, Johns Hopkins [19], Center for Disease Control [11] and the Mayo Clinic [22].

recommendations. We study the properties of unobserved individual behavior; the estimated residuals from a model of mobility conditional on policy restrictions for each state in the sample. The residual represents the unobserved mobility preferences of individuals and is an indicator how closely individual and policy maker preferences align.

Our interest are the time series properties of estimated residual. Deviation in estimates on policy prescriptions indicates differences in estimated utility on policy recommendations versus individual preferences. The residual of individual mobility behavior that exceeds or falls short of policy recommendations incorporates individual's degree of risk-aversion or risk-taking and whether or not individuals are complying with policy, or either under- or over-comply with policy. We evaluate if individual response residual eventually aligns to policy preferences and how long does it take to converge. Mean reversion theory suggests that regret, fear, or risk will converge towards normal state of responses over time. Can we say that individual behavior follows stochastic process with sporadic drift close around the mean that eventually converges to utility prescribed by the policy? Put another way, can we say that behavior and policy preferences eventually converge.

Using both structural break and rolling unit root tests, we reject a unit root process if individual behavior eventually converges as prescribed by the policy. Our estimates cannot determine whether or not individuals are risk-taking or risk-averse, we only observe risk attitudes changes relative to policy recommendation. This adjustment of behavior can be interpreted by myopic loss aversion (Thaler et al. [29]), where evaluation on prospects and associated risk attitudes changes cause preference reversal and evolution in behavior as new information and learning emerge (Berg et al. [7]). Put into the context of the more micro level literature, such as Czeisler et al. [12], Goolsbee and Syverson [14], and Alfaro et al. [2]; our results suggest that for the most states the lack of short run convergence to policy recommendations is due to risk attitudes changes. Individuals response to lock-down varies from early stage of shock where anxiety and fear are salient effects causing non-consistent heuristic behavior (Alfaro et al. [2]). The closest antecedent to this analysis is by Gupta et al. [15] and Goolsbee and Syverson [14], most of the reduction in mobility occurred prior to specific state policy actions due to high uncertainty and high sensitivity to losses at the early stage of pandemic. They argue that voluntary individual behavior led to greater immobility compared to conduct dictated by policy recommendations, implying individuals were more risk averse than policy makers at the first wave of pandemic.

Our estimation results differs across the states. Population density and social disparities Best [8], together with the structural differences between the states like the capacity in hospitals, social sensitivity on school closing, unemployment and income level or neighboring states' prevention directly reflect the policies stringency. The role of US federalism in states policy differences and division of authority between the central government and individual states played a central role in coordinating interventions and policies for COVID response (Singer et al. [27]). According to Adolph et al. [1] the most important driver for social distancing policy adaptations among the states were political, Republican-led governors were more reluctant, on average, to implement stringency policies during the early stage of COVID-19. Government responses differed in time and level of stringency policy based on specific trade-off between public health and societal, economic and political costs. On average, Democrat led states have more stringent policy recommendations and impose higher impact on resident behavior that Republican-led states. Political costs, voter preferences and public opinion played an important role in shaping the policy. For simplicity, we can think of policy utility in terms of "policy loss function" with trade off between cases and unemployment. In our study, effectiveness of policy alignment with resident behavior directly represent preference alignment between policy makers and residents.

Desired public behavior on policy recommendation is not stationary on a short run. The tendency and period respond before return to "normal" pattern of behavior are heterogeneous across the states. Ceteris paribus, behavior varies as resident experience about pandemic emerges. Believes involve probabilities people think negative consequences related to disease will occur conditional on available information and public trust. This implies that the reference point is changing over time as marginal utility on policy compliance adjusts. With more experience, people tend to return to normal and to optimize. Over the long run, the policy and behavior preferences do converge.

2. Methodology

2.1. Estimating individual behavior

Individual mobility can be decomposed into two components, one observed and the other not. Because some consumer behavior is dictated by policies designed to restrict movements with the goal to minimize risk associated with COVID this component can be estimated. Any behavior that differs from that which is "dictated" by policy represents individual choice and can be estimated using the following ARDL specification,

$$MI_t = \alpha \ + \rho MI_{t-7} + \beta(L) OxSI_t + \mathbf{X}_t' \gamma + \eta_t, \ t = 0, \ldots, T \tag{1}$$

where MI is the mobility index and $OxSI$, for stringency index, is a state specific policy variable. Lagged MI represents inertia in individual behavior and L is a lag operator. We include concurrent and 7 and 14 day lags of $OxSI$ for most state in our specification. We include the one and two week lags because most COVID-19 data was presented to the public as a rolling 7 day moving average and 14 days corresponds to the length of quarantine should a test be positive. The inclusion of lags also allows individuals to adjust to new policy interventions. There might be concern that policy prescriptions and behavior might be endogenous. That is, policy makers respond to individual voter preferences when determining pandemic restrictions. However, this would require very frequent changes in policy making behavior. Moreover, Sonora [28] demonstrates that the level of stringency is determined through minimizing an "economy-health" loss function and can, therefore, be considered exogenous. $\mathbf{X}$ is a vector of other state specific and national controls, discussed in the data section below.

Of primary interest are the contemporaneous, "short run", estimates behavioral elasticity with respect to policy, $\hat{\beta}_0$ and the the longer term adjustment elasticity estimates $\hat{\rho}$ and $\hat{\beta}(L)$. Given that people require some time to fully incorporate policy into their decisions, we anticipate the short run estimate to more inelastic than the longer term elasticity. This reflects the benefits of using very high frequency data in empirical modeling, it allows us to better understand individual behavior in the immediate short run and compare it to individual choices as the time horizon lengthens. In an urgent situation, such as a pandemic, the initial response to new restrictive policy is not likely to be as effective as policy makers may believe.

The second interest for understanding unobservable human behavior with respect to policy intended to inhibit social interaction is the time series properties of η which are the estimated residuals from equation (1), $\hat{\eta}$. In this context, $\hat{\eta}$ represents state level "animal spirits" and is the manifestation of individual preferences with respect to risk, politics, behavioral norms, information, etc., for example Allcott et al. [3], Barrios and Hochberg [6], and Poletti et al. [24].

To study the convergence of actual behavior to that preferred by the policy maker, we can apply unit root tests to the estimated residual, $\hat{\eta}$, which reflects individual unobservable choices. $\hat{\eta} \sim I(0)$ is interpreted as the revealed preferences of residents (r) converging over time to those of the policy makers (p) in state j

$$\lim_{t \to T} U_p^j(M^\tau) = U_r^j(M^\tau). \tag{2}$$

where M^τ is the target level of mobility. There are two observations about the equation above. First, the preferences of both individuals and state policy makers different across all states. And secondly, the above equation does not imply that the policy maker is making optimal health policy to reduce the spread of the disease, but rather the decision is made based on her idiosyncratic set of preferences, as in Sonora [28]. We can interpret this preference matching, too, within the context of the Tiebout [31] hypothesis.

We employ two models that test for the presence of unit roots. The first is the standard augmented Dickey-Fuller test. Secondly, we employ the one- and two-break tests proposed by Clemente Lopez, Montas and Reyes [10] (CMR). These tests are well established in the literature and therefore we only highlight the basic mechanics of these tests. CMR extend the Perron and Vogelsang [23] model by allowing for up to two breaks. Structural break unit root are useful when analyzing time series which could be subject to deterministic breaks which could be misinterpreted as permanent stochastic processes that bias standard unit root tests towards nonstationarity leading to size distortions.

This model can handle both "long" run innovative and "short" run additive, or level, shocks. For our purposes, breaks in the data represent structural changes in individual behavior as it relates to health policy and could therefore be either of these types of inflection points. For example, long term changes are the reduction in going out in, or returning to, public whereas a one time change could be working from home or school closures. Alternatively, innovations represent changes in behavioral trends. Note, both of these breaks could manifest in either less or more mobility as the social environment changes.

It is important to note that the unit root tests do not detect whether or not behavior is overly cautious or heedless of policy, just whether or not $\hat{\eta} \sim I(0)$ and reverts to the policy restrictions. For each test, we allowed for up to 14 days of lagged differenced terms to account for serial correlation. The suggested rule of thumb for choosing the maximum number of lagged data, suggested by Schwert [26], is about 18 days, but we chose to use 14 as it corresponds to the number of days required for quarantine.

3. Data

Our data is daily and the sample is the two year period February 17, 2020 to February 14, 2022, which roughly coincides with the end of the omicron variant spike. To simplify the presentation of the analysis we choose the ten largest states in terms of gross state product (GSP). In descending order the states are: California (D), Texas (R), New York (D), Florida (R), Illinois (D), Pennsylvania (D), Ohio (R), Washington (D), Georgia (R), and New Jersey (D). There are differences in the way Democrat (D) and Republican (R) Governors responded to the pandemic and our sample includes a relatively balanced mix of the two parties, with six of the states having Democrat governors.

To proxy for individual mobility behavior we use recently available cell phone data to observe movements. Google's Community Mobility Report (Google [13]) indexes six different "types" of mobility, by state: Grocery and pharmacy, retail and recreation, parks, residential, work, and transit. We use the daily mean of *five* of these indices to derive an overall index of mobility, the Google mobility index (GMI).[2] The index is defined as the percentage difference between the mobility on any given day based on pre-pandemic mobility, $GMI \in$

[2]We also tried a principle components approach. Because the correlation between the mean and principle component index was over 0.94 for all states we chose to use the mean as the calculation is more transparent.

$(-100\%, 100\%)$. We exclude "park" mobility – defined as "…trends for places like national parks, public beaches, marinas, dog parks, plazas, and public gardens, see Google [13] – for two reasons. First, stringency measures are designed to *restrict* movement and with many areas being closed off, outdoor mobility will rise as there are few other locations to go outside the home. Secondly, from a data standpoint, the park index varies greatly across the states. In Ohio, mobility in parks averaged about 130% above pre-pandemic levels.

Two other mobility indices were also considered. The first is the Dallas Federal Reserve Bank's Mobility and Engagement Index derived from cell phone tracking data. However this data was discontinued in late-March of 2021 which limits the time series length and would exclude some important structural changes in national COVID policy, such as the vaccine being readily available to all residents over the age of 18. A second alternative is produced by Apple called the Mobility Trends Report, Apple [5]. We chose not to use this index because it is constructed from users of iPhones, which could bias results given that iPhone users are generally more urban, higher income, and better educated, see Hixon [18].

COVID restriction policies are from the Oxford Coronavirus Government Response Tracker (OxCGRT) Stringency Index (denoted $OxSI$) (Hale et al. [16]). This one of the several indices constructed by OxCGRT but was chosen as it quantifies the degree to which states restrict individual movements, via "lockdown" policies, rather than economic support or health containment policies. The index is between 0 and 100 with a value of 100 being the most restrictive. The index represents all containment and closure policy indicators, such as school and workplace closing, public events cancellation and restrictions on gathering, public transport access restrictions, and travel restriction incorporated with stay at home promotions including the proxy for recorded public information campaigns. OxCGRT daily data captures the intensity and modality on policy interventions for COVID-19. The stringency index shows variation within regions and states, correlated with political determination of particular state government.

The control variables are both state and national level indicators. First, we include the number of new weekly cases per state, available from Johns Hopkins [19]. This is used as an information variable, spikes in new recent cases indicate COVID spread which lead to less mobility. Secondly, we include the current and one-week lagged percentage of eligible population which has been vaccinated, $\%VAX$, at least once.

The first national level control is the risk premium, $Risk$, defined as the AAA corporate bond yield minus the three month treasury yield and is included as a control for systemic economic risk. Changes to the AAA corporate bond corresponds to greater economic uncertainty as a wave of anti-COVID "austerity" measures were imposed to prevent the virus from spreading. Moreover, individual behavior contributes to economic uncertainty as individual voluntarily restrict their own movements as a preventative measure. The three month bond yield reflects both financial uncertainty as investors move funds to safer instruments and the FED's anti-recessionary monetary policy. Prior to the declaration of the pandemic, the risk premium hovered in the 1.2-1.5% range before spiking in early 2020 to over 4.0% and then settling in between 2.0-3.0%. We also used the CBOE VIX index as a measure of risk with no change in the results.

We add a time fixed effect effective the date of the Biden administration's inauguration on January 20, 2021 which signaled a shift in national COVID policy, discussed before the election by Malakoff [21] with differences later highlighted in Kates et al. [20]. Finally, seasonal fixed effects are also included: The first is a summer indicator, June 1 to September 1 for both 2020 and 2021. The second is a winter holiday season variable extending from October 1 to January 15. The 2021 Texas power crisis, February 10-27, 2021 is also included in the Texas model as this had a significant impact on mobility in the state.

Table 1 provides the mean, standard deviation (SD), the minimum and maximum of the GMI and $OxSI$ data by party of the governor and for each state. States with a Republican governor had fewer restrictions and more overall mobility. The GMI averaged –15.6 and –11.2 and the $OxSI$ averaged 53.1 and 47.7 for Democrat and Republican governor states respectively. The state with highest mobility is Oho and the least mobility is in New York while the least/most stringent states are Florida and New York. It is also worth noting that states governed by either party followed similar policy responses across party lines, indicated by comparable $OxSI$ standard deviations. However, there is greater differences in mobility behavior in Democrat states than Republican ones. Figure 1 shows the average movement of both indices by states with Democrat and Republican governors. Republican led states have more mobility and less stringency than Democrat states. Interestingly, at the end of the sample period Democrat led states have less stringency than Republican states. This is likely because of the surge in new cases in non-vaccinated residents due to the Delta variant of COVID which primarily impacted Republican states. We can also see the effect of severe weather that precipitated the Texas power problem in February, 2021. Holiday mobility can also be seen to decline.

4. Results

Results for the model presented in equation (1) for each state can be found in Table 2. Newey-West standard errors using 7 lags are in parenthesis, with $^{*}, ^{**}$, and *** representing statistical significance at the 5%, 1%, and 0.1% levels respectively. Estimates of most interest are at the top of the table shaded in gray. The first line, denoted $OxSI_t$, is the instantaneous, or day-to-day, response of behavior to the current state of policy. The two week response elasticity is calculated from the estimates of ρ and $\beta(L)$ and is denoted "*Response*". We also present the results for vaccination rate, weekly cases, the Biden policy effect, and the Texas weather crisis, we do not include the seasonal fixed effects.

We focus our discussion on estimated responses coefficient rather than the immediate estimated elasticities. If individual and policy maker preferences align estimated *Response* will be in the neighborhood of $|-1|$. If $Response > |-1|$ residents are more risk-averse than policy dictates and are risk-takers if $Response < |-1|$. It is important to recall that the policy implemented by each state may not be optimal for reducing the spread of the virus, rather it reflects the individual preferences of the policy maker's trade-off between the economy and the health of the state's residents.

As discussed we would anticipate the short run response, denoted $OxSI$ in Table 2, to be more inelastic than estimates of *Response*. We find this to be true across the ten states in our sample, and all estimates are highly statistically significant, with exception of Florida. This is notable because Florida was more "laissez-faire" about its COVID policy under Governor Rod DeSantis, see Thompson [30]. Because the COVID policy had less "bite" to it in Florida, it is understandable that estimates of individual behavior to policy would demonstrate a relatively weak empirical relationship.

For states other than Florida, the short run elasticity for California is the closest to zero, with an estimated coefficient of -0.112 and the state with most elastic *Response* is New York, -0.611. There are obvious differences between the two states. First, about half of New Yorkers live in the New York City MSA which relatively dense compared to California implying the costs of not complying with policy could lead to a more rapid spread of COVID. Secondly, New York is relatively old compared to California. Because this demographic is more likely to experience severe illness or death from the pandemic, there is likely to be a higher degree of risk aversity in New York. Lastly, New York imposed many more pandemic policies than California, roughly twice as many according to Johns Hopkins [19], which suggests that New York residents either took the policies more seriously, risk aversity; became more accustomed to the introduction of policy; or policy makers in California had less credibility than those in New York.[3]

To answer the question about how individual preferences changed over the course of the pandemic we conducted 270 day rolling regressions of equation (1) for each state, presented in Figures 2(a)-(j). The longer term *Response* estimates are the blue line and the immediate response is the red line. Changes in policy elasticities reflect changing attitudes by residents as they become weary of lockdowns and restrictions as well as responses to the dynamics of the disease – as cases and deaths rise, individuals react by reducing their mobility.

As might be expected, in the earlier stages of the pandemic, elasticities were closer to -1, but as individuals became more accustomed to living with COVID, their mobility vis-á-vis policy recommendations began to loosen, increasing the estimated policy elasticity to 0. We can also differences in responses to the various variant waves. The increase in cases and deaths associated with delta and omicron reduced mobility in the short term, but had relatively little effect on longer term behavior, particularly as the pandemic began to stretch into its later stages, reflecting acceptance and exhaustion of individuals. However, it is also due to the reduction of restrictions imposed by policy makers over time. The gradual return to a "normal" set of policy restrictions, i.e. zero, means that it's less costly comply with policy reducing the magnitude of estimated elasticity coefficients.

Next we estimate whether or not individual's preferences align with the desired behavior

[3]In the early stages of the pandemic New York Governor Andrew Cuomo became one of leading voices for health policy both in New York and the nation as a whole.

dictated by policy stringency. Results of the unit root tests can be found in Table 3. We present the Studentized $t-$statistics for each of the tests. For the break tests, we provide the estimated break dates. If no break dates are tabulated, no statistically significant break date was estimated. Beginning with ADF results we see that, across the board, $\hat{\eta} \sim I(0)$ over the full sample suggesting that policy and individual behavior *do* converge over the entire sample period, which is roughly 700 days.

Because we find evidence for convergence across the sample, the CMR tests are helpful in identifying specific dates where behavior changed. Interestingly, the single break test only finds a minority of significant structural breaks. And many of the breaks happen around holidays, rather than easily identifiable policy and/or behavioral changes. It is worth noting that the timing of Texas breaks found were during the power crisis.

5. Dynamic Alignment Analysis

While these tests can inform discrete dates of changes in behavior, these tests do not necessarily let us observe the evolution of behavior as new information or policy becomes available day-to-day Put another way, are there short term changes in behavior? To investigate these more changes in behavior we conduct rolling augmented Dickey-Fuller (ADF) unit root tests.

We chose a rolling window of 270 days for each state and allow for maximum of 14 days of lagged first differenced variables to control for serial correlation. The first window begins on February 17, 2020 and last window begins on May 21, 2021. This window should give us the necessary observations to have confidence in our estimates and to give sufficient time for individuals to adjust their behavior either becoming more or less mobile than policy dictates. Also presented is the rejection rate of the null hypothesis of nonstationarity.

Results of the rolling unit root tests are in Figures 3(a) – 3(j). The figures show the $ADF - t$ statistic from each rolling unit root test and the horizontal red line is the 5% critical value. Periods when the $ADF - t$ falls below the critical value are interpreted as periods when individual behavior conforms to state policy. Focusing on the two states with the highest and lowest rejections rates, Ohio (g) and Pennsylvania (h), we find that over a 270 window, Ohio residents complied with state policy makers 88% of the sample period whereas Pennsylvania residents only did so 14% of the time. In Table 2 we note that the *Response* elasticity to policy is more inelastic in Pennsylvania than in Ohio, which explains why the comply less, the are more risk taking, relative to PA policy, than Ohioans. Thus they are more likely to respond less to new policy than in Ohio. This could also suggest that residents of PA believe that policy in that state is overly restrictive, even though it is *less* stringent than OH over the sample period, as we can see in the summary statistics in Table 1.

Given the relatively low percentage of behavioral convergence over 270 days we next consider the number of days required to achieve near 100% policy compliance. We conduct a range of rolling window unit root tests beginning with 60 days and ending with a 650 days. Results can be found in Figure 4. The results demonstrate the percentage of each window length that $\hat{\eta} \sim I(0)$. If we again consider OH and PA we estimate that in the shortest window less than 5% of the 60 day windows reject a unit root behavioral process. However, as the length of the window increases rejection rates begin to increase, by 180 days about 30% of the windows reject a unit root in OH and 10% in PA. Within one year, or 360 days, the rejection rate is 100% in OH but only 25% in PA. We find that it takes about 620 days for PA to reach 100% behavioral compliance. The remaining eight states fall between these two extremes.

There are a number of reasons for the length of time for full behavioral convergence and the disparities between the states. First, our results suggest that in a fast moving crisis, like a pandemic, policy makers cannot expect residents to fully comply with mandates in a short period of time. Only TX experienced over 50% compliance within 180 days and roughly half the states required one year before reaching 50% behavioral alignment.

A second reason is that the shorter time required to comply with policies is that policy makers may better understand their constituents. If voters believe that policy is either too lax or restrictive their own behavior will deviate from recommendations. States like Texas and Washington both achieved relatively rapid near 100% convergence. We saw in Figures 3(i) – 3(j) that for these states, a unit root was rejected 79% and 85% of the time for Texas and Washington respectively, two states at either end of stringency distribution, see Table 1. Therefore, despite have considerably different policies, the policies themselves were more effective given the preferences of each states' constituents. For policy to be effective, constituents must believe in the policy, highly correlated with credibility, and the policy must be attainable. A recently example can be found when the Chines government ended the zero COVID policy in December of 2022. This policy pivot was a response to the breakdown of these two criteria as residents endured extended lockdowns and other heavy-handed measures intended to reduce the spread of infection. Ultimately, China's COVID restrictions, and their retraction, could have significant costs. Recent research predicts that as a result of shifting from zero COVID to zero restrictions could result in up to 1.5 million, or more, deaths in the country in 2023, Cai et al. [9]. An unintended consequence overly Draconian policy is it potentially leads to a more grave than would have occurred in a country with a less restrictive policy.

6. Summary

The COVID-19 pandemic represents a rare opportunity to empirically study behavior vis-á-vis during a time of duress. Using state cell phone data a proxy for social interaction in the ten largest US states to analyze if behavior converges to policy dictated at the state level. Previous research at the micro level demonstrated that most individuals are more risk-averse than policy recommendations. Using unit root tests, our results confirm these that individual behavior does not converge to policy recommendations. While the unit root tests cannot identify whether or not the source of nonstationarity is because residents are risk-averse or risk-takers, estimated coefficients corroborate previous research for the majority of the states in our sample – residents are generally risk-averse *relative* to state specific policy recommendations.

We demonstrate that time to achieve compliance can be somewhat lengthy and differs across states. Reasons for the heterogeneous responses to policy include mis-matching policy to preferences, potentially confusing policy messaging, and the believability and attainability criteria.

While this paper focuses on a specific event, the COVID pandemic, the results of this research can be easily applied to other similar negative shocks, and will help policy makers in the future formulate recommendations in the future. Researchers are already looking for the next pandemic to hit, see Yong [32], and health institutes, such as the CDC, have begun re-considering their approach to promulgating policy to more clear, concise, and definitive to avoid mis-communication and improve the signal-to-noise ratio.

References

[1] Adolph, C., Amano, K., Bang-Jensen, B., Fullman, N., Wilkerson, J., 2021. Pandemic politics: Timing state-level social distancing responses to covid-19. Journal of Health Politics, Policy and Law 46, 211–233.

[2] Alfaro, L., Faia, E., Lamersdorf, N., Saidi, F., 2020. Social Interactions in Pandemics: Fear, Altruism, and Reciprocity. Technical Report 27134. National Bureau of Economic Research.

[3] Allcott, H., Boxell, L., Conway, J., Gentzkow, M., Thaler, M., Yang, D., 2020. Polarization and public health: Partisan differences in social distancing during the coronavirus pandemic. Journal of Public Economics 191, 104254. doi:`https://doi.org/10.1016/j.jpubeco.2020.104254`.

[4] Anthes, E., 2021. lurching between crisis and complacency: Was this our last Covid surge? The New York Times .

[5] Apple, 2021. Mobility trends report. URL: `https://covid19.apple.com/mobility`.

[6] Barrios, J.M., Hochberg, Y., 2020. Risk Perception Through the Lens of Politics in the Time of the COVID-19 Pandemic. Working Paper 27008. National Bureau of Economic Research.

[7] Berg, J.E., Dickhaut, J.W., Rietz, T.A., 2010. Preference reversals: The impact of truth-revealing monetary incentives. Games and Economic Behavior 68, 443–468. URL: `https://www.sciencedirect.com/science/article/pii/S0899825609001626`, doi:`https://doi.org/10.1016/j.geb.2009.07.008`.

[8] Best, J., 2020. The social problems of COVID-19, in: Muschert, G.W., Budd, K.M., Christian, M., Lane, D.C. (Eds.), Social Problems in the Age of COVID-19: US Perspectives. Bristol University Press, Bristol, UK. volume 1. chapter 13, pp. 134–142.

[9] Cai, J., Deng, X., Yang, J., Sun, K., Liu, H., Chen, Z., Peng, C., Chen, X., Wu, Q., Zou, J., Sun, R., Zheng, W., Zeyao, Lu, W., Liang, Y., Zhou, X., Ajelli, M., Yu, H., 2022. Modeling transmission of SARS-CoV-2 Omicron in China. Nature Medicine 28, 1468 – 1475. doi:`https://doi.org/10.1038/s41591-022-01855-7`.

[10] Clemente Lopez, J., Montas, A., Reyes, M., 1998. Testing for a unit root in variables with a double change in the mean. Economics Letters 59, 175–182.

[11] Center for Disease Control, 2023. COVID data tracker. URL: `https://covid.cdc.gov/covid-data-tracker/#datatracker-home`.

[12] Czeisler, M.E., Howard, M.E., Robbins, R., Barger, L.K., Facer-Childs, E.R., Rajaratnam, S.M.W., Czeisler, C.A., 2021. Early public adherence with and support for stay-at-home COVID-19 mitigation strategies despite adverse life impact: a transnational cross-sectional survey study in the united states and australia. BMC Public Health 21. doi:`https://doi.org/10.1186/s12889-021-10410-x`.

[13] Google, 2021. Community mobility reports. URL: https://www.google.com/covid19/mobility.

[14] Goolsbee, A., Syverson, C., 2021. Fear, lockdown, and diversion: Comparing drivers of pandemic economic decline 2020. Journal of Public Economics 193, 104311.

[15] Gupta, S., Wing, C., Nguyen, T., Bento, A.I., Rojas, F.L., Lee, B., Rahman, S., Simon, K., 2021. Tracking public and private responses to the COVID-19 epidemic: Evidence from state and local government actions. American Journal of Health Economics 0, 222–224.

[16] Hale, T., Angrist, N., Goldszmidt, R., Kira, B., Petherick, A., Phillips, T., Webster, S., Cameron-Blake, E., Hallas, L., Majumdar, S., Tatlow, H., 2021. A global panel database of pandemic policies (oxford COVID-19 government response tracker). Nature Human Behaviour 5, 529–538.

[17] Heuring, E.D., 2021. Empirical analysis of COVID-19 in the U.S. Presentation. Institute for Defense Analyses. Alexandria, VA.

[18] Hixon, T., 2014. What kind of person prefers an iPhone? Forbes URL: https://www.forbes.com/sites/toddhixon/2014/04/10/what-kind-of-person-prefers-an-iphone.

[19] Johns Hopkins, 2021. Coronavirus Resource Center. https://coronavirus.jhu.edu/map.html. Accessed: 2022-2-10.

[20] Kates, J., Michaud, J., Levitt, L., Pollitz, K., Neuman, T., Long, M., Rudowitz, R., Musumeci, M., Freed, M., Cubanski, J., . Comparing Trump and biden on COVID-19. KFF: Coronavirus (COVID-19) .

[21] Malakoff, D., 2020. What if biden wins? Science 370, 284–285.

[22] Mayo Clinic, 2021. U.S. COVID-19 vaccine tracker. https://www.mayoclinic.org/coronavirus-covid-19/vaccine-tracker. Accessed: 2021-9-17.

[23] Perron, P., Vogelsang, T.J., 1992. Nonstationarity and level shifts with an application to purchasing power parity. Journal of Business and Economic Statistics 10, 301–20.

[24] Poletti, P., Ajelli, M., Merler, S., 2011. The effect of risk perception on the 2009 H1N1 pandemic influenza dynamics. PLOS ONE 6, 1–7.

[25] Potter, C.C., Harries, J., 2006. The determinants of policy effectiveness. Bulletin of the World Health Organization 84.

[26] Schwert, G.W., 1989. Tests for unit roots: A Monte Carlo Investigation. Journal of Business and Economic Statistics 7, 147–160.

[27] Singer, P.M., Willison, C.E., Moore-Petinak, N., Greer, S.L., 2021. Anatomy of a failure: COVID-19 in the united states, in: Greer, S.L., King, E.J., da Fonseca, E.M., Peralta-Santos, A. (Eds.), The Comparative Politics and Policy of COVID-19. University of Michigan Press, Ann Arbor, MI. chapter 26, pp. 478–493.

[28] Sonora, R.J., 2022. Estimates of a Pandemic Taylor rule loss function. Applied Economics Letters 0, 1–6. doi:10.1080/13504851.2022.2089619.

[29] Thaler, R.H., Tversky, A., Kahneman, D., Schwartz, A., 1997. The effect of myopia and loss aversion on risk taking: An experimental test. The Quarterly Journal of Economics 112, 647–661.

[30] Thompson, D., 2021. The curious case of florida's pandemic response. The Atlantic .

[31] Tiebout, C., .

[32] Yong, E., 2022. The pandemic's legacy is already clear: All of this will happen again. The Atlantic URL: https://www.theatlantic.com/health/archive/2022/09/covid-pandemic-exposes-americas-failing-systems-future-epidemics/671608/.

Tables

Table 1: Descriptive statistics

		Google Mobility Index (GMI)				Oxford Stringency Index (OxSI)			
Variable	Obs	Mean	Std. Dev	Min	Max	Mean	Std. Dev	Min	Max
DEM	729	-15.67	6.69	-32.87	3.43	51.49	17.15	5.56	82.41
REP	729	-11.08	5.76	-28.15	4.18	46.01	17.15	2.09	79.17
CA (D)	729	-18.74	5.67	-31.83	2.57	54.68	15.57	8.33	82.41
FL (R)	729	-14.82	6.54	-33.11	3.69	43.51	19.68	0.00	81.94
GA (R)	729	-10.75	5.76	-28.31	4.26	47.34	17.40	0.00	74.07
IL (D)	729	-13.18	6.63	-29.63	5.34	48.27	20.10	5.56	82.41
NY (D)	729	-18.55	7.35	-38.83	3.06	47.81	18.17	5.56	82.41
NJ (D)	729	-15.42	7.95	-38.57	3.11	58.38	16.85	8.33	82.41
OH (R)	729	-7.44	6.13	-24.49	5.54	48.10	18.29	5.56	82.41
PA (D)	729	-12.84	6.86	-32.34	3.97	46.50	17.57	5.56	85.19
TX (R)	729	-11.32	5.90	-33.06	5.06	45.31	16.08	2.78	78.24
WA (D)	729	-15.27	6.54	-31.66	4.11	53.43	17.82	0.00	79.63

(Note: Political party is state governor's party affiliation. GMI is in percent.)

Table 2: Behavior regression results: GMI dependent variable

	CA	FL	GA	IL	NJ	NY	OH	PA	TX	WA
$OxSI_t$	-0.112**	-0.047	-0.147**	-0.229***	-0.255***	-0.198*	-0.298***	-0.342***	-0.143***	-0.171***
	(0.034)	(0.035)	(0.050)	(0.046)	(0.034)	(0.089)	(0.080)	(0.065)	(0.039)	(0.034)
$Response$	-0.422***	-0.310*	-0.318***	-0.386***	-0.495***	-0.611***	-0.525***	-0.484***	-0.358***	-0.418***
	(0.034)	(0.125)	(0.049)	(0.066)	(0.046)	(0.081)	(0.078)	(0.043)	(0.024)	(0.041)
GMI_{t-7}	0.498***	0.772***	0.554***	0.690***	0.589***	0.634***	0.542***	0.471***	0.170	0.618***
	(0.096)	(0.078)	(0.096)	(0.088)	(0.067)	(0.078)	(0.081)	(0.089)	(0.165)	(0.057)
$OxSI_{t-7}$	-0.184*	-0.211**	-0.247**	-0.086	-0.176***	-0.280*	-0.041	-0.098	-0.212***	-0.170**
	(0.079)	(0.065)	(0.088)	(0.051)	(0.051)	(0.110)	(0.082)	(0.073)	(0.060)	(0.062)
$OxSI_{t-14}$	0.085	0.188**	0.252***	0.195***	0.228***	0.255***	0.099	0.184**	0.058	0.181***
	(0.061)	(0.065)	(0.070)	(0.050)	(0.036)	(0.072)	(0.068)	(0.066)	(0.048)	(0.051)
$\%VAX_t$	0.792**	-0.057	-0.238	0.389	1.013***	-0.314	1.528***	0.685**	0.754**	1.267***
	(0.248)	(0.343)	(0.332)	(0.236)	(0.299)	(0.194)	(0.351)	(0.215)	(0.290)	(0.281)
$\%VAX_{t-7}$	-0.812**	0.045	0.228	-0.406	-1.011***	0.287	-1.551***	-0.683**	-0.783**	-1.275***
	(0.252)	(0.337)	(0.327)	(0.233)	(0.292)	(0.188)	(0.354)	(0.210)	(0.293)	(0.280)
Weekly cases$_t$	-0.087	-0.034	-0.270**	-0.224*	-0.289***	-0.201*	-0.449***	-0.340**	-0.420***	-0.148*
	(0.050)	(0.061)	(0.096)	(0.106)	(0.072)	(0.086)	(0.121)	(0.104)	(0.108)	(0.067)
TX Outage	–	–	–	–	–	–	–	–	-7.091*	–
									(3.337)	
Biden	-1.762**	0.694	0.231	-1.431	-2.937*	-1.402	-4.643***	-4.268***	-0.309	-2.224*
	(0.636)	(0.796)	(0.761)	(0.940)	(1.145)	(0.738)	(1.136)	(1.047)	(0.580)	(0.870)
Obs	715	713	708	708	715	715	715	715	713	714
F	115.19	111.50	86.98	114.97	154.82	114.76	90.01	122.49	72.06	232.68

Note: Newey-West standard errors are in parenthesis. $^{*}, ^{**}, ^{***}$ represent statistical significance at the 5%, 1% and 0.1% respectively. Other controls include the AAA-BAA risk spread, summer, and the Thanksgiving-Christmas-New Years holiday travel season.

Table 3: DFGLS and CMR one and two break unit root tests

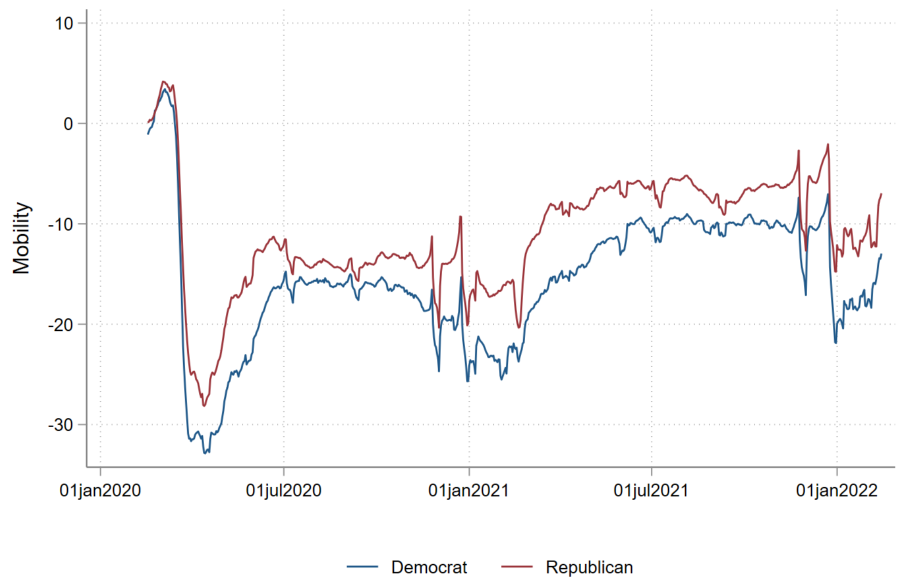
	CA	FL	GA	IL	NJ	NY	OH	PA	TX	WA
ADF test										
$Resid_{t-1}$	-0.115***	-0.148***	-0.166***	-0.152***	-0.190***	-0.140***	-0.149***	-0.154***	-0.132***	-0.141***
t_{DF}	-4.773	-4.493	-4.951	-5.003	-4.934	-4.732	-5.782	-4.691	-5.689	-5.589
One break tests										
AO model										
t_{AO}	-5.078	-4.069	-3.994	-3.908	-3.247	-5.062	-3.363	-3.991	-3.482	-5.890
Break	4/17/2020	NA	NA	NA	5/2/2020	5/2/2020	NA	NA	2/12/2021	5/4/2020
5% cv	**-3.56**									
IO model										
t_{IO}	-4.743	-4.053	-4.085	-4.17	-3.262	-4.24	-3.494	-3.988	-5.449	-4.584
Break	NA	NA	NA	NA	NA	NA	NA	NA	2/13/2021	NA
5% cv	**-4.27**									
Two break tests										
Break 1	4/14/20	11/27/20	12/25/20	1/18/21	NA	4/21/20	5/14/21	1/17/21	2/8/21	1/24/21
Break 2	12/25/20	4/14/21	2/20/21	2/21/21	NA	11/12/20	5/21/21	2/24/21	2/20/21	2/20/21
t_{AO}	-8.808	-6.495	-9.379	-9.981	-8.050	-7.485	-8.087	-8.920	-9.702	-8.534
5% cv	**-5.49**									
Break 1	NA	11/28/20	12/19/20	1/10/21	1/14/21	NA	1/10/21	5/9/21	2/10/21	1/14/21
Break 2	NA	4/7/21	2/15/21	2/17/21	2/3/21	NA	2/17/21	5/19/21	2/16/21	2/14/21
t_{IO}	-9.406	-7.535	-9.899	-10.189	-9.066	-7.129	-8.436	-9.226	-11.090	-8.889
5% cv	**-5.49**									

Note: "NA" means no statistically significant breaks estimated at the 10% level or better.

Figures

Figure 1: Mobility and Stringency: Democrat vs Republican Governors

(a) Google Mobility Index



(b) Oxford Stringency Index

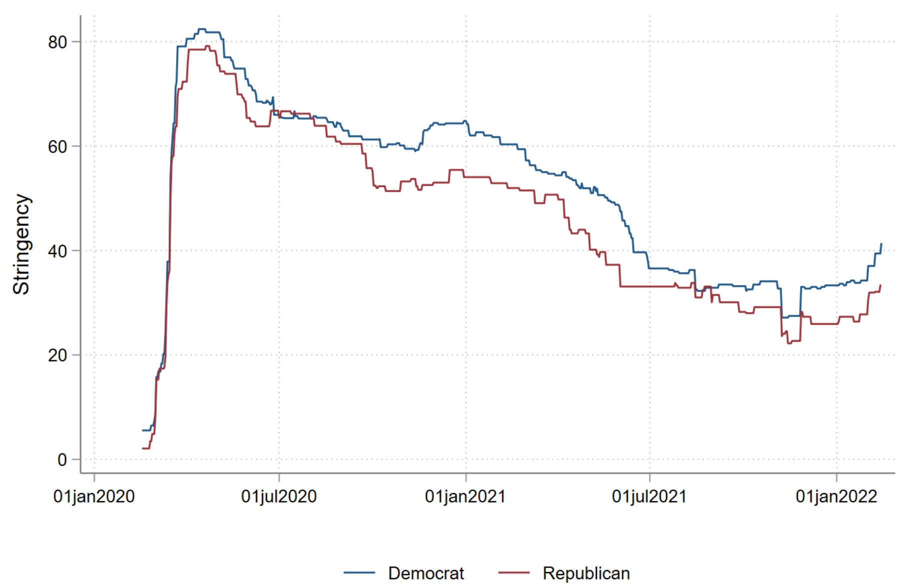
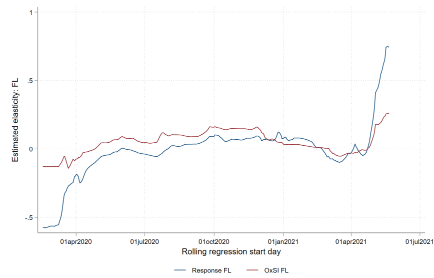


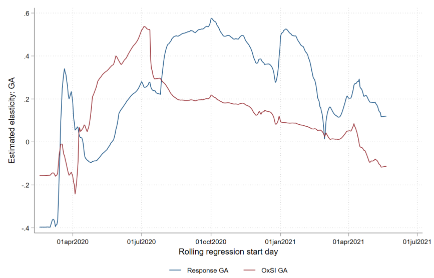
Figure 2: 270 day rolling response elasticities

(a) CA

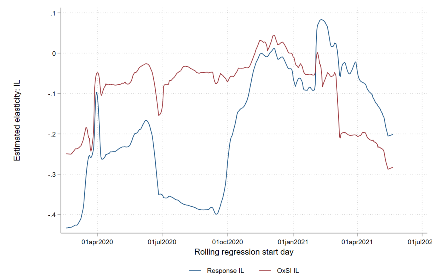
(b) FL



(c) GA

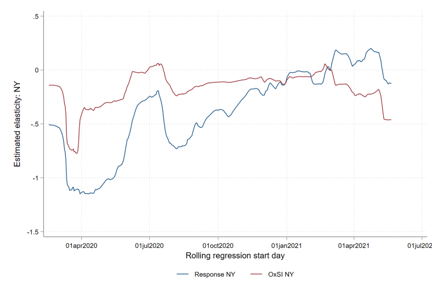


(d) IL



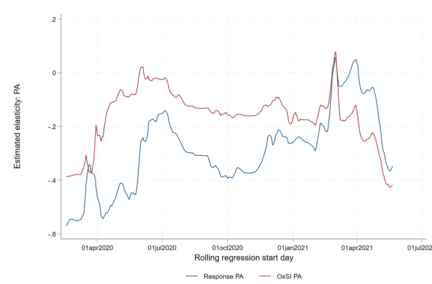
(e) NJ

(f) NY

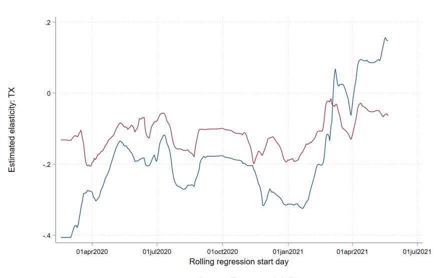


(g) OH

(h) PA



(i) TX



(j) WA

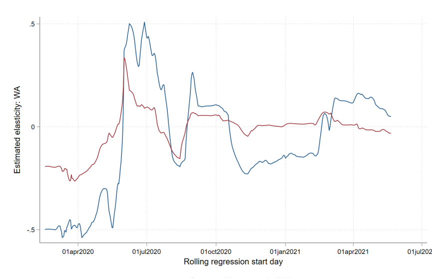
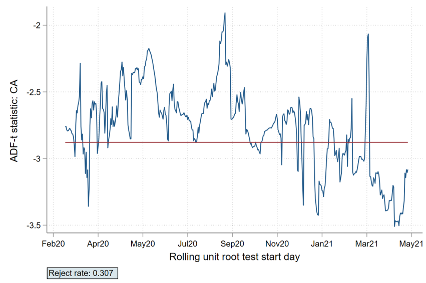
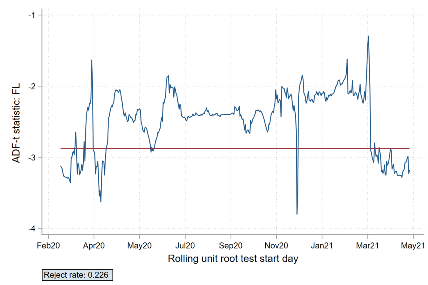


Figure 3: 270 day rolling ADF tests

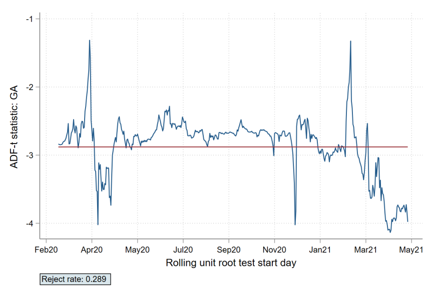


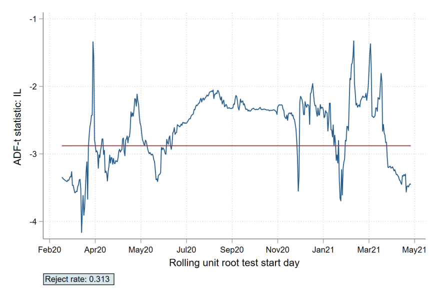
(a) FL



(b) GA

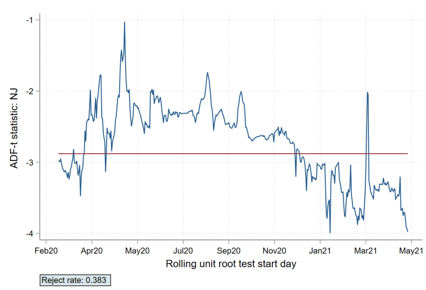
(c) IL

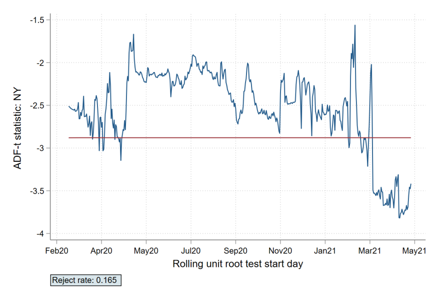




(d) NJ

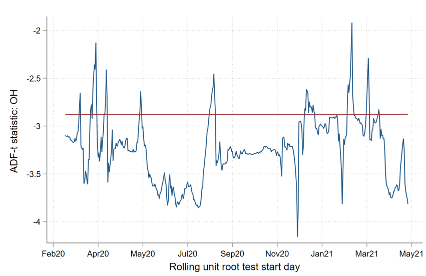
(e) NY

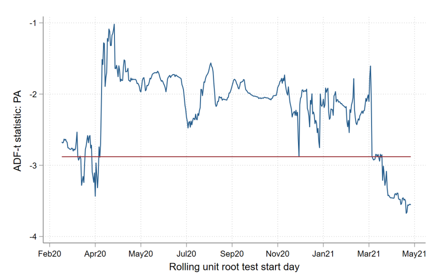




(f) OH

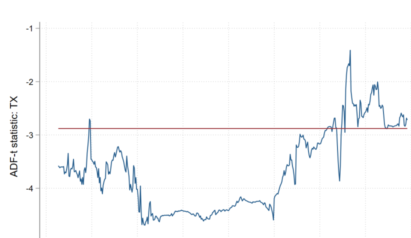
(g) PA





(h) TX

(i) WA



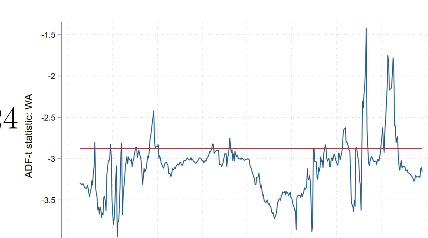
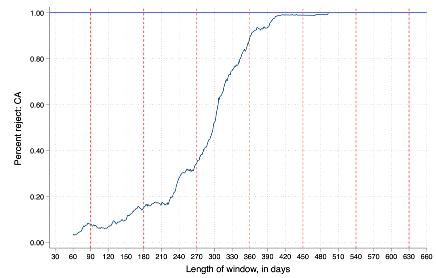
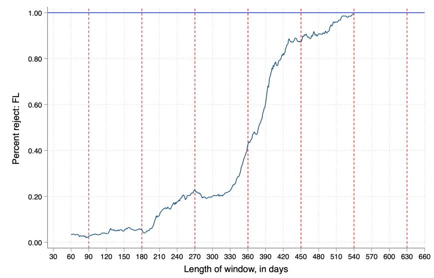


Figure 4: 60 – 540 day rolling ADF rejection rate

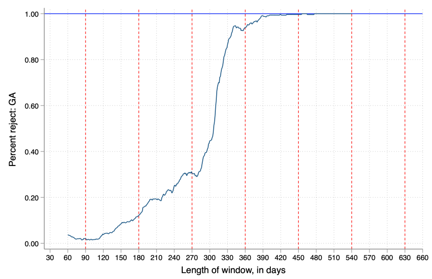
(a) CA



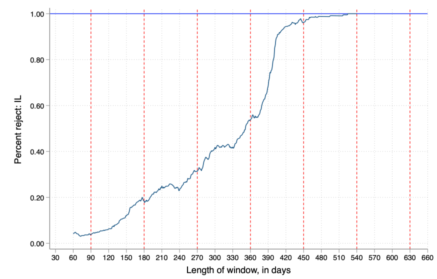
(b) FL



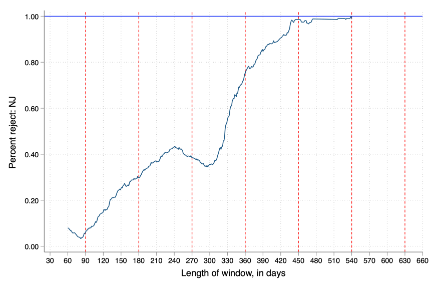
(c) GA



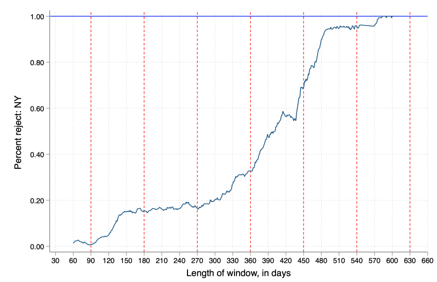
(d) IL



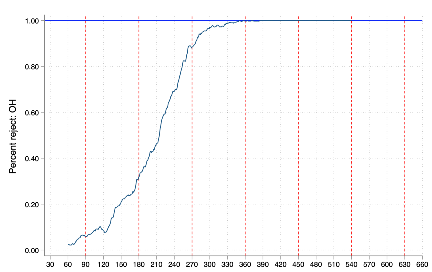
(e) NJ



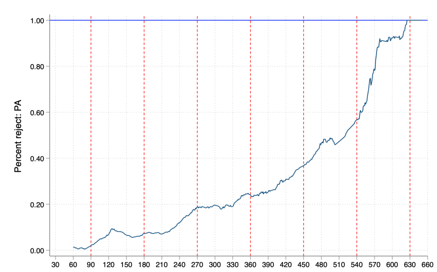
(f) NY



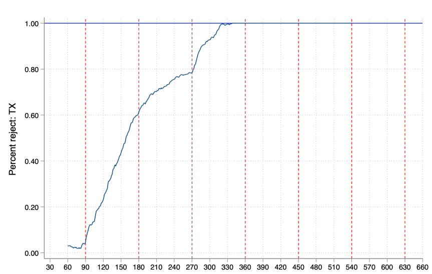
(g) OH



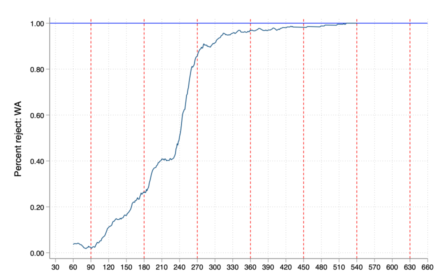
(h) PA



(i) TX



(j) WA



(Note: red dashed lines at 90 day intervals. NY and PA have a maximum of 600 and 650 days, respectively)