Policy Evaluation with Nonlinear Trended Outcomes: Covid-19 Vaccination Rates in the US*

Lynn Bergeland Morgan[a], Peter C. B. Phillips[b,c,d] and Donggyu Sul[f]

[a] *University of Texas at Dallas,* [b] *Yale University,* [c] *University of Auckland,*
[d] *Singapore Management University,*
[f] *University of Texas at Dallas*

April 2, 2025

Abstract

This paper discusses pitfalls in two way fixed effects (TWFE) regressions when the outcome variables contain nonlinear and possibly stochastic trend components. If a policy change shifts trend paths of outcome variables, TWFE estimation can distort results and invalidate inference, especially in a context of evolving policy decisions. A robust solution is proposed by allowing for dynamic club membership empirically using a relative convergence test procedure. The determinants of respective club memberships are assessed by panel ordered logit regressions. The approach allows for policy evolution and shifts in outcomes according to a convergence cluster framework with transitions over time and the possibility of eventual convergence to a single cluster as policy impacts mature. The long run impact of a policy can thus be examined via its impact on convergence club membership. An application to new weekly US Covid-19 vaccination policy data reveals that federal level vaccine mandates produced a merger of state vaccination rates into a single convergence cluster by mid-September 2021.

Keywords: Two Way Fixed Effects Regression, Robust Clustering Algorithm, Relative Convergence, Automatic Clustering Mechanism, Panel logit Regression

JEL Classification: I18, C33, H51.

---

*Thanks go to the Editor and two referees for many helpful comments and suggestions that assisted in revising an earlier version of the paper. Phillips acknowledges research support from the NSF (Grant No. SES 18-50860) at Yale University and a Kelly Fellowship at the University of Auckland. Email: Lynn.Morgan@utdallas.edu; peter.phillips@yale.edu; d.sul@utdallas.edu.

1 Introduction

Two way fixed effects (TWFE) regressions are commonly used in empirical work, particularly to evaluate the effectiveness of an exogenous policy change. Rapidly developing difference in difference (DiD) and event-study literature has revealed that the traditional TWFE regression is not an appropriate method to estimate the average treatment effect on treated units (ATT) in particular treatment settings. Callaway (2023), De Chaisemartin and D'haultfœuille (2023), Roth et al. (2023) and Arkhangelsky and Imbens (2024) provide surveys on this ongoing literature. To implement DiD or event-study estimation, some basic conditions are required: first, controlled and treated units (or groups) need to be clearly defined for DiD estimation. If there is no control unit, then at least treatment timing should be well defined for event-study estimation. Second, the parallel trends condition should be satisfied prior to treatment. If these conditions are not satisfied, ATT is not well defined and hard to identify (Roth (2022)).

The present paper examines a recent empirical case where these specified conditions were not met, namely Covid-19 vaccination policies. And the paper proposes an alternative approach to estimate policy efficacy that takes into account the time path of responses to treatment. Indeed, in attempts to slow the spread of Covid-19 and to increase vaccination rates, a variety of policies were implemented at the federal, state, and local levels in the United States in 2021 following the availability of Covid-19 vaccines. Some such policies were cash lotteries for vaccination, mask mandates, and federal and employer vaccination mandates. While some states were implementing policies to incentivize vaccination and slow the spread of Covid-19, other states responded with contrary policies, such as bans on the proof of vaccination, and mask mandate bans. With such a vast range of policy implementations, it is important to disentangle the multiple policies to identify those which were effective and those which were not in order to understand their impact and inform future policies.

DiD estimation is used often to evaluate the efficacy of a singular binary policy. In our analysis data are at the state level, but some policies were implemented at city or county levels, thereby impacting only a portion of the state. The policy variables may therefore be treated as continuous because they account for the fraction of a state's population exposed to the policy and they were in practice inevitably time varying.[1] Another difficulty is that multiple vaccine and mask policies were implemented with overlapping timing during this period and there were federal policies simultaneously active that impacted all states. Moreover, as is shown later, the parallel trends assumption does not hold, making it nearly impossible to define a control group for a panel DiD analysis, or to define sharp event timing for each vaccination policy. Under the absence of a clearly defined control group, overlapping timing of policies, and continuous time varying policy variables, the commonly

[1]Vaccine mandates for state employees, bans on proof of vaccination, and mask mandate bans are an exception. These three policies were effective state-wide and were therefore binary.

employed approach to estimation is TWFE. But serious drawbacks arise with TWFE implementation when outcome variables include stochastic trends or other nonstationay components.

To fix ideas suppose y_{it} is a relevant outcome variable comprising a sequence of panel observations of interest across individuals $(i = 1, \cdots, n)$ and over time $(t = 1, \cdots, T)$ and let x_{it} be a relevant policy variable. A typical TWFE regression specification for analyzing such data takes the form

$$y_{it} = a_i^o + \eta_t^o + x_{it}'\beta + v_{it}, \tag{1}$$

where a_i^o and η_t^o are individual and time specific effects. When x_{it} is a single dummy variable, we will denote it as D_{it}. Otherwise, we assume that x_{it} is a vector of continuous stationary random variables.[2] The present paper considers general and empirically relevant cases where the policy variable x_{it} may itself evolve over time and where the outcome variable y_{it} may have nonstationary characteristics sourced beyond the policy input. In complex applied work where the nonstationarity in outcomes is not fully captured by policy inputs or time specific effects, the regression residual v_{it} becomes nonstationary and the adequacy of TWFE regressions in accurately measuring policy effectiveness may be called into question. We address the impact of nonstationary outcome variables both with continuous policy variables, as is the case in our empirical example, and with binary policy variables which more readily lend themselves to DiD estimation.

Previous studies often bypass such complications as nonstationary output variables and evolving policy variables by assuming a small fixed T, often simply $T = 2$ observations, and by allowing a large cross section sample size n to deliver asymptotics. In cases like the recent Covid-19 pandemic experience successive policy changes over time need to be accommodated in the data analysis and longer time series samples are available for assessing the evidence of successive policy effects in the observed outcomes. Our empirical study has $T = 40$ and $n = 51$ observations, so that cross section and time series sample sizes are comparable.

We study cases where trends are homogeneous and heterogeneous across individuals under the null of no treatment effect or no policy effectiveness. Under the alternative, policy variables influence outcome variables either in levels or in trends. If a stationary policy variable affects a nonstationary outcome variable in levels, the policy effectiveness based on x_{it} does not last long, holding only temporarily. In this case, a TWFE regression with x_{it} is valid only under homogeneous trends with consistent estimation and valid standard errors as $n \to \infty$. But the properties of estimation and inference are less favorable when the outcome variables are nonstationary with heterogeneous trend behavior. A partial solution is straightforward under the assumption that policy variables affect outcomes *only in levels*. Indeed, level effects can be consistently estimated either under

[2] A set of control variables usually enter in (1) with x_{it}.

homogeneous or heterogeneous trends using the differenced regression

$$\Delta y_{it} = a_i + \Delta\eta_t^o + \Delta x_{it}'\beta + \Delta v_{it}, \tag{2}$$

However, it is unrealistic to expect that policy variables affect outcomes only in levels. When a policy affects trend behavior of the outcome variables, the first difference TWFE regression (2) is misspecified, leading to inconsistent estimation of policy effects.

Our study considers several issues that bear directly on TWFE and DiD treatment effect analysis. The primary concern examines outcome variables when stochastic trend components are omitted from the regression and when heterogeneous trend behavior over time or across individuals affects outcomes. In both cases, the TWFE system is misspecified, leading to growing uncertainty and inconsistent estimation. Our approach allows, in addition, for policy evolution and shifts in outcomes over time according to a convergence cluster framework with transitions over time and the possibility of eventual convergence to a single cluster as policy impacts mature.

One of the main contributions of the paper is to provide a methodology that organizes nonstationary panel data into ordered panel multinomial variables by means of a dynamic clustering mechanism that allows for shifts in clusters over time. To fix ideas, define C_{it} to be the convergence club membership in period t of the i-th individual and suppose all individuals within a certain convergence club share the same stochastic trend. If an individual grows faster over time and joins another convergence club that attains higher outcomes, then C_{it} membership changes over time. The convergence clustering mechanism (CCM hereafter) proposed in earlier work by Phillips and Sul (2007a, hereafter, P-S) transforms statistics from panel observations to clustered cross-sections, but potential dynamic changes amongst club memberships are ignored. In practice, club membership can change over time, particularly as relevant policies are introduced or evolve over time, and such evolution can itself be a natural focus of interest concerning policy impacts. Our approach in the present paper is to develop the CCM of P-S into a dynamic version that accommodates such possibilities. Once dynamic group membership is estimated, panel logit (or multinomial logit) regression enables estimation and inference concerning driver variables and the determining mechanisms of the groups.

A second significant contribution of the paper is to create a new weekly database that tracked state and District of Columbia announcements of the numerous vaccination policies implemented over the period from March 2021 to February 2022.[3] This database enables a detailed empirical study of the impact of federal and state vaccination policies on state vaccination rates.

A final contribution is to address how nonstationary and nonlinear trended outcome variables

[3]These policies included lotteries, cash for vaccination incentives, community outreach programs, vaccine mandates for state employees and or healthcare workers, indoor vaccine mandates or mandates for gatherings over a certain number of people, mask mandates, bans on proof of vaccination, and bans on mask mandates.

impact DiD estimation. Depending on the type of DiD estimator, nonstationary or nonlinear trended outcome variables can affect both identification and inference. While this is not an immediate concern in our empirical example, it deserves attention in other settings and we discuss how DiD is impacted by each outcome type at the end of each subsection in Section 3, which may be of interest to readers using DiD methods. Additionally, our proposed approach for evaluating nonstationary outcome variables – the dynamic club clustering mechanism (DCCM) – is directly related to parallel trends tests in the DiD literature. Accordingly, in this context we address some potential issues and other testing methods available in the panel data econometrics literature.

The rest of the paper is organized as follows. Section 2 considers issues arising from nonstationary outcome variables and pitfalls in the use of first differences in policy evaluation. This section has three subsections. The first two subsections deal with problems of TWFE regressions when stationary policy variables affect the level and trend of nonstationary outcomes. The third subsection provides a short discussion of how nonstationary outcomes can affect various DiD estimators. Dynamic mechanisms for club membership are developed in Section 3, which also has three subsections. The first subsection provides a short review of the relative convergence test and how it can be used to test for a homogeneous (stochastic) trend. We also discuss how this relative convergence test can be utilized for testing parallel trends in DiD estimation. The second subsection provides a step-by-step procedure for implementing the dynamic convergence clustering mechanism (DCCM). The final subsection explains how to use the estimated club membership to evaluate policy effectiveness. Section 4 provides an application of the DCCM to state Covid-19 vaccination rates in the United States. Panel logit modeling is applied to assess the effects of U.S. federal level mandate announcements and various state level Covid-19 vaccine policies on state convergence club membership. Section 5 concludes. Technical background, derivations, proofs, federal and state level mandate policies, additional logit regression specification results, and further simulations are provided in the Online Supplement to this paper.

2 Pitfalls of TWFE Regressions with Nonstationary Outcomes

Throughout the paper we assume that the dependent or outcome variables are nonstationary, whereas the policy variable is stationary.[4] We consider the following three general models allowing nonstationary outcomes when a policy variable does not influence y_{it} (i.e., under the null hypoth-

[4]As shown later in the empirics, key policy variables often tend to change in a staggered manner over time. After removing the steps these variables are typically flat or stationary about certain levels. As such, they impact outcome variables only temporarily. On the other hand a policy that switched on and off at different levels of intensity could produce nonstationarity in mean.

esis):

$$y_{it} = \begin{cases} a_i + t + \xi_{it},\ \xi_{it} = \xi_{it-1} + u_{it} & \text{for M1,} \\ a_i + b_i t + \xi_{it},\ \xi_{it} = \xi_{it-1} + u_{it} & \text{for M2,} \\ a_i + b_{it} t + \xi_{it},\ \xi_{it} = \xi_{it-1} + u_{it} & \text{for M3,} \end{cases} \tag{3}$$

where the nonstationarity includes heterogeneous deterministic trend functions as well as a stochastic trend. Specifically: in M1, y_{it} has a homogeneous linear trend; in M2, y_{it} has a heterogeneous linear trend with time invariant individual coefficients b_i; and in M3, the individual trend coefficients b_{it} are time varying, which nests M1 and M2. A policy variable, x_{it}, is assumed to be stationary. One may introduce a staggered or continuous trend to x_{it}[5], but such modifications do not alter the main results of the paper.

Under this framework there are two mechanisms by which the stationary policy change can affect the nonstationary outcome under the alternative: (i) x_{it} can affect the level of y_{it} (designated C1); and (ii) x_{it} can affect the trend coefficient b_{it} (designated C2). Under the null of no policy effectiveness, the outcome variables are given by (3). In this case the parameter β, which measures the policy impact in the following equations, is zero. The two alternative cases are written as follows.

$$y_{it} = a_i + \eta_t + \beta x_{it} + \xi_{it} \ \text{ for C1,} \tag{4}$$

$$b_{it} = b_o + \beta x_{it} + e_{it} \ \text{ for C2,} \tag{5}$$

where η_t in (4) may be a linear or more general deterministic trend and b_{it} are the trend coefficients in M3. We consider C1 in the next subsection, showing that if the true data generating process is either M1 or M2, a first difference regression can account for the nonstationary process in C1. C2, however, is a more realistic formulation that has been used by labor and health economists in situations where a stationary policy affects the long run behavior of outcome variables. For example, Abadie et al. (2010) showed that California's tobacco control program changed the slope of the trend in tobacco consumption. In such situations first difference regression fails, as shown in Section 2.2. An alternative solution is provided in Section 3.

2.1 Policy influence only on levels: C1

This model has been popularly considered in the treatment effects literature. Without a homogeneous trend assumption, difference-in-difference estimation fails to identify the treatment effects. Even with homogeneous trends, when a policy variable x_{it} is stationary in a model such as (4), it is not a long run determinant of the outcome variable and impacts the outcome only temporarily. An empirical example of this phenomenon was the 'Cash for Clunkers' program implemented in the

[5] As shown in Figure 3 there were staggered increases in the Federal vaccine mandate, whereas the vaccine lottery had been in decline continuously after July in 2021.

United States in 2009. Under that program, eligible car owners were given a substantial rebate to trade in their old cars for new ones. The purpose of the program was twofold: to stimulate a weak economy by encouraging car sales, and to reduce CO_2 emissions by getting rid of older, less fuel efficient cars. The program caused a temporary, moderate increase in new car sales during the two months it was in place, however there was no increase in overall car sales for the year. Essentially, some people who would have bought a new car anyway that year did so in the couple of months the program was in place rather than spreading their purchases across the year. The end result in new car sales was no different than it would have been in the absence of the program (Li et al. (2013)).

Against this background consider the following TWFE regression frequently used in applied research

$$y_{it} = a_i + \eta_t + \beta x_{it} + v_{it}, \tag{6}$$

where η_t is a time fixed effect and the regression error v_{it} in (6) is the nonstationary component ξ_{it}. Let $\hat{\beta}_l$ be the TWFE estimator of β in (6). To simplify notation, rewrite (6) in differenced form as

$$\dot{y}_{it} = \beta \dot{x}_{it} + \dot{v}_{it}, \tag{7}$$

where $\dot{y}_{it} = y_{it} - \frac{1}{T}\sum_{t=1}^{T} y_{it} - \frac{1}{n}\sum_{i=1}^{n} y_{it} + \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T} y_{it}$, with similar definitions for the other variables. This within-group transformation does not remove stochastic nonstationarity but it does eliminate heterogeneous individual fixed effects and any homogeneous (over individuals) trend. The consequences of using this type of TWFE regression with a policy variable x_{it} in place (such as the 'Cash for Clunkers' program) arise from the fact that the regression is ill-balanced because stochastic trend effects remain in $\dot{v}_{it}$ and are transmitted to $\dot{y}_{it}$, making policy effects only temporary. The lack of balance induces a reduction in the signal to noise ratio that affects the accuracy of the TWFE estimator, $\hat{\beta}_l$ in (7). The properties of $\hat{\beta}_l$ depend in this case primarily on the cross section sample size n, so that increases in the number of time series observations do not help to shrink the variance of $\hat{\beta}_l$.[6] More formally, as $n, T \to \infty$, when x_{it} affects y_{it} only in levels as in (4), the limit distribution of $\hat{\beta}_l$ is given by

$$\sqrt{n}(\hat{\beta}_l - \beta) \to \mathcal{N}(0, V_\beta), \tag{8}$$

with asymptotic variance V_β shown in the Online Supplement (Theorem 1). A panel robust variance estimator, clustering by time, can be used to consistently estimate V_β.

Next, consider the heterogeneous trends case M2. Many dependent variables used in empirical TWFE regressions can be expected to involve heterogeneous trends in the data. For example, if the unit of observation is counties, and the dependent variable is obesity levels, the slope of obesity growth almost certainly varies county by county. This heterogeneity in trends is a complication

[6] In this regression, the signal from the sample variance of $\dot{x}_{it}$ is dominated by the sample covariance of $\dot{x}_{it}$ and $\dot{v}_{it}$ because the regression error is nonstationary. See the Online Supplement for further analysis.

that affects asymptotic behavior and finite sample performance even when some of the regressors themselves include trends. As long as time-homogeneous coefficients are assumed in estimation, the regression error carries the effects of heterogeneous trends, leading to failure in standard limit theory and inference. To illustrate, suppose y_{it} is generated from

$$y_{it} = a_i + b_i t + \beta x_{it} + \xi_{it}, \tag{9}$$

where ξ_{it} is nonstationary.[7] Then the induced TWFE residual in (7) includes a trend term under the null $\beta = 0$, i.e.,

$$\dot{v}_{it} = (b_i - \frac{1}{n}\sum_{i=1}^{n} b_i)(t - \frac{1}{T}\sum_{t=1}^{T} t) + \dot{\xi}_{it}, \tag{10}$$

so that cross section variation of $\dot{v}_{it}$ grows at an $O(t^2)$ rate even when ξ_{it} is stationary. When x_{it} affects y_{it} in levels under heterogeneous trends, the asymptotic behavior as $(T, n) \to \infty$ of $\hat{\beta}_l$ are shown in Theorem 3 of the Online Supplement to have the following form

$$\sqrt{n}(\hat{\beta}_l - \beta) = O_p(T^{1/2}) + \mathcal{N}(0, V). \tag{11}$$

The following first difference (FD) regression can be used to bypass this problem

$$\Delta y_{it} = a_i + \Delta\eta_t + \beta\Delta x_{it} + e_{it}, \tag{12}$$

or in two-way transformed form as

$$\Delta\dot{y}_{it} = \beta\Delta\dot{x}_{it} + \dot{e}_{it}. \tag{13}$$

Importantly, the FD regression includes two way fixed effects. The individual fixed effect a_i captures potential heterogeneous trend coefficients, while the time fixed effect $\Delta\eta_t$ captures common time changes. The regression error e_{it} is now stationary. Let $\hat{\beta}_{\text{fd}}$ be the TWFE estimator in (13). It is straightforward to show that as $T, n \to \infty$,

$$\sqrt{nT}(\hat{\beta}_{\text{fd}} - \beta) \to \mathcal{N}(0, V_{\beta_{\text{fd}}}), \tag{14}$$

where $V_{\beta_{\text{fd}}}$ is the asymptotic variance of $\hat{\beta}_{\text{fd}}$, which is smaller than V_{β_l} in (8). This limit theory is valid only if x_{it} is known to influence just the level of y_{it}. Tables 1 and 2 in the Online Supplementary Appendix report Monte Carlo simulation results that detail these effects.

In practice, it may be challenging to assess whether x_{it} affects the level or the trend behavior of y_{it} or both. For when the outcome variable itself has a trend, any level change is temporary and

[7] The asymptotic results do not change even when ξ_{it} is stationary under heterogeneous trends

permanent effects or long run impacts are inevitably influenced by the trend or trend coefficients. It may therefore appear reasonable to assess how a policy change affects growth rates rather than level outcomes, suggesting a TWFE regression in differences as in (2). However, if the outcome variable has nonlinear or time-variable trend behavior then a TWFE regression in differences also fails to capture the impacts of a policy change. This issue is now discussed.[8]

2.2 Policy influence on trends: C2

For further analysis of potential pitfalls in TWFE estimation consider the data generating process (DGP)

$$y_{it} = a_i + b_{it}t + \xi_{it}, \text{ with } \xi_{it} = \rho\xi_{it-1} + e_{it}, \tag{15}$$

where intercepts a_i affect the level of y_{it}, the b_{it} are heterogeneous time varying trend coefficients, and the ξ_{it} are time series with a unit root when $\rho = 1$.[9] Now suppose that a policy variable x_{it} influences b_{it} where x_{it} is a simple dummy variable. In this case, the policy changes the slope of the trend in the outcome variable, y_{it}, impacting its long run trajectory. Assume $x_{it} = 0$ for all i except for $i = 1$ and $t \geq \tau$, which is a typical setting in the synthetic control literature.

To avoid spurious effects and restore test power, as mentioned in subsection 2.1, common empirical practice is to take the first difference of the outcome variable, y_{it} when it is nonstationary, as it is in (15). The pitfalls that can arise from the practice of taking first differences when a policy impacts long run trend behavior of an outcome variable are a prime focus of the present paper. In C2 where x_{it} affects b_{it}, regressions either in first differences or growth rates of y_{it} fail to deliver a satisfactory proxy for the temporal impact b_{it}. For instance, taking differences of (15) gives

$$\Delta y_{it} = b_{it}t - b_{it-1}(t-1) + \Delta\xi_{it} = b_{it} + \Delta b_{it}(t-1) + \Delta\xi_{it}, \tag{16}$$

The reason for taking first differences is to eliminate the trend from y_{it}, rendering Δy_{it} stationary. But when the impact of a policy is nonlinear, such as when b_{it} changes at some point in time, taking first differences does not eliminate the time trend effects. Hence, regression of Δy_{it} on fixed effects (both individual and time specific) and x_{it} produces misleading findings about the impact of the policy because of the missing trend and policy effects in the component $\Delta b_{it}(t-1)$.

Instead of a linear trend, we can use a general nonlinear trend function. To provide an intuitive demonstration of how a time varying coefficient affects the regression, we consider a simple time

[8] A more technical discussion is given in the Online Supplement.

[9] If the DGP (9) changes to (15), the asymptotic effects are similar but the reported simulation results deteriorate and these now depend on the variance of b_{it}.

series without any stochastic random variables given by

$$y_t = a + b_t \eta_t, \text{ with } b_t = x_t,$$

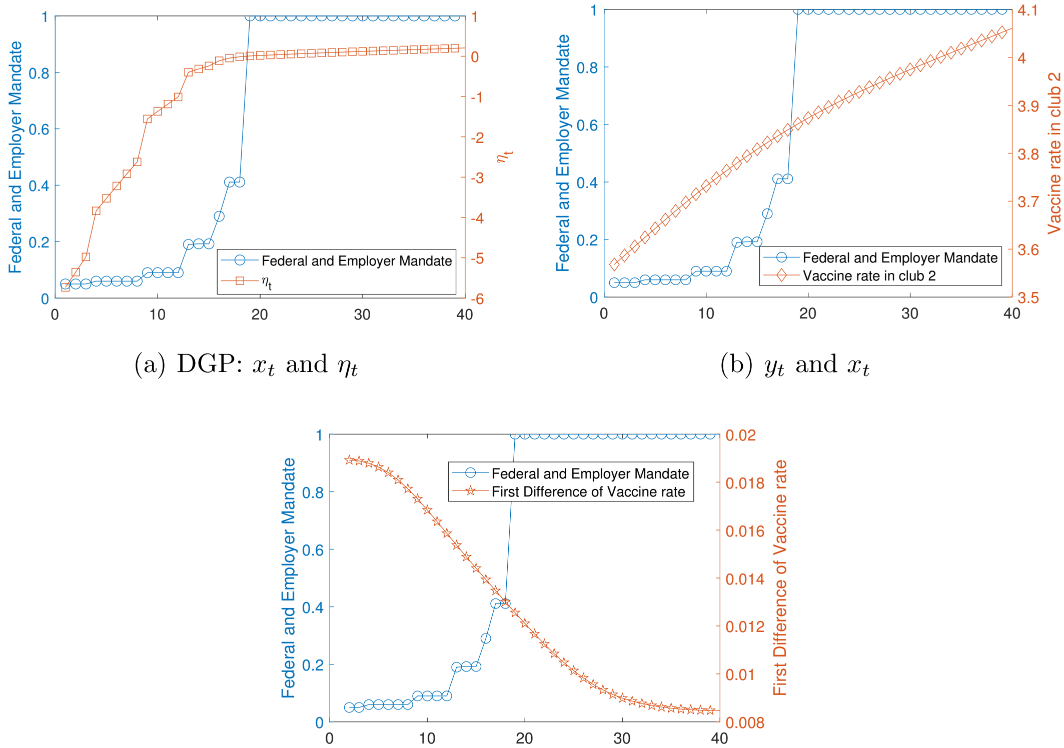
where η_t is a unknown nonlinear trend, x_t is a policy variable, and b_t is the coefficient on the unknown nonlinear trend. As we will show later, the time series of federal and employer vaccine mandates (FEVM) for Covid-19 vaccinations in the United States is positively correlated with the state trends in cumulative vaccination rates. To fix ideas, let y_t be the cross sectional average of vaccination rates in these states, which we will later call 'Club 2'; and x_t is FEVM. After eliminating a by taking the time series mean of y_t, we can calculate η_t by dividing $y_t - \frac{1}{T}\sum_{t=1}^{T} y_t$ by x_t. This hypothetical trend, η_t, and x_t are plotted in Figure 1(a). As the figure shows, the calculated η_t does not have a simple linear trend but rather a nonlinear trend. Next, Figure 1(b) shows FEVM and y_t. Obviously, FEVM seems to have a positive relationship with y_t, but due to the nonstationary or nonlinear trending nature of y_t, this relationship is not stable in the long run as we discussed in the previous subsection. Figure 1(c) displays FEVM and Δy_t. As the figure shows, Δy_t has a negative relationship with FEVM even though η_t is generated under the restriction of $b_t = x_t$.

To summarize, we examine the use of TWFE in nonstationary outcome variables with three different types of trends: homogeneous trends (M1), heterogeneous linear trends (M2), and our main focus, heterogeneous nonlinear trends (M3). The use of TWFE with M1 will result in misleading results because any impact from x_{it} is only temporary and will not alter the long run outcomes of the variable of interest. Running a TWFE regression with an M2 outcome variable results in divergent asymptotic behavior and there is little test power to detect policy impacts. Taking first differences can alleviate the lack of power that affects TWFE regressions when the DGP of y_{it} contains linear trends. But TWFE regressions formulated with first differenced outcomes are not suited to evaluate the effectiveness of policy changes if y_{it} involves nonlinear trend effects. In such circumstances the problem of how to evaluate the effectiveness of policy changes is of considerable empirical interest. Two solutions are discussed in section 3.

2.3 Influence of Nonstationary Outcomes on DiD Estimation

TWFE estimation is often used in DiD analysis but the DiD approach covers a broad range of estimators. The impact of nonstationarity, particularly with stochastic trends, of the outcome variables on the estimated treatment effect depends on the type of DiD estimation. DiD identifies causal treatment effects of a policy change by comparing a treated group with an untreated control group. Identification hinges on the parallel trends assumption, which requires that the two groups have parallel trends prior to policy implementation so that, in the absence of the policy change, the trajectory or time trend of the treated and untreated groups would continue to be parallel. Any

Figure 1: Data Generating Process and Observed Series



(a) DGP: x_t and η_t

(b) y_t and x_t

(c) Δy_t and x_t

Notes: y_t is the cross sectional average of the log cumulative vaccination rates for states initially in 'Club 2', the states with relatively low initial cumulative vaccination rates. x_t is the FEVM, and we decompose $\tilde{y}_t = y_t - T^{-1}\sum_{i=1}^{T} y_t = b_t\eta_t$ where $b_t = x_t$. Given this relationship the unknown trend, η_t, is calculated as $\tilde{y}_t/b_t$. Thus, x_t has a one to one relationship with b_t. Figure 1(a) displays the FEVM variable, x_t and the calculated hypothetical trend variable, η_t. The figure clearly shows the nonstationary nature of η_t. Figure 1(b) plots the mean cumulative vaccinations y_t alongside x_t. The two are clearly positively correlated but, due to the nonstationary and nonlinear trending nature of y_t, this relationship is not stable in the long run, as discussed in subsection 2.1. Figure 1(c) displays the FEVM and the first-differenced average log cumulative vaccination rate, Δy_t. The pitfalls discussed in subsection 2.2 are readily apparent in this figure, as there is a clear negative relationship between the FEVM and Δy_t even though η_t is generated using $x_t = b_t$.

subsequent deviation from the pre-trend can therefore be attributed to the policy. Literature on DiD estimation has grown rapidly the recent years. De Chaisemartin and D'haultfœuille (2023), Roth et al. (2023), and Arkhangelsky and Imbens (2024) survey some of this rapidly developing literature.

Impact of Nonstationary Outcomes on ATT

Before considering the impact of nonstationary outcomes we address how TWFE estimation fails to identify the ATT in a traditional DiD setting based on the following TWFE regression

$$y_{it} = a_i + \eta_t + \beta D_{it} + e_{it}, \tag{17}$$

Under the homogeneous trend model (M1) or parallel trend setting in (3), the TWFE estimator is known to be inconsistent. To see this, rewrite (7) as

$$\dot{y}_{it} = \beta \dot{D}_{it} + \dot{e}_{it}, \tag{18}$$

where $\dot{y}_{it}$ is fully demeaned y_{it} defined following (7). It is easy to see that $\dot{D}_{it}$ includes negative numbers due to two-way within group transformation. Naturally, the estimated treatment effect based on (18) becomes inconsistent. (See Goodman-Bacon (2021) and De Chaisemartin and d'Haultfoeuille (2020) for detailed discussions.) There are two alternative estimators. The first type uses Δy_{it} and directly calculates weights for the treatment effects. (See De Chaisemartin and d'Haultfoeuille (2020) and Callaway and Sant'Anna (2021)). For example, De Chaisemartin and d'Haultfoeuille (2020)'s estimator ($\mathrm{DiD_M}$) is the weighted average

$$\mathrm{DiD_M} = \frac{n_{(1,0),t}}{n_s} \sum_{t=2}^{T} \mathrm{DiD}_{+,t} + \frac{n_{(0,1),t}}{n_s} \sum_{t=2}^{T} \mathrm{DiD}_{-,t} \tag{19}$$

of the two estimators

$$\mathrm{DiD}_{+,t} = \frac{1}{n_{(1,0),t}} \sum_{D_{it}=1, D_{it-1}=0} \Delta y_{it} - \frac{1}{n_{(0,0),t}} \sum_{D_{it}=0, D_{it-1}=0} \Delta y_{it},$$

$$\mathrm{DiD}_{-,t} = \frac{1}{n_{(1,1),t}} \sum_{D_{it}=1, D_{it-1}=1} \Delta y_{it} - \frac{1}{n_{(0,1),t}} \sum_{D_{it}=0, D_{it-1}=1} \Delta y_{it},$$

where $n_{(j,s),t}$ is the total number of units with $D_{it} = j$ and $D_{it-1} = s$ for each t and $t-1$ with $j = 0, 1$ and $s = 0, 1$. And n_s is the total number of cases when $n_{(0,1),t}$ and $n_{(1,0),t}$ occur during the same time periods. Callaway and Sant'Anna (2021)'s estimator can be expressed in a similar way based on Δy_{it}. We call them 'direct DiD' (D-DiD) estimators. The second type of estimator uses D_{it} instead of $\dot{D}_{it}$ in (18), giving the regression equation

$$\dot{y}_{it}^{N} = \beta D_{it} + e_{it}^{*}, \tag{20}$$

where $\dot{y}_{it}^{N} = y_{it} - \hat{a}_i^N - \hat{\eta}_t{}^N$, $\hat{a}_i^N$ and $\hat{\eta}_t^N$ are individual and time fixed effect estimators using only non-treated observations. Roth et al. (2023) call the resulting estimator the 'imputation DiD (I-DiD)' estimator. (See Borusyak et al. (2024), Gardner (2022) and Liu et al. (2024) for more discussion).

Now, assume y_{it} follows M1. In this case it is evident that D-DiD does not have any estimation and inference issue since it is constructed based on Δy_{it}. Denote $n_{I,t}$ and $n_{N,t}$ as the total number of treated and non-treated units at time t. Further let T_I be the total number of treatment times, and $n_m = \min[n_{I,t}, n_{N,t}]$ for all t. Then as $n_{m,t}, T_I \to \infty$, we have

$$\sqrt{n_m T_I}(\hat{\beta}_{\text{D-DiD}} - \beta) \to_d \mathcal{N}(0, \sigma^2_{\text{D-DiD}}) \tag{21}$$

On the other hand the I-DiD estimator may suffer from inference problems since the regression error e^*_{it} is nonstationary, leading to a lack of balance in the regression as discussed earlier in relation to (7). So I-DiD estimation of (20) means that the resulting estimator $\hat{\beta}_{\text{I-DiD}}$ has similar asymptotic behavior as in (8). That is,

$$\sqrt{n_m}(\hat{\beta}_{\text{I-DiD}} - \beta) \to_d \mathcal{N}(0, \sigma^2_{\text{I-DiD}}), \tag{22}$$

where there is no panel gain asymptotically by considering more T_I observations.

Violation of Parallel Trend Assumption

The parallel trends assumption is the cornerstone of the identification of the ATT using DiD. Several recent papers deal with the problems of pre-trend testing. Roth (2022), Rambachan and Roth (2023), Kahn-Lang and Lang (2020), and Bilinski and Hatfield (2018) provide critiques of the conventional pre-parallel trend test and its alternative. In the presence of a heterogeneous trends (model M2) the parallel trends assumption no longer holds and both D-DiD and I-DiD estimators are influenced by trend effects. In the dynamic heterogeneous trends model (M3), as long as the parallel trends assumption holds in the pre-treatment period, the D-DiD estimator can consistently capture the treatment effect. But inference based on the I-DiD estimator remains dependent on whether the regression errors are stationary. Test power is particularly sensitive: in stationary cases power approaches unity when n or T rises, whereas in nonstationary cases power rises only as n rises and quickly stabilizes as T becomes large for a given value of n. A simulation study in the online supplementary appendix illustrates these findings.

In the next section we demonstrate how a relative convergence test can be used for testing for parallel trends in more complex settings that allows for a general evolution of the trend environment that helps to address these issues.

3 A Dynamic Clustering Approach

This section describes a clustering approach to the evaluation of policy impacts in the presence of trending outcomes, allowing for coalescence or divergence. The first part introduces the relative convergence test proposed by P-S, which can be used to test for a common nonlinear or stochastic trend in outcome data. If a panel of interest shares a common nonlinear or stochastic trend, then the TWFE regression in (13) is valid since x_{it} cannot influence the homogeneous trend coefficient heterogeneously.[10] The second part provides a methodology for transforming nonstationary panel data into stable multinomial club membership using a recursive clustering algorithm.

3.1 Testing for Homogeneous Trends

As in P-S, the starting point is to represent trending multidimensional data in terms of a panel components model as

$$y_{it} = b_{it}\eta_t, \tag{23}$$

where η_t is an unknown trend which may be either a stochastic process or a nonlinear deterministic time trend. The representation in (23) is general and typically has unidentified multiple components. For instance, if the DGP were $y_{it} = a_i + b_i t + \xi_{it}$ where $\xi_{it} = \xi_{it-1} + e_{it}$, then we could rewrite the model in the form of (23) as $y_{it} = (a_i t^{-1} + b_i + \xi_{it} t^{-1})t = b_{it} t$.

If $y_{it}/y_{jt} \to 1$ as $t \to \infty$, then we say that y_{it} relatively converges to y_{jt} over time. Let $\hat{\mu}_t$ be the sample cross section average of y_{it}. If $y_{it}/\hat{\mu}_t \to 1$ as $t \to \infty$ for all i, then the panel y_{it} is said to be relatively convergent to its cross section average. The ratio $h_{it} := y_{it}/\hat{\mu}_t$ traces out a *transition path* over time that manifests convergence when $h_{it} \to 1$. In this case, y_{it} shares the same (stochastic) trend, which is factored out in the ratio, thereby enabling analysis and inference about convergence. Figure 2 shows the relative transition paths for two groups in our empirical example. 'Club 1' and 'Club 2' are the cross sectional averages of the relative transition curves for high and low vaccination states, respectively. In this case there was no overall convergence early in the period, as evident in the initially diverging paths of the two clubs. If there were initial convergence, we would have only a single club and the ratios would converge with $\frac{y_{it}}{\mu_t} \to 1$ for all states i. Instead, up until the 10th week the distance between two clubs increased, showing clear evidence of divergence. Following the 10th week the distance between them decreases but remains larger than the initial distance for the next 9 weeks. After the 19th week, however, the distance between the two groups continues to decrease and is soon less than the original distance, suggesting a merger of the two clubs to a larger single convergence club.

The test for relative convergence in P-S relies on the following (so-called $\log t$) regression

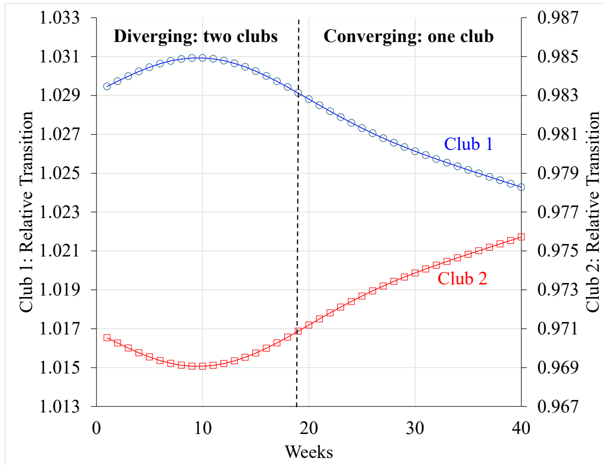
$$\log \frac{H_1}{H_t} - 2\log\left(L\left(t\right)\right) = a + b\log t + e_t, \tag{24}$$

which is estimated by ordinary least squares and where

$$H_t = \frac{1}{n}\sum_{i=1}^{n}\left(h_{it} - 1\right)^2, \text{ for } h_{it} = \frac{y_{it}}{\hat{\mu}_t}, \tag{25}$$

[10] Identification of the long run determinant then becomes of interest because in this case x_{it} does not affect the long run trend of y_{it}. Analysis of this case is a separate matter and left for future research.

Figure 2: Relative Transitions for Club 1 and 2



Notes: The blue line with circles is the relative transition path of states that were initially in Club 1, where the Club 1 relative transition path is defined as the ratio of the average cumulative vaccination rate of states in Club 1 (C_1) at time t, divided by the national average cumulative vaccination rate at time t, $\frac{1}{n_1}\sum_{i\in C_1}^{n_1}\frac{y_{it}}{\hat{\mu}_t}$. Similarly, the relative transition path of Club 2 (C_2), $\frac{1}{n_2}\sum_{i\in C_2}^{n_2}\frac{y_{it}}{\hat{\mu}_t}$, displayed in red with squares, charts the ratio of the average vaccination rate of states belonging to Club 2 to the national average. As the figure shows, initially the ratio of Club 1 vaccination rates to the national average was increasing, meaning that the states in Club 1 had cumulative vaccination rates that were increasing relatively faster than the national average. Necessarily, this means that the vaccination rates in Club 2 were increasing relatively slower than the national average. This trend continued through week 10, at which point the largest distance between the relative transition paths of Club 1 and Club 2 was reached. From weeks 10 onward the distance between the two paths decreases indicating convergence to a single club toward the end of the period.

$L(t) = \log t$, $t = p+1, \ldots, T$, $p = \lfloor r \times T \rfloor$ with $r = 1/3$, and $\lfloor \cdot \rfloor$ is the integer floor function.

Under the null of relative convergence, H_t is asymptotically converging to zero over time since $h_{it} \to 1$ as $t \to \infty$. Hence, $\log(H_1/H_t)$ is increasing over time. If the t-value for $\hat{b}$ exceeds -1.65, then the null of relative convergence is not rejected in the test at the 5% level. Note that in finite samples the term involving $2\log(L(t))$ serves as a penalty function in the regression (24), as explained in P-S. Under relative divergence, H_t and $\log(H_1/H_t)$ should increase and decrease over time, respectively. Under fluctuations over time, H_t simply fluctuates, but in view of the penalty function of $-2\log(L(t))$, the dependent variable in (24) decreases over time. Hence, the fitted OLS coefficient $\hat{b}$ becomes significantly less than zero in this case.

For present purposes in the empirical evaluation of policy effects under trending outcomes, if the null of convergence is not rejected, then the TWFE regression with Δy_{it} is well justified since in the long run the panel y_{it} is identified as having a homogeneous (stochastic or nonlinear) trend. If the null is rejected, then data analysis may be conducted as described in the next section.

Parallel Trends Tests

The relative convergence test can be used to test for parallel trends. To be specific let τ be the treatment timing. Then if the panel y_{it} relatively converges for all i up to $t = \tau$, then the outcome variable shares the same common (nonlinear) stochastic trend during the pre-treatment period. After τ, if a staggered treatment becomes effective in the long run, then the treated y_{it} should diverge from the original convergent club or develop to form another convergent cluster.

But if the panel y_{it} is relatively converging over the entire period then this evidence supports the hypothesis that there is no treatment effect on the trend coefficients or no long run treatment effect. Relative convergence means that while all the y_{it} might not be converging to the same value, their relative relationship to each other becomes closer over time. Thus, if treatment is randomly assigned and the treated and untreated are relatively converging, we can conclude that the treatment was ineffective in the long run. As will be shown in the next section, if the y_{it} form two convergent clubs from $t = 1$ to $t = \tau$, then the null of parallel trends is rejected. In this case, any DiD estimators are distorted due to bias from nonparallel (stochastic) trends.[11]

3.2 Dynamic Clustering Mechanism

One possible outcome is that there are few subgroup convergence clubs, but each of these clubs diverges from the others over time, in which case a null of overall club convergence would be rejected. P-S suggested how to find convergent subgroups by using a *convergence clustering mechanism* (CCM). This mechanism transforms the full $(n \times T)$ panel dataset into a club membership structure that features each individual member $(n \times 1)$. The CCM requires finding a core convergence club within the panel. Once a core club is identified, each individual time series is compared with the core group and is added to the convergence group if it relatively converges. Otherwise, the individual is classified to another group. Successive repetition of this procedure identifies members of the first convergent club. The clustering algorithm is then repeated with non-members of the first convergent club. The approach allows empirical researchers to explore the underlying determinants of club membership through multinomial logit regression of club membership on driver variables, as suggested in Phillips and Sul (2007b) and Phillips and Sul (2009).[12]

$$y_{it} \sim I(1) \rightarrow \boxed{CCM} \rightarrow C_i$$

The present paper utilizes this approach to design a robust method of clustering club membership over time. The proposed method is straightforward and involves recursive implementation of the CCM algorithm over time to identify the clubs and cluster evolution over time. By allowing club

[11] There is an alternative method to test parallel trends by means of the weak $\sigma-$convergence test proposed in Kong et al. (2019), which makes use of cross sectional variance rather than $\log H_t$. Kong et al. (2019) considered the null of no $\sigma-$convergence, but their null hypothesis can readily be converted to the null of no divergence. If the cross sectional variance either shrinks or fluctuates over time, the parallel trends condition holds. Compared with the relative convergence test, the weak $\sigma-$convergence test is more attractive. This is because the relative convergence test needs to discard some initial sample observations and requires large T since the limit theory reasoning relies on $T \rightarrow \infty$ asymptotics. Weak $\sigma-$convergence testing need not discard any initial data and works fairly well in a large n but small T settings. More thorough asymptotic analysis and simulations will be provided in future work.

[12] See Sul (2019) for more detailed discussion.

membership to evolve over time, empirical researchers can use the DCCM method to examine the impact of policy changes on club membership, thus allowing for the evaluation of the long run policy impacts. As we will show in the next section, various patterns of dynamic evolution over time can be identified by recursively estimating club memberships in this way.

This dynamic version of the CCM approach (viz., DCCM) transforms nonstationary y_{it} to stationary panel ordered variables, i.e.,

$$y_{it} \sim I(1) \rightarrow \boxed{DCCM} \rightarrow C_{it} \sim I(0).$$

DCCM employs some modifications of the original algorithm including a fixed rule for initialization in the recursive regressions[13] and a fixed rule for core member detection.[14] To clarify, we provide the following step-by-step procedure showing how to perform the DCCM.[15]

Step by Step Procedure for DCCM

Step 0 *Data Preparation:* Let Y_{it} be a raw level variable.

Step 0-1 *Measurement unit invariance:* Multiply Y_{it} by 10^k to get Y_{it}^* where k is a large constant such as 10. Take the log of Y_{it}^*. If some of Y_{it}^* are zero, add one.

Stpe 0-2 *Subtract minimum:* Define $y_{it}^* = \log Y_{it}^*$ and $y_{\min}^* = \min_{1\leq i\leq n, 1\leq t\leq T} y_{it}^*$. Further let $y_{it} = y_{it}^* - y_{\min}^*$. See Section 3.3 in the Online Supplementary Appendix for how this method eliminates any measurement unit issues.

Step 1 *Fixed rule for initial discarding:* Delete the first 5 or 6 time series samples. See Section 3.4 in the online supplementary appendix for further discussion.

Step 2 *Robust selection for core group:* To distinguish if each y_{it} is relatively converging to a single convergent club, we need a baseline core group. Using the full sample, find a core group using the $\max t$ rule described in P-S. Usually the minimum number of a core group is 3, and is typically less than 10. Using the full sample ensures robustness of the core group membership.

Step 3 *Initial clustering:* With the fixed core from Step 2, run the CCM with a sub-sample of the data, $t = 1, ..., T_1$ where T_1 should be larger than 10. Otherwise, the core group selection fails due to lack of degrees of freedom. Record the club membership for T_1. Assign $\hat{C}_{i,1} = \hat{C}_{i,T_1} = 1$ if $i \in G_1$ where G_1 is the first convergent club.

[13]The $\log t$ test in P-S requires initialization of the regressions, eliminating some early observations. The discard rule used in P-S removed the first 1/3 observations. The rule is problematic in the present implementation because the sample size changes in recursive regression. Instead, a fixed rule is used here in which the first 5 or 6 observations are discarded. The Online Supplement provides further discussion.

[14]The CCM algorithm estimates the initial core members based on the sample observations. To maintain the core membership in the recursive approach, the core members are fixed in the recursion by employing the entire sample in their initial detection.

[15]Stata code for the DCCM is available upon request.

Step 4 *Recursive clustering:* Increase the subsample size to $t = 1, ..., T_1 + 1$ and re-run CCM using the fixed core from Step 2. Record the club membership, $\hat{C}_{i,T_1+1}$. Repeat this procedure by adding one additional observation at a time until the whole sample is included.

Step 5 *Constructing second convergence club:* Collect $y_{it} \notin G_1$. Run the $\log t$ regression in (24). If the null of relative convergence is not rejected, then $y_{it} \notin G_1$ forms a separate convergent club, G_2. If the null is rejected, then repeat Step 2 through Step 4 with $y_{it} \notin G_1$. Continue this process until all y_{it} belong to a convergence club, or are classified as divergent members (see P-S for more details).

If the time series sample is large enough and the initial T_1 is not small (say, $T_1 > 25$), then the estimated core group is rather robust, so that it does not need to be fixed in Step 2. After modifying the data, one can use 'logtreg' command in Stata for clustering at each $t \geq T_1$. See Section 3.5 in the Online Supplementary Appendix for a more detailed discussion.

The asymptotic justification for the clustering method is given in Appendix C of P-S. As long as the number of core members is not large relative to the time series sample, consistency of the clustering mechanism is easily achieved. Consistency of the dynamic version of the CCM approach can be achieved in the same way. The Online Supplementary Appendix provides details of this method and reports findings of its finite sample performance from Monte Carlo simulations.

If we knew the specific functional form of dependence and values of b_{it} it would be straightforward to evaluate its determinants by running a regression of b_{it} on the relevant function $f(z_i, \theta_{x,t}, x_{it})$. In practice, finding a specific functional form is challenging and consistent estimation of all of the idiosyncratic trend coefficients b_{it} is not possible using only the data y_{it}.[16] To avoid these issues, we utilize the clustering method in P-S, as now explained.

Define J as the number of convergent sub-groups: $j = 1, \cdots, J$. The original algorithm was designed to provide club memberships based on descending order of the final observation values, y_{iT}. Hence, the first convergent sub-group always dominates the remaining convergent sub-groups. Define $\hat{C}_{it}$ as the estimated membership emerging from the application of DCCM for the ith individual from 1 to t. This transformation changes nonstationary outcomes y_{it} to stationary club memberships, $\hat{C}_{it}$.

Policy Evaluation Based on DCCM

As discussed in the Introduction and previous subsection, if y_{it} forms a single convergent club before the implementation of any treatment, then DiD estimation could be employed to evaluate multiple binary policies by using the estimation method in De Chaisemartin and D'haultfœuille (2023). If policy variables are not binary, then in practice researchers may assign zero and one based on

[16] P-S simplifies this challenge by approximating b_{it} using a relative transition curve h_{it} as explained earlier in (25). However, in this approach the sample cross-sectional mean is often not a robust measure and can be sensitive to outliers, so it is not used here.

various threshold points. Such threshold assignment is an ad hoc device, but at least it allows for the evaluation of various policies jointly. Problems arise, as we will show shortly, when as in our application a federal policy (or policy that is common to all units) is the only effective policy. In this case there are no control or nontreated units. Additionally, it is not straightforward to design an event study analysis when various policies are implemented at different but overlapping times.

When the parallel trends assumption does not hold, or y_{it} forms multiple convergent clubs (or clusters) before treatment, then DiD estimators are not well defined for y_{it}. Instead, club membership, C_{it} grouping is a simple way to deal with this problem. For example, consider two units (y_{1t} and y_{2t}) over three time periods $(-1, 0, 1)$. Assume y_{1t} is increasing over time continuously, but y_{2t} decreased initially from $t = -1$ to $t = 0$, so that the parallel trend assumption does not hold. The club membership becomes $C_{1t} = 1$ but $C_{2t} = 2$ if $t \leq 0$. Further assume that a treatment is given only to y_{2t} at $t = 1$ and at $t = 1$, y_{2t} starts to increase and catch up with y_{1t}. Then, the club membership of y_{2t} changes from 2 to 1 at $t = 1$. In this case, the treatment is effective. However, accurate estimation of the ATT is difficult unless a suitable synthetic control could be constructed. This interesting and challenging task is beyond the scope of this study, but will be considered in future work.

Next, assume that a treatment is given to both convergent clubs, but the dynamic patterns are the same as before: y_{1t} increases continuously but y_{2t} shows a V-shape pattern, with inflection at the time of treatment. Even though the same treatment is given, the treatment for y_{2t} is more effective, which is a violation of the random treatment assumption. Nonetheless, the change in club membership provides evidence of effectiveness of the treatment. In our empirical example, initially states with lower vaccination rates may have had larger vaccine hesitant populations. As shown later, the federal vaccine mandate policy is associated with movement in which the lower vaccinated states converge with the higher vaccinated states. In this case, therefore, even though the federal vaccine mandate pertained to all states equally, it had a stronger impact on states with lower vaccination rates. Quantifying the treatment effect from the federal mandate is not straightforward, however. Similar to the previous case, it is difficult to measure the ATT accurately since the parallel trends assumption does not hold, and more importantly there is no comparison group. We discuss how to address this delicate issue in the next subsection.

3.3 Panel Ordered or Logit Regression

If the membership does not change over time, one does not need to run panel ordered logit, but instead simply run an ordered logit regression with the final club memberships at time T. Otherwise, the next panel ordered logit regression can be run. The j-th ordered logit model is given by

$$\text{logit}\left[\Pr(\hat{C}_{it} \leq j)\right] = a_j + z_i'\gamma_j + \theta_{x,t}'\lambda_j + x_{it}'\beta_j \text{ for } j = 1, \cdots, J-1, \tag{26}$$

where $\theta_{x,t}$ is a vector of known common policy variables, such as macro factors including market interest and inflation rates, and federal policy changes. Once the known common factors are included in the regression, panel fixed effects cannot be identified.[17] Note that all variables on the right hand side influence the trend coefficients in y_{it}. Variables not affecting the trend behaviors of y_{it} must not be significantly different from zero.

In the case of two sub-convergent clubs, as is the case in the empirical study of the next section, instead of a panel ordered logit regression one needs to run a panel logit regression with random effects. In this case, $\hat{C}_{it} = 1$ or 2, and the ordered logit regression becomes

$$2 - \hat{C}_{it} = 1\{a + z_i'\gamma + \theta_{x,t}'\lambda + x_{it}'\beta + e_{it} \geq 0\}, \tag{27}$$

Note that neither conditional logit nor ordered logit regressions can identify λ since $\theta_{x,t}$ is common across individuals so that the conditional likelihood function eliminates $\theta_{x,t}'\lambda$ automatically.[17]

The economic interpretation of β in (26) and (27) is different from that of conventional TWFE regression. Since the dependent variable is club membership, the marginal effect of x_{it} is of interest, which indicates the change in probability when x_{it} increases by one unit. The unconditional logit regression in (27) also provides potentially complex causal effects to explain club membership. To see this, rewrite (27) with a single variable, ignoring x_{it} and e_{it}. Further, let $\gamma = \lambda = 1$ and $a = 0$. If the i-th state is in the first convergent sub-group at time t, then assign $\hat{C}_{it} = 1$; if it is in the second convergent sub-group, then assign $\hat{C}_{it} = 2$. Assume that $\theta_{x,t}$ is a federal policy which implements after $t \geq \tau$ (so $\theta_{x,t} = 0$ if $t \leq \tau$, otherwise $\theta_{x,t} = 1$), and z_i is a particular state characteristic variable. Then, depending on the values of z_i, the federal policy may influence the club membership differently.

Importantly, the proposed methodology is designed to identify determinants of club memberships rather than estimate overall treatment effects. If y_{it} includes a nonlinear trend and exogenous policy variables cause changes in the trend behavior, the proposed method can identify the relevant exogenous variables but not estimate the overall treatment effects on y_{it}. But it is possible to deduce relative information about the implied treatment effects. For instance, if y_{it} forms a single convergence club before treatment but forms two convergence sub-groups after treatment, define $n_{\mathcal{G}_{a,t}}$ as the number of individuals in $\mathcal{G}_a$ at time t for $a \in \{1, 2\}$. Then the average outcomes for

[17]Consider, for example, the following conditional logit model with a single common factor and a single policy variable with two individuals for notational convenience : $2-\hat{C}_{it} = 1\{a_i+\lambda\theta_{x,t}+\beta x_{it}+e_{it} \geq 0\}$ with $i = 1, 2$. The conditional probability at time $t = 1$ becomes $\frac{\exp(\lambda\theta_{x,1}+\beta x_{11})}{\exp(\lambda\theta_{x,1}+\beta x_{11})+\exp(\lambda\theta_{x,1}+\beta x_{21})} = \frac{\exp(\beta x_{11})}{\exp(\beta x_{11})+\exp(\beta x_{21})}$. So λ cannot be identified with observed $\theta_{x,t}$.

each sub-group can be estimated directly from the data by

$$\hat{\mu}_{\mathcal{G}_1} = \frac{1}{n_{\mathcal{G}_{1,t}}} \sum_{i \in \mathcal{G}_1} \frac{1}{T - \tau - 1} \sum_{t \geq \tau, \hat{C}_{it}=1}^{T} y_{it}, \tag{28}$$

$$\hat{\mu}_{\mathcal{G}_2} = \frac{1}{n_{\mathcal{G}_{2,t}}} \sum_{i \in \mathcal{G}_2} \frac{1}{T - \tau - 1} \sum_{t \geq \tau, \hat{C}_{it}=2}^{T} y_{it}. \tag{29}$$

The difference between these two averages provides an estimate of the difference between the two average overall outcomes, thereby giving information about the relative impact of treatments or policies. In fact, if x_{it} were all binary policy variables, then either D-DiD or I-DiD estimation could be used to estimate ATT accurately in this case.

If these two convergent sub-groups eventually merged into a single convergent group from some point $t > \tau$ and more importantly such an event were caused by policy changes, then the implied treatments in (28) and (29) would become the same in the long run.[18] In this case, the implied treatment effects in the long run could be measured by the difference of the cross sectional averages of y_{it} between the two clubs at $t = \tau$.

$$\hat{\mu}_{\mathcal{G}_1,\tau} - \hat{\mu}_{\mathcal{G}_2,\tau} = \frac{1}{n_{\mathcal{G}_{1,\tau}}} \sum_{i \in \mathcal{G}_1} y_{i,\tau} - \frac{1}{n_{\mathcal{G}_{2,\tau}}} \sum_{i \in \mathcal{G}_2} y_{i,\tau}. \tag{30}$$

We finally note that in any convergence sub-group empirical analysis some individuals may end up outside any of the identified groups. Such individuals are treated in P-S as outliers or divergent members of the population. In this case the relevant individuals display behavior outside the framework of the convergence analysis and these outcomes may need separate empirical study to explain their behavior.

4 Covid-19 Vaccination in the US

As an empirical illustration, our recursive club clustering methodology was applied to state level Covid-19 vaccination rates in the US and panel logit regressions were employed to explore the impact of vaccination policies on actual vaccination rates. By late spring of 2021 COVID-19 vaccinations were widely available in the US, but vaccination rates began to plateau even though only roughly 45% of the targeted US population were fully vaccinated by mid-May 2021.[19] There

[18]In our covid-19 application reported in the next section, there were initially two separate groups that diverged and later merged into a single convergent group following a White House announcement of an upcoming federal vaccine mandate – a situation where the parallel trend assumption was violated.

[19]The term *fully vaccinated* was defined at the time as two doses of the Pfizer or Moderna vaccine, or a single dose of the Johnson & Johnson vaccine.

was also substantial variation in state vaccination rates: Maine had the highest vaccination rate at 49% in mid-May 2021; and Mississippi had the lowest at the time, with only 26% of residents fully vaccinated.

Determinants of this variation in state vaccination rates are naturally of considerable interest to policy makers, epidemiologists, and social scientists. Some preliminary research conducted over the summer of 2021 pointed to partisanship having a strong association with vaccination rates. Specifically, it was found that the percentage of votes cast for Donald Trump in the 2020 presidential election was a primary predictor of vaccination rates: the higher the Trump vote, the lower the vaccination rate, on average. Around the same time in 2021, cities, counties, and states attempted to bolster their waning vaccination rates by implementing various vaccine incentive campaigns such as vaccine lotteries and cash for vaccination. Numerous studies have examined the efficacy of such incentives in various states and counties across the United States, and have come to differing conclusions. Some found modest increases in vaccinations resulting from vaccine lotteries or cash incentives, while others found no statistical evidence that these lotteries or cash incentives increased vaccinations, even finding small negative impacts in some cases. Table 1 provides reference details for some of these explicit findings in the literature.

Table 1: Findings of Vaccination Incentives

Small Positive Effect	Zero or Small Negative Effect
Barber and West (2022)	Chang et al. (2021)
Brehm et al. (2022)	Dave et al. (2021)
Sehgal (2021)	Lang et al. (2022)
Wong et al. (2022)	Thirumurthy et al. (2022)
	Walkey et al. (2021)

By late summer 2021, policies mandating vaccinations for particular sub-populations were announced and implemented at the state and the federal levels. To examine the impact of vaccination incentives and policies on vaccination rates, we assembled a dataset of vaccination policies and incentives at the state level, including policies that were implemented in large cities or counties within a state. Section 5 in the Online Supplement explains how the state policy dataset was constructed and Table 8 in the appendix provides summary statistics for the various state level policies.

In addition, we created a separate federal level vaccine mandate variable. This variable includes information from a combination of vaccine mandate announcements that were national in scale. More specifically, it includes announcements signaling upcoming military and federal mandates, the military vaccine mandate announcement, the various vaccine mandates announced by President Biden on September 9, 2021, and mandates by private employers that typically were national in scope.[20] Details on the construction of the federal vaccine mandate variable are given in the Online Supplement: Section 5.2 and Table 9 in the appendix displays the relevant events and dates that were used.

Our state level vaccination data came from the publicly available county level data from the Centers for Disease Control and Prevention (CDC), spanning the period from December 13, 2020 to February 9, 2022. Data prior to May 12, 2021 was discarded because Covid-19 vaccines were initially in short supply and difficult to obtain, meaning that discrepancies in vaccination rates across states during this early period may not have been voluntary but simply due to availability. By mid-May 2021 Covid-19 vaccines were easily accessible in most areas of the United States. Daily county level data were converted to weekly state level data and logarithms of the resulting vaccination rates were recorded. There were a few points of decreasing cumulative vaccination rates for a short period in a small number of states in the data, which was likely due to state or county reporting errors. To correct for these, we applied Stata's HP filter[21] with a smoothing parameter of 1600,[22] and then subtracted y_{min}, as suggested in Section 3.1 of the Online Supplement.

As Figure 3 shows, the implementation percentages of both federal level mandates, and state level mandates are very low through late July 2021, at which point they both sharply increase and then stabilize. State vaccination mandates stabilize around 0.45, meaning about 45% of states had some form of state level vaccination mandate in place by September 2021. Federal vaccine mandates sharply increase over roughly the same time period, although the increase is more stairways than a single sharp jump, as is the case with state level mandates. The federal level mandates level off at an implementation rate of 100% since all states were impacted by this federal level mandate. Vaccination lotteries also show a sharp increase, but the increase is several weeks earlier than the vaccine mandate increases. This shows that states initially tried to incentivize people to get vaccinated through positive incentives. After peaking in June of 2021, the use of lotteries to incentivize vaccination started to steadily fall until converging at zero.

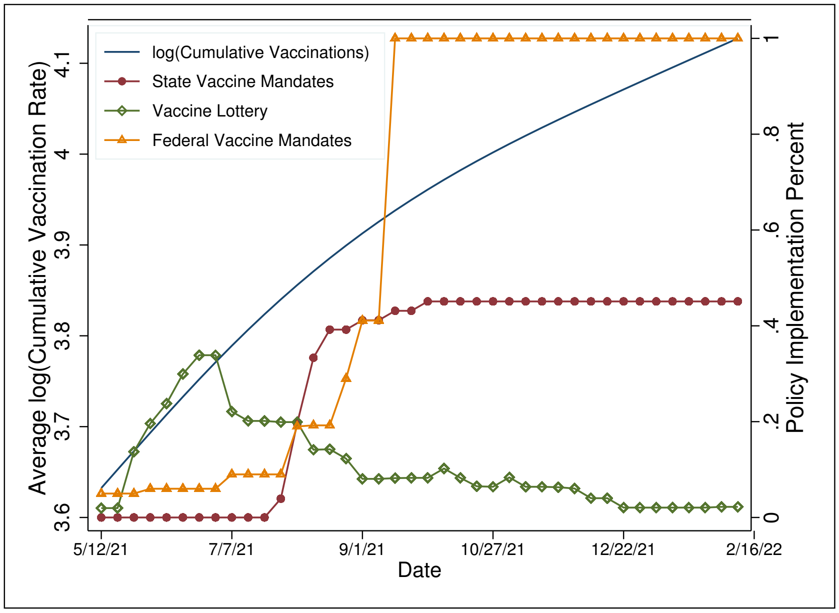
Cumulative vaccination rates are partial sum time series and therefore typically stochastically

[20]Not all of the federal mandates announced by President Biden on September 9, 2021 were ultimately implemented. The mandate on large employers was struck down in court, and the mandate for employees of federal contractors was blocked for months before an August 31, 2022 announcement from the federal government that it would not be enforcing the mandate.

[21]The HP filter is by far the most commonly used filter in empirical studies that have employed the P-S CCM algorithm. A boosting iteration (Mei et al., 2024; Phillips and Shi, 2021) to the filter is known to estimate deterministic and stochastic trends consistently (even in cases of nonstationary fractional process trends, as shown in Biswas et al. (2024)) and this modification is often well captured by a single extra iteration or by suitable changes in the smoothing parameter. We considered other filters in the Online Supplementary Appendix with little changes in the results.

[22]Since the data are weekly, higher values than the quarterly smoothing parameter 1600 are sometimes preferred. In the present case our goal is to smooth the series only moderately because use of a smoothing parameter that is too large produces a filtered trend that is almost linear. Provided the smoothing parameter is neither very small nor very large, the empirical results were not sensitive to the specific choice. Details are provided in the Online Supplement.

Figure 3: log(Cumulative Vaccination Rate) and Select Policy Common Trends.



Notes: The left vertical axis scale is the logarithm of the national average cumulative vaccination rate after filtering out the growth components using the HP filter, which corresponds with the smooth blue line. The right vertical axis scale is the adoption percentage of vaccination policies. The maroon, green, and orange lines are measured on the right axis, and each is the national average of the policy at each time t. The maroon and green lines are sample state level vaccination policies, vaccine mandates for state employees and/or healthcare workers, and vaccination lotteries, respectively. The orange line is federal level vaccine mandates. Log cumulative vaccinations do not show any jump or discontinuity over time and the path of this variable appears impervious to the policy variables being enacted.

nonstationary, and as previously noted in comparing Maine's cumulative vaccination rate in May 2021 to that of Mississippi, states had heterogeneous trends in vaccination rates from the initial roll out in December 2020 through the beginning of our sample period in May 2021. For the reasons explained in Section 2, these characteristics suggest caution in the use of TWFE to assess the impact of vaccination policies on cumulative vaccination rates. Using first differences (new vaccination numbers) as the dependent variable does not resolve the potentially misleading findings from TWFE regressions since the vaccination data involve nonlinear heterogeneous trends (see Section 2.2).

To address heterogeneity and nonlinearity in the trend behavior of the vaccination data, our empirical approach was to classify state vaccination rates into groups where homogeneous trends were manifest. The groupings were obtained by applying the DCCM described in Section 3 to log cumulative vaccination rates, y_{it}, producing individual club membership data $\hat{C}_{it}$, denoting the estimated convergence club that state i belonged to at time t. As previously noted, $\hat{C}_{it}$ takes on the value of 1 or 2 when there are two convergence clubs. Also, as noted earlier, when vaccination rates began to plateau in each state a variety of vaccination incentive schemes and policies were announced and implemented. A dynamic version of the CCM was necessary to explore the evolutionary relationship between club membership and individual state policy implementation.

We first applied the CCM on the full sample period from May 12, 2021 to February 9, 2022, using $m = 6$, where m is the first sample observation used in the regression, a setting that matches findings from the simulations reported in Section 4 of the Online Supplement. The t-ratio from the initial $\log t$ regression of the entire data set was 7.24, so the null of convergence was not rejected, implying that vaccination rates of all fifty states plus the District of Columbia were converging to the same long run national average in February, 2022. The core group from from the P-S CCM of the full sample consisted of eight members: Connecticut, the District of Columbia, Maine, Maryland, Massachusetts, New York, Rhode Island, and Vermont. This club membership outcome is a static full sample result that is uninformative regarding the actual process of convergence and, in particular, the important empirical question of whether state convergence may have occurred without intervention or whether vaccination policies impacted state convergence over time.

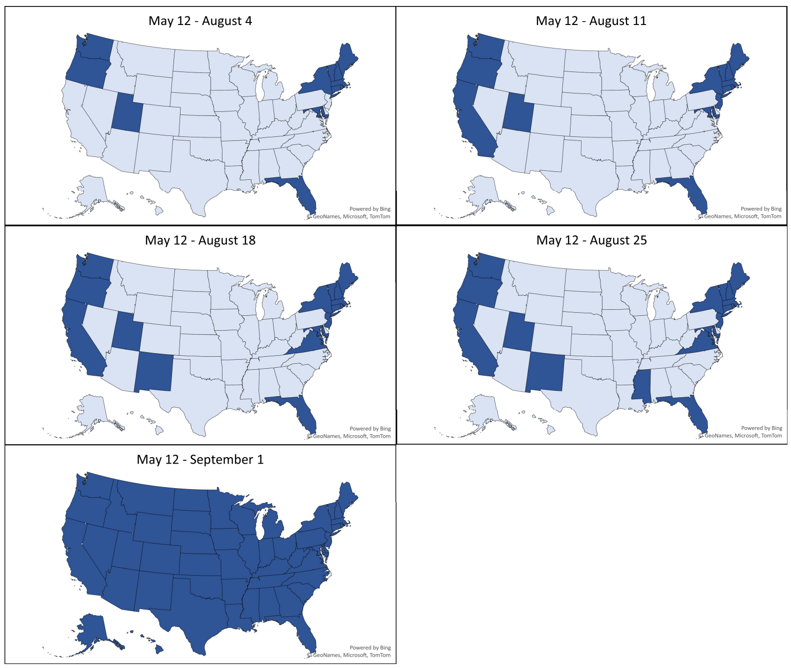
To address this question the DCCM was employed with recursive sampling to estimate club membership evolution over time. Core membership was fixed to the aforementioned seven states plus the District of Columbia. From this core membership we applied the remaining steps of the clustering process using the first thirteen weeks of data from May 12, 2021 to August 4, 2021, as described in Section 3. Interestingly, with this shortened dataset involving vaccination rates only from the late spring to the mid-summer of 2021, there was no evidence of convergence to a single long run average. Instead, in the mid-summer of 2021 there were two distinct clubs: one comprising thirteen members (twelve states plus the District of Columbia) with relatively high vaccination rates, and a second club consisting of 38 states with relatively low vaccination rates.

Working from the given initial club membership obtained for the original (May 12 - August 4) sample, a recursive analysis was commenced by adding a further week to the original sub-sample, making the new sub-sample span from May 12, 2021 to August 11, 2021. The clustering algorithm was re-applied, using the same fixed core from the initial CCM on the full sample. The resulting outcome again produced two clubs, but Club 1 had all original thirteen members in the higher vaccination group that were in the May 12, 2021 - August 4, 2021 sample plus two additional states. This process was continued adding one week at a time and re-running the CCM. With the initial sub-samples, the addition of each additional week to the sub-sample resulted in more states joining the relatively high vaccination club, Club 1, each week. This pattern continued until the sub-sample included data from May 12, 2021 - September 1, 2021. At this point, all of the states relatively converged, forming a single convergence club. Relative convergence to a single club continued to hold for each additional week included in the DCCM until the recursion covered the entire sample, May 12, 2021 to February 9, 2022.

Figure 4 shows the dynamic membership evolution of Club 1 from May 12 through September 1, 2021. The top left panel shows the twelve states plus the District of Columbia in Club 1 from

May 12, 2021 - August, 2021. Each subsequent panel shows the states belonging to Club 1 as the sub-sample recursively expands. Club 1 membership evidently grows over time, as can also be seen by the diamond markers in Figure 5. The last panel of Figure 4, in the lower left position, shows that when the sub-sample includes the weeks from May 12, 2021 through September 1, 2021, all the US states are seen to have the same club membership and full convergence applies. (See Figure 2 for the relative transition paths of Clubs 1 and 2 as they move toward convergence over time.)

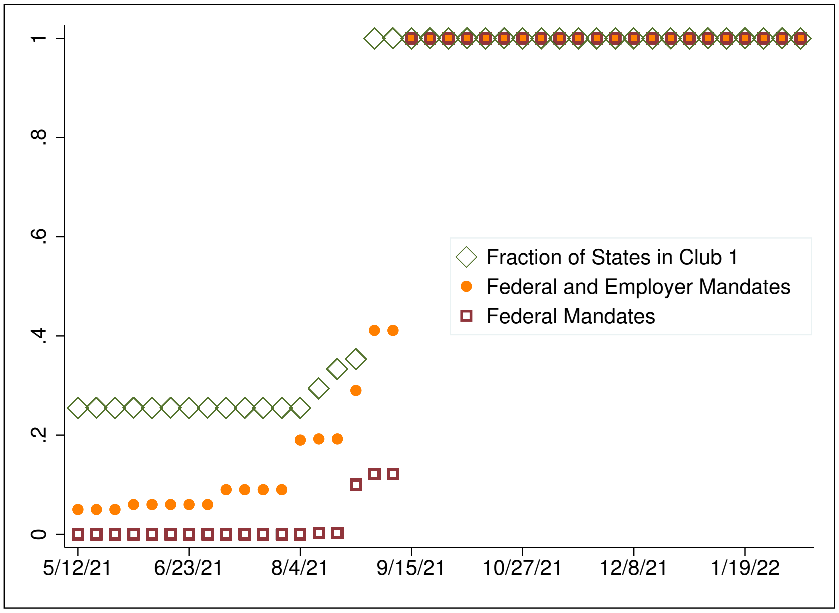
Figure 4: Dynamic State Membership in Club 1



Notes: The states shaded dark blue are in Club 1 and the light blue states are in Club 2. As the sample added weeks the membership of Club 1 continued to grow until September 1, 2021, at which point Club 1 included all states giving a single convergence club.

Figure 5 shows federal level vaccine mandate variables plotted alongside the evolving Club 1 membership. Two federal level vaccine mandate variables are displayed in the graph: one includes private employer vaccine mandates and the other does not. By design the fraction of states in Club 1 is held constant through August 4, at which point it began to rise steadily for several weeks merging into a single convergence club in September. In July 2021, 9% of employees nationwide reported working for an employer with a vaccination mandate according to a monthly Gallup poll conducted during this time period. One month later, in August 2021, that figure jumped by 10 percentage points to 19%. Additionally, on August 23, 2021, President Biden's press secretary announced that more stringent national vaccine mandates were coming. See Table 9 in the online appendix for further details on the construction of FEVM and its timing. The fraction of state membership of Club 1 closely tracks the course of these federal level mandates.

Figure 5: National-Level Mandates and Dynamic Club 1 Membership



Notes: The federal mandates variable is plotted in maroon, this variable combined with employer mandated vaccinations is plotted in orange, and the fraction of states belonging to Club 1 is plotted in green. The fraction of states in Club 1 and the federal vaccine mandate variable are flat until early August 2021. The federal mandates variable, the one capturing both the federal mandates and the employer mandates, slowly increased in early summer 2021 before increasing dramatically from July to August 2021, and again from August to September 2021. All three variables end up at unity by September 9, 2021.

The federal mandates variable, the state policy dataset, and the dynamics of club membership offer the opportunity to explore the impact of federal-level mandate announcements and state-level policies on club membership. An unconditional panel logit (random effects) regression in (27) was used to examine some of the effects of these time-varying policies. We used the combined federal and employer mandates as a common factor $\theta_{x,t}$, and various state specific variables as z_i including political trends, state demographic characteristics, such as population density, median household income, education, and the percentage of people employed by industry for each state. Various state-level vaccination policy variables comprise x_{it}, including state incentives for vaccinations and state vaccination mandates.

Table 2 column (5) shows the results of the preferred specification from the panel unconditional logit (random effects) regression.[23] The large coefficient on the federal level mandate variable shows that the probability of being in Club 1 given the federal mandate is extremely high. We also ran the model for column (5) from Table 2 with a two-week lag in each of the policy variables.

[23] Tables 9, 10, and 11 in the Online Supplement report results of other specifications used in the unconditional panel logit regressions.

While lagging each policy two weeks weakened the coefficient on the federal level mandate variable, it was still highly significant and remained the only policy variable with a correlation with club membership that was statistically different from zero. Thus, including the two week lag did not impact our overall findings, but we conclude that the association with vaccination rates and the federal mandates is stronger simultaneously than it is with a two week lag. See Table 14 in the online supplementary appendix for details. This aligns with Figure 5 where transition to Club 1 moves very closely with the implementation of the federal level mandates.

Interestingly, no state level vaccination policies or incentives had any significant impact on club membership. Also, contrary to our findings in the summer of 2021, by February 2022 our results suggest that once population density, median household income, the percentage of the population that is foreign born, and the industry composition of a state are all controlled for, political party is not associated with club membership in a statistically significant way. States with a larger percent of foreign born individuals were more likely to be in Club 1 initially. States with higher numbers of health care and social assistance workers as well as higher percentages of people employed in the retail trade industry were also initially more likely to be in Club 1, whereas states with higher percentages of employees working in wholesale trades were less likely to be in Club 1. The McFadden pseudo R^2 value is very high for our preferred specification, showing that the improvement in our model from the based intercept only specification is substantial, indicating the fitted model regression almost fully explains club membership.

In sum, as shown in Table 2, the only significant variables are FEVM and some state specific characteristics. These state specific variables (a_i) are associated with heterogeneous vaccine uptake initially across states, which led to divergent behaviors of y_{it}. As the FEVM were implemented, we saw an overall convergence of the y_{it}. As discussed in (30), we might deduce the implied treatment effect. Table 3 shows the national average of the log cumulative vaccine rates across convergent clubs. The maximum difference between the two clubs was 0.158, which occurred during the 14th week in our sample, the week of August 11, 2021. After this point, the difference in vaccination rates between the two clubs began to shrink as the two clubs merged into one convergent club. Hence, the implied ATT in the long run might be 0.158, but in the short run at the 40-th week it is 0.024 (=0.158-0.135).

For comparison with the above findings linear random effects regressions were run according to the following specifications using levels and differences as the dependent variable:

$$y_{it} = a + z_i'\gamma + \lambda\theta_{x,t} + x_{it}'\beta + \epsilon_{it}, \tag{31}$$

$$\Delta y_{it} = a + z_i'\gamma + \lambda\theta_{x,t} + x_{it}'\beta + \epsilon_{it}, \tag{32}$$

$$y_{it} = a_i + \lambda\theta_{x,t} + x_{it}'\beta + \epsilon_{it}, \tag{33}$$

$$\Delta y_{it} = a_i + \lambda\theta_{x,t} + x'_{it}\beta + \epsilon_{it}, \tag{34}$$

where y_{it} is the logarithm of state i's cumulative vaccination rate, Δy_{it} is the log of the number of new vaccinations per 10,000 people (the first difference of vaccinations), $\theta_{x,t}$ is the federal level mandates, z_i are state fixed effects, and x_{it} is a vector of state level policies and the number of new infections per 10,000 people in a state each week. The results are displayed in columns 1-4 of Table 2. The R^2 values show that the policies explain significantly more of the variation in cumulative vaccination rates than they do new vaccinations. The coefficients and levels of significance are very similar to the random effects model and fixed effects model for each of the dependent variables. For this reason we limit our discussion of the regression results here to the fixed effects model for both dependent variables.

When the log cumulative vaccination rate is the dependent variable, new infections and federal level vaccine mandates are both positive and highly significant. But, state level policies have no significant impact on cumulative vaccinations, with the exception of bans on proof of vaccination. The positive coefficient result suggests that if a state implemented a ban on proof of vaccination, cumulative vaccinations in that state would increase, which is a curious outcome that may be the spurious result of the nonlinear trend effects discussed earlier.

In the regressions with new (first differenced) vaccinations as the dependent variable there are several anomalous signs in the fitted coefficients. For instance, the signs on new infections, federal level vaccine mandates, and state level mandates on state employees are all negative and counter intuitive as higher infection rates and vaccination mandates are more likely to increase than reduce new vaccinations. Further, the empirical results imply that the only state level policy that increased new vaccinations per 10,000 people was a ban on mask mandates. It might be argued that banning mask mandates led people with high risk aversion to Covid-19 infection to get vaccinated because they felt less secure, but those people were already most likely to be vaccinated. As discussed earlier in Section 2.2, use of first differences does not eliminate time trend effects in the data and these counter intuitive results are again the likely outcome of misspecification and failure to capture separate group behavior in the data.

In sum, comparing the regressions results across Table 2, the panel logit regressions seem to provide the most plausible and intuitive findings. The natural explanation is that the panel logit models provide well specified formulations that take account of club membership arising from nonlinear trend effects and separate group convergence behavior that together determine cumulative vaccinations.

Table 2: Regression Results

	Linear Models				Logit Model
Dependent Variable:	y_{it}		Δy_{it}		$(2 - \hat{C}_{it})$
	(1)	(2)	(3)	(4)	(5)
	FE	RE	FE	RE	RE
New Infections per 10,000 People (x_{it})	0.005	0.005	-0.004	-0.004	
	(0.0001)	(0.0001)	(0.002)	(0.002)	
Federal Mandate and Employer Mandates ($\theta_{x,t}$)	0.251	0.252	-0.438	-0.442	101.0
	(0.012)	(0.012)	(0.053)	(0.053)	(22.39)
State Incentives (x_{it})					
Lottery	0.003	0.003	0.041	0.037	-1.468
	(0.012)	(0.013)	(0.047)	(0.046)	(2.038)
Cash	0.016	0.016	0.090	0.089	-0.084
	(0.011)	(0.011)	(0.049)	(0.050)	(3.405)
Community Outreach	0.028	0.029†	-0.043	-0.047	-2.440
	(0.015)	(0.014)	(0.051)	(0.045)	(3.372)
State Policies (x_{it})					
Vaccine Mandate State Employees	0.016	0.017	-0.199	-0.190	-1.919
	(0.012)	(0.011)	(0.056)	(0.055)	(2.258)
Indoor Vaccine Mandate	0.016	0.015	0.043	0.041	14.33
	(0.011)	(0.011)	(0.125)	(0.121)	(36.00)
Mask Mandate	0.007	0.007	-0.051	-0.051	2.096
	(0.012)	(0.012)	(0.094)	(0.090)	(3.567)
Ban on Proof of Vaccination	0.086	0.071	-0.221	-0.201	-3.164
	(0.020)	(0.016)	(0.129)	(0.099)	(5.166)
Mask Mandate Ban	0.013	0.006	0.193	0.129	-1.810
	(0.059)	(0.045)	(0.075)	(0.062)	(6.160)
Political (a_i)					
Percent of State House that is Republican		-0.550		-0.130	7.037
		(0.104)		(0.212)	(24.41)
Percent of Vote for Trump 2020		-0.042		-0.262	-13.66
		(0.073)		(0.104)	(10.53)
State Characteristics (a_i)					
Population Density		-0.010		-0.002	11.36
		(0.020)		(0.019)	(6.153)
Median Household Income		0.026		0.018	2.379
		(0.014)		(0.018)	(2.497)
Percent Foreign Born		-0.001		0.126	1.327
		(0.003)		(0.005)	(0.609)
Percent of People Employed by Industry (a_i)					
Health Care and Social Assistance		0.048		0.004	3.626
		(0.011)		(0.017)	(1.412)
Government and Government Enterprises		-0.006		0.002	-2.055
		(0.066)		(0.008)	(1.291)
Retail Trade		0.029		0.020	11.90
		(0.025)		(0.043)	(3.968)
Wholesale Trade		-0.015		-0.018	-13.07
		(0.036)		(0.058)	(5.901)
Transportation and Warehousing		-0.019		-0.017	-9.513
		(0.026)		(0.041)	(6.242)
n	51	51	51	51	51
T	40	40	40	40	40
R^2	0.832	0.808	0.664	0.632	
McFadden's R^2					0.936

Notes: Numbers in parentheses are standard errors. Median household income is measured in tens of thousands of dollars and population density is per 1,000 square miles. The dependent variable for the linear level model, y_{it}, is log(cumulative vaccination rate) and the dependent variable for the linear first differenced model, Δy_{it}, is log(new vaccinations per 10,000 people). The binary club membership obtained from the P-S club clustering technique, $\hat{C}_{it}$, is the dependent variable in the logit regressions. The coefficient on the federal level mandates in the unconditional logit model is large and all state level policies had no impact on the likelihood of being in the high vaccination club.

Table 3: Cross sectional averages for Club 1 and 2

No. of Weeks	$\hat{\mu}_{1t}$	$\hat{\mu}_{2t}$	Difference
1	3.629	3.527	0.102
14	4.016	3.859	0.158
17	4.057	3.904	0.153
40	4.247	4.111	0.135

5 Conclusion

When outcome variables have nonlinear and possibly stochastic trends, evaluating the effectiveness of policy changes by using TWFE regressions can be problematic. This paper shows the underlying reasons for this empirical problem and proposes an alternative approach. The key idea is a simple method to transform panel nonstationary outcome data into panel multinomial data by using a dynamic clustering method based on the relative convergence test of Phillips and Sul (2007a). This approach allows researchers to use panel logit regressions to investigate how policies that are implemented can impact convergence club membership or the long run behavior of the dependent variables over time.

The dynamic convergence clustering mechanism is applied to state level Covid-19 vaccination rates in an empirical example of this methodology. Our findings indicate that there were initially two distinct convergence clubs, but over time all states converged to a single club. We use panel logit to show how national and state policies were correlated with club membership over time, and that the timing of the convergence was very highly correlated with the timing of federal level vaccination mandates. Finally, we demonstrate how the regression results from panel logit regressions appear to give more realistic results than linear models.

There are two drawbacks of the proposed method. First, the number of time series observations cannot be too small. To estimate time varying convergent club membership the time series sample size T should be large enough to capture the evolution of club membership. In the Covid-19 vaccination empirical example $T = 40$ observations were available for the time series sample and the recursive sampling procedure was initiated with sample size $T = 13$. This choice complies with the minimum sample size $T = 10$ for the clustering algorithm that was used in Phillips and Sul (2007a). When T is smaller we suggest not using the dynamic CCM but instead static CCM and running a cross-sectional logit or multinomial logit regression.

Second, to use the proposed method, the panel should be balanced since the dynamic CCM tracks each individual club membership over time. If some data are missing within time periods, then interpolation and filtering to smooth out the series can be employed. Since the proposed method is designed for analyzing long run effects, small modifications of this type typically do not affect the membership findings. If data are missing at the beginning or end of the sample, then backward or forward forecasting would be required and the accuracy of such modifications is not studied in the paper.

References

Abadie, Alberto, 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.

Arkhangelsky, Dmitry and Guido Imbens, "Causal models for longitudinal and panel data: A survey," *The Econometrics Journal*, 2024, *27* (3), C1–C61.

Barber, Andrew and Jeremy West, "Conditional cash lotteries increase COVID-19 vaccination rates," *Journal of Health Economics*, 2022, *81*, 102578.

Bilinski, Alyssa and Laura A Hatfield, "Nothing to see here? Non-inferiority approaches to parallel trends and other model assumptions," *arXiv preprint arXiv:1805.03273*, 2018.

Biswas, Eva, Farzad Sabzikar, and Peter C B Phillips, "Boosting the HP filter for trending time series with long-range dependence," *Econometric Reviews*, 2024, *44* (1), 41–79.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess, "Revisiting event-study designs: robust and efficient estimation," *Review of Economic Studies*, 2024, p. rdae007.

Brehm, Margaret E., Paul A. Brehm, and Martin Saavedra, "The Ohio Vaccine Lottery and Starting Vaccination Rates," *American Journal of Health Economics*, 2022, *8* (3), 387–411.

Callaway, Brantly, "Difference-in-differences for policy evaluation," *Handbook of Labor, Human Resources and Population Economics*, 2023, pp. 1–61.

—— **and Pedro HC Sant'Anna**, "Difference-in-differences with multiple time periods," *Journal of econometrics*, 2021, *225* (2), 200–230.

Chaisemartin, Clément De and Xavier d'Haultfoeuille, "Two-way fixed effects estimators with heterogeneous treatment effects," *American Economic Review*, 2020, *110* (9), 2964–96.

Chaisemartin, Clement De and Xavier D'haultfœuille, "Two-way fixed effects and differences-in-differences estimators with several treatments," *Journal of Econometrics*, 2023, *236* (2), 105480.

Chang, Tom, Mireille Jacobson, Manisha Shah, Rajiv Pramanik, and Samir B Shah, "Financial Incentives and Other Nudges Do Not Increase COVID-19 Vaccinations among the Vaccine Hesitant," Working Paper 29403, National Bureau of Economic Research October 2021.

Dave, Dhaval, Andrew I. Friedson, Benjamin Hansen, and Joseph J. Sabia, "Association Between Statewide COVID-19 Lottery Announcements and Vaccinations," *JAMA Health Forum*, 10 2021, *2* (10), e213117–e213117.

Gardner, John, "Two-stage differences in differences," *arXiv preprint arXiv:2207.05943*, 2022.

Goodman-Bacon, Andrew, "Difference-in-differences with variation in treatment timing," *Journal of Econometrics*, 2021, *225* (2), 254–277.

Kahn-Lang, Ariella and Kevin Lang, "The promise and pitfalls of differences-in-differences: Reflections on 16 and pregnant and other applications," *Journal of Business & Economic Statistics*, 2020, *38* (3), 613–620.

Kong, Jianning, Peter C B Phillips, and Donggyu Sul, "Weak σ-convergence: Theory and applications," *Journal of Econometrics*, 2019, *209* (2), 185–207.

Lang, David, Lief Esbenshade, and Robb Willer, "Did Ohio's Vaccine Lottery Increase Vaccination Rates? A Pre-Registered, Synthetic Control Study," *Journal of Experimental Political Science*, 2022, p. 1–19.

Li, Shanjun, Joshua Linn, and Elisheba Spiller, "Evaluating "Cash-for-Clunkers": Program effects on auto sales and the environment," *Journal of Environmental Economics and management*, 2013, *65* (2), 175–193.

Liu, Licheng, Ye Wang, and Yiqing Xu, "A practical guide to counterfactual estimators for causal inference with time-series cross-sectional data," *American Journal of Political Science*, 2024, *68* (1), 160–176.

Mei, Ziwei, Peter C B Phillips, and Zhentao Shi, "The boosted HP filter is more general than you might think," *Journal of Applied Econometrics*, 2024, *39* (7), 1260–1281.

Phillips, Peter C B and Donggyu Sul, "Transition modeling and econometric convergence tests," *Econometrica*, 2007a, *75* (6), 1771–1855.

____ **and** ____ , "Some empirics on economic growth under heterogeneous technology," *Journal of Macroeconomics*, 2007b, *29* (3), 455–469.

____ **and** ____ , "Economic transition and growth," *Journal of applied econometrics*, 2009, *24* (7), 1153–1185.

——— **and Zhentao Shi**, "Boosting: Why you can use the HP filter," *International Economic Review*, 2021, *62* (2), 521–570.

Rambachan, Ashesh and Jonathan Roth, "A more credible approach to parallel trends," *Review of Economic Studies*, 2023, p. rdad018.

Roth, Jonathan, "Pretest with caution: Event-study estimates after testing for parallel trends," *American Economic Review: Insights*, 2022, *4* (3), 305–322.

———**, Pedro HC Sant'Anna, Alyssa Bilinski, and John Poe**, "What's trending in difference-in-differences? A synthesis of the recent econometrics literature," *Journal of Econometrics*, 2023.

Sehgal, Neil K.R., "Impact of Vax-a-Million Lottery on COVID-19 Vaccination Rates in Ohio," *The American Journal of Medicine*, 2021, *134* (11), 1424–1426.

Sul, Donggyu, *Panel data econometrics: Common factor analysis for empirical researchers*, Routledge, 2019.

Thirumurthy, Harsha, Katherine L. Milkman, Kevin G. Volpp, Alison M. Buttenheim, and Devin G. Pope, "Association between statewide financial incentive programs and COVID-19 vaccination rates," *PLOS ONE*, 2022, *17* (3), e0263425.

Walkey, Allan J., Anica Law, and Nicholas A. Bosch, "Lottery-Based Incentive in Ohio and COVID-19 Vaccination Rates," *JAMA*, 08 2021, *326* (8), 766–767.

Wong, Charlene A., William Pilkington, Irene A. Doherty, Ziliang Zhu, Hattie Gawande, Deepak Kumar, and Noel T. Brewer, "Guaranteed Financial Incentives for COVID-19 Vaccination: A Pilot Program in North Carolina," *JAMA Internal Medicine*, 01 2022, *182* (1), 78–80.