Implementing the Panel Event Study

Autores:

Damian Clarke
Kathya Tapia Schythe

Implementing the Panel Event Study

Damian Clarke
Department of Economics
University of Chile
dclarke@fen.uchile.cl

Kathya Tapia Schythe
Department of Economics
University of Santiago de Chile
kathya.tapia@usach.cl

Abstract. Many studies estimate the impact of exposure to some quasi-experimental policy or event using a panel event study design. These models, as a generalized extension of 'difference-in-differences' or two-way fixed effect models, allow for dynamic lags and leads to the event of interest to be estimated, while also controlling for fixed factors (often) by area and time. In this paper we discuss the set-up of the panel event study design in a range of situations, and lay out a number of practical considerations for its estimation. We describe a Stata command `eventdd` that allows for simple estimation, inference, and visualization of event study models in a range of circumstances. We then provide a number of examples to illustrate `eventdd`'s use and flexibility, as well as its interaction with various native Stata routines, and other relevant user-written libraries such as `reghdfe` and `boottest`.

Keywords: , event studies, difference-in-differences, estimation, inference, visualization

1 Introduction

Recent developments in quasi-experimental methods have brought increasing attention to panel event study models. Using data covering a panel of observations (such as states) over time, the design seeks to estimate the impact of some event which occurs, or 'switches on' in certain units and certain time periods.[1] They seek to use as counterfactuals the areas in which the policy or event does not occur, or has not yet occurred. By considering the variation in outcomes around the adoption of the event compared with a baseline reference period, both event lags and leads are estimated, allowing for a clear visual representation of the event's causal impact provided that key identifying assumptions are met.

These methods have been borne out of older difference-in-differences, or two-way fixed effect models. These models often seek to examine the impact of natural experiments, where events are assigned to certain units due to some process beyond the control of the analyst but owing to environmental or political factors (among others), and as such, generally do not assume that assignment is random. Indeed, as we lay

---

[1]Throughout this paper, for expositional purposes we will refer to a geographical and a temporal dimension to these panel event study designs. However, we note that such models are suitable to other settings where one wishes to estimate the impact of an event which is assigned to certain units, and where there is some temporal variation in outcomes to be studied. For example, rather than being based on different geographical units at different times, these models could be based on an event occurring in different age groups at different times. All results discussed in this paper extend to settings such as these.

out at more length in the following section, the key assumption underlying consistent estimation in event study models is that the occurrence of the event in a particular area is not systematically related to the changes in levels that will occur in the future.

These models are widely used in empirical analyses in a range of contexts, having been applied to (among many others themes) automotive plant closures and opioid overdoses (Venkataramani et al. 2020), family planning access and childhood economic circumstance (Bailey et al. 2018), health care reform and ambulatory care usage (Dimitrovová et al. 2020) and university reform and intergenerational mobility (Suhonen and Karhunen 2019). These cases suggest usage across a range of fields including social sciences, medicine and public health, and additional reviews of their frequency of use in a number of economic journals are provided in Abraham and Sun (2018); Roth (2019). A burgeoning literature has laid out a number of identification requirements in this setting (Freyaldenhoven et al. 2019; Borusyak and Jaravel 2018; Abraham and Sun 2018; Athey and Imbens 2018; Schmidheiny and Siegloch 2019). These methods can be used, with some restrictions, both in cases where events occur at the same time period in each unit, and in cases where the adoption of events is staggered. Indeed, Athey and Imbens (2018) refer to these as "Staggered Adoption Designs", although here we follow the more common nomenclature of panel event studies.[2] Additionally, these methods are related to a much broader literature on staggered adoption of policies and the estimation of a single-coefficient model (de Chaisemartin and D'Haultfoeuille 2019; Callaway and Sant'Anna 2018; Goodman-Bacon 2018). While we briefly discuss these models in the methods section, our principal interest is on full panel-event study specifications which come with their own considerations.

In this paper we provide a discussion of these panel-event study models, and practical issues related to their estimation and to inference in these settings. We also present the `eventdd` command, a Stata command allowing for estimation and inference in event studies, as well as the graphical presentation of their results.[3] This command can flexibly interact with both in-built Stata commands such as `regress` and `xtreg`, as well as the user-written regression command `reghdfe` which is highly convenient in two-way fixed effect models such as those described in this paper (Correia 2016). We provide discussion of both estimation and inference in event study models. As well as standard inference procedures such as robust and cluster-robust inference, the `eventdd` command allows for wild-bootstrap based inference respecting the clustered nature of the occurrence of events, and specifically the user-written `boottest` command (Roodman et al. 2019). After reviewing the theory behind panel event study models in section 2, we discuss the command syntax in section 3, before documenting the command's usage, applied to a particular empirical example, in Section 4.

---

[2]An older literature, often related to the analyses of financial events and information shocks, refers to event studies. These event studies in finance are generally based on time-series observations, and have quite different properties to the panel event studies used in policy analysis that we discuss in this paper. A useful discussion of these finance-style event studies, and their application in Stata, is provided in Pacicco et al. (2018).

[3]An earlier version of this command is available as Clarke and Tapia Schythe (2020).

2 Methods

2.1 Estimation

Consider a panel covering states s and time periods t. We are interested in estimating the impact of the passage of an event which may occur at different times in different states. We will denote as $Event_s$ a variable recording the time period t in which the event is adopted in state s. Denoting the outcome of interest as y_{st}, the panel event study specification can be written as[4]:

$$y_{st} = \alpha + \sum_{j=2}^{J} \beta_j (\text{Lag } j)_{st} + \sum_{k=1}^{K} \gamma_k (\text{Lead } k)_{st} + \mu_s + \lambda_t + X'_{st}\Gamma + \varepsilon_{st}. \qquad (1)$$

Here μ_s and λ_t are state and time fixed effects, X_{st} are (optionally) time-varying controls, and ε_{st} is an unobserved error term. In equation 1, lags and leads to the event of interest are defined as follows:

$$\begin{align}
(\text{Lag } J)_{st} &= \mathbb{1}[t \le Event_s - J], & (2)\\
(\text{Lag } j)_{st} &= \mathbb{1}[t = Event_s - j] \text{ for } j \in \{1, \dots, J-1\}, & (3)\\
(\text{Lead } k)_{st} &= \mathbb{1}[t = Event_s + k] \text{ for } k \in \{1, \dots, K-1\}, & (4)\\
(\text{Lead } K)_{st} &= \mathbb{1}[t \ge Event_s + K]. & (5)
\end{align}$$

Lags and leads are thus binary variables indicating that the given state was a given number of periods away from the event of interest in the respective time period. J and K lags and leads are included respectively, and, as indicated in equations 2 and 5, final lags and leads "accumulate" lags or leads beyond J and K periods. A single lag or lead variable is omitted to capture the baseline difference between areas where the event does and does not occur. In specification 1, as standard, this baseline omitted case is the first lag, where $j = 1$.

---

[4]There are a number of ways to specify such a model. Slightly different notations are used by Schmidheiny and Siegloch (2019) who define the model as:

$$y_{st} = \sum_{j=\underline{j}}^{\overline{j}} \beta_j b_{st}^j + \mu_s + \lambda_t + \varepsilon_{st},$$

where

$$b_{st}^j = \begin{cases} \mathbb{1}[t \le Event_s + j] & \text{if } j = \underline{j} \\ \mathbb{1}[t = Event_s + j] & \text{if } \underline{j} < j < \overline{j} \\ \mathbb{1}[t \ge Event_s + j] & \text{if } j = \overline{j}, \end{cases}$$

and where $\underline{j}$ is equivalent to our definition of J and $\overline{j}$ is equivalent to our L. In the case of Freyaldenhoven et al. (2019), they define a version of this model as:

$$y_{st} = \delta_{-K+}(1 - z_{s,t+(K-1)}) + \delta_{L+} z_{s,t-L} + \sum_{k=-(L-1)}^{K-1} \delta_{-k} \Delta z_{s,t+k} + \mu_s + \lambda_t + \varepsilon_{st},$$

where $z_{st} \equiv PostEvent_{st}$ as defined in Table 1, $z_{s,t+k}$ and $z_{s,t-k}$ refer to lags and leads of this variable respectively, and Δ refer to the first difference of these lag/lead terms. These models, and that laid out in equations 1-5 are equivalent.

State (s)	Year (t)	Event	Post Event	Time to Event	Lag 4	Lag 3	⋯	Lead 0	Lead 1	⋯	Lead 4
State A	2000	2004	0	-4	1	0	⋯	0	0	⋯	0
State A	2001	2004	0	-3	0	1	⋯	0	0	⋯	0
State A	2002	2004	0	-2	0	0	⋯	0	0	⋯	0
State A	2003	2004	0	-1	0	0	⋯	0	0	⋯	0
State A	2004	2004	1	0	0	0	⋯	1	0	⋯	0
State A	2005	2004	1	1	0	0	⋯	0	1	⋯	0
State A	2006	2004	1	2	0	0	⋯	0	0	⋯	0
State A	2007	2004	1	3	0	0	⋯	0	0	⋯	0
State A	2008	2004	1	4	0	0	⋯	0	0	⋯	1
State A	2009	2004	1	5	0	0	⋯	0	0	⋯	1
State B	2000	2005	0	-5	1	0	⋯	0	0	⋯	0
State B	2001	2005	0	-4	1	0	⋯	0	0	⋯	0
State B	2002	2005	0	-3	0	1	⋯	0	0	⋯	0
State B	2003	2005	0	-2	0	0	⋯	0	0	⋯	0
State B	2004	2005	0	-1	0	0	⋯	0	0	⋯	0
State B	2005	2005	1	0	0	0	⋯	1	0	⋯	0
State B	2006	2005	1	1	0	0	⋯	0	1	⋯	0
State B	2007	2005	1	2	0	0	⋯	0	0	⋯	0
State B	2008	2005	1	3	0	0	⋯	0	0	⋯	0
State B	2009	2005	1	4	0	0	⋯	0	0	⋯	1
State C	2000	.	0	.	0	0	⋯	0	0	⋯	0
State C	2001	.	0	.	0	0	⋯	0	0	⋯	0
State C	2002	.	0	.	0	0	⋯	0	0	⋯	0
State C	2003	.	0	.	0	0	⋯	0	0	⋯	0
State C	2004	.	0	.	0	0	⋯	0	0	⋯	0
State C	2005	.	0	.	0	0	⋯	0	0	⋯	0
State C	2006	.	0	.	0	0	⋯	0	0	⋯	0
State C	2007	.	0	.	0	0	⋯	0	0	⋯	0
State C	2008	.	0	.	0	0	⋯	0	0	⋯	0
State C	2009	.	0	.	0	0	⋯	0	0	⋯	0
State D	2000	2007	0	-7	1	0	⋯	0	0	⋯	0
State D	2001	2007	0	-6	1	0	⋯	0	0	⋯	0
State D	2002	2007	0	-5	1	0	⋯	0	0	⋯	0
State D	2003	2007	0	-4	1	0	⋯	0	0	⋯	0
State D	2004	2007	0	-3	0	1	⋯	0	0	⋯	0
State D	2005	2007	0	-2	0	0	⋯	0	0	⋯	0
State D	2006	2007	0	-1	0	0	⋯	0	0	⋯	0
State D	2007	2007	1	0	0	0	⋯	1	0	⋯	0
State D	2008	2007	1	1	0	0	⋯	0	1	⋯	0
State D	2009	2007	1	2	0	0	⋯	0	0	⋯	0

Table 1: A Stylized Example

A stylized example of such a setting is provided in Table 1. We consider four states forming a balanced panel of years from 2000-2009. The $Event_s$ variable occurs at different times in different states, and in the case of one state, does not occur. Here both four lags and four leads are included, such that $J = K = 4$. Lag and Lead 4 (exclusively) are switched on for periods in which the "Time to Event" exceeds 4 lags or leads respectively.

States in which the event never occurs (such as State C in Table 1) act as pure controls. These units have 0s in all lag and lead terms, and act as the counterfactual on which the estimation of impacts is based. Differences between these pure controls states and states which adopt the event of interest are anchored at 0 in the omitted base period, ie the first lag in equation 1. Hence, lags and leads capture the difference between treated and control states, compared to the prevailing difference in the omitted base period. Unbiased estimation of post-event treatment effects thus relies fundamentally on the so called "parallel trends assumption". In the absence of treatment, it is assumed that treated and control states would have maintained similar differences as in the baseline period. For this reason, these models have been demonstrated to be under-identified, or identified only up to a linear trend, when all units adopt treatment at some point in time (Schmidheiny and Siegloch 2019; Borusyak and Jaravel 2018). Schmidheiny and Siegloch (2019) show that in this case, it is necessary to bin lags and leads beyond certain maximum lag (J) and lead (K) periods.

The panel event study is an extension of the standard two-way fixed effect (sometimes called difference-in-differences) model, where a single "Post Event" indicator is included for all periods posterior to the occurrence of the event in treated states. This is simply:

$$y_{st} = \alpha + \beta \text{PostEvent}_{st} + \mu_s + \lambda_t + X'_{st}\Gamma + \varepsilon_{st}, \tag{6}$$

where following the notation from (2)-(5), $\text{PostEvent}_{st} = \mathbb{1}[t \geq Event_s]$. Estimation of event specification 1 provides two key pieces of information not observable in this single-coefficient model. Firstly, the full set of event leads allows for the inspection of parallel trends in the *pre*-treatment period. While this does not provide evidence that the units in which the event was adopted and not adopted would have necessarily followed similar trends in the post-reform period (Kahn-Lang and Lang 2019) (which is the identifying assumption of these models), if trends in treated and untreated areas were not parallel even pre-event, it is unlikely that they would be parallel post-event. Secondly, the policy lags allow for inspection of the temporal nature of treatment effects, noting for example, any dynamics in the appearance of effects, for example growing or shrinking over time, and whether effects are transitory or permanent.

A developing literature including papers by de Chaisemartin and D'Haultfoeuille (2019); Callaway and Sant'Anna (2018); Goodman-Bacon (2018) point to challenges in interpreting the estimated $\widehat{\beta}$ from two-way fixed effects models when treatment effects are heterogeneous (across either groups or time periods). Goodman-Bacon (2018) for example, demonstrates that treatment effects which are heterogeneous in time since treatment in contexts where treatments are adopted in different time periods in different groups, can result in estimates which are biased away from a weighed average of the

average treatment effect on the treated, a problem which is resolved in the panel event study design. However, results from Abraham and Sun (2018) suggest that specific types of heterogeneity concerns remain even in panel event study models examined here. In particular, they note undesired weighting of treatment effects if there is heterogeneity across treatment groups in particular lag and lead terms. Other concerns exist in event study designs, such as possible inferential problems related to selective survival of models based on pre-trend tests (Roth 2019). The `eventdd` command will not account for corrections raised in these particular settings, as these are inherent to empirical estimation of panel event study designs. We do however note that there are a number of alternative estimators which are complementary to panel event study designs, and which should be considered as part of a complete estimation and testing procedure, such as the stacked DD procedure of Abraham and Sun (2018), sensitivity tests described in Roth (2019); Rambachan and Roth (2019), and alternative models to account for dynamic paths of treatment effects, such as those described in de Chaisemartin and D'Haultfoeuille (2019); Callaway and Sant'Anna (2018). As many of these have existing estimation libraries in some languages, when discussing the command syntax of `eventdd` in section 3 and examples of use in section 4 we discuss ways which `eventdd` and its returned objects have been designed to facilitate interaction with these other commands.

2.2   Inference

A standard inference concern where policies are assigned by some unit such as state, and outcomes are followed over time within these units, is related to potential serial-correlation in the outcome variable over time (Bertrand et al. 2004). While the derivations from (Bertrand et al. 2004) are based on single-coefficient models of the form of equation 6, the crux of the concern relates to high serial correlation in the outcome variable of interest, and relatively little change in the independent variables of interest. This setting is replicated in event study models described in equations 1-5. It is thus fundamental to account for this within-cluster correlation when conducting inference in such models.

The standard solution is to allow for within-cluster auto-correlation by using a cluster-robust variance-covariance estimator (CRVE) to estimate standard errors and confidence intervals on regression parameters. Such an estimator is provided as standard in Stata by specifying the `vce(cluster` *clustvar*`)` option in e(class) models.[5] However, as has been extensively documented, standard CRVEs are only asymptotically valid, where the asymptotic behavior depends on the number of clusters $S \to \infty$ (see for

---

[5]Denoting the matrix of k independent variables as $\boldsymbol{X}$, the dependent variable vector as $\boldsymbol{y}$ and the vector of regression coefficients as $\boldsymbol{\beta}$, this CRVE is estimated in Stata for a linear regression model with N observations in S clusters as:

$$\widehat{V}_{CR}(\widehat{\boldsymbol{\beta}}) = (\boldsymbol{X}'\boldsymbol{X})^{-1} \left(\sum_{s=1}^{S} \boldsymbol{X}_s \tilde{\boldsymbol{u}}_s \tilde{\boldsymbol{u}}_s' \boldsymbol{X}_s' \right) (\boldsymbol{X}'\boldsymbol{X})^{-1}.$$

Here $\tilde{\boldsymbol{u}}_s = \sqrt{c}\hat{\boldsymbol{u}}_s$, where c is a small sample correction $c = (S/(S-1)) \times ((N-1)/(N-k))$ and $\hat{\boldsymbol{u}} = \boldsymbol{y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}$ are standard regression residuals (Cameron et al. 2008).

example the comprehensive review in Cameron and Miller (2015)). When standard clustering is used based on 'too few' clusters, the CRVE is generally downward-biased, resulting in over-rejection of null hypotheses. This bias can be severe (Cameron and Miller 2015; Mackinnon and Webb 2018).

In practice, knowing how many clusters it 'too few' depends on a number of factors. While there are rules of thumb such as the rule of 42 laid out in Angrist and Pischke (2008) which suggests that standard clustering provides a good approximation if $S \geq 42$ clusters, the performance of these methods under simulation has been shown to depend also on the relative size of clusters (Mackinnon and Webb 2017). A range of results surveyed in Cameron and Miller (2015) leads to their suggestion that if one is analyzing data with fewer than 50 clusters in a state-year panel (such as the case with panel-event studies), alternative inference methods should be considered.

In this case where the quasi-experimental set-up is based on fewer than around 50 clusters, the wild cluster bootstrap has been documented to be a successful resampling-based method to take account of auto-correlation in variables underlying panel event studies, even in cases with fewer clusters (see *eg* Cameron et al. (2008); Cameron and Miller (2015); Roodman et al. (2019)). This has been efficiently implemented in Stata as described in Roodman et al. (2019), and programmed for Stata as `boottest` (Roodman 2015). Finally, note that in the case of very few clusters, and in particular few clusters where an event occurs, inference is further complicated. In cases such as this a number of potential solutions have been proposed, such as those described in Mackinnon and Webb (2018); Conley and Taber (2011). As we lay out in the following sections, the `eventdd` command allows simple access to various inference options depending on the context of interest, including standard clustering, bootstrap, and wild cluster bootstrap in various guises based on both Stata's native CRVE procedures, as well as the user-written `boottest` command.

3   The eventdd command

Panel event studies can be implemented in Stata using the following command syntax:

`eventdd` *varlist* $\big[$ *if* $\big]$ $\big[$ *in* $\big]$ $\big[$ *weight* $\big]$, `timevar`(*timevar*) `ci`(*string*) $\big[$ *method*
    `baseline`(#) `level`(#) `accum` `lags`(#) `leads`(#) `noend` `noline`
    `keepbal`(*varname*) `absorb`(*varname*) `wboot` `wboot_op`(*string*) `balanced`
    `inrange` `graph_op`(*string*) `ci_op`(*string*) `coef_op`(*string*)
    `endpoints_op`(*string*) * $\big]$

The required *varlist* should first specify the dependent variable of interest, and then all controls including fixed effects to be included in the panel event study model (equation 1), but *not* including lags and leads, that should be entered in the regression. The *method* option specifies the estimation procedure for the underlying model, and can be `ols` (ordinary least squares), `fe` (fixed effects) or `hdfe` (absorbing multiple levels of fixed

effects with the user-written `reghdfe` command). If no estimation method is specified, `ols` is used by default. In the case of fixed effects (`fe`) or high-dimension fixed effect (`hdfe`) models, fixed effects can be absorbed (as discussed in the options below), and as such need not be entered in the standard *varlist* syntax. In the case of `fe` specifications, data must first be `xtset` in Stata. Based on this syntax, `eventdd` takes care of the generation of all lag and lead terms, estimation and inference, and the production of an event study plot. Examples are provided in section 4 of this document.

Options

`timevar` is a required option. The time variable specified should contain a standardized value, where 0 corresponds to the time period in which the event of interest occurs for a given unit, -1 refers to one year prior to the event, 1 refers to one year following the event, and so forth. For any units in which the event does not occur (pure controls), this variable should contain missing values.

`ci`(*string*) is a required option indicating the type of graph the command will produce. The types of graphs available for the confidence intervals are *rarea* for an interval with area shading (see `twoway_rarea`), *rcap* for an interval with capped spikes (see `twoway_rcap`) and *rline* for an interval with lines (see `twoway_rline`). Only one type can be specified and all intervals will be the same type. The appearance can be modified with `ci_op()`.

`baseline`(*#*) Specifies the reference period for the event study, which is a baseline omitted category to which all other periods should be compared on the event study output. By default this value is set at -1 as in equation 1.

`level`(*#*) Specifies the confidence level, as a percentage, for confidence intervals. The default is `level(95)` or as set by `set level`. This sets the levels for confidence intervals in regression output, as well as the event study plot and matrices. This will also be passed to `boottest` if wild clustered standard errors are requested.

`accum` Specifies that all periods beyond some specified values should be accumulated into final points, indicated as J and K in equation 1. For example if `accum` is specified and `lags(#)` and `leads(#)` are both set equal to 10, a single coefficient will be displayed in regressions and graphical output capturing 10 or more periods prior/post reform. If `accum` is not specified, all possible lags and leads will be included in models and graphical output.

`lags`(*#*) Indicates the maximum amount of pre-event periods to consider in the event study. This can only be specified if either `accum`, `keepbal` or `inrange` are also specified. Only integer values are permitted.

`leads`(*#*) Indicates the maximum amount of post-event periods to consider in the event study. This can only be specified if either `accum`, `keepbal` or `inrange` are also specified. Only integer values are permitted.

`noend` Requests that accumulative end points are not shown on graphical output when

the accum option is specified.

keepbal(*varname*) Specifies that only units which are balanced in the panel should be kept for estimation. Here varname indicates the panel variable (eg State) which indicates units. In this case "balance" refers to balance over calendar time. An alternative option (`balanced`), discussed below, allows for only balanced lags and leads relative to treatment to be considered in graphical output.

ols Requests that the event study model underlying graphical output should be estimated by OLS using Stata's `regress` command. In this case, unit-specific fixed effects and time-specific fixed effects must be included in the *varlist* indicated in the command syntax. This is the default estimation method.

fe Requests that the event study model underlying graphical output should be estimated by fixed-effects (within) estimation, using Stata's `xtreg, fe` command. In this case the data must be `xtset` prior to use, and unit-specific fixed effects should not be included in the varlist indicated in the command syntax. Time-specific fixed effects still need to be included in the *varlist* indicated in the command syntax.

hdfe Requests that the event study model underlying graphical output should be estimated using the user-written `reghdfe` command of Correia (2016) (if installed). If this option is specified, the `absorb`(*varlist*) option should also be specified to indicate which fixed effects should be controlled in the regression. Any fixed effects indicated in `absorb`(*varlist*) should not be included in the *varlist* indicated in the command syntax. This option cannot be used in combination with the `wboot` option.

absorb(*varlist*) This option is only required when specifying the `hdfe` estimation option. The *varlist* identifies fixed effects to be absorbed (such as unit fixed effects). Refer to `reghdfe` (if installed) for additional details.

wboot Indicates that inference in the event study plot produced by the command should be based on wild cluster bootstrapped standard errors. When indicated, confidence intervals for each lag and lead term will the be calculated using a wild cluster bootstrap. This requires the user-written `boottest` command of Roodman (2015) (if installed). This option may not be combined with the `hdfe` estimation option.

wboot_op(*string*) Allows for the inclusion of any other wild bootstrap option permitted in `boottest`, including `seed(#)` to set the seed for simulation based calculations and replicate the confidence intervals, and `bootclust(varname)` to specify which variable(s) to cluster the wild bootstrap upon, among others. When setting the level (which is 95 by default), this should be indicated in the level option of the command, and this will be passed to `wboot_op()`.

balanced Requests that only "balanced" lags and leads are plotted. This will produce a graph only showing lags and leads for which each treated unit has data, and as such, all coefficients plotted will be based on all units in the data. While only balanced lags and leads will be plotted, all units and time periods will be included in the estimation of the event study.

inrange Requests that only the specified lags and leads are plotted. While only lags and

leads indicated in `lags(#)` and `leads(#)` will be plotted, all units and time periods will be included in the estimation of the event study.

`noline` Requests that the line before the event on the x-axis is not shown on graphical output.

`graph_op(`*string*`)` allows for the inclusion of any other graphing option permitted in `twoway_options`, including `title_options`, `added_lines_options`, `axis_label_options`, among others. This also allows for the use of alternative labels for graph axes. If not specified, a standard graphical output will be provided.

`ci_op(`*string*`)` Allows for the inclusion of any graphing option for the confidence intervals permitted in `twoway_rarea`, `twoway_rcap` or `twoway_rline` depending on the type of CI indicated in `ci()`; including `area_options`, `line_options` and `connect_options`, respectively. This does not allow the use of the general options of `graph_op()`. If not specified, a standard graphical output will be provided.

`coef_op(`*string*`)` Allows for the inclusion of any graphing option for the coefficients permitted in the scatter plot including `marker_options` and `marker_label_options` among others. This does not allow the use of the general options of `graph_op()`. If not specified, a standard graphical output will be provided.

`endpoints_op(`*string*`)` Allows for the inclusion of any graphing option for the end points coefficients permitted in `scatter` including `marker_options`, `marker_label_options`, among others. This is only available if specifying the `accum` and does not allow the use of the general options of `graph_op()`. If not specified, a standard graphical output will be provided.

∗ Any other `estimation options` permitted by `regress`, `xtreg`, or `reghdfe` can be included, and will be passed to the specified estimation command. This allows for the inclusion of clustered standard errors or other variance estimators such as bootstrap or jackknife estimators (see `vce_option`), as well as any other option relevant to the underlying estimation or inference in the panel event study regression model.

Returned Objects

`eventdd` stores the following in `e()`:

Macros:
`e(cmd)`	eventdd
`e(cmdline)`	command as typed
`e(depvar)`	name of dependent variable
`e(wtype)`	weight type
`e(clustvar)`	name of cluster variable
`e(vce)`	vcetype specified in `vce()`
`e(vcetype)`	title used to label Std. Err.

Matrices:

`e(b)`	coefficient vector
`e(V)`	variance-covariance matrix of the estimators
`e(leads)`	all event leads, their lower bound, the point estimate, and their upper bound
`e(lags)`	all event lags, their lower bound, the point estimate, and their upper bound
`e(V_lags_leads)`	variance-covariance matrix of the lags and leads estimators

Note that methods related to event study models such as that described by Rambachan and Roth (2019) rely on access to point estimates and standard errors of lag and lead terms, which are available through the matrices returned here.

4 Examples based on an Empirical Application

We now provide a number of illustrations of the performance of `eventdd` to estimate the panel event study in empirical applications. We use data from Stevenson and Wolfers (2006) of the no-fault divorce reforms and female suicide in United States. These data have been used in other papers to demonstrate the functionality of recent advances in two-way fixed effect models (see Goodman-Bacon (2018)) and are drawn from examples used in documenting such methods when used in Stata (Goodman-Bacon et al. 2019).[6] The data consists of a balanced panel with 49 states observed from 1964 to 1996 with different timing of unilateral divorce reforms across the states.

The specification of the baseline two-way fixed effect "difference-in-differences" (DD)-style model of female suicide on no-fault divorce reforms used is:

$$asmrs_{st} = \gamma_s + \lambda_t + \tau post_{st} + X'_{st}\Gamma + \varepsilon_{st}. \tag{7}$$

This is the analogue of equation 6 applied to this case in particular. Here, *asmrs* refers to the female suicide rate for all women in state *s* at time *t*, γ is a fixed effect by state, λ is a temporal (year) fixed effect, *post* takes the value of 1 after the implementation of a no-fault divorce reform and ε is a stochastic error. The controls (X_{st}) include Per-Capita Income (*pcinc*), Homicide Mortality (*asmrh*) and the Aid to Families with Dependent Children (AFDC) rate for a family of four (*cases*). Here τ is the parameter which captures the average impact of unilateral divorce on suicide rate assuming a standard DD parallel-trends assumption.[7]

4.1 Estimation of the Panel Event Study

In order to estimate a panel event study specification corresponding to the no fault divorce reform, the first step is to create the standardized version of the time to reform

---

[6]This data is available online at the following URL: http://www.damianclarke.net/stata/bacon_example.dta, and can be imported directly into Stata typing `webuse set www.damianclarke.net/stata/` and then `webuse bacon_example.dta`.

[7]In reality, as laid out in section 2, in the case that this effect is heterogeneous over time, this will not capture the ATE, but rather a weighted average of pair-wise comparisons (Goodman-Bacon 2018; de Chaisemartin and D'Haultfoeuille 2019; Callaway and Sant'Anna 2018).

variable, presuming such a variable is not already available in the data. In this case in particular, the creation of the variable in Stata simply requires subtracting the reform period, called $Event_s$ in section 2 (and `_nfd`, for 'no fault divorce', in the data) from the time period t, (called `year` in the data):

```
. webuse set www.damianclarke.net/stata/
(prefix now "http://www.damianclarke.net/stata")
. webuse bacon_example.dta, clear
(Stevenson and Wolfers (2006) divorce example, provided by Goldring et al.)
. gen timeToTreat = year - _nfd
(429 missing values generated)
```

Note that as expected missing values are generated for states in which the reform is not adopted at any point in this period, and which act as pure controls in the panel event study.

The second step is to estimate the event study, as per equations 1-5. In this example, the general form of event study model including all lags and leads available is:

$$
\begin{aligned}
asmrs_{st} &= \alpha + \beta_{21}(\text{Lag } 21)_{st} + \ldots + \beta_2(\text{Lag } 2)_{st} \\
&\quad + \gamma_0(\text{Lead } 0)_{st} + \ldots + \gamma_{27}(\text{Lead } 27)_{st} \\
&\quad + X'_{st}\Gamma + \mu_s + \lambda_t + \varepsilon_{st}
\end{aligned} \tag{8}
$$

where as above *asmrs* is the female suicide rate for all women, and a series of $J = 21$ Lags and $K = 27$ Leads are considered relative to the event of interest (fully saturating the model). As is generally standard, the reference period is set as -1: the period immediately preceding the adoption of the event in each state. Fixed effects for state and time are included as μ and λ, respectively.

The `eventdd` command provides a simple syntax to generate all necessary lags and leads for equation 8, estimate the event study model, and plot point estimates and confidence intervals. The command requires two options: the standardized variable in `timevar`(*timeToTreat*) and also the type of confidence interval graph, for instance, we specify `ci`(*rcap*) for intervals indicated by capped error bars. Below we request that the command runs quietly (`qui`), however later in this section we document an example where full regression output is displayed:

```
. #delimit ;
delimiter now ;
. qui
> eventdd asmrs pcinc asmrh cases i.year i.stfips, timevar(timeToTreat) ci(rcap)
> cluster(stfips) graph_op(ytitle("Suicides per 1m Women") xlabel(-20(5)25));
. #delimit cr
delimiter now cr
```

The command stores all event lags, their lower bound, the point estimate, and their upper bound. For example, if we wish to visualize the estimates on the full set of lags, as well as their upper and lower confidence intervals, we can simply examine the returned `lags` matrix:

```
. mat list e(lags)
e(lags)[21,4]
              Lag         LB         Est          UB
r21            1           0           0           0
r20            2   -5.6533713  -.51573968    4.621892
r19            3   -8.3800554  -2.3125875   3.7548807
r18            4   -4.6691194   .22835743   5.1258345
r17            5    -8.129283  -2.7754233   2.5784359
r16            6   -6.8603821  -.75055814   5.3592658
r15            7   -10.123232  -1.2564343   7.6103635
r14            8   -10.710829  -2.7376504   5.2355275
r13            9   -12.330478  -5.0007019   2.3290732
r12           10   -11.226175  -1.1506662   8.9248428
r11           11    -17.51306  -9.3819475  -1.2508357
r10           12   -14.165311  -.04345011   14.078411
 r9           13   -10.856544  -1.3885684   8.0794067
 r8           14   -6.3317404   4.3279953   14.987731
 r7           15   -7.7231503   .84775668    9.418664
 r6           16   -8.3638954   -1.022577   6.3187418
 r5           17    -17.12311  -4.4348736   8.2533636
 r4           18   -10.074435  -.51595128   9.0425329
 r3           19    -3.328351   8.8427277   21.013805
 r2           20    -34.54834  -12.084179   10.379983
 r1           21   -31.114813  -22.920727   -14.72664
```

Since we do not specify the estimation method, `eventdd` uses Stata's `regress` command to estimate the model by ordinary least-squares regression (if the `ols` option is specified, the same result is obtained). We can also request others estimators for the underlying event study model; if we specify the `fe` option, the model should be estimated with the fixed-effects (FE) estimator.[8]

In the same way, we can estimate the results efficiently absorbing multiple levels of fixed effects via the `reghdfe` command by specifying the `hdfe` option, which is quite useful when we have to control for many fixed effects. Note that in this case, the fixed effects of interest must be indicated using the `absorb()` option which is passed to the `reghdfe` command. For instance, if we wish to absorb the temporal and geographic fixed effects, the necessary syntax is as follows:

```
. #delimit ;
delimiter now ;
. eventdd asmrs pcinc asmrh cases, hdfe absorb(i.stfips i.year) timevar(timeToTreat)
> ci(rcap) cluster(stfips) graph_op(ytitle("Suicides per 1m Women") xlabel(-20(5)25));
```

─────────────────────────────

[8]Assuming the data has been `xtset`, this would simply be:

```
. #delimit ;
delimiter now ;
. qui
> eventdd asmrs pcinc asmrh cases i.year, fe timevar(timeToTreat) ci(rcap)
> cluster(stfips) graph_op(ytitle("Suicides per 1m Women") xlabel(-20(5)25));
Note: with FE option do not include in varlist the categorical variables
      that identify the individual fixed effect

. #delimit cr
delimiter now cr
```

where note we no longer include the state fixed effects, as these will be taken care of given the panel fixed-effect (`xtreg`) estimator used to estimate the regression.

```
Note: with HDFE option do not include in varlist the categorical variables
      that identify the fixed effects to be absorbed in absorb()
(MWFE estimator converged in 2 iterations)
warning: missing F statistic; dropped variables due to collinearity or too few clusters

HDFE Linear regression                        Number of obs   =      1,617
Absorbing 2 HDFE groups                       F( 51,    48) =          .
Statistics robust to heteroskedasticity       Prob > F        =          .
                                              R-squared       =     0.7212
                                              Adj R-squared   =     0.6964
                                              Within R-sq.    =     0.0731
Number of clusters (stfips)  =          49    Root MSE        =    10.8116

                              (Std. Err. adjusted for 49 clusters in stfips)
```

| | | Robust | | | | |
asmrs	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
pcinc	-.0011046	.0004071	-2.71	0.009	-.0019232	-.0002861
asmrh	1.08064	.5968879	1.81	0.076	-.119483	2.280764
cases	-190.3716	134.4991	-1.42	0.163	-460.8002	80.05698
lag21	-22.92073	4.011063	-5.71	0.000	-30.98551	-14.85594
lag20	-12.08418	10.99637	-1.10	0.277	-34.19385	10.0255
lag19	8.842727	5.957829	1.48	0.144	-3.13629	20.82174
lag18	-.5159513	4.678946	-0.11	0.913	-9.923602	8.891699
lag17	-4.434874	6.210982	-0.71	0.479	-16.92289	8.053142
lag16	-1.022577	3.593627	-0.28	0.777	-8.248049	6.202895
lag15	.8477567	4.19552	0.20	0.841	-7.587901	9.283414
lag14	4.327995	5.218016	0.83	0.411	-6.163529	14.81952
lag13	-1.388568	4.634641	-0.30	0.766	-10.70714	7.930001
lag12	-.0434501	6.912751	-0.01	0.995	-13.94247	13.85557
lag11	-9.381948	3.980237	-2.36	0.023	-17.38475	-1.379145
lag10	-1.150666	4.932033	-0.23	0.817	-11.06718	8.765851
lag9	-5.000702	3.587977	-1.39	0.170	-12.21481	2.213409
lag8	-2.73765	3.902927	-0.70	0.486	-10.58501	5.10971
lag7	-1.256434	4.34036	-0.29	0.773	-9.983313	7.470445
lag6	-.7505582	2.990802	-0.25	0.803	-6.763968	5.262852
lag5	-2.775423	2.620752	-1.06	0.295	-8.044798	2.493952
lag4	.2283574	2.39735	0.10	0.925	-4.591837	5.048552
lag3	-2.312587	2.970068	-0.78	0.440	-8.28431	3.659136
lag2	-.5157397	2.514907	-0.21	0.838	-5.572299	4.54082
lead0	.2507466	2.722144	0.09	0.927	-5.222491	5.723984
lead1	-1.619351	2.941537	-0.55	0.585	-7.533709	4.295006
lead2	-1.687107	3.898178	-0.43	0.667	-9.524919	6.150706
lead3	-.7444709	2.862572	-0.26	0.796	-6.500058	5.011116
lead4	-2.956354	2.832628	-1.04	0.302	-8.651735	2.739027
lead5	-2.377841	2.75474	-0.86	0.392	-7.916617	3.160935
lead6	-3.311888	3.568157	-0.93	0.358	-10.48615	3.862372
lead7	-5.136502	3.401946	-1.51	0.138	-11.97657	1.70357
lead8	-6.991146	3.086374	-2.27	0.028	-13.19672	-.785575
lead9	-4.82321	3.089481	-1.56	0.125	-11.03503	1.388607
lead10	-8.814158	3.6746	-2.40	0.020	-16.20244	-1.42588
lead11	-7.27331	3.631759	-2.00	0.051	-14.57545	.0288304
lead12	-6.151559	4.089512	-1.50	0.139	-14.37407	2.070957
lead13	-8.276837	3.946249	-2.10	0.041	-16.2113	-.3423707
lead14	-6.593221	3.867273	-1.70	0.095	-14.3689	1.182453
lead15	-7.850839	4.070836	-1.93	0.060	-16.0358	.3341258
lead16	-7.234422	4.270836	-1.69	0.097	-15.82151	1.35267
lead17	-8.516898	4.344278	-1.96	0.056	-17.25166	.2178592
lead18	-9.991582	3.758781	-2.66	0.011	-17.54912	-2.434046
lead19	-11.53613	3.861769	-2.99	0.004	-19.30074	-3.771526

```
    lead20 │   -9.219165    4.501869    -2.05   0.046   -18.27078    -.167551
    lead21 │   -10.79088    4.417864    -2.44   0.018   -19.67359   -1.908173
    lead22 │   -10.65478    4.608349    -2.31   0.025   -19.92049   -1.389076
    lead23 │   -12.08658     5.29214    -2.28   0.027   -22.72714   -1.446016
    lead24 │   -10.67796    6.147523    -1.74   0.089   -23.03838    1.682466
    lead25 │   -10.26777    7.459044    -1.38   0.175   -25.26518    4.729644
    lead26 │   -16.69255    10.54234    -1.58   0.120   -37.88934    4.504239
    lead27 │   -.4344752    8.147106    -0.05   0.958   -16.81533    15.94638
     _cons │    85.59069    10.90519     7.85   0.000    63.66433    107.5171
```

```
Absorbed degrees of freedom:
```

```
Absorbed FE │  Categories  - Redundant  = Num. Coefs
────────────┼──────────────────────────────────────────
     stfips │          49           49            0  *
       year │          33            0           33
```

```
* = FE nested within cluster; treated as redundant for DoF computation

. #delimit cr
delimiter now cr
```

The standard command output consists of the regression output, and the event study lag and lead coefficients along with their confidence intervals are plotted as in Figure 1. As discussed in Stevenson and Wolfers (2006), the event study plot provides evidence of a reduction in rates of female suicide following the passage of no fault divorce laws, with significant declines observed 8 years following reform passage. We note that in this specification where all possible lags and leads are included (the default behavior of `eventdd`), we do observe a number of significant differences in the *pre*-reform period, in lag 11, and lag 21. It is important to note however that these lags are sufficiently far from the time-period of treatment that not all treated states are observed, and so these significant declines likely owe to compositional changes in these variables. We discuss this further below, and limit analysis to balanced periods when discussing the `balanced` option of the command. Nevertheless, if desired, we can also formally test the joint significance of all the lag terms simultaneously with the hypothesis:

$$H_0 : \beta_{21} = \beta_{20} = \cdots = \beta_2 = 0 \qquad \text{versus} \qquad H_1 : H_0 \text{ does not hold.}$$

This can be simply assessed post-estimation using Stata's `test` command:

```
. #delimit ;
delimiter now ;
. test lag21 lag20 lag19 lag18 lag17 lag16 lag15 lag14 lag13 lag12 lag11 lag10
> lag9 lag8 lag7 lag6 lag5 lag4 lag3 lag2;

 ( 1)  lag21 = 0
 ( 2)  lag20 = 0
 ( 3)  lag19 = 0
 ( 4)  lag18 = 0
 ( 5)  lag17 = 0
 ( 6)  lag16 = 0
 ( 7)  lag15 = 0
 ( 8)  lag14 = 0
 ( 9)  lag13 = 0
 (10)  lag12 = 0
 (11)  lag11 = 0
```

```
(12)  lag10 = 0
(13)  lag9 = 0
(14)  lag8 = 0
(15)  lag7 = 0
(16)  lag6 = 0
(17)  lag5 = 0
(18)  lag4 = 0
(19)  lag3 = 0
(20)  lag2 = 0
      F( 20,   48) =   32.13
            Prob > F =    0.0000
. #delimit cr
delimiter now cr
```
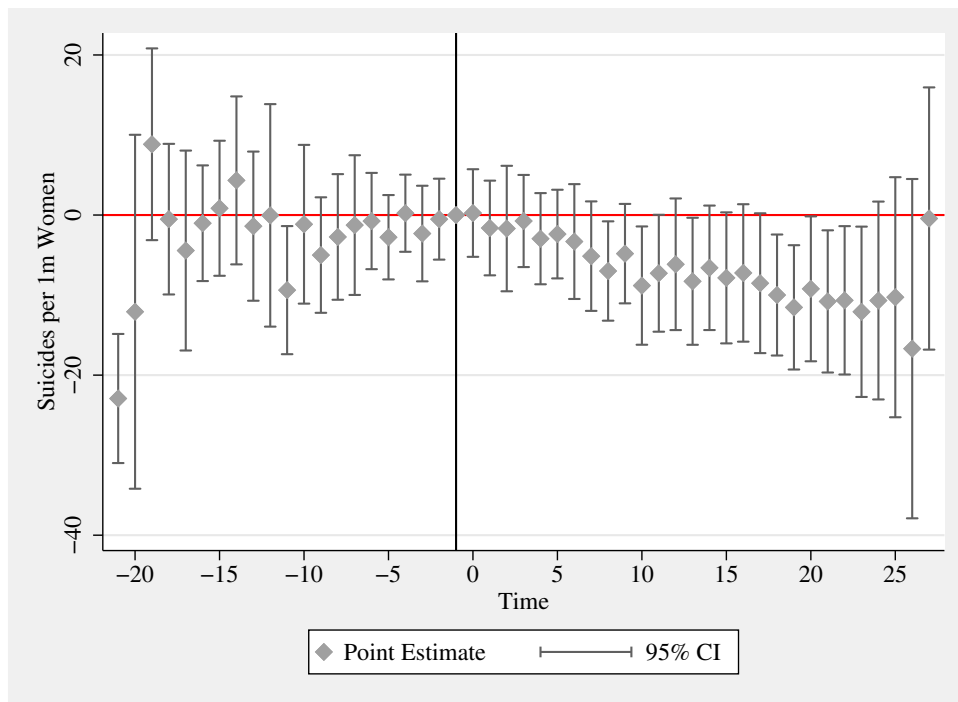

Figure 1: Event Study Example Based on No-fault Divorce Reforms

Notes: Event study model follows the No-Fault Divorce analysis described in (Stevenson and Wolfers 2006), and replication/extension of Goodman-Bacon (2018). Point estimates are displayed along with their 95% confidence intervals as described in equation 8. The baseline (omitted) base period is 1 year prior to the adoption of the reform in each reforming state, indicated by the solid vertical line in the plot.

This 'fully-saturated' model where all possible lags and leads are plotted is the default output in the eventdd command. However many alternative estimation procedures are permitted, and indeed are likely preferred, for example to avoid the behavior observed

above where lags and leads far from treatment will not be balanced given that only states adopting in certain early or late time periods will be observed in these lead/lag terms. Here we discuss a number of such alternatives, documenting their syntax in the `eventdd` command. Graphical output in each case is summarized in Figure 2.

Limiting Visualized Lags and Leads It may be a matter of interest to only show some lag/lead periods in the plot. For example, one such case discussed below relates to only plotting those lag/lead terms in which each treated state is observed. Generically, the `inrange` option allows for specifying that only certain coefficients and CIs should be included in plot. We note here that in this case, the underlying regression model *will* include all periods as in the first case, and as such, these lag/lead terms will simply correspond to a restricted range from Figure 1. For instance, if we want to show only the results between the time periods -10 and 10, the command will be:

```
. #delimit ;
delimiter now ;
. qui
> eventdd asmrs pcinc asmrh cases i.year, fe timevar(timeToTreat) ci(rcap) inrange
> lags(10) leads(10) cluster(stfips) graph_op(ytitle("Suicides per 1m Women"));
Note: with FE option do not include in varlist the categorical variables
      that identify the individual fixed effect
. #delimit cr
delimiter now cr
```

The output in this case is displayed in Figure 2(a). A special case of plotting limited lags/leads consists of the case in which one only wishes to show coefficients and CIs for which all states have a lag and lead term. We refer to this as a balanced plot, which can be produced quite simply using the `balanced` option. In this case, while all lags and leads are included in the underlying panel event study model, and only certain periods are plotted on the graph (like `inrange`), we do not need to know *a priori* which periods are balanced, as `eventdd` automatically identifies them. As panel (b) shows, in our case the balanced periods comprise periods between 5 years pre-reform and 11 years post-reform.[9] In this case, the syntax simply requires indicating the `balanced` option:

```
. #delimit ;
delimiter now ;
. qui
> eventdd asmrs pcinc asmrh cases i.year, fe timevar(timeToTreat) ci(rcap)
> balanced cluster(stfips) graph_op(ytitle("Suicides per 1m Women"));
Note: with FE option do not include in varlist the categorical variables
      that identify the individual fixed effect
. #delimit cr
delimiter now cr
```

---

[9]This can also be seen quite simply by tabulating (`tab`) the time variable used to indicate time to treatment. For example, in the case of the `timeToTreat` variable indicating relative time to the arrival of the divorce reform, we observe that for each of lags -5 up until lead 11, all periods are observed in the 36 states in which the event occurs. However, earlier lags and later leads are only observed in a sub-sample of states given that for others, these lags (leads) are earlier (later) than the first (last) year in the data sample.

Restricting Samples or Accumulating Lags/Leads In contrast to simply focusing on particular coefficients in the un-altered baseline model, one may wish to work with particular sub-samples which meet inclusion criteria, or accumulate lags and leads into periods which exceed some defined time, as an alternative way to avoid unbalanced lags and leads, as well as to avoid problems related to under-identification where all units are treated (Schmidheiny and Siegloch 2019). Consider the case where we wish to include 15 lags and 10 leads, but to only estimate the model with units that effectively have data for each of these periods. In the case of these data in particular which are yearly from 1964-1996, any units adopting no fault divorce reform between 1978 and 1996 will have (at least) 15 lags and 10 leads. Units adopting prior to 1978 will have fewer than 15 lags, and units adopting after 1996 will have fewer than 10 leads. To implement an estimation based on a balanced panel of observations with these lead/lag terms, the keepbal(*varname*) option is available, where *varname* indicates the panel unit over which balance should be applied (stfips in this case where the treatment unit is states). It is additionally necessary to explicitly indicate the period of interest for plotting within the balanced panel, for instance lags(15) and leads(10). This is all implemented in the below command.

```
. #delimit ;
delimiter now ;
. eventdd asmrs pcinc asmrh cases, hdfe timevar(timeToTreat) ci(rcap) cluster(stfips)
> absorb(i.stfips i.year) keepbal(stfips) lags(15) leads(10) graph_op(ytitle("Suicides
> per 1m Women"));
Note: with HDFE option do not include in varlist the categorical variables
      that identify the fixed effects to be absorbed in absorb()
(MWFE estimator converged in 5 iterations)
warning: missing F statistic; dropped variables due to collinearity or too few clusters
```

```
HDFE Linear regression                            Number of obs   =        507
Absorbing 2 HDFE groups                           F( 28,     15) =          .
Statistics robust to heteroskedasticity           Prob > F        =          .
                                                  R-squared       =     0.4758
                                                  Adj R-squared   =     0.3831
                                                  Within R-sq.    =     0.1416
Number of clusters (stfips)   =          16       Root MSE        =    10.1003
                                (Std. Err. adjusted for 16 clusters in stfips)
```

asmrs	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf. Interval]	
pcinc	-.0019977	.0007185	-2.78	0.014	-.0035291	-.0004663
asmrh	2.287467	1.540722	1.48	0.158	-.996504	5.571437
cases	-390.4487	180.673	-2.16	0.047	-775.544	-5.353367
lag15	1.292386	6.839513	0.19	0.853	-13.28569	15.87046
lag14	4.959406	8.015248	0.62	0.545	-12.12469	22.0435
lag13	4.003705	9.322968	0.43	0.674	-15.86773	23.87514
lag12	-3.997656	5.732168	-0.70	0.496	-16.21548	8.22017
lag11	-2.629473	4.276906	-0.61	0.548	-11.74548	6.486536
lag10	-3.987131	5.759912	-0.69	0.499	-16.26409	8.289831
lag9	-7.773707	4.653677	-1.67	0.116	-17.69278	2.145371
lag8	-6.673529	6.393574	-1.04	0.313	-20.30111	6.954051
lag7	-3.861924	4.692105	-0.82	0.423	-13.86291	6.13906
lag6	-6.44329	3.524894	-1.83	0.088	-13.95642	1.069845
lag5	-7.389526	3.375333	-2.19	0.045	-14.58388	-.1951743
lag4	-11.18006	6.114788	-1.83	0.087	-24.21342	1.853305

```
        lag3 │  -2.146437   5.282028   -0.41   0.690   -13.40481    9.111939
        lag2 │  -2.254288   3.599554   -0.63   0.541   -9.926557     5.41798
       lead0 │  -3.458685   2.757283   -1.25   0.229   -9.335696    2.418325
       lead1 │  -2.050697   3.988558   -0.51   0.615   -10.55211    6.450712
       lead2 │     3.8915   4.477937    0.87   0.399   -5.652997      13.436
       lead3 │     -5.949   5.387559   -1.10   0.287   -17.43231     5.53431
       lead4 │   1.151061   3.715023    0.31   0.761   -6.767323    9.069445
       lead5 │   4.283462   6.478434    0.66   0.519   -9.524993    18.09192
       lead6 │   -3.08539   2.849672   -1.08   0.296   -9.159321    2.988541
       lead7 │   1.776614    7.27749    0.24   0.810   -13.73499    17.28822
       lead8 │   .7996726   5.648138    0.14   0.889   -11.23905    12.83839
       lead9 │  -1.215031   4.864573   -0.25   0.806   -11.58362     9.15356
      lead10 │  -2.948988   4.776411   -0.62   0.546   -13.12967    7.231691
       _cons │   98.91732   14.70189    6.73   0.000    67.58098    130.2536
```

```
Absorbed degrees of freedom:
```

Absorbed FE	Categories	− Redundant	= Num. Coefs	
stfips	16	16	0	*
year	33	0	33	

```
* = FE nested within cluster; treated as redundant for DoF computation

. #delimit cr
delimiter now cr
```

Given that we now restrict to only certain states based on their period of adoption (as well as un-adopting states) the lag and lead estimates will differ to those from the fully saturated model discussed previously. In the output of the above command, we observe that the estimation sample consists only of 507 observations for adopting states with balance in the indicated lags/leads, as well as states which do not adopt (versus 1617 observations in the full sample specification). The corresponding event study plot is presented in the panel (c), where we note that the considerable change in estimation sample (chosen simply for expositional reasons) produces quite different results.

An alternative way to work with the imbalance in standardized time periods is to stipulate that all periods beyond some specified values should be accumulated into final lag and lead points, as indicated in equation 2 and 5. This is implemented with the `accum` option. When this is specified, the panel event study is provided based on the number of leads and lags indicated in the `lead(#)` and `lag(#)` option respectively, accumulating all periods beyond these periods into the final lag and lead term. For instance, if we specify `lags(-15)` and `leads(10)`, a single coefficient will capture the period −15 and earlier and the period 10 and later. This is illustrated in the following syntax, with the resulting graphical output presented in panel (d) of Figure 2.

```
. #delimit ;
delimiter now ;
. qui
> eventdd asmrs pcinc asmrh cases i.year, fe timevar(timeToTreat) ci(rcap)
> cluster(stfips) accum lags(15) leads(10) graph_op(ytitle("Suicides per 1m Women"));
Note: with FE option do not include in varlist the categorical variables
        that identify the individual fixed effect
```

```
. #delimit cr
delimiter now cr
```

As these end points have a different interpretation to additional lags and leads, acting as an estimate of long-term impacts of the event for all periods beyond intermediate lags/leads, by default the end point estimates will be plotted in an alternative color. This behavior can be controlled fully using the `endpoints_op()` options, allowing for options such as marker styles and colors to be passed to the underlying scatter plot (additional discussion is provided in sub-section 4.3 of this paper). Alternatively, as documented below, the `noend` option can be invoked which will omit these final accumulative end points from graphical output, as shown in panel (e):

```
. #delimit ;
delimiter now ;
. qui
> eventdd asmrs pcinc asmrh cases i.year, fe timevar(timeToTreat) ci(rcap)
> cluster(stfips) accum lags(15) leads(10) noend graph_op( ytitle("Suicides per
> 1m Women"));
Note: with FE option do not include in varlist the categorical variables
      that identify the individual fixed effect
. #delimit cr
delimiter now cr
```
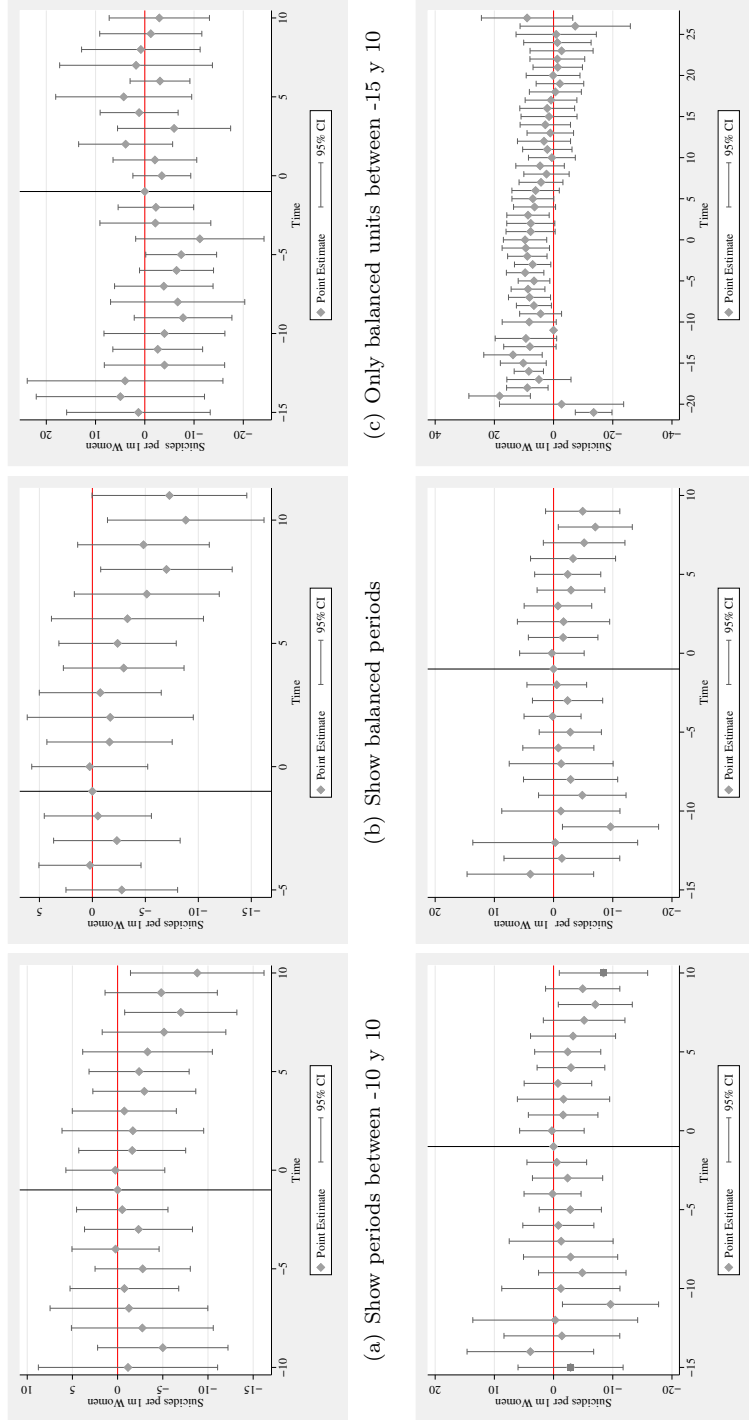
Finally, as discussed in section 2, the reference period for any panel event study estimated will be assumed to be the period immediately prior to the occurrence of the event in each state, unless otherwise indicated. This can be simply changed via the `baseline(#)` option. While the choice of -1 as the baseline period is arbitrary, it is frequently adopted, and so alternative baseline periods should be based on some empirical or theoretical consideration, although both models will be equivalent up to a single constant shift. Below we provide the syntax setting an alternative baseline period, with all coefficients and standard errors referring to differences relative to 11 years prior to the event of interest. By default, the `eventdd` ado places a vertical reference line at period -1 to visually indicate the period immediately prior to the passage of the event. However, if this reference line is not desired, the `noline` option can be specified, as documented in Figure 2f. If one wishes to provide alternative reference periods, these can be passed directly to the graphing command. For example, to add an alternative reference line in period 0 one should specify `graph_op(xline(0))`.

```
. #delimit ;
delimiter now ;
. qui
> eventdd asmrs pcinc asmrh cases i.year, fe timevar(timeToTreat) ci(rcap) noline
> baseline(-11) cluster(stfips) graph_op(ytitle("Suicides per 1m Women")
> xlabel(-20(5)25));
Note: with FE option do not include in varlist the categorical variables
      that identify the individual fixed effect
. #delimit cr
delimiter now cr
```

(a) Show periods between -10 y 10

(b) Show balanced periods

(c) Only balanced units between -15 y 10

(d) Accumulating into end points -15 y 10

(e) Not showing end points

(f) Baseline -11

Figure 2: Event Study Plots for No-fault Divorce Reforms: Output with Alternative Estimation Options

Notes: Refer to Figure 1 for notes. Panels here provide output under alternative options for the **eventdd** command including limiting lags and leads to certain periods (panels a and b), limiting to only states where all indicated leads are observed (panel c), accumulating all lags and leads beyond a certain point (panel d), not showing these end points (panel e), or based on alternative baseline reference periods (panel f).

4.2   Inference Options

The previous sub-section describes a number of alternative estimation procedures which are potentially of relevance in the estimation of a panel event study design. However, as discussed in section 2 of this paper, there are a number of inference considerations which must be weighed when implementing a panel event study model. Up until this point, the command has always been implemented with `cluster(stfips)`, indicating that a CRVE should be estimated, where clusters are based at the level of the state — the level at which the event is assigned in this case. As discussed in section 2.2, in this example based on 49 states, and hence 49 clusters, a CRVE is likely the appropriate inference mode for this model.

   However, the `eventdd` command allows for inference using a wild clustered bootstrap as post-estimation procedure, via its interaction with the `boottest` command (provided this command is installed on the user's system). This is indicated by the `wboot` option, which by default assumes that a clustered wild bootstrap is desired, with the cluster variable indicated in the `cluster()` option. This is especially useful when there are few clusters in the panel. However, note that given that this procedures is based around bootstrap resampling, the inference procedure likely will take longer than inference based on Stata's native CRVE, and additionally, that the `wboot` option may not be combined with the `hdfe` estimation option. However, `boottest` offers considerable other benefits, including the option to undertake inference with two way clustering which may exhibit preferable size properties in the case of very few clusters (Mackinnon and Webb 2018). Any option which should be passed directly to boottest can be indicated in the `wboot_op()` option, as illustrated with the `seed` option below, ensuring replicability in pseudo-random bootstrap resamples if desired. Figure 3 contrasts the differences between the previous CRVE-based inference procedure with the wild cluster bootstrap inference procedure illustrated here.

```
. #delimit ;
delimiter now ;
. qui
> eventdd asmrs pcinc asmrh cases i.year, fe timevar(timeToTreat) ci(rcap)
> cluster(stfips) accum lags(10) leads(10) wboot wboot_op(seed(1303)) graph_op(ytitle(
> "Suicides per 1m Women"));
Note: with FE option do not include in varlist the categorical variables
      that identify the individual fixed effect
. #delimit cr
delimiter now cr
```

   Finally, note that as standard, `eventdd` provides 95% confidence intervals in the command's output, returned objects, and the resulting graph and legend. The `level()` option allows for alternative levels to be indicated, where for example 90% CIs are requested below. Graphical output only differs in the confidence intervals provided (Figure 5(a) versus (b)).

```
. #delimit ;
delimiter now ;
. qui
> eventdd asmrs pcinc asmrh cases i.year, fe timevar(timeToTreat) ci(rcap)
```

```
> level(90) cluster(stfips) accum lags(15) leads(10) graph_op(ytitle("Suici
> des per 1m Women"));
Note: with FE option do not include in varlist the categorical variables
      that identify the individual fixed effect
. #delimit cr
delimiter now cr
```
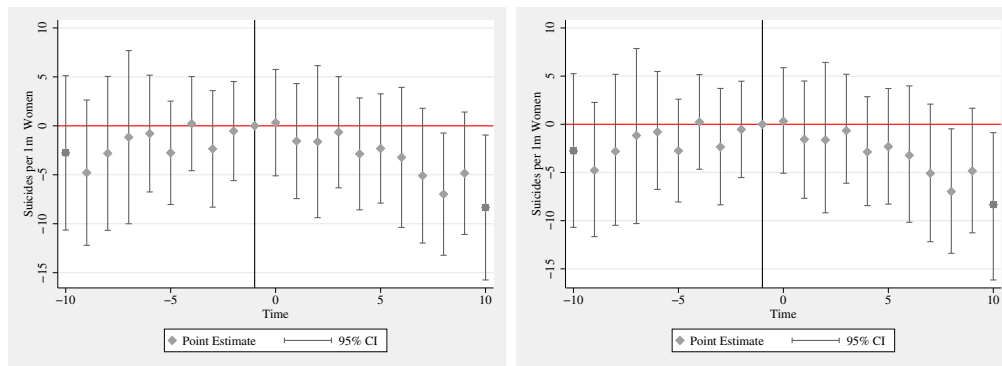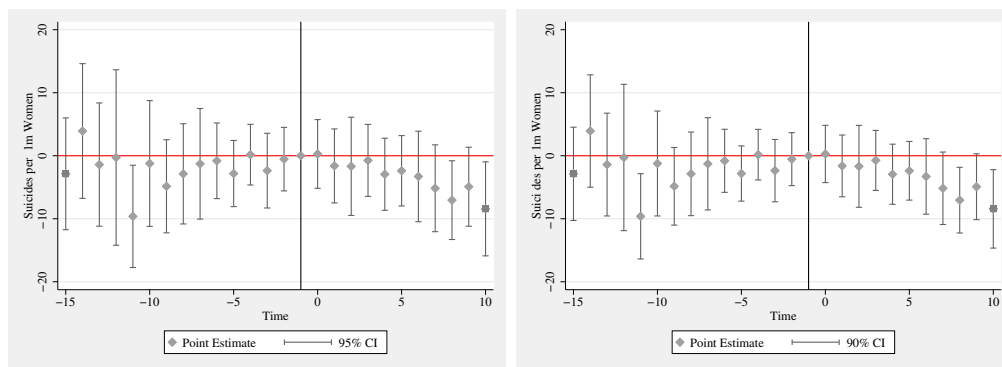

(a) Cluster-Robust VCE   (b) Wild cluster boot-strapped, `seed(1303)`

Figure 3: Visualizing Alternative Inference Procedures for Event Study Models

(a) 95% Confidence level   (b) 90% Confidence level

Figure 4: Default Event Study Plots with Alternative Confidence Intervals

4.3 Altering Standard Appearance

`eventdd`, allows for a number of ways to visualize the confidence intervals using a range of Stata's standard `twoway graph` types. The command requires that the user specify one of the following types of confidence intervals by specifying `ci`(*rarea*) for an

interval with area shading, ci(*rcap*) for an interval with capped peaks, and ci(*rline*) for an interval with lines. Figure 5 shows the initial event study from Figure 1, however now with the three alternative types of plots available. One type must be specified, and this will be apply for all intervals displayed.

These graph types can be fully controlled using the ci_op() option, though the options included must be compatible with the actual type of confidence interval requested. The compatibility of options can be confirmed in Stata's help files for twoway rcap, twoway rarea, or twoway rline for each of the accepted ci() options. Similarly, we can specify any options desired for the graphing of the the coefficients in the plot with the coef_op() option and if we are accumulating periods into final points, we can specify graphing options for these points in endpoints_op(). In both cases, these accept any valid options for Stata's twoway scatter plot type. Finally, a graph_op() option allows for the inclusion of any general graphing options, such as alternative labelling schemes, graph schemes or title options. In Figure 6 we compare a standard output (left) with an alternative output (right) taking advantage of Stata's transparency options and alternative color schemes. The eventdd syntax used to generate Figure 7(b) is provided below, followed by the resulting output.

```
. #delimit ;
delimiter now ;
. qui
> eventdd asmrs pcinc asmrh cases i.year, fe timevar(timeToTreat) ci(rarea)
> cluster(stfips) accum lags(15) leads(20) graph_op(xlabel(-15 "{&le} -15"
> -10 "-10" -5 "-5" 0 "0" 5 "5" 10 "10" 15 "15" 20 "{&ge} 20") scheme(s1mono)
> ytitle("Suicides per 1m Women")) ci_op(fcolor(ltblue%45)) coef_op(msymbol(Oh))
> endpoints_op(msymbol(O));
Note: with FE option do not include in varlist the categorical variables
      that identify the individual fixed effect

. #delimit cr
delimiter now cr
```
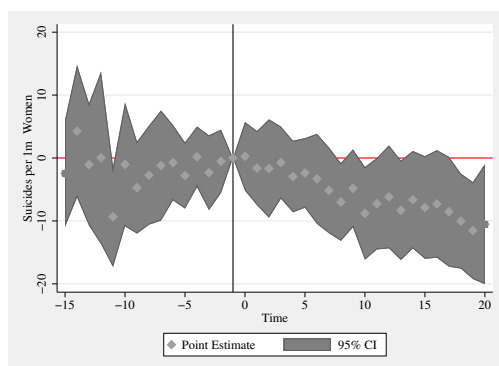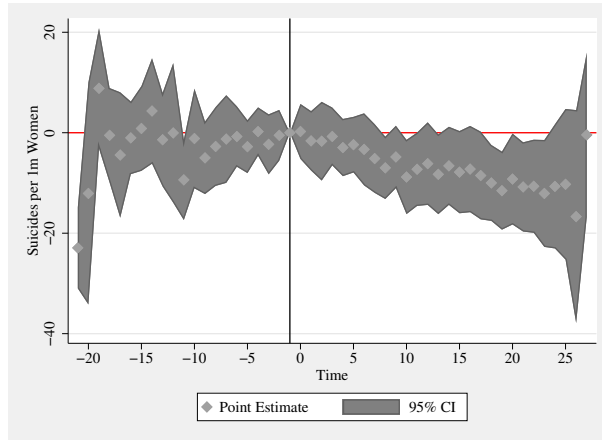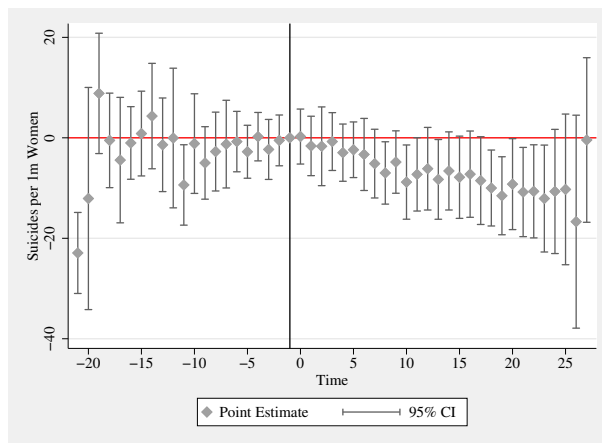

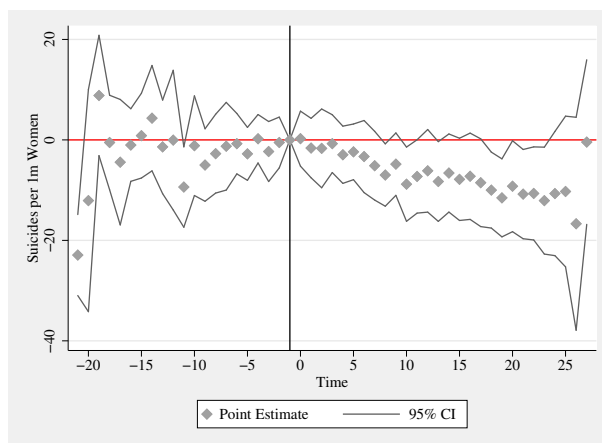(a) Plot with standard appearance   (b) Plot with alternative appearance options
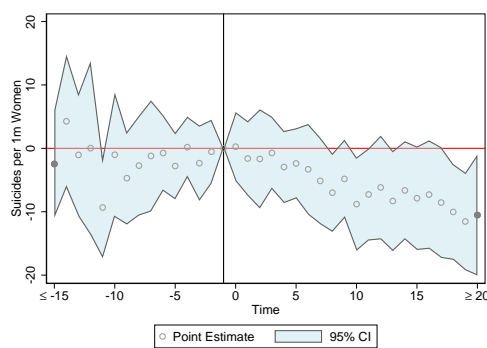
Figure 6: Event study plots No-fault Divorce Reforms: Appearance options

(a) `ci(rarea)`: Interval with area shading

(b) `ci(rcap)`: Interval with capped peaks

(c) `ci(rline)`: Interval with lines

Figure 5: Alternative Visualization Options for Event Study Confidence Intervals

5   Conclusions

The panel event study is an increasingly frequently used tool in the applied analysts' toolbox. It allows for the clear presentation of estimated impacts in quasi-experimental (observational) contexts, when one wishes to consider the impact of some event which occurs at (potentially) different times in different geographical areas. What's more while the discussion and examples provided in this paper are structured around geographical clustering of events (such as the application of divorce reforms studied in Stevenson and Wolfers (2006), and applied to demonstrate other two-way fixed effects methods (Goodman-Bacon 2018), this setting can similarly be applied where there is the temporal arrival of some event of interest in other dimensions, such as by age or other demographic groups.

In this paper we discuss a growing literature laying out panel event study designs, and introduce a flexible command `eventdd` which allows for their estimation and visualization in Stata. We introduce a number of estimation and inference concerns, and show how the command can simply deal with such concerns in an applied setting. While `eventdd` can be based on Stata's native routines such as `regress` or `xtreg` and cluster robust variance-covariance estimators, it is also able to interact with a number of extremely powerful user-written commands, allowing for extensions such as the efficient estimation of high-dimensional fixed effects equations, and the use of a wild cluster bootstrap for inference.

6   References

Abraham, S., and L. Sun. 2018. Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects.

Angrist, J. D., and J.-S. Pischke. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Athey, S., and G. W. Imbens. 2018. Design-based Analysis in Difference-In-Differences Settings with Staggered Adoption. Working Paper 24963, National Bureau of Economic Research.

Bailey, M. J., O. Malkova, and Z. M. McLaren. 2018. Does Access to Family Planning Increase Children's Opportunities? Evidence from the War on Poverty and the Early Years of Title X. *Journal of Human Resources* . http://jhr.uwpress.org/content/early/2018/07/03/jhr.55.1.1216-8401R1.abstract.

Bertrand, M., E. Duflo, and S. Mullainathan. 2004. How Much Should We Trust Differences-In-Differences Estimates? *The Quarterly Journal of Economics* 119(1): 249–275.

Borusyak, K., and X. Jaravel. 2018. Revisiting Event Study Designs, with an Application to the Estimation of the Marginal Propensity to Consume. mimeo. https://scholar.harvard.edu/files/borusyak/files/event_studies_may8_website.pdf.

Callaway, B., and P. H. C. Sant'Anna. 2018. Difference-in-Differences with Multiple Time Periods and an Application on the Minimum Wage and Employment. DETU Working Papers 1804, Department of Economics, Temple University. https://ideas.repec.org/p/tem/wpaper/1804.html.

Cameron, A., and D. Miller. 2015. A Practitioner's Guide to Cluster-Robust Inference. *Journal of Human Resources* 50(2): 317–372.

Cameron, A. C., J. B. Gelbach, and D. L. Miller. 2008. Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics* 90: 414–427.

de Chaisemartin, C., and X. D'Haultfoeuille. 2019. Two-way Fixed Effects Estimators with Heterogeneous Treatment Effects. NBER Working Papers 25904, National Bureau of Economic Research, Inc. https://ideas.repec.org/p/nbr/nberwo/25904.html.

Clarke, D., and K. Tapia Schythe. 2020. EVENTDD: Stata module to panel event study models and generate event study plots. https://EconPapers.repec.org/RePEc:boc:bocode:s458737.

Conley, T. G., and C. R. Taber. 2011. Inference with "Difference in Differences" with a Small Number of Policy Changes. *The Review of Economics and Statistics* 93(1): 113–125.

Correia, S. 2016. Linear Models with High-Dimensional Fixed Effects: An Efficient and Feasible Estimator. Technical report. Working Paper. http://scorreia.com/research/hdfe.pdf.

Dimitrovová, K., J. Perelman, and M. Serrano-Alarcón. 2020. Effect of a national primary care reform on avoidable hospital admissions (2000–2015): A difference-in-difference analysis. *Social Science & Medicine* 252: 112908. http://www.sciencedirect.com/science/article/pii/S0277953620301271.

Freyaldenhoven, S., C. Hansen, and J. M. Shapiro. 2019. Pre-event Trends in the Panel Event-Study Design. *American Economic Review* 109(9): 3307–38. http://www.aeaweb.org/articles?id=10.1257/aer.20180609.

Goodman-Bacon, A. 2018. Difference-in-Differences with Variation in Treatment Timing. Working Paper 25018, National Bureau of Economic Research. http://www.nber.org/papers/w25018.

Goodman-Bacon, A., T. Goldring, and A. Nichols. 2019. BACONDECOMP: Stata module to perform a Bacon decomposition of difference-in-differences estimation. Statistical Software Components, Boston College Department of Economics. https://ideas.repec.org/c/boc/bocode/s458676.html.

Kahn-Lang, A., and K. Lang. 2019. The Promise and Pitfalls of Differences-in-Differences: Reflections on 16 and Pregnant and Other Applications. *Journal of Business & Economic Statistics* 0(0): 1–14.

Mackinnon, J., and M. Webb. 2017. Wild Bootstrap Inference for Wildly Different Cluster Sizes. *Journal of Applied Econometrics* 32: 233–254.

———. 2018. The wild bootstrap for few (treated) clusters. *The Econometrics Journal* 21: 114–135.

Pacicco, F., L. Vena, and A. Venegoni. 2018. Event Study Estimations Using Stata: The Estudy Command. *The Stata Journal* 18(2): 461–476.

Rambachan, A., and J. Roth. 2019. An Honest Approach to Parallel Trends.

Roodman, D. 2015. BOOTTEST: Stata module to provide fast execution of the wild bootstrap with null imposed. Statistical Software Components, Boston College Department of Economics. https://ideas.repec.org/c/boc/bocode/s458121.html.

Roodman, D., M. Ø. Nielsen, J. G. MacKinnon, and M. D. Webb. 2019. Fast and wild: Bootstrap inference in Stata using boottest. *The Stata Journal* 19(1): 4–60.

Roth, J. 2019. Pre-test with Caution: Event-study Estimates After Testing for Parallel Trends.

Schmidheiny, K., and S. Siegloch. 2019. On Event Study Designs and Distributed-Lag Models: Equivalence, Generalization and Practical Implications. IZA Discussion Papers 12079, Institute of Labor Economics (IZA).

Stevenson, B., and J. Wolfers. 2006. Bargaining in the Shadow of the Law: Divorce Laws and Family Distress*. *The Quarterly Journal of Economics* 121(1): 267–288.

Suhonen, T., and H. Karhunen. 2019. The intergenerational effects of parental higher education: Evidence from changes in university accessibility. *Journal of Public Economics* 176: 195 – 217. http://www.sciencedirect.com/science/article/pii/S0047272719300945.

Venkataramani, A. S., E. F. Bair, R. L. O'Brien, and A. C. Tsai. 2020. Association Between Automotive Assembly Plant Closures and Opioid Overdose Mortality in the United States: A Difference-in-Differences Analysis. *JAMA Internal Medicine* 180(2): 254–262. https://doi.org/10.1001/jamainternmed.2019.5686.

About the authors

Damian Clarke is an Assistant Professor at The Department of Economics of The Universidad de Chile, a Research Fellow at IZA and an Associate at the Millennium Nucleus for the Study of the Life Course and Vulnerability.

Kathya Tapia Schythe is an Adjunct Professor in econometrics at The University of Santiago de Chile, and a young researcher associated with the Millennium Nucleus for the Study of the Life Course and Vulnerability.