# The Stata Journal



The *Stata Journal* publishes reviewed papers together with shorter notes or comments, regular columns, book reviews, and other material of interest to Stata users. Examples of the types of papers include 1) expository papers that link the use of Stata commands or programs to associated principles, such as those that will serve as tutorials for users first encountering a new field of statistics or a major new technique; 2) papers that go "beyond the Stata manual" in explaining key features or uses of Stata that are of interest to intermediate or advanced users of Stata; 3) papers that discuss new commands or Stata programs of interest either to a wide spectrum of users (e.g., in data management or graphics) or to some large segment of Stata users (e.g., in survey statistics, survival analysis, panel analysis, or limited dependent variable modeling); 4) papers analyzing the statistical properties of new or existing estimators and tests in Stata; 5) papers that could be of interest or usefulness to researchers, especially in fields that are of practical importance but are not often included in texts or other journals, such as the use of Stata in managing datasets, especially large datasets, with advice from hard-won experience; and 6) papers of interest to those who teach, including Stata with topics such as extended examples of techniques and interpretation of results, simulations of statistical concepts, and overviews of subject areas.

The *Stata Journal* is indexed and abstracted by *CompuMath Citation Index*, *Current Contents/Social and Behavioral Sciences*, *RePEc: Research Papers in Economics*, *Science Citation Index Expanded* (also known as *SciSearch*), *Scopus*, and *Social Sciences Citation Index*.

For more information on the *Stata Journal*, including information for authors, see the webpage

http://www.stata-journal.com

**Subscriptions** are available from StataCorp, 4905 Lakeway Drive, College Station, Texas 77845, telephone 979-696-4600 or 800-STATA-PC, fax 979-696-4601, or online at

http://www.stata.com/bookstore/sj.html

**Subscription rates** listed below include both a printed and an electronic copy unless otherwise mentioned.

| U.S. and Canada | | Elsewhere | |
|---|---|---|---|
| **Printed & electronic** | | **Printed & electronic** | |
| 1-year subscription | $115 | 1-year subscription | $145 |
| 2-year subscription | $210 | 2-year subscription | $270 |
| 3-year subscription | $285 | 3-year subscription | $375 |
| 1-year student subscription | $ 85 | 1-year student subscription | $115 |
| 1-year institutional subscription | $345 | 1-year institutional subscription | $375 |
| 2-year institutional subscription | $625 | 2-year institutional subscription | $685 |
| 3-year institutional subscription | $875 | 3-year institutional subscription | $965 |
| **Electronic only** | | **Electronic only** | |
| 1-year subscription | $ 85 | 1-year subscription | $ 85 |
| 2-year subscription | $155 | 2-year subscription | $155 |
| 3-year subscription | $215 | 3-year subscription | $215 |
| 1-year student subscription | $ 55 | 1-year student subscription | $ 55 |

Back issues of the *Stata Journal* may be ordered online at

http://www.stata.com/bookstore/sjj.html

Individual articles three or more years old may be accessed online without charge. More recent articles may be ordered online.

http://www.stata-journal.com/archives.html

The *Stata Journal* is published quarterly by the Stata Press, College Station, Texas, USA.

Address changes should be sent to the *Stata Journal*, StataCorp, 4905 Lakeway Drive, College Station, TX 77845, USA, or emailed to sj@stata.com.



The *Stata Journal* (ISSN 1536-867X) is a publication of Stata Press. Stata, Stata, Stata Press, Mata, Mata, and NetCourse are registered trademarks of StataCorp LP.

The Stata Journal (2015)
**15**, Number 3, pp. 796–808

# didq: A command for treatment-effect estimation under alternative assumptions

Ricardo Mora
Carlos III University of Madrid
Madrid, Spain
ricmora@eco.uc3m.es

Iliana Reggio
Carlos III University of Madrid
Madrid, Spain
ireggio@eco.uc3m.es

**Abstract.** When several pretreatment periods are available, identification of the treatment effect in a difference-in-differences framework requires an assumption relating dynamics for controls and treated in absence of treatment. Mora and Reggio (2012, Working Paper 12-33, Universidad Carlos III de Madrid) define a family of alternative identifying assumptions and propose a model that, contrary to the usual econometric specifications, allows one to identify the treatment effect for any given assumption in the family. In this article, we introduce a command, `didq`, that implements the model presented in Mora and Reggio, reports the estimated effect under alternative assumptions, and performs tests for the equivalence of the estimates. We also explain how to use the command to obtain the standard difference-in-differences estimator with or without polynomial trends.

**Keywords:** st0405, didq, difference-in-differences, treatment effect, identification, fully flexible model

## 1 Introduction

Difference-in-differences (DID) methods are widely used to evaluate the impact of policy interventions or other specific treatments on different outcomes of interest. DID estimators require an assumption relating dynamics for controls and treated in absence of treatment. The most common assumption for this is the "parallel paths" assumption when there is only one pretreatment period. The parallel paths assumption requires that, in absence of treatment, the average change in the outcome variable for the treated equals the observed average change in the outcome variable for the controls. This assumption implies that differences between the controls and the treated if untreated are assumed to be time-invariant.

When several pretreatment periods are available, the assumption equivalent to parallel paths is usually referred to as "common trends". Under the common trends assumption, in the absence of treatment, the average outcome change from any pretreatment period to any posttreatment period for the treated is equal to the equivalent average outcome change for the controls. This assumption is appealing if trends do not differ between the treated and the controls before treatment. Many researchers use the absence-of-pretreatment trend differentials between the controls and the treated as an argument in favor of the common trends assumption. In the presence of pretreatment-trend differentials, it is customary to adjust the econometric specification to try to ac-

 

commodate for those differences. Mora and Reggio (2012) (henceforth, MR) show that the inclusion of trend polynomials is not innocuous. Different trend-modeling strategies generally imply different assumptions regarding how trends for the controls and the treated in the absence of treatment are related.

More specifically, it is common to introduce linear trends to account for trend differences between the treated and the controls. Researchers usually associate the parameter for the interaction of a posttreatment dummy and the treated indicator with the treatment effect. This practice, which is no longer consistent with common trends, is correct if one assumes that the average acceleration for the treated under no treatment would have been equal to the observed average acceleration for the controls. MR refer to this assumption as parallel-2. They propose a family of alternative parallel-$q$ assumptions where $q$ is, at most, the number of pretreatment periods. They introduce a model that they call the "fully flexible" model, and they identify the treatment effect under each parallel-$q$ assumption. A critical result in MR is that the treatment effect $s$ periods after treatment under any given parallel-$q$ assumption can be expressed as the solution of an equation in differences with the parameters of the fully flexible model.

Implementation of DID under common trends requires only standard least-squares estimation of a very simple model. Because the treatment-effect estimate is identified as the parameter of one of the regressors, testing its significance is straightforward. In Stata, the user just needs to set the correct variable specification and employ the command `regress`. In contrast, estimating treatment effects with the fully flexible model in MR under alternative parallel assumptions requires two steps: first, standard least-squares estimation of the fully flexible model is conducted, and second, the solution of the equation in differences identifies the estimates. Computation of the standard errors of the treatment-effect estimates must take into account that the solution of the equation in differences is a linear combination of the parameters of the fully flexible model.

In this article, we show how the command `didq` implements this two-step procedure in Stata. The `didq` command fits the fully flexible model, and then it computes using default treatment effects under all parallel assumptions from parallel-1 to parallel-$q^{\max}$, where $q^{\max}$ is set by the user. Treatment effects are evaluated for a period set by the user. In addition, `didq` implements tests: a) for the equivalence of all parallel-$q$ assumptions between parallel-1 and parallel-$Q$, where $Q$ is the number of pretreatment periods; b) for each $q = 2, 3, \ldots, q^{\max}$, the equivalence of parallel-$q$ and parallel-$(q-1)$; and c) for each $q = 1, 2, \ldots, q^{\max}$, the absence of dynamics in treatment effects. `didq` also offers the option to report the DID standard model with flexible common dynamics and extensions that include a linear and a quadratic trend.

The rest of this article is structured as follows. In section 2, we briefly discuss MR and state identification conditions of the treatment effect under alternative parallel-$q$ assumptions. In section 3, we describe the syntax of `didq` and illustrate its use through several simulated examples. In section 4, we conclude.

# 2 Identification conditions under alternative parallel assumptions

Here we overview the identification of the treatment effect under alternative parallel assumptions in DID applications. A more detailed explanation is found in MR.

## 2.1 The basic setup

In the simplest empirical DID application, we have information on the variable of interest in at least two periods: before and after the treatment. Generally, we have information for $T \geq 2$ periods, and treatment starts sometime after the last pretreatment period, $t^*$, and finishes before the first posttreatment period, $t^* + 1$. There are at least two periods before treatment and one period after treatment. The effect of the treatment is evaluated $s$ periods after treatment, with $1 \leq s \leq T - t^*$.

Following conventional notation, we define $Y_{it}$ as the observed outcome variable at period $t$ for individual $i$. Let $Y_{it}^0$ denote the outcome in period $t$ when the individual receives no treatment, and let $Y_{it}^1$ denote the outcome in period $t$ when the individual receives treatment. For a given individual, either $Y_{it}^0$ or $Y_{it}^1$ is observed. Let $D_i = 1$ if the individual receives treatment and let $D_i = 0$ otherwise. Finally, let $X_i = (X_{i1}', \ldots, X_{iT}')'$, where $X_{it}$ is a vector of time-varying individual characteristics. The average treatment effect $s$ periods after treatment on the treated given $X$ is

$$\alpha\left(s|X\right) = E\left(Y_{i,t^*+s}^1 - Y_{i,t^*+s}^0 \,|X_i = X, D_i = 1\right)$$

To estimate the average counterfactual $E(Y_{it^*+s}^0|X_i = X, D_i = 1)$, one needs an assumption on how the trend behavior of the treated if untreated compares with the observed trend behavior of the untreated.

Let $\Delta_s \equiv (1 - L^s)$, $s \geq 2$, denote the $s$-period difference operator. MR propose a family of alternative nonnested assumptions. For a given positive integer $q \leq t^*$, parallel-$q$ implies that, for any $s$, $1 \leq s \leq T - t^*$,

$$E\left(\Delta_s \Delta^{q-1} Y_{i,t^*+s}^0 \,|X_i = X, D_i = 1\right) = E\left(\Delta_s \Delta^{q-1} Y_{i,t^*+s}^0 \,|X_i = X, D_i = 0\right)$$

where $\Delta^{q-1} \equiv (1 - L)^{q-1}$. For example, parallel-1 implies that, in the absence of treatment, average changes in outcome among those treated are equal to the average changes among comparable controls. By contrast, parallel-2 implies that, in the absence of treatment, average accelerations in outcomes among those treated are equal to the average accelerations among comparable controls. When the number of pretreatment periods is equal to 2, the common trends assumption implies parallel-1 and parallel-2 simultaneously.

Let us define $\text{DID}\,(q, s)$ as the difference-in-$q$-differences operator $s$ periods ahead,

$$\text{DID}\,(q, s) \equiv E\left(\Delta_s \Delta^{q-1} Y_{i,t^*+s} \,|X_i = X, D_i = 1\right) - E\left(\Delta_s \Delta^{q-1} Y_{i,t^*+s} \,|X_i = X, D_i = 0\right)$$

MR show that under parallel-$q$,

$$\Delta^{q-1}\alpha\left(s|X\right) = \text{DID}\,(q, s) \tag{1}$$

where $\alpha(s|x) = 0$ for all $1 \leq s \leq T - t^*$. This result can be used to obtain recursively $\alpha(s|X)$ for any value of $s$ under any parallel-$q$.

Define the operator $\alpha^q(s|X)$ as the solution to (1). MR show that for any $q \in (2, \ldots, t^*)$ and $s$ such that $1 \leq s \leq T - t^*$, $\alpha^q(s|X) = \alpha^{q-1}(s|X)$ if and only if

$$E\left(\Delta^{q-1} Y_{it^*} | X_i = X, D_i = 1\right) = E\left(\Delta^{q-1} Y_{it^*} | X_i = X, D_i = 0\right) \tag{2}$$

This result sets pretreatment-trend conditions under which assumptions parallel-$q$ and parallel-$(q-1)$ are equivalent.

## 2.2 Alternative modeling strategies

The conventional DID estimator is obtained using standard linear regression techniques. In the simplest case with only two periods, the treatment effect can be estimated under parallel paths from a regression that includes a constant, the treatment indicator $D_i$, a dummy variable for the posttreatment period ($\text{Post}_t$), and the interaction term $\text{Post}_t \times D_i$. In this setup, the treatment effect is identified by the parameter associated with the interaction term.

With several pretreatment periods, the assumption equivalent to parallel paths is common trends. MR show that common trends implies the validity of parallel-1 to parallel-$t^*$. Under common trends, the treatment effect can be estimated by ordinary least squares in the standard model

$$E(Y_{it} | X_i, D_i) = \beta X_{it} + \delta + \sum_{\tau=t_2}^{T} \delta_\tau I_t^\tau + \gamma^D D_i + \gamma_P^D \text{Post}_t \times D_i \tag{3}$$

where $I_t^\tau$ is a dummy for period $\tau$. More-flexible specifications in applied work involve the inclusion in the standard model of an interaction between $D_i$ and a linear or a quadratic time trend. Including a linear time trend implies the validity of parallel-2 to parallel-$t^*$, while a quadratic time trend implies the validity of parallel-3 to parallel-$t^*$.

By contrast, MR consider the fully flexible model with group-specific, fully flexible pretreatment and posttreatment dynamics:

$$E(Y_{it} | X_i, D_i) = \beta X_{it} + \delta + \sum_{\tau=t_2}^{T} \delta_\tau I_t^\tau + \gamma^D D_i + \sum_{\tau=t_2}^{T} \gamma_\tau^D \times I_t^\tau \times D_i \tag{4}$$

MR show that the treatment effect is identified under any parallel-$q$ assumption,

$$\Delta^{q-1} \alpha(s) = \Delta_s \Delta^{q-1} \gamma_{t^*+s}^D \tag{5}$$

Equation (5) implies that the identification strategy of the treatment effect will generally differ under alternative parallel assumptions. Suppose that we have three pretreatment periods and one posttreatment period. Consider the treatment effect under parallel-1, that is, $q = 1$. Because $\Delta^0 = 1$, then $\alpha(1) = \Delta \gamma_{t^*+1}^D = \gamma_{t^*+1}^D - \gamma_{t^*}^D$. In contrast,

under parallel-2, $\alpha(1) = \gamma^D_{t^*+1} - 2\gamma^D_{t^*} + \gamma^D_{t^*-1}$, while under parallel-3, $\alpha(1) = \gamma^D_{t^*+1} - 3\gamma^D_{t^*} + 3\gamma^D_{t^*-1} - \gamma^D_{t^*-2}$. In general, only when $\gamma^D_\tau = 0$ for all $\tau \leq t^*$—that is, only when pretreatment trends are equal between the treated and the controls—all parallel assumptions are equivalent and $\alpha(1) = \gamma^D_{t^*+1}$. Therefore, the test of the null hypothesis of common pretreatment trends ($H_0 : \gamma^D_\tau = 0$ for all $\tau \leq t^*$) is a test for the simultaneous equivalence of all parallel-$q$ assumptions.

The fully flexible model also allows for the comparison of any two consecutive parallel-$q$ assumptions. Using conditions (2) and (5), we see that testing the null $H_0 : \Delta^{q-1}\gamma^D_{t^*} = 0$ versus the alternative $H_a : \Delta^{q-1}\gamma^D_{t^*} \neq 0$ with $1 < q \leq t^*$ is a test for the equivalence of parallel-$q$ and parallel-$(q-1)$. For example, in the case of parallel-1 and parallel-2, the test would be $H_0 : \gamma^D_{t^*} = \gamma^D_{t^*-1}$ versus $H_a : \gamma^D_{t^*} \neq \gamma^D_{t^*-1}$.

Finally, the inclusion of fully flexible posttreatment-trend differentials also allows us to implement tests on the dynamics of the treatment effect under any parallel-$q$ assumption. For example, under parallel-1, testing the null $H_0 : \gamma^D_{t^*+s} = \gamma^D_{t^*+s+1}$ with $s = 1, \ldots, S-1$ is a test for the effect to be constant in the posttreatment period.

# 3 Stata implementation

In this section, we describe the command `didq`, which performs DID estimates under alternative parallel-$q$ assumptions. The following are minimum data requirements for executing the command. First, the data must contain at least two observations per group and period combinations. Second, there must be at least one period before treatment starts and one period after treatment ends. The dataset must contain a variable that identifies the period from which each observation is drawn as well as a time-invariant treatment variable that signals treatment. The output and the treatment variables must be numeric for the command to run. The time variable must be an integer (that is, must be either byte, int, or long). In addition, for the computations in `didq` to be meaningful, the difference between any two consecutive periods should be equal to 1.

The command `didq` first estimates an auxiliary regression using `regress` and then computes—in Mata—the treatment effects and test statistics as linear combinations of the estimates of the auxiliary regression. `didq` is by-able, allows weights, and is an e-class ado. In addition to the treatment effects and their standard errors, `didq` also saves the vector of coefficient estimates of the auxiliary regression and their variance–covariance matrix.

By default, the command will stop with an error message if the fully flexible model is not implementable, that is, if there is a problem of multicollinearity in the fully flexible model. This is intended behavior. All interactions between $D_i$ and the time dummies for the pretreatment periods should be estimated to compute the effects under any given parallel-$q$ assumption. If the `regress` command in Stata automatically drops one of the interactions to avoid the perfect multicollinearity issue, the interpretation of the remaining coefficients changes, and the algorithm used to compute the effect under a given parallel-$q$ assumption is no longer valid.

## 3.1 Syntax

`didq` *depvar* [ *indepvars* ] [ *if* ] [ *in* ] [ *weight* ], `treated(`*treatvar*`)` `time(`*timevar*`)`
[ `begin(`#*b*`)` `end(`#*e*`)` `q(`#*q*`)` `ff` `standard` `linear` `quadratic` `force`
`cluster(`*varname*`)` `detail` `level(`#`)` `auxiliary` ]

*depvar* is the variable on which the effects are to be estimated, and it must be numeric. The list *indepvars* is an optional variable list for the inclusion of controls in the model.

## 3.2 Options

`treated(`*treatvar*`)` specifies variable *treatvar* as the variable that signals treatment. *treatvar* is time-invariant and must take value `0` for observations from the control group and value `1` for observations from the treated group. `treated()` is required.

`time(`*timevar*`)` sets numeric variable *timevar* as the time variable. *timevar* specifies the discrete periods from which the observations are taken and must be of type byte, int, or long. Two consecutive periods should differ by 1 for the computations to be meaningful. `time()` is required.

`begin(`#*b*`)` and `end(`#*e*`)` set the first and the last posttreatment periods on which we want to evaluate the effects $t^* + 1$ and $t^* + S$, respectively. They take integers as arguments only. When $timevar = t^* + 1$, then $s = 1$. Values $t^* + 1$ and $t^* + S$ must be such that

$$\min(timevar) < t^* + 1 \leq t^* + S \leq \max(timevar)$$

By default, `begin()` and `end()` are set equal to $\max(timevar)$; that is, the last period in *timevar* is assumed to be the only posttreatment period. If one of the two options is not specified, then the missing option is set equal to the one specified.

`q(`#*q*`)` sets the highest parallel-$q$ assumption, $q^{\max}$, to be used in the estimations (relevant only with option `ff`). It must lie between 1 and $t^*$, the number of pretreatment periods. For example, with only two pretreatment periods, $q^{\max}$ may be, at most, equal to 2. To compute estimates under both parallel-1 and parallel-2, we must set $q^{\max} = 2$. If we set $q^{\max} = 1$, only the estimates under parallel-1 will be obtained. If `q()` is not specified or if $q^{\max}$ is set equal to a value larger than $t^*$, then `didq` sets $q^{\max} = t^*$.

`ff`, `standard`, `linear`, and `quadratic` refer to the model to be used, and hence, only one of them may be specified.

`ff` fits the fully flexible model from (4); this is the default if no model is specified. When the fully flexible model is chosen, `didq` computes all estimates $\widehat{\alpha}^q(s)$, $q = 1, \ldots, q^{\max}$, and $s = 1, \ldots, S$, and their standard errors. In addition, three types of tests are conducted: first, the test for the equivalence of all parallel-$q$ assumptions between 1 and $t^*$; second, for any $q \in (2, \ldots, q^{\max})$, the equivalence of parallel-$q$ and

parallel-$(q-1)$; and third, if $S > 1$, for any $q \in (1, \ldots, q^{\max})$, the test of absence of dynamics in treatment effects, that is, $H_0 \colon \alpha(q, s) = \alpha(q, s-1)$ for $s = 2, \ldots, S$.

`standard` fits the standard model with common flexible dynamics from (3). `linear` and `quadratic` fit the extended standard model with a linear and quadratic trend interaction with $D_i$, respectively. The output displayed under each of these three options includes the estimate and standard error of the DID estimator as well as tests for the absence of dynamics in treatment effects between $t^* + 1$ and $t^* + S$ under the respective model. In addition, the output includes the test of the equivalence of all parallel-$q$ assumptions implicitly assumed in each model. Hence, it reports the test of the equivalence of all parallel assumptions between parallel-1 and parallel-$t^*$ with `standard`, between parallel-2 and parallel-$t^*$ with `linear`, and between parallel-3 and parallel-$t^*$ with `quadratic`.

`force` computes estimates even with perfect multicollinearity issues in the auxiliary regression. If the `regress` command drops one of the interactions to avoid the perfect multicollinearity issue, the algorithm used to compute the effect under a given parallel-$q$ assumption is no longer valid. However, if the variable or variables automatically dropped are not the interactions between the pretreatment time dummies and $D_i$, the effect estimates are still valid.

`cluster(`*varname*`)` specifies the clustered sandwich estimator. The default is the Huber/White/sandwich estimator.

`detail` is relevant only when the fully flexible model is fit. When this option is chosen, `didq` additionally displays $t$ ratios, $p$-values, and confidence intervals of all effect estimates.

`level(#)` sets the confidence level, as a percentage, for confidence intervals. The default is `level(95)`, and it is relevant only with the `detail` option.

`auxiliary` displays the auxiliary regression. This option is important when there is perfect multicollinearity. In this case, `didq` stops with an error message by default. Using the option `auxiliary` combined with the option `force` checks whether the perfect multicollinearity problem affects any of the parameter estimates used to estimate the treatment effect. If so, the results obtained with the `force` option are invalid and should not be used. If the perfect multicollinearity issue arises from the additional controls, the `didq` estimates obtained with the `force` option are valid.

## 3.3 Stored results

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

Scalars

| | |
|---|---|
| `e(N)` | number of observations |
| `e(common_trend)` | Wald test of the joint significance of all interactions of pretreatment time dummies and the treatment dummy |

Matrices

| | |
|---|---|
| `e(alpha)` | $q^{\max} \times S$ matrix where element alpha $(q, s)$ corresponds to $\widehat{\alpha}(q, s)$ |
| `e(std_alp)` | $q^{\max} \times S$ matrix where element std_alp $(q, s)$ corresponds to $\widehat{\text{std}}\{\widehat{\alpha}(q, s)\}$ |
| `e(beta)` | vector of estimates in auxiliary regression. The first elements of `e(beta)` are the estimates of the coefficients for the interactions between the treatment variable and the time dummies in the fully flexible model. For the standard, linear, and quadratic models, the first elements are the estimates of the coefficients of the interactions between the treatment variable and the corresponding polynomial elements (that is, constant, linear, and quadratic terms). In all models, the estimate of the coefficient of the treatment dummy is next. Then, the estimates for the coefficients for the common time dummies follow. Finally, `e(beta)` includes estimates for the coefficients for the additional controls (when available) and the constant. |
| `e(Vbeta)` | (co)variance estimates in auxiliary regression |
| `e(tests)` | equivalence tests and tests on the equality of the effect on all posttreatment periods |
| `e(p_values)` | $p$-values for the equivalence tests and tests on the equality of the effect on all posttreatment periods |

## 3.4 Some examples

Consider the simulated data `didq_examples.dta`. Variable `t` records the observation period and ranges from 1 to 5. Variable `output` is the outcome on which we want to estimate the treatment effect, and `D` is the treatment indicator. The data, with 250 observations in each of 5 periods, were generated from a particular case of the standard model

$$y_{it} = \delta + x_{it} + \sum_{\tau=2}^{5} \delta_\tau I_{\tau,t} + \gamma^D D_i + \gamma_P^D \text{Post}_t \times D_i + u_{it}$$

where $x_{it} \sim N(0, 0.25)$, $I_{\tau,t} = \mathbb{1}(t = \tau)$, $\text{Post}_t = \mathbb{1}(t \geq 4)$, $u_{it} \sim N(0, 1)$, $\Pr(D_i = 1) = 0.5$, $D_i \perp\!\!\!\perp x_{it}$, $\delta = \gamma^D = \gamma_P^D = 1$, and $\delta_t = t$ for all $t = 2, \ldots, 5$. In this model, conditional on exogenous $x_{it}$, controls and treated outcomes are subject to a common linear trend. The treated differ on average from controls before treatment by a constant $\gamma^D$ and after treatment by $\gamma^D + \gamma_P^D$. Under all parallel-$q$, $q = 1, 2, 3$, the treatment effect is identified as $\gamma_P^D = 1$.

### Example 1

Assume that you want to use observations from periods 3 and 4 only. The following example estimates the treatment effect at period 4 with the standard model without additional controls under the parallel paths assumption:

```
. use didq_examples
(Simulated data for illustration purposes)
. preserve
. drop if t<3 | t>4
(750 observations deleted)
. didq output, treated(D) time(t) standard
    Unconditional Standard Model
    Output: output                                    Number of obs =    500
    Sample Period: 3:4                  H0: Common Pre-dynamics =    n/a
    Treatment Period: 4:4
```

| | All s | H0: s=s-1 |
|---|---|---|
| All q | 1.083984 | n/a |
| | (0.2005) | |

```
    Robust Standard Errors in parenthesis
. restore
```

The heading of the output display provides basic information on the model, the dependent variable, and the sample. Because there is only one pretreatment period, no test for common pretreatment dynamics is applicable. The standard model assumes the equivalence of all parallel-$q$ assumptions, and it assumes that the effect has no dynamics. The estimated effect, 1.08, is thus presented under `All q` and `All s` categories. Because there is only one posttreatment period, the test for dynamics of the treatment effect is also not applicable.

## Example 2

Suppose that we also want to use the data from period 2. In the fully flexible model with two pretreatment periods, there are two alternative assumptions (parallel-1 and parallel-2) that lead to two alternative estimates.

```
. didq output if (t>1 & t<5), treated(D) time(t)
    Unconditional Fully Flexible Model
    Output: output                                    Number of obs =    750
    Sample Period: 2:4                  H0: Common Pre-dynamics =  .6931
    Treatment Period: 4:4                                  p-value =  .4051
```

| | s=1 | H0: q=q-1 | H0: s=s-1 |
|---|---|---|---|
| q=1 | 1.083984 | | n/a |
| | (0.2005) | | |
| q=2 | 1.250342 | -.1663581 | n/a |
| | (0.3453) | [0.4051] | |

```
    Robust Standard Errors in parenthesis
    p-values in brackets
```

By default, the beginning and end of the treatment period represent the last period in the estimating sample. Because there are two pretreatment periods, the test for common pretreatment dynamics is displayed. Each alternative estimate of the treatment effect is displayed under the corresponding `q` line. Line `q=1` corresponds to assuming parallel-1,

and 1.08 is the estimate under parallel-1. This is the same estimate (and the same standard error) as the estimate of the treatment effect with the standard model using only periods 3 and 4. The estimate under parallel-2, which is displayed in line `q=2`, is slightly larger at 1.25. The display also includes the test of the equivalence between parallel-2 and parallel-1 (at line `q=2` and column `H0: q=q-1`). With two pretreatment periods, this test is equivalent to the test on common predynamics, so the $p$-value is the same.[1] The conclusion of the test is that both parallel assumptions are equivalent, meaning that the controls and the treated have common pretreatment dynamics.

### Example 3

With only three periods, the treatment-effect estimate under parallel-2 is equivalent to the estimate of the treatment effect with the standard model and linear deterministic trends.

```
. didq output if (t>1 & t<5), treated(D) time(t) linear
    Unconditional Linear Trend Model
    Output: output                                    Number of obs =    750
    Sample Period: 2:4                  H0: Common Pre-dynamics =    n/a
    Treatment Period: 4:4
```

| | All s | H0: s=s-1 |
|---|---|---|
| All q | 1.250342<br>(0.3453) | n/a |

```
    Robust Standard Errors in parenthesis
```

Given that the linear model implies parallel-2 and beyond, the test for common predynamics requires at least three pretreatment periods. In this example, there are only two pretreatment periods, so the test is not applicable.

### Example 4

Consider the full sample, that is, three pretreatment periods ($t = 1, 2, 3$) and two posttreatment periods ($t = 4, 5$). With multiple posttreatment periods, options `begin()` and `end()` should be used to identify the interval in which to obtain the effects estimates. Under the fully flexible model, we can obtain three alternative estimates for the effect in period 4 and three alternative estimates for the effect in period 5.

1. The statistic of the equivalence of the parallel assumptions is the estimated effect on the last pretreatment period under parallel-2. The test statistic on the common dynamics is the Wald test of the joint significance of all pretreatment $\gamma_t^D$ in (4).

```
. didq output, treated(D) time(t) begin(4) end(5)
    Unconditional Fully Flexible Model
    Output: output                                      Number of obs =   1250
    Sample Period: 1:5                        H0: Common Pre-dynamics =  1.359
    Treatment Period: 4:5                                    p-value =   .507
```

| | s=1 | s=2 | H0: q=q-1 | H0: s=s-1 |
|---|---|---|---|---|
| q=1 | 1.083984 | 1.002169 | | .1756106 |
| | (0.2005) | (0.1926) | | [0.6752] |
| q=2 | 1.250342 | 1.334885 | -.1663581 | .0915826 |
| | (0.3453) | (0.5245) | [0.4051] | [0.7622] |
| q=3 | 1.365318 | 1.679811 | -.1149753 | .1244636 |
| | (0.6308) | (1.4784) | [0.7388] | [0.7242] |

```
    Robust Standard Errors in parenthesis
    p-values in brackets
```

The three alternative estimates for the effect at period 4 are shown under the heading `s=1`, while those for period 5 are shown under the heading `s=2`.[2] With three pretreatment periods, the test for common pretreatment dynamics is a test for the joint equivalence of parallel-1, parallel-2, and parallel-3. Two tests for the equivalence of parallel assumptions are additionally shown under column `H0: q=q-1`: the first is the equivalence of parallel-1 and parallel-2 in line `q=2`, and the second is the equivalence of parallel-2 and parallel-3 in line `q=3`. Because there are multiple posttreatment periods, we can conduct, for any given parallel assumption, a test on the equality of the effect on all posttreatment periods. These tests are shown in column `H0: s=s-1`.

## Example 5

Additional controls can be added to improve the accuracy of the estimates.

```
. didq output x1, treated(D) time(t) begin(4) end(5)
    Conditional Fully Flexible Model
    Output: output                                      Number of obs =   1250
    Sample Period: 1:5                        H0: Common Pre-dynamics =  2.721
    Treatment Period: 4:5                                    p-value =   .2566
```

| | s=1 | s=2 | H0: q=q-1 | H0: s=s-1 |
|---|---|---|---|---|
| q=1 | 1.123735 | 1.081014 | | .0608667 |
| | (0.1820) | (0.1723) | | [0.8051] |
| q=2 | 1.325644 | 1.484831 | -.2019087 | .400746 |
| | (0.3148) | (0.4776) | [0.2678] | [0.5267] |
| q=3 | 1.445622 | 1.844765 | -.1199781 | .2430778 |
| | (0.5743) | (1.3449) | [0.7015] | [0.6220] |

```
    Robust Standard Errors in parenthesis
    p-values in brackets
```

2. If there are more than three posttreatment periods, the default display reports the effects for only `s=1`, `s=2`, and `s=3`. To display all the effects, the option `detail` should be used.

**Example 6**

The tests in examples 4 and 5 suggest that the standard model is appropriate for these simulated data because we cannot reject common pretreatment dynamics and equal dynamic effects.

```
. didq output x1, treated(D) time(t) begin(4) end(5) standard
    Conditional Standard Model
    Output: output                                    Number of obs =   1250
    Sample Period: 1:5                  H0: Common Pre-dynamics =  2.721
    Treatment Period: 4:5                                 p-value =  .2566
```

| | All s | H0: s=s-1 |
|---|---|---|
| All q | .9404042 | .0611879 |
| | (0.1131) | [0.8046] |

```
    Robust Standard Errors in parenthesis
    p-values in brackets
```

Using the standard model with the full sample and additional controls, we obtain a reduction in the standard error of the estimated effect. In contrast with example 1, we can test both pretreatment dynamics and equal dynamic effects. The common predynamics test is the test of the joint equivalence of parallel-1, parallel-2, and parallel-3, and it is the same as the test in example 5. The test of equal dynamic effects (in column `H0: s=s-1`) is a Wald test of $H_0 : \gamma_4^D = \gamma_5^D$ in a standard model where the treatment effects can differ by period.

$$y_{it} = \delta + x + \sum_{\tau=2}^{5} \delta_\tau I_{\tau,t} + \gamma^D D_i + \gamma_4^D I_4 \times D_i + \gamma_5^D I_5 \times D_i + u_{it}$$

# 4 Conclusions

Identification of treatment effects using cross-sections when the dataset contains multiple pretreatment periods depends on specific assumptions about pretreatment dynamics and how they inform the counterfactual for the treated in the absence of treatment. MR discuss the most popular models used in the empirical literature and present the fully flexible model. For all these models, they derive the identification conditions of the treatment effect in terms of alternative assumptions.

In this article, we present the new command `didq`, which performs DID estimations under alternative assumptions as proposed by MR. We illustrate how to use the command `didq` through several examples. `didq` is a helpful tool for analyzing the robustness of estimated effects to alternative identifying assumptions and dynamic specifications. Moreover, equivalence and dynamics tests can be used to validate alternative models.

Other methods are available for estimating treatment effects in Stata. `teffects` provides six estimators of potential-outcome means, average treatment effects, and average treatment effects on the treated using observational data. In addition, the package

`diff` performs several DID estimations of the treatment effect in the standard model (Villa 2012).

# 5 References

Mora, R., and I. Reggio. 2012. Treatment effect identification using alternative parallel assumptions. Working Paper 12-33, Universidad Carlos III de Madrid.

Villa, J. M. 2012. Simplifying the estimation of difference in differences treatment effects with Stata. MPRA Paper 43943, University Library of Munich. http://mpra.ub.uni-muenchen.de/43943/1/MPRA_paper_43943.pdf.

### About the authors

Ricardo Mora is an associate professor of economics at Universidad Carlos III de Madrid in Spain.

Iliana Reggio is an assistant professor of economics at Universidad Carlos III de Madrid in Spain.