

# Discussion Papers

No. 5

February 2019



flexpaneldid

A Stata Command for Causal Analysis with Varying Treatment  
Time and Duration

Eva Dettmann, Alexander Giebler, Antje Weyh

## Authors

### **Eva Dettmann**

Halle Institute for Economic Research (IWH) –  
Member of the Leibniz Association,  
Centre for Evidence-based Policy Advice  
(IWH-CEP)  
E-mail: [eva.dettmann@iwh-halle.de](mailto:eva.dettmann@iwh-halle.de)  
Tel +49 345 7753 855

### **Alexander Giebler**

Halle Institute for Economic Research (IWH) –  
Member of the Leibniz Association,  
Department of Structural Change and  
Productivity  
E-mail: [alexander.giebler@iwh-halle.de](mailto:alexander.giebler@iwh-halle.de)  
Tel +49 345 7753 794

### **Antje Weyh**

Institute for Employment Research of the  
Federal Employment Agency (IAB)  
E-mail: [antje.weyh@iab.de](mailto:antje.weyh@iab.de)  
Tel +49 371 9118 642

## Editor

Halle Institute for Economic Research (IWH) –  
Member of the Leibniz Association

Address: Kleine Maerkerstrasse 8  
D-06108 Halle (Saale), Germany  
Postal Address: P.O. Box 11 03 61  
D-06017 Halle (Saale), Germany

Tel +49 345 7753 60  
Fax +49 345 7753 820

[www.iwh-halle.de](http://www.iwh-halle.de)

ISSN 2194-2188

The responsibility for discussion papers lies solely with the individual authors. The views expressed herein do not necessarily represent those of IWH. The papers represent preliminary work and are circulated to encourage discussion with the authors. Citation of the discussion papers should account for their provisional character; a revised version may be available directly from the authors.

Comments and suggestions on the methods and results presented are welcome.

IWH Discussion Papers are indexed in RePEc-EconPapers and in ECONIS.

*flexpaneldid*

# A Stata Command for Causal Analysis with Varying Treatment Time and Duration\*

## Abstract

The paper presents a modification of the matching and difference-in-differences approach of Heckman et al. (1998) and its Stata implementation, the command *flexpaneldid*. The approach is particularly useful for causal analysis of treatments with varying start dates and varying treatment durations (like investment grants or other subsidy schemes). Introducing more flexibility enables the user to consider individual treatment and outcome periods for the treated observations.

The *flexpaneldid* command for panel data implements the developed flexible difference-in-differences approach and commonly used alternatives like CEM Matching and difference-in-differences models. The novelty of this tool is an extensive data preprocessing to include time information into the matching approach and the treatment effect estimation. The core of the paper gives two comprehensive examples to explain the use of *flexpaneldid* and its options on the basis of a publicly accessible data set.

*Keywords: causal analysis, CEM matching, statistical distance function, conditional fixed effects difference-in-differences, nonparametric conditional difference-in-differences*

*JEL classification: A11, D61, H20, Z0*

\* A completely revised version of this paper has been published as: Dettmann, Eva; Giebler, Alexander; Weyh, Antje: *flexpaneldid*. A Stata Toolbox for Causal Analysis with Varying Treatment Time and Duration. IWH Discussion Paper 3/2020. Halle (Saale) 2020.

# 1 Introduction

This paper describes a Stata tool for the causal effects estimation of treatments within panel data sets, especially individual treatments characterized by varying start dates and different treatment durations. This treatment structure is typical for subsidy schemes that are common in industrial and placed based policy (e. g. investment subsidies and R&D funding), labor market programs (e. g. support for start-ups or training vouchers) or research funding (e. g. EU funding for scientists). But the tool can be relevant also for other research areas like economics of education or labor market research.

The estimation tool implements an approach based on a combination of matching and difference-in-differences (DID) and some alternative standard approaches. To gain more flexibility we modify the conditional difference-in-differences approach of [Heckman et al. \(1998\)](#) in three ways. First, we include individual treatment time information from the panel into the matching process. Second we introduce a combined statistical distance function for matching. Third, we incorporate flexible observation durations into the difference-in-differences estimation. This flexible conditional DID approach ensures that varying treatment phases can be accounted for in an appropriate way and that the point in time an individual is compared to his 'statistical twin' can be exactly determined.

The remainder of the paper is organized as follows. In section two, the special data structure and related challenges resulting from the treatment at different times and of different durations are explained in more detail. Section three gives an overview on empirical approaches dealing with this data structure when estimating causal effects. In section four the flexible conditional difference-in-differences approach is explained. Subject of section five is the presentation and explanation of *flexpaneldid* and the implemented options using two comprehensive examples. In section six, the syntax of the command and some general instructions are given, section seven concludes with an outlook.

## 2 Special characteristics of the data structure

The data structure we focus on is characterized by panel information on treated and non-treated observations, where treatment can start and end basically at every time. In this case, the treatment period consists of a mix of individual pre-treatment, treatment and post-treatment phases. Figure 1 gives an example for an unbalanced panel data set of firms for the years 2004-2014 with varying dates of application and individual durations from application to the start of the treatment ('waiting phases') as well as individual treatment durations. Within the observation period, the treatment phase of treated firm 1 is in the pre-treatment phase of firm 2 and the post-treatment phase of firm 3. Observing e. g. the outcome development from treatment start to one year later (denoted by O1 in figure 1) or from treatment start until one year after the treatment is finished (denoted by O2) means in this case we have to observe different 'sub-periods' for the three treated firms.

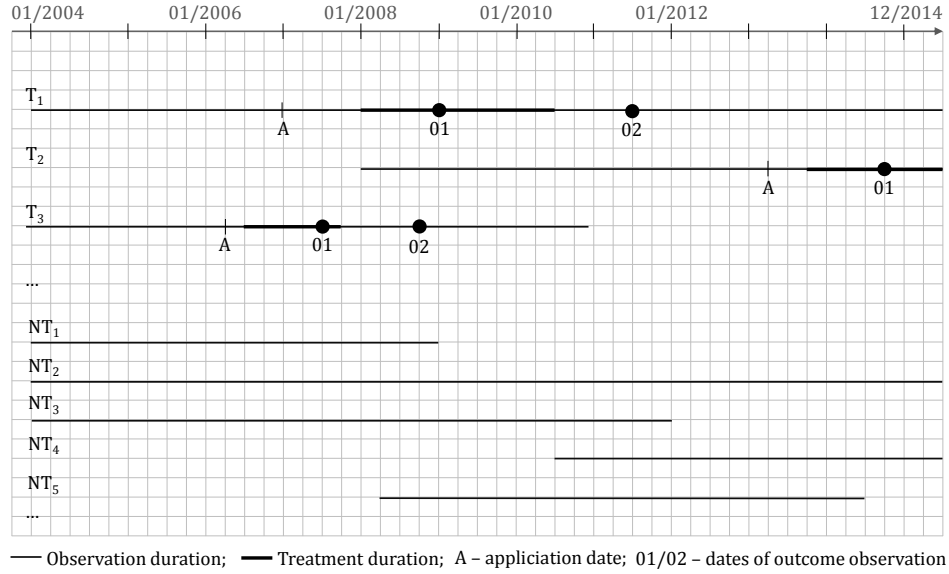


Figure 1: sketch of the typical data structure

As we know from Heckman et al. (1999), the economic environment influences the economic performance of persons or firms and should be considered when analyzing treatments effects. In a dynamic environment, this applies not only to the place but also to the time of observation. The example in figure 1 can illustrate that: in the year 2008, the financial crisis 'arrived' in real economy. In the following recession, the economic environment strongly deteriorated – resulting in a worse economic performance of many firms and also persons. The subsequent upswing phase caused a significant performance improvement for firms and persons. If we would ignore this development in the economic environment and would compare e. g. a firm before to a firm after the crisis, we would compare them in a different economic environment. This is denoted by 'time bias' in the following.

Also the treatment effect itself is influenced by the economic situation and can be heterogeneous over time (Bergemann et al. 2009).

Both, the time bias and the effect heterogeneity can be considered by including the information on individual 'sub-periods' into the estimation. This is possible with the flexible conditional DID approach that will be introduced in the paper.

Another phenomenon is referred to as *Ashenfelter's dip* (Ashenfelter 1978) or *fallacy of alignment* (Heckman et al. 1999): if the anticipation of a treatment leads to a (short-term) change in the behavior of the applicants, the estimated effect will be distorted. Examples for such behavioral changes are mainly discussed with regard to active labor market programs; e. g. unemployed persons may reduce their search effort for a new

job when they anticipate their participation in a training program (Bergemann et al. 2009). Pellegrini and Centra (2006) give another example: in the presence of a positive temporary shock to a specific sector, firms in this sector might be more willing to invest and to apply for a subsidy. In this case there is a positive relation between the application probability and the short term market perspectives in this sector, and the estimator tends to overestimate the effect of the policy intervention.

We do not contribute to the theoretical discussion on the best timing for matching and DID to account for an *Ashenfelter's dip*. We are rather convinced that the flexible conditional DID helps to account for this behavioural change in that it enables the user to consider expectations on the duration of the 'dip' when exactly determining the matching and the outcome observation time in relation to the treatment start.

### 3 Estimation approaches in empirical literature

In recent empirical literature we find a growing number of causal analyses of treatments in various economic fields that are based on data sets with a similar structure to the above described one. This literature review contains examples from labour market economics and health economics as well as evaluation studies of active labour market policies and place based policies. In this overview, we focus on causal effects estimation on the basis of panel data with more than two observation times.

Following comprehensive overviews on different evaluation approaches (e. g. Abadie 2005; Blundell and Costa Dias 2009; Imbens and Wooldridge 2009) and controversial discussions in the literature,<sup>1</sup> the idea of a combined control for selection resulting from observable and unobservable heterogeneity found a widespread adoption in empirical literature. In this context we can distinguish three estimation strategies. They are all based on the DID estimator and crucially depend on the assumption that unobservable individual characteristics influencing the outcome are invariant over time. This assumption is referred to as *common trend assumption* (Blundell and Costa Dias 2009) or *bias stability assumption* (Baltagi 2015) in the literature.

The *first* approach is the conditional DID introduced by Heckman et al. (1998). It combines a conventional DID estimation and matching. Here the compared outcome changes are defined conditional on matched samples instead of the whole samples of treated and non-treated persons or firms. The virtue of this nonparametric approach is the flexibility regarding distribution of the characteristics and the kind of influence on the observed outcome. Based on this idea, Bandick et al. (2014) evaluate the causal effect of foreign acquisition on R&D intensity in manufacturing firms with at least 50 employees operating in Sweden. Merging administrative data from different sources for the period 1993–2002 they combine a DID estimator with a Propensity score nearest neighbor matching including pre-treatment characteristics into the matching.

Bergemann et al. (2009) use a combination of kernel matching and DID to analyze causal effects of training in East Germany on transition from non-employment to em-

<sup>1</sup>See the discussion of Dehejia and Wahba (1999, 2002); Dehejia (2005) and Smith and Todd (2005a,b) as the most prominent example.

ployment. Based on the Labor Market Monitor Saxony Anhalt for the years 1997, 1998, and 1999, they allow for heterogeneous treatment effects depending on the time the training program takes place, i.e. they estimate different effects of training programs on employment for distinct fixed time periods. Based on their results they emphasize the importance of considering labour market dynamics. Additionally, they discuss different ideas to control for a potential *Ashenfelter's dip*.

The purpose of the paper of [Pellegrini and Centra \(2006\)](#) is to evaluate the impact of subsidies allocated by Law 488/1992 on subsidized finished projects starting until 1999 on the basis of a combined data set. Using a DID estimator augmented by a matching procedure, they compare subsidized and rejected firms in manufacturing and firm services sectors in the Mezzogiorno region in the year before the project started and the year after the project is finished. [Pellegrini and Centra \(2006\)](#) consider different treatment durations. Here, the 'virtual project duration' for the controls are defined with the help of the start date of the projects and the average project duration of different auctions.

Similarly, [Bernini and Pellegrini \(2011\)](#) analyze subsidies allocated to manufacturing firms in the southern Italian regions under the Law 488/1992. Using information on finished projects over the period 1996–2004, they apply a nonparametric MDID estimator (with kernel and stratification matching) to compare subsidized and rejected firms.

[Caliendo and Künn \(2011\)](#) provide empirical evidence on the effectiveness of start-up subsidies for unemployed persons. They use administrative data on entries into Start-up subsidies and Bridging Allowance in the third quarter of 2003 as the basis for Propensity Score matching. The subsequent difference-in-difference estimation of effects on different outcomes reported in repeated computer assisted telephone interviews allow them to show long-term effects and effect heterogeneity among the interviewed persons.

The *second* strategy, the DID model, incorporates the idea of difference-in-difference into a conventional panel regression model. The virtue of this parametric approach is that individual time-invariant characteristics are captured by the fixed effect, and that additional covariates and time information (e. g. via year dummies) can be included. Besides the mean treatment effect, the development of the effect over time can be estimated. As is the case in conventional DID, without any data preprocessing, the estimation results are based on the whole sample and may be not very meaningful for the treated individuals that commonly represent a rather special subgroup in the sample. Another problem is the model dependence of the estimation ([Ho et al. 2007](#)) and the usual, rather strict model assumptions (homoscedasticity, no autocorrelation, etc.) of the used panel models.

Within the DID model framework, [Autor \(2003\)](#) assesses the contribution of unjust dismissal doctrine on employment outsourcing in the U.S. Temporary Help Services between 1979 and 1995. In order to control for a potential anticipation effect – or Ashenfelter's dip –, he includes 'leads' of the dependent variable for one and two years before treatment. Besides the average treatment effect, he additionally includes 'lags' to verify dynamics of the treatment effect over time.

Based on linked treatment and corporation data, [Bronzini and de Blasio \(2006\)](#) evaluate the causal effects of subsidies under the Law 488/1992 on firm performance. For the period 1994–2001 they compare firms with subsidized and firms with rejected projects

within a DID model including yearly post-treatment dummies to capture the time varying treatment effect.

Ham et al. (2011) use a fixed effects DID model comparing treated Census tracts with their (in geographical sense) nearest neighbor tracts to measure the impact of different place-based policies on local labor markets on the base of data from the 1980, 1990 and 2000 Censuses.

In a similar way, Neumark and Kolko (2010) evaluate the effectiveness of California's enterprise zone programs on employment with panel data from different sources for the period 1992-2004. Taking advantage of the rich yearly information in their data, they also include leads and lags of the dependent variable.

Marcus and Siedler (2015) use monthly administrative data from a random sample of all hospitalizations during the years 2007–2011 in Germany to inspect the effect of the late-night alcohol sales ban on alcohol-related hospitalization rates. They apply fixed-effects DID models with various control variables to estimate the mean effect of the ban as well as the development of the effect over time.

*Third*, in more recent literature, panel regression models are combined with matching. Heyman et al. (2007) use a combination of matching and the DID model to estimate the influence of foreign ownership on wages in Swedish firms. Based on a combination of different register data from Statistics Sweden for the years 1990-2000, they implement year-by-year Propensity Score Nearest Neighbor Matching. The matched sample is then used in a DID model to examine the dynamic effects of takeovers on wages.

Greenaway et al. (2005) use matching and difference-in-differences analysis to investigate causal links between exports and firm performance for a panel of Swedish manufacturing firms over the period 1980–1997. They perform Propensity Score caliper matching at each year of the observation period for the entrants into export markets. The matched firms are then pooled in a panel data set for the subsequent random effects DID model estimation.

In their paper, Gustafsson et al. (2016) combine matching and a DID model to evaluate Swedish innovation subsidies. Based on combined administrative and register data for the years 1997-2011 they use CEM Matching to prepare the sample for fixed effects DID estimations of different outcomes.

Freier et al. (2015) estimate the effect of graduating from university with an honors degree on subsequent earnings. Based on the initial waves of the University Graduates Panel in Germany, they use (inter alia) a difference-in-differences research design combined with entropy balancing of Hainmueller and Xu (2013).

The next chapter 4 presents a nonparametric approach based on the conditional DID that especially focuses on treatment periods with mixed individual pre-treatment, treatment and post-treatment phases. The novelty of the flexible conditional DID approach is the flexibility in the definition of treatment start and treatment duration and the possibility to consider time information in the matching process.



## 4 The flexible conditional DID

The starting point is the idea of the nonparametric conditional difference-in-differences approach introduced by Heckman et al. (1998). The matching process can be seen in this context as a nonparametric data preprocessing in the sense of Ho et al. (2007) leading to more reliable causal effect estimates by reducing bias and variance. A special focus in our approach is to exclude potential 'time bias' resulting from comparing individuals at different points in time (see section 2). That means we have to incorporate the observation time information from the panel data into the matching process.

The first step of the flexible conditional DID approach is an extensive data reorganization to incorporate the observation date of all matching variables and outcomes.<sup>2</sup> Hence, we limit the set of potential partners for every treated unit to those observed just at the individual matching date, e. g. the treatment start. Then the matching algorithm selects one or more statistical twins among these pre-selected firms.

The second step is matching. Basically, each matching process that allows for (at least partial) exact matching is suitable. This exact matching option is required to consider the time information from the pre-selection process.<sup>3</sup> As a novelty, we add a nearest neighbor matching based on a combined statistical distance function.<sup>4</sup> This distance function follows an idea of Kaufmann and Pape (1996) and can be described as the weighted average of scale-specific distance functions. It belongs to the group of linear-homogeneous aggregations (Opitz 1980). For our analysis, we combine the mean absolute difference for continuous and the generalized matching coefficient for categorical variables.

When combining scale-specific distance functions, the functions usually have to be normalized and transformed (Diday and Simon 1976). In our case, the differences in the continuous variables are normalized by the maximal observed differences of the respective variables. The similarity information of the generalized matching coefficient is transformed into a distance measure. Weighting the functions by the respective number of variables, the distance function for a treated firm  $i$  and a non-treated firm  $j$  can be described as follows:

$$Dist_{ij} = \frac{1}{N} [N_m \cdot AD_{ij} + N_n \cdot (1 - GMC_{ij})]. \quad (1)$$

The terms  $Dist_{ij}$ ,  $AD_{ij}$  and  $GMC_{ij}$  denote the aggregated distance function and the scale-specific distances,  $N$  is the total number of variables with  $N = N_m + N_n$ , where  $N_m$  is the number of continuous variables and  $N_n$  that of the categorical ones.

<sup>2</sup>For example, when a firm receives investment subsidies in January 2007, we consider its characteristics in this month and will assign a firm which has similar characteristics in January 2007.

<sup>3</sup>Potential matching procedures are Mahalanobis metric matching (i. e. the *nnmatch* option of the Stata command *teffects*) and the Coarsened exact matching procedure of Blackwell et al. (2009). The second option is currently implemented in the *flexpaneldid* command.

<sup>4</sup>In a previous simulation study, statistical distance functions proved to be superior to Propensity Score matching in small samples, see Dettmann et al. (2011). Besides, King and Nielsen (2016) and King and Zeng (2006) doubt the Propensity Score to be suitable for empirical studies, when the score itself has to be estimated.

The mean difference of the continuous variables is calculated using the normalized absolute difference:

$$AD_{n,ij} = \frac{1}{N_m} \sum_{n=1}^{N_m} \frac{|x_{ni} - x_{nj}|}{diff_{max}(x_n)}$$

where  $||$  denotes absolute values, and  $diff_{max}(x_n)$  is the maximum observed difference of variable  $x_n$ .

The generalized matching coefficient  $GMC_{ij}$  can be defined as the share of covariates with equal values in all categorical variables:

$$GMC_{ij} = \frac{1}{N_n} \sum_{n=1}^{N_n} Q(x_{ni}, x_{nj}) \quad \text{with} \quad Q(x_{ni}, x_{nj}) = \begin{cases} 1 & \text{if } x_{ni} = x_{nj} \\ 0 & \text{else.} \end{cases}$$

As can be observed from the equation, using the generalized matching coefficient allows for different numbers of possible values in the covariates. The variables with coincident values are equally weighted irrespective of the number of possible values.

Based on this matching process, the average treatment effect for the treated  $ATT$  is estimated. Within the framework of the conditional DID model, usually the mean outcome developments in the treated and the control group are compared. Different from the standard model, the flexible conditional DID compares the mean of the individual differences in outcome development between the treated firms  $i$  and their respective controls  $j$ .

$$ATT = \frac{1}{I} \sum_{i=1}^I (Y_{i,t_{0i}+\beta_i} - Y_{i,t_{0i}}) - (Y_{j,t_{0i}+\beta_i} - Y_{j,t_{0i}}). \quad (2)$$

As can be observed from equation 2, we include individual treatment start dates, denoted by index  $t_{0i}$ , and a flexible number of time units, e. g. months,  $t_{0i} + \beta_i$ , reflecting the individual duration from treatment start to outcome observation.  $Y$  denotes the outcome. Due to heterogeneous treatment durations, the observed periods can be heterogeneous among the treated individuals.

## 5 Two detailed application examples

We implement the above described flexible conditional DID approach in Stata. The new command is called *flexpaneldid*. The command allows users to estimate the average treatment effect for the treated with different combinations of matching and DID approaches. It combines an extensive data rearrangement to consider time information for treatment effect estimation with matching, difference-in-differences approaches and some quality tests.

In this section, we present two comprehensive examples to illustrate the use and the resulting outputs of the available options. In order to reproduce the data characteristics described in section 2 (see figure 1), we need a panel data set with more than two observation times, individual treatment starts and treatment durations, possible

multiple treatments and differently scaled variables characterizing the observed units. We start with a publicly available data set, the 'patent' data provided by Wooldridge (2010)<sup>5</sup> and generate some additional variables: a fictive treatment variable and some categorically scaled variables by manipulating existing categorical variables and generating categorical variables from continuous ones. The result is a small example data set that exhibits a similar structure to the above described one and is provided at our home-page (<https://cloud.iwh-halle.de/index.php/s/flexpaneldid>).<sup>6</sup> Under the same link, one can find the file `flexpaneldid.ado` and the respective help file `flexpaneldid.sthlp`. Before starting work with the command, one also has to install the Stata ado-files `psmatch2`, `pstest` and `cem`, which are used by `flexpaneldid`.

The example panel data set consists of yearly information on uniquely identified firms that are characterized by categorical and continuously scaled variables. The observation period is 1972 to 1981. The treatment can occur within the first five years of observation, the start and duration can vary among the treated firms. Also multiple treatments are possible. For the treated firms, the treatment variable equals one during the whole treatment period and zero before and after the treatment; for non-treated firms it is always zero. In case of multiple treatments, the treatment variable equals one from the start of the first treatment until the end of the last treatment. Suppose now we want to estimate the causal effect of a certain treatment on the number of patents at the firm level. `flexpaneldid` offers different opportunities to do this.

In the first example we present an estimation based on the flexible conditional DID approach described in section 4 and two alternative estimations based on the DID model. We will include an exact matching variable and the number of patents two years before treatment as an additional matching variable. The matching time should be one year before treatment, and we want to observe the change in the number of patents from treatment start to three years after the treatment has started. Furthermore, we want to get displayed the results of the quality tests for matching.

When using `flexpaneldid`, we always have to define the dependent variable, in our case `patents`. Compulsory details are also the individual identification of the observed units, `id(cusip)`, the definition of the treatment variable, `treatment(treatment)` and the variable identifying the time units in the panel, in our sample `time(year)`. Besides yearly data, also e.g. monthly or quarterly data can be the basis for estimations with `flexpaneldid`.

Because we want to estimate the treatment effect according to the matching approach described in section 4, we select option `statmatching` and distinguish between continuous and categorical matching variables: `statmatching(con(employ stockpr rnd sales) cat(pats_cat rndstck_cat))`. Next, we define characteristics that are used for exact matching, in our example `matchvarsexact(sic_cat)`. Additionally, we consider the pre-treatment outcome, i. e. the number of patents two years prior to the treatment start for matching: `outcome2dev(-2)`.<sup>7</sup> We also define the matching time

<sup>5</sup>The data set can be found at <http://www.stata.com/data/jwooldridge/eacsap/patent.dta>.

<sup>6</sup>The data generation is presented in the appendix.

<sup>7</sup>It is important to note that the observation time of the pre-treatment outcome is defined relative

in relation to the treatment start. For the example we choose `matchtimerel(-1)`, meaning that the (individual) matching time for treated firms is one year before the individual treatment starts.<sup>8</sup> In a similar way, we define the period of outcome development that should be compared between treated and controls. The starting point of the observed outcome development coincides to the start of the treatment. We can decide if the end point of the observed development should be related to the treatment start or the treatment end. For the example we choose `outcometimerelstart(3)`, which means that we compare the outcome development from the treatment start to three years later.<sup>9</sup> Finally, we want to test the quality of the the matching procedure and the results. Summing up all the information, the command looks like this:

```
flexpanelidid patents, id(cusip) treatment(treatment) time(year)
  statmatching(con(employ stckpr rnd sales) cat(pats_cat rndstck_cat))
  outcometimerelstart(3) matchvarsexact(sic_cat) matchtimerel(-1)
  outcomedev(-2) test
```

After having submitted all the required information, we get the following output for the first example. The first part consists of a summary of all our submitted information and selected options:

```
*****
***** flexpanelidid *****
*****

outcome:          patents
id:               cusip
treatment:        treatment
time:             year
exact:            sic_cat
match_time:       -1
outcome_time_start: 3
outcome_time_end: .
outcome_dev:      -2
cemmatching:
statmatch_con:    employ stckpr rnd sales
statmatch_cat:    pats_cat rndstck_cat
test:             test
```

The first step of the estimation is the data preprocessing. For each of the preprocessing steps, we get a message when it is finished. This is because this procedure may take a while when the data set becomes large (especially, when we have many treated units and/or many treatment times).

---

to the treatment start and refers to the variable indicating the time in the panel, in our example `time(year)`.

<sup>8</sup>Also the relative matching time refers to the variable indicating the time in the panel, `time(year)`.

<sup>9</sup>The dimension of the parameter in brackets again depends on the dimension of time that is defined previously.

```

*****
***** Preprocessing *****
*****
Preprocessing 1 of 61 treated.
Preprocessing 2 of 61 treated.
Preprocessing 3 of 61 treated.
...
Preprocessing 61 of 61 treated.

```

After preprocessing, the matching procedure is executed – taking into account the individual (internal generated) indication of possible partners for each treated unit. We get a short information on the number of the matched treated units and the number of the control units used for matching. In our example, for 47 out of the 61 treated units, the matching procedure finds a partner. In the matching process, 36 non-treated units are used as partners. That means, some of the non-treated units are used as partner for more than one treated, which is typical for the implemented nearest neighbor matching with replacement.

```

*****
***** flexpaneldid - Matching Summary *****
*****

```

|                | NT  | T  |
|----------------|-----|----|
| All            | 165 | 61 |
| Matched sample | 36  | 47 |

To assess the comparability of the matched groups, the results of different tests are presented next. For every matching procedure, we get the tests provided in the Stata command *pstest* by [Leuven and Sianesi \(2003\)](#) and a graphical proxy for the balance of the variable distributions in the treated and the control group. The tests are made at matching time, in our example one year before the treatment starts. (Accordingly, the names of the tested variables have the suffix `_pre1`.)

For each of the matching variables, we find the means in the treated and in the control group, a measure for the standardized percentage difference – or bias – between the means in both groups, and a test if the means in the control group equal the ones in the treated group. Additionally we get an information on the similarity of the variances in the treated and the control group. We would conclude that the means of all the matching variables are balanced, but the variances of `stckpr_pre1`, `sales_pre1` and `outcome_dev` are not.

```
*****
***** ps-test *****
*****
```

| Variable         | Mean    |         | %bias | t-test |       | V(T)/<br>V(C) |
|------------------|---------|---------|-------|--------|-------|---------------|
|                  | Treated | Control |       | t      | p> t  |               |
| employ_pre1      | 26.197  | 27.207  | -1.3  | -0.06  | 0.951 | 0.64          |
| stckpr_pre1      | 25.779  | 19.272  | 19.8  | 0.96   | 0.339 | 5.47*         |
| rnd_pre1         | 38.116  | 44.949  | -4.2  | -0.20  | 0.840 | 0.82          |
| sales_pre1       | 1201    | 1474.4  | -6.2  | -0.30  | 0.764 | 0.34*         |
| pts_cat_pre1     | 1.9362  | 1.8511  | 6.4   | 0.31   | 0.757 | 1.04          |
| rndstck_cat_pre1 | 3.4255  | 3.1702  | 10.5  | 0.51   | 0.613 | 1.09          |
| outcome_dev      | 50.574  | 21.149  | 26.1  | 1.26   | 0.210 | 9.97*         |

\* if variance ratio outside [0.56; 1.80]

| Ps R2 | LR chi2 | p>chi2 | MeanBias | MedBias | B     | R     | %Var |
|-------|---------|--------|----------|---------|-------|-------|------|
| 0.045 | 5.85    | 0.557  | 10.6     | 6.4     | 45.6* | 2.19* | 43   |

\* if B>25%, R outside [0.5; 2]

The quantile-quantile plots of the matching variables give a graphical impression on the comparability of the matched groups. They compare the distributions in both groups by means of the plotted quantiles. The 45°-line represents identical distributions. From the figure we see small deviations from the 45°-line for all continuous variables, especially the distribution of `outcome_level` (which is defined as 'patents two years before treatment' in our example) is not similar. The plots for the categorical variables `pts_cat` and `rndstck_cat` should be interpreted as rather rough guidelines for similarity.

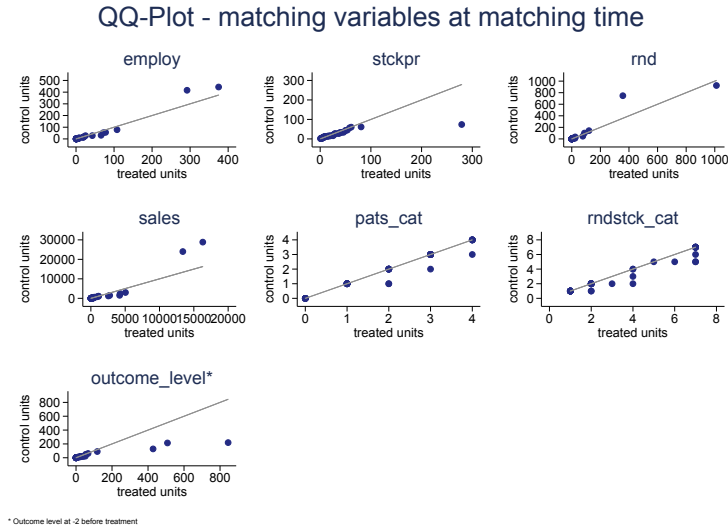


Figure 2: Quantile-quantile plots of the matched samples, example 1

In case of matching based on the statistical distance function, additional scale-specific test statistics for the included variables are displayed. For all continuous variables, the results of a Kolmogorov-Smirnov test are presented, for the categorical variables, the results of chi-square tests are available. For reasons of space, we only present here the tests for the pre-treatment-outcome and the variables **employ** (continuous) and **pats\_cat** (categorical) as an example. In all three cases, the tests indicate no significant differences in the variable distributions between the treated and the control group. In case of the KS test, the corrected p-values of 0.595 for the pre-treatment-outcome and 0.28 for employment (**employ\_pre1**) tell us that the variable distributions between the treated and the control group are not significantly different. In case of the  $\chi^2$  test, the p-value of 0.99 – and also a look at the respective number of observations in the categories – also indicate balanced samples.

```
*****
***** KS-Test *****
*****
ksmirnov outcome_dev , by(treated)
Two-sample Kolmogorov-Smirnov test for equality of distribution functions
+-----+-----+-----+-----+
Smaller group      D          P-value   Corrected
+-----+-----+-----+-----+
0:                  0.1489    0.353
1:                  0.0000    1.000
Combined K-S:      0.1489    0.674      0.595
Note: Ties exist in combined dataset;
      there are 34 unique values out of 94 observations.
```

```

ksmirnov employ_pre1 , by(treated)
Two-sample Kolmogorov-Smirnov test for equality of distribution functions
+-----+-----+-----+-----+
Smaller group      D          P-value   Corrected
+-----+-----+-----+-----+
0:                  0.1915    0.178
1:                 -0.0638    0.826
Combined K-S:       0.1915    0.355        0.280
Note: Ties exist in combined dataset;
      there are 89 unique values out of 94 observations.

...
(Output for stckpr_pre1, rnd_pre1 ans sales_pre1 omitted.)

*****
***** Chi2-Test *****
*****
tabulate pats_cat_pre1 treated, chi2
+-----+-----+-----+-----+
_pre1      |      treated      |      |      |
pats_cat    |      0      1      |      |      |
+-----+-----+-----+-----+
0           |      8      8      |      |      |
1           |     14     12      |      |      |
2           |      8      9      |      |      |
3           |     11     11      |      |      |
4           |      6      7      |      |      |
+-----+-----+-----+-----+
Total       |     47     47      |      |      |
Pearson chi2(4) = 0.2896   Pr = 0.990

...
(Output for rndstck_cat_pre1 omitted.)

```

In the last step, the results of different DID approaches discussed in the literature are presented. First, the result of the flexible conditional DID – a comparison of individual differences between treated and their controls (according to equation 2) – is displayed.<sup>10</sup>

One important hint: as the data base contains only an arbitrarily defined treatment, the following outputs of the DID approaches have no meaningful interpretation. We present them for illustrative purpose only.

```

*****
***** Conditional Diff-in-Diff *****
*****
Outcome      Obs      |      Mean Diff      |      Diff      |      t-test      |
                   |      T      NT      |      |      t      p-value      |
+-----+-----+-----+-----+
patents      94      |     -10.0213     -6.4710      |     -3.5503      |     0.5100     0.6113      |
* Consistent bias-corrected estimator as proposed in Abadie & Imbens (2006,2011).

```

<sup>10</sup>In order to get consistent estimates of the treatment effect, we implement the bias-correction of Abadie and Imbens (2006, 2011) at the beginning and the end of the outcome observation.



In our example, we observe a negative development of the number of patents for the period from the start of the treatment until three years afterward, both for the treated ( $-10.0213$ ) and the controls ( $-6.4710$ ). The mean difference in the patents development between treated and controls is  $-3.5503$ , the p-value of the t-tests ( $0.6113$ ) indicates that this difference is not significant.

The second output is the mean treatment effect estimation within a fixed-effects DID model for the defined 3-years-period starting at individual treatment start. For the estimation of the treatment effect we use the *xtreg* command in Stata, including a constant and time dummies for the previously determined period, but no further covariates. The included time dummies are defined according to the dimension of the *time* identifier. In our case, the time units are years, and *flexpaneldid* defines year dummies for the regression. If we want to have more flexibility in the specification of the fixed-effects DID model, we can additionally apply the *xtreg* command to the current data (i.e. a panel data set for the matched sample that is kept in the memory) as soon as the *flexpaneldid* procedure is finished.<sup>11</sup>

Also from the results of the second approach we would conclude that the treatment has no significant effect on the development of the number of patents: The interesting coefficient here is  $1.083$ , indicating a positive effect, but the corresponding p-value of  $0.655$  indicates that this effect is not significant.

---

<sup>11</sup>For a later version of the *flexpaneldid* command we plan to include more options for the specification of the fixed-effects DID model.

```
***** Fixed Effects Diff-in-Diff - mean treatment effect *****
*****
```

xtreg patents i.treatment##i.post\_treat\_dummy i.year if post\_treat\_dummy\_rel\_time <= 3,  
fe vce(cluster panel\_id)

Fixed-effects (within) regression                      Number of obs        =        607  
Group variable: panel\_id                              Number of groups    =        83  
R-sq:    within    = 0.0744                            Obs per group: min =        6  
          between   = 0.0200                                        avg    =        7.3  
          overall    = 0.0000                                        max    =        9  
  
   F(11,82)                    =        2.10  
corr(u\_i, Xb)    = -0.0330                            Prob > F                =        0.0289  
  
   (Std. Err. adjusted for 83 clusters in panel\_id)

| patents                    | Coef.     | Robust<br>Std. Err.               | t     | P> t  | [95% Conf. Interval] |           |
|----------------------------|-----------|-----------------------------------|-------|-------|----------------------|-----------|
| 1.treatment                | -.655248  | 1.464566                          | -0.45 | 0.656 | -3.568736            | 2.25824   |
| 1.post_treat_dummy         | -2.927036 | 5.006574                          | -0.58 | 0.560 | -12.88671            | 7.032635  |
| treatment#post_treat_dummy |           |                                   |       |       |                      |           |
| 1 1                        | 1.083032  | 2.417029                          | 0.45  | 0.655 | -3.725208            | 5.891272  |
| year                       |           |                                   |       |       |                      |           |
| 73                         | .7951807  | 1.398711                          | 0.57  | 0.571 | -1.987302            | 3.577663  |
| 74                         | 2.130467  | 1.906603                          | 1.12  | 0.267 | -1.662374            | 5.923307  |
| 75                         | 5.816296  | 4.904775                          | 1.19  | 0.239 | -3.940864            | 15.57345  |
| 76                         | 3.201816  | 5.501172                          | 0.58  | 0.562 | -7.741768            | 14.1454   |
| 77                         | 3.621212  | 6.312409                          | 0.57  | 0.568 | -8.936179            | 16.1786   |
| 78                         | 1.698973  | 5.894385                          | 0.29  | 0.774 | -10.02684            | 13.42478  |
| 79                         | -.2114756 | 4.742809                          | -0.04 | 0.965 | -9.646434            | 9.223482  |
| 80                         | -18.58481 | 8.305863                          | -2.24 | 0.028 | -35.10781            | -2.061801 |
| _cons                      | 37.12422  | 1.642992                          | 22.60 | 0.000 | 33.85578             | 40.39265  |
| sigma_u                    | 110.99724 |                                   |       |       |                      |           |
| sigma_e                    | 16.979395 |                                   |       |       |                      |           |
| rho                        | .97713482 | (fraction of variance due to u_i) |       |       |                      |           |

The third DID output displays the estimation result of yearly treatment effects within the specified 3-years-period. We apply the same Stata command and the same model – with a different treatment effect definition (`post_treat_dummy_rel_time`)<sup>12</sup>. Also the estimated yearly treatment effects are not significant. Here, the interesting coefficients of 3.365, -2.398 and 1.389 point to a positive effect in the first year and negative effects in the two following years, but the p-values (0.233, 0.419, 0.758) again indicate no significant effects.

<sup>12</sup>Again, if we want more flexibility, we can apply the *xtreg* command to the current data.

```
*****
***** Fixed Effects Diff-in-Diff - dynamic treatment effect *****
*****
```

xtreg patents i.treatment##i.post\_treat\_dummy\_rel\_time i.year if post\_treat\_dummy\_rel\_time <= 3,  
fe vce(cluster panel\_id)

Fixed-effects (within) regression                      Number of obs        =        607  
Group variable: panel\_id                              Number of groups    =        83  
R-sq:    within    = 0.0784                            Obs per group: min =        6  
          between   = 0.0263                                        avg    =        7.3  
          overall    = 0.0002                                        max    =        9  
  
   F(15,82)                    =        2.01  
corr(u\_i, Xb) = -0.0570                                Prob > F                    =        0.0236  
  
   (Std. Err. adjusted for 83 clusters in panel\_id)

| patents                             | Coef.     | Robust Std. Err.                  | t     | P> t  | [95% Conf. Interval] |          |
|-------------------------------------|-----------|-----------------------------------|-------|-------|----------------------|----------|
| 1.treatment                         | .2241196  | 1.684504                          | 0.13  | 0.894 | -3.126895            | 3.575134 |
| post_treat_dummy_rel_time           |           |                                   |       |       |                      |          |
| 1                                   | -3.773297 | 5.074651                          | -0.74 | 0.459 | -13.86839            | 6.321801 |
| 2                                   | .2068019  | 5.618524                          | 0.04  | 0.971 | -10.97023            | 11.38384 |
| 3                                   | 1.389394  | 5.817341                          | 0.24  | 0.812 | -10.18315            | 12.96194 |
| treatment#post_treat_dummy_rel_time |           |                                   |       |       |                      |          |
| 1 1                                 | 3.364883  | 2.801935                          | 1.20  | 0.233 | -2.209059            | 8.938824 |
| 1 2                                 | -2.398169 | 2.949858                          | -0.81 | 0.419 | -8.266376            | 3.470037 |
| 1 3                                 | -1.02086  | 3.304151                          | -0.31 | 0.758 | -7.593869            | 5.55215  |
| year                                |           |                                   |       |       |                      |          |
| 73                                  | .7951807  | 1.403437                          | 0.57  | 0.573 | -1.996702            | 3.587063 |
| 74                                  | 1.971545  | 1.901826                          | 1.04  | 0.303 | -1.811793            | 5.754882 |
| 75                                  | 5.718231  | 4.89971                           | 1.17  | 0.247 | -4.028853            | 15.46531 |
| 76                                  | 2.029779  | 5.645405                          | 0.36  | 0.720 | -9.20073             | 13.26029 |
| 77                                  | 1.524809  | 6.564609                          | 0.23  | 0.817 | -11.53429            | 14.58391 |
| 78                                  | -.7137019 | 6.487056                          | -0.11 | 0.913 | 13.61852             | 12.19112 |
| 79                                  | -4.518972 | 5.736604                          | -0.79 | 0.433 | -15.93091            | 6.892961 |
| 80                                  | -23.83004 | 10.69908                          | -2.23 | 0.029 | -45.11392            | -2.54616 |
| _cons                               | 37.22324  | 1.658672                          | 22.44 | 0.000 | 33.92361             | 40.52286 |
| sigma_u                             | 111.11762 |                                   |       |       |                      |          |
| sigma_e                             | 17.008644 |                                   |       |       |                      |          |
| rho                                 | .97710633 | (fraction of variance due to u_i) |       |       |                      |          |

For the second example we want to estimate the treatment effect for the period from treatment start to two years after the treatment is finished. The matching time should be the treatment start. Again want to consider the pre-treatment-outcome, therefore we include the development of patents from two years before treatment until the start of the treatment. We apply *CEM* matching<sup>13</sup> and include an exact matching variable. Furthermore, we want to get displayed the results of the quality tests. For

<sup>13</sup>This approach refers to the *Coarsened Exact Matching* of Blackwell et al. (2009).

the treatment effect estimation we will again get the results of three DID approaches, the flexible conditional DID approach and two estimations within the framework of the DID model.

Besides the compulsory inputs (`depvar`, `id`, `treatment`, `time`), we now select `cemmatching`. In this option, we can manually choose strata – either as a fixed number of equally spaced strata or by explicitly defining the cutpoints, e. g. `cemmatching(employ (#5) stckpr (100 200 300) rnd pats_cat (#0) rndstck_cat (#0))`. If no stratification is defined (as is the case for the outcome development, `rnd` and `sales`), the default number of strata (13) is used. For the categorical variables, we must retain the number of values: `pats_cat (#0) rndstck_cat (#0)`. See Blackwell et al. (2009) for a more detailed description. The following command lines sum up all our selections:

```
flexpanelid patents, id(cusip) treatment(treatment) time(year)
cemmatching(employ (#5) stckpr (100 200 300) rnd pats_cat (#0) rndstck_cat (#0))
outcometimerelend(2) matchvarsexact(sic_cat) outcomedev(-2 0) test
```

The first part of the output again gives a summary of the submitted details. The preprocessing produces the same output as in the first example.<sup>14</sup>

```
*****
***** flexpanelid *****
*****
```

---

|                     |   |
|---------------------|---|
| outcome:            | patents   |
| id:                 | cusip   |
| treatment:          | treatment   |
| time:               | year  |
| exact:              | sic_cat   |
| match_time:         | 0   |
| outcome_time_start: | .   |
| outcome_time_end:   | 2   |
| outcome_dev:        | -2 0  |
| cemmatching:        | employ(#5) stckpr(100 200 300) rnd sales pats_cat(#0) rndstck_cat(#0) |
| statmatch_con:      |   |
| statmatch_cat:      |   |
| test:               | test  |

---

The matching output for `cemmatching` differs from the `statmatching` output in that some details of the matching procedure are displayed: We find the stratification cutpoints for each variable, an alternative matching summary (containing the number of matched and unmatched treated and control units (in terms of observations, not uniquely identified units)) and alternative quality checks for the matching results. At the end of the displayed output we find again the summary on the number of matched treated and controls (in terms of uniquely identified units like in example one). If we compare both outputs, we see that the matched sample is remarkably smaller in example two – may be due to the longer observation period, the different matching process or the slightly different matching variables.

<sup>14</sup>For reasons of space we omit the display of the preprocessing output.

```

*****
***** Matching: CEM *****
*****

Cutpoints:
selection_group: (user)
0
employ_post0: (user)
1

1 .0849999785
2 119.886748
3 239.6884961
4 359.4902441
5 479.2919922

stckpr: (user)
1

1 100
2 200
3 300

rnd_post0: (sturges)
1

1 0
2 97.56640625
3 195.1328125
4 292.6992188
5 390.265625
6 487.8320313
7 585.3984375
8 682.9648438
9 780.53125
10 878.0976563
11 975.6640625
12 1073.230469
13 1170.796875

sales_post0: (sturges)
1

1 1.221999168
2 3154.578499
3 6307.934999
4 9461.291499
5 12614.648
6 15768.0045
7 18921.361
8 22074.7175
9 25228.074
10 28381.4305
11 31534.787
12 34688.1435
13 37841.5

pats_cat: (user)
0
rndstck_cat: (user)

```

```

0
outcome_dev: (sturges)
1

```

|    |              |
|----|--------------|
| 1  | -151         |
| 2  | -134.1666667 |
| 3  | -117.3333333 |
| 4  | -100.5       |
| 5  | -83.6666667  |
| 6  | -66.8333333  |
| 7  | -50          |
| 8  | -33.1666667  |
| 9  | -16.3333333  |
| 10 | .5           |
| 11 | 17.3333333   |
| 12 | 34.1666667   |
| 13 | 51           |

Matching Summary:

```

-----
Number of strata: 1283
Number of matched strata: 26

```

|           | 0    | 1  |
|-----------|------|----|
| All       | 2129 | 46 |
| Matched   | 26   | 26 |
| Unmatched | 2103 | 20 |

Multivariate L1 distance: .53846154

Univariate imbalance:

|                 | L1     | mean    | min     | 25%    | 50%     | 75%     | max     |
|-----------------|--------|---------|---------|--------|---------|---------|---------|
| selection_group | 0      | 0       | 0       | 0      | 0       | 0       | 0       |
| employ_post0    | .11538 | -3.723  | -.105   | .001   | .685    | -2.046  | -33.2   |
| stckpr          | .26923 | -2.2742 | 2       | 1.375  | -.625   | -7.375  | -19.125 |
| rnd_post0       | .11538 | -4.2496 | -.02601 | .01517 | -.38406 | .54775  | -25.331 |
| sales_post0     | .07692 | -115.84 | -6.601  | -7.881 | 2.291   | -6.9106 | -126.48 |
| pats_cat        | 0      | 0       | 0       | 0      | 0       | 0       | 0       |
| rndstck_cat     | 0      | 0       | 0       | 0      | 0       | 0       | 0       |
| outcome_dev     | .07692 | .46154  | 2       | 0      | 0       | 0       | 3       |

```

*****
***** flexpanelidid - Matching Summary *****
*****

```

|                | NT  | T  |
|----------------|-----|----|
| All            | 165 | 61 |
| Matched sample | 24  | 26 |

Because we selected the `test` option, also the information of *pstest* and quantile-quantile plots are displayed.

```
*****
***** ps-test *****
*****
```

| Variable          | Mean    |         |       | t-test |       | V(T)/<br>V(C) |
|-------------------|---------|---------|-------|--------|-------|---------------|
|                   | Treated | Control | %bias | t      | p> t  |               |
| employ_post0      | 6.4512  | 10.174  | -19.5 | -0.70  | 0.486 | 0.41*         |
| stckpr_post0      | 14.279  | 16.553  | -15.8 | -0.57  | 0.571 | 0.60          |
| rnd_post0         | 3.6857  | 7.9353  | -31.1 | -1.12  | 0.267 | 0.20*         |
| sales_post0       | 276.05  | 391.89  | -17.0 | -0.61  | 0.542 | 0.62          |
| pats_cat_post0    | 1.3846  | 1.3846  | 0.0   | 0.00   | 1.000 | 1.00          |
| rndstck_cat_post0 | 2.5769  | 2.5769  | 0.0   | 0.00   | 1.000 | 1.00          |
| outcome_dev       | -.34615 | -.80769 | 8.1   | 0.29   | 0.771 | 0.99          |

\* if variance ratio outside [0.45; 2.23]

| Ps R2 | LR chi2 | p>chi2 | MeanBias | MedBias | B     | R     | %Var |
|-------|---------|--------|----------|---------|-------|-------|------|
| 0.068 | 4.89    | 0.673  | 13.1     | 15.8    | 59.9* | 0.41* | 29   |

\* if B>25%, R outside [0.5; 2]

### QQ-Plot - matching variables at matching time

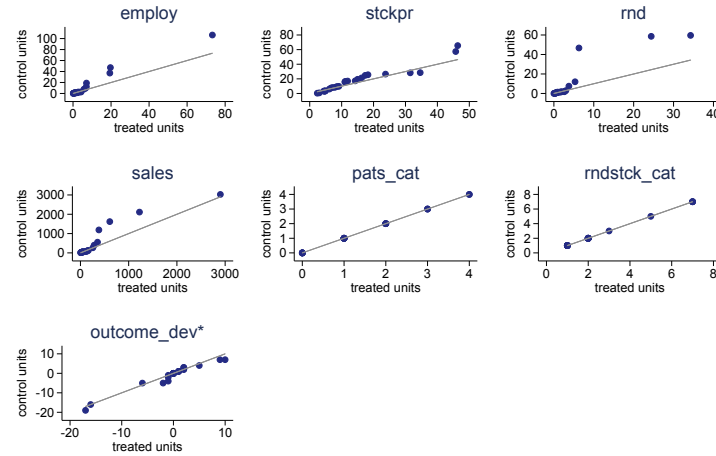


Figure 3: Quantile-quantile plots of the matched samples, example 2

Even though the data contain no 'real' treatment information and some arbitrarily defined variables, also a comparison of the test results of both examples show that

different matching approaches and estimation details produce different control groups for the treated units – that will result in different estimated treatment effects in practice. Thus, the different matching procedures included in the `flexpaneldid` command should/can be used as a robustness check.

Also for the `cemmatching`, we get the three different estimates of the treatment effect for the treated. Because the treatment variable does not contain any 'real' information, an output comparison between the two examples would not have any informative value. For reasons of space, we therefore present only the first DID output.

```
*****
***** Conditional Diff-in-Diff *****
*****
```

| Outcome | Obs | Mean Diff |        | Diff    | t-test |         |
|---------|-----|-----------|--------|---------|--------|---------|
|         |     | T         | NT     |         | t      | p-value |
| patents | 52  | -1.8077   | 0.4272 | -2.2348 | 1.6741 | 0.1004  |

```
* Consistent bias-corrected estimator as proposed in Abadie & Imbens (2006,2011).
```

As expected, also the estimation output differs strongly from the results in the first example.

## 6 The flexpaneldid command

Before running `flexpaneldid` one has to install or update the Stata ado-files `psmatch2`, `pstest` and `cem`, which are used by this command.

### 6.1 Syntax

```
flexpaneldid depvar, id(varname) treatment(varname) time(varname)
    (cemmatching(varname1 [(cutpoints1)] [varname2 [(cutpoints2)]...])|
    statmatching(con(varlist) cat(varlist))
    (outcometimerelstart(integer) | outcometimerelend(integer))
    [matchvarsexact(varlist) matchtimerel(integer)
    outcomedev(integer [integer]) test]
```

### 6.2 Description

*Flexpaneldid* is a tool for causal analysis of treatments with varying start dates and varying treatment durations within panel data with more than two observation times. The command estimates the average treatment effect for the treated with different combinations of matching and difference-in-differences approaches. *Flexpaneldid* combines an extensive data rearrangement to consider treatment time information in the effects estimation with different conditional DID approaches. It is written for Stata version 13.



The *flexpaneldid* command contains many relative time definitions. The following graph illustrates the relationships between treatment start and related time definitions and treatment end and related time.

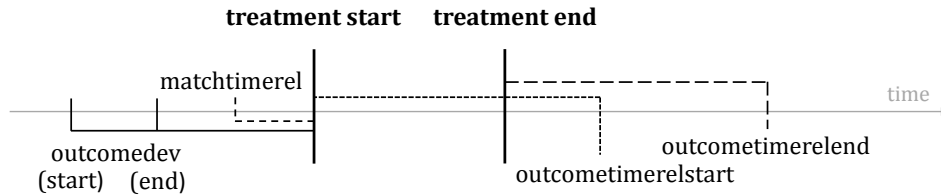


Figure 4: sketch of the relative time definitions in *flexpaneldid*

### 6.3 Inputs and Options

The following information are compulsory for *flexpaneldid* to work.

The *depvar* variable defines the analyzed outcome; input must be numerical.

The *id(varname)* uniquely identifies objects in the panel dataset; the variable must be an integer or string.

The *treatment(varname)* contains the variable defining the treatment; input must be in 0-1 format.

IMPORTANT NOTE: The variable must equal to one for the whole treatment phase. In case of repeated treatments for one unit (identified by a unique id), the repeated treatments are handled as one treatment phase.

The *time(varname)* identifies the time information in the panel; input must be an integer indicating an absolute time, e. g. year, month, quarter.

IMPORTANT NOTE: If the data contain only information in 'date' format, this information must be converted into an integer.

Next, the selection of one of the two options for the matching procedure and the definition of the matching variables is required.

*cemmatching(varname1 [(cutpoints1)] [varname2 [(cutpoints2)] ...])* indicates that the *Coarsened Exact Matching* of Blackwell et al. (2009)<sup>15</sup> will be executed. Including cutpoints for the matching variables is possible. Like in the *cem* command, either formats *[(#integer)]* or *(numlist)* are possible; see Blackwell et al. (2009) for more details.

*statmatching(con(varlist) cat(varlist))* indicates that a nearest neighbor matching with replacement based on the statistical distance function according to equation 1 is

<sup>15</sup>To ensure reproducible results, a 'hard coded' seed value is set in the ado-file before running *cem*.

used for matching. This option refers to `psmatch2`, `neighbor(1)` `pscore(statistical distance)`. The variable names included in `con(varlist)` indicate the continuous matching variables, which must be numerical variables; `cat(varlist)` contains the categorical variables, which must be integers.

`flexpaneldid` enables the user to define the period of outcome development that should be compared between treated and controls. (The starting point of the observed outcome development coincides to the start of the treatment.) One of both options must be selected:

`outcometimerelstart(integer)` is a relative time specification that defines the end of the outcome development in relation to the treatment start; default = 0. In case of repeated treatments, the relative time refers to the start of the first treatment.

IMPORTANT NOTE: The dimension of the parameter in brackets depends on the dimension of time that is defined for `time`.

For example, `outcometimerelstart(3)` means we observe the outcome development from the individual treatment start to three years after the start of the treatment, if the dimension of the time variable is years.

`outcometimerelend(integer)` is a relative time specification that defines the end of the outcome development in relation to the treatment end; default = 0. In case of repeated treatments, the relative time refers to the end of the last treatment.

IMPORTANT NOTE: The dimension of the parameter in brackets depends on the dimension of time that is defined for `time`.

For example, `outcometimerelend(2)` means that we compare the outcome development from the treatment start to two years after the treatment is finished, if the dimension of the time variable is years.

All the following options are 'really optional'. None of the information are required for `flexpaneldid` to work.

`matchvarsexact(varlist)` indicates the variables that are used for exact matching; this option is independent from the chosen matching algorithm.

`matchtimerel(integer)` is a relative time specification (in relation to the treatment start) that defines the time of matching; default = 0 (if no matching time is defined). In this case, `flexpaneldid` uses variable values observed at the treatment start.

IMPORTANT NOTE: The dimension of the parameter in brackets depends on the dimension of time that is defined for `time`.

For example, `matchtimerel(-1)` means that the matching process is conducted one year before the treatment starts, if the dimension of the time variable is years.

`flexpaneldid` provides two options for considering the pre-treatment outcome in the matching process. `outcomedev(integer)` selects the level value of the outcome at a time defined in relation to the treatment start. `outcomedev(integer integer)` defines the outcome development, the two integers give the start and the end of the development in relation to the treatment start.

IMPORTANT NOTES: The dimension of the parameter(s) in brackets depend on

the dimension of time that is defined for `time`. Both parameters are required to be *integer*  $\leq 0$ .

For example, `outcomedev(-3 -1)` means the outcome development from three to one year(s) before the individual treatment starts, if the dimension of the time variable is years. `outcomedev(-3)` considers the outcome level three years before treatment starts as additional matching variable.

Choosing the option `test` executes quality tests after matching. The tests conducted in `pstest` and *quantile-quantile plots* are presented. Further test are presented depending on the matching process that is selected: For `cemmatching` additionally the overall imbalance measure  $L_1$  and univariate imbalance measures described in Blackwell et al. (2009) are displayed. For `statmatching`, *KS-tests* for continuous variables and *chi-square tests* for the categorical variables are executed in addition.

One note regarding the specification of the DID models: As default, the fixed-effects DID models is specified including a constant and time dummies, but no further covariates. The included time dummies are defined according to the dimension of the `time` identifier. In case of yearly data, this is a common specification in empirical research. If more flexibility in the specification of the fixed-effects DID model is required, one can additionally apply the `xtreg` command to the current data (i.e. a panel data set for the matched sample that is kept in the memory) as soon as the *flexpaneldid* procedure is finished.

## 7 Further steps

The *flexpaneldid* command is a fully functional Stata tool for causal effects estimation within the framework of panel data. It provides a solution to deal with individual treatments that are typical especially for subsidy programs like investment or wage subsidies or R&D and research funding. This paper describes the theoretical background, the tool and gives comprehensive application examples.

Nevertheless, this paper represents work in progress. The very next step will be to improve the data preprocessing time. Further steps like including additional assignment procedures besides the implemented nearest neighbor matching with replacement (i.e. radius matching and ties) and implementing more flexibility for the did models (and others...) will follow.

## 8 References

- Abadie, A. (2005), ‘Semiparametric Difference-in-Differences Estimators’, *Review of Economic Studies* **72**(1), 1–19.
- Abadie, A. and Imbens, G. W. (2006), ‘Large Sample Properties of Matching Estimators for Average Treatment Effects’, *Econometrica* **74**(1), 235–267.
- Abadie, A. and Imbens, G. W. (2011), ‘Bias-Corrected Matching Estimators for Average Treatment Effects’, *Journal of Business and Economic Statistics* **29**(1), 1–11.
- Ashenfelter, O. (1978), ‘Estimating the Effect of Training Programs on Earnings’, *Review of Economics and Statistics* **60**(1), 47–57.
- Autor, D. H. (2003), ‘Outsourcing at Will: The Contribution of Unjust Dismissal Doctrine to the Growth of Employment Outsourcing.’, *Journal of Labor Economics* **21**(1), 1–42.
- Baltagi, B. H., ed. (2015), *The Oxford Handbook of Panel Data*, Oxford University Press.
- Bandick, R., Görg, H. and Karpaty, P. (2014), ‘Foreign Acquisitions, Domestic Multinationals, and R&D’, *The Scandinavian Journal of Economics* **116**(4), 1091–1115.
- Bergemann, A., Fitzenberger, B. and Speckesser, S. (2009), ‘Evaluating the dynamic employment effects of training programs in East Germany using conditional difference-in-differences’, *Journal of Applied Econometrics* **24**(5), 797–823.
- Bernini, C. and Pellegrini, G. (2011), ‘How are growth and productivity in private firms affected by public subsidy? Evidence from a regional policy’, *Regional Science and Urban Economics* **41**, 253–265.
- Blackwell, M., Iacus, s., King, G. and Porro, G. (2009), ‘cem: Coarsened exact matching in stata’, *The Stata Journal* **9**(4), 524–546.
- Blundell, R. and Costa Dias, M. (2009), ‘Alternative Approaches to Evaluation in Empirical Microeconomics’, *The Journal of Human Resources* **44**(3), 565–640.
- Bronzini, R. and de Blasio, G. (2006), ‘Evaluating the impact of investment incentives: The case of Italy’s Law 488/1992’, *Journal of Urban Economics* **60**, 327–349.
- Caliendo, M. and Künn, S. (2011), ‘Start-up subsidies for the unemployed: Long-term evidence and effect heterogeneity’, *Journal of Public Economics* **95**, 311–331.
- Dehejia, R. (2005), ‘Practical Propensity Score Matching: A Reply to Smith and Todd’, *Journal of Econometrics* **125**, 355–364.
- Dehejia, R. H. and Wahba, S. (1999), ‘Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs’, *Journal of the American Statistical Association* **94**(448), 1053–1062.

- Dehejia, R. H. and Wahba, S. (2002), ‘Propensity Score-Matching Methods for Nonexperimental Causal Studies’, *The Review of Economics and Statistics* **84**(1), 151–161.
- Dettmann, E., Becker, C. and Schmeißer, C. (2011), ‘Distance functions for matching in small samples’, *Computational Statistics and Data Analysis* **55**(5), 1942–1960.
- Diday, E. and Simon, J. (1976), Clustering Analysis, in K. S. Fu, ed., ‘Digital Pattern Recognition’, Springer-Verlag, Berlin, chapter 3, pp. 47–94.
- Freier, R., Schumann, M. and Siedler, T. (2015), ‘The earnings returns to graduating with honors — Evidence from law graduates’, *Labour Economics* **34**, 39–50.
- Greenaway, D., Gullstrand, J. and Kneller, R. (2005), ‘Exporting May Not Always Boost Firm Productivity’, *Review of World Economics* **141**(4), 561–582.
- Gustafsson, A., Stephan, A., Hallman, A. and Karlsson, N. (2016), ‘The ‘sugar rush’ from innovation subsidies: a robust political economy perspective’, *Empirica* **43**, 729–756.
- Hainmueller, J. and Xu, Y. (2013), ‘ebalance: a Stata package for entropy balancing’, *Journal of Statistical Software* **54**(7), 1–18.
- Ham, J. C., Swenson, C., İmrohoroglu, A. and Song, H. (2011), ‘Government programs can improve local labor markets: Evidence from State Enterprise Zones, Federal Empowerment Zones and Federal Enterprise Community’, *Journal of Public Economics* **95**, 779–797.
- Heckman, J. J., Ichimura, H., Smith, J. A. and Todd, P. E. (1998), ‘Characterizing Selection Bias Using Experimental Data’, *Econometrica* **66**(5), 1017–1098.
- Heckman, J. J., LaLonde, R. J. and Smith, J. A. (1999), The Economics and Econometrics of Active Labor Market Programs, in O. Ashenfelter and D. E. Card, eds, ‘Handbook of Labor Economics’, Vol. III, Elsevier Science B.V., Amsterdam, pp. 1865–2097.
- Heyman, F., Sjöholm, F. and Tingvall, P. G. (2007), ‘Is there really a foreign ownership wage premium? Evidence from matched employer–employee data’, *Journal of International Economics* **73**, 355–376.
- Ho, D. E., Imai, K., King, G. and Stuart, E. A. (2007), ‘Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference’, *Political Analysis* **15**, 199–236.
- Imbens, G. W. and Wooldridge, J. M. (2009), ‘Recent Developments in the Econometrics of Program Evaluation’, *Journal of Economic Literature* **47**(1), 5–86.
- Kaufmann, H. and Pape, H. (1996), Clusteranalyse, in L. Fahrmeir, A. Hamerle and G. Tutz, eds, ‘Multivariate statistische Verfahren’, 2 edn, Verlag de Gruyter, Berlin, pp. 437–536.
- King, G. and Nielsen, R. (2016), Why Propensity Scores Should Not Be Used for Matching, Working paper, Harvard University.

- King, G. and Zeng, L. (2006), ‘The Dangers of Extreme Counterfactuals’, *Political Analysis* **14**, 131–159.
- Leuven, E. and Sianesi, B. (2003), PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing, Technical report, University of Oslo. access: 2016-05-06.  
**URL:** <http://ideas.repec.org/c/boc/bocode/s432001.html>
- Marcus, J. and Siedler, T. (2015), ‘Reducing binge drinking? The effect of a ban on late-night off-premise alcohol sales on alcohol-related hospital stays in Germany’, *Journal of Public Economics* **123**, 55–77.
- Neumark, D. and Kolko, J. (2010), ‘Do enterprise zones create jobs? Evidence from California’s enterprise zone program’, *Journal of Urban Economics* **68**, 1–19.
- Opitz, O. (1980), *Numerische Taxonomie*, Fischer-Verlag, Stuttgart, New York.
- Pellegrini, G. and Centra, M. (2006), Growth and efficiency in subsidized firms, in ‘Workshop ’The Evaluation of Labour Market, Welfare and Firms Incentive Programmes’, Istituto Veneto di Scienze, Lettere ed Arti, Venezia.
- Smith, J. A. and Todd, P. E. (2005a), ‘Does matching overcome LaLonde’s critique of nonexperimental estimators?’, *Journal of Econometrics* **125**(1-2), 305–353.
- Smith, J. A. and Todd, P. E. (2005b), ‘Does matching overcome LaLonde’s critique of nonexperimental estimators? Rejoinder’, *Journal of Econometrics* **125**(1-2), 365–75.
- Wooldridge, J. M. (2010), *Econometric Analysis of Cross Section and Panel Data*, 2nd edn, MIT Press, Cambridge, Massachusetts.

## Appendix

### Generation of the example data set

We start with a publicly available data set with a similar structure to the one described in section 2 (see figure 1).

```
. use http://www.stata.com/data/jwooldridge/eacsap/patent.dta, clear
```

Based on this, we generate some additional variables. First, we add a fictive treatment variable that can occur within the first five years of the observation period. If treatment equals zero, this indicates a non-treated unit, *treatment* = 1 marks the duration of the fictive treatment of the treated units.

```
. set seed 13
. gen random = runiform()
. sort random
. gen treatment = 0
. replace treatment = 1 if random>=0.95 & year>=73 & year<=77
. sort cusip year
. replace treatment=1 if random<=0.5 & treatment[_n-1]==1 & year>=73 & year<=77
> & cusip[_n-1]==cusip
. lab var treatment "treatment in 73 to 77"
. drop random
. order cusip year treatment
```

Additionally, we generate some categorical variables, i. e. we manipulate some of the existing categorical variables and generate new categorical variables from continuous ones.

```
. by cusip: egen merger_cat=max(merger)
. gen sic_cat=2000 if sic>=2000 & sic<2300
. replace sic_cat=2300 if sic>=2300 & sic<2600
. replace sic_cat=2600 if sic>=2600 & sic<2900
. replace sic_cat=2900 if sic>=2900 & sic<3200
. replace sic_cat=3200 if sic>=3200 & sic<3500
. replace sic_cat=3500 if sic>=3500 & sic<3800
. replace sic_cat=3800 if sic>=3800
. lab var sic_cat "sector categories"
. gen pats_cat=0 if patentsg==0
. replace pats_cat=1 if patentsg>=1 & patentsg<=3
. replace pats_cat=2 if patentsg>=4 & patentsg<=9
. replace pats_cat=3 if patentsg>=10 & patentsg<=50
. replace pats_cat=4 if patentsg>=51
. lab var pats_cat "patents categories"
. gen rndstck_cat=0 if rndstck==.
. replace rndstck_cat=1 if rndstck>0 & rndstck<=5
. replace rndstck_cat=2 if rndstck>5 & rndstck<=10
```

```

. replace rndstck_cat=3 if rndstck>10 & rndstck<=15
. replace rndstck_cat=4 if rndstck>15 & rndstck<=20
. replace rndstck_cat=5 if rndstck>20 & rndstck<=40
. replace rndstck_cat=6 if rndstck>40 & rndstck<=60
. replace rndstck_cat=7 if rndstck>60
. lab var rndstck_cat "RnDstock categories"
. gen rndeflt_cat=0 if rndeflt==0
. replace rndeflt_cat=1 if rndeflt>0 & rndeflt<=0.5
. replace rndeflt_cat=2 if rndeflt>0.5 & rndeflt<=1
. replace rndeflt_cat=3 if rndeflt>1 & rndeflt<=5
. replace rndeflt_cat=4 if rndeflt>5 & rndeflt<=10
. replace rndeflt_cat=5 if rndeflt>10
. lab var rndeflt_cat "RnDexpenditures categories"

```

The result is the dataset 'example\_data.dta' provided at our homepage (<https://cloud.iwh-halle.de/index.php/s/flexpaneldid>).





Halle Institute for Economic Research –  
Member of the Leibniz Association

Kleine Maerkerstrasse 8  
D-06108 Halle (Saale), Germany

Postal Address: P.O. Box 11 03 61  
D-06017 Halle (Saale), Germany

Tel +49 345 7753 60  
Fax +49 345 7753 820

[www.iwh-halle.de](http://www.iwh-halle.de)

ISSN 2194-2188