**Increasing retirement age and mental health of older workers: the role of working conditions**

**Update 05/04/2024 – Robust DID estimators**

Linear regressions with period and group fixed effects used in DID designs to estimate policy effects may produce misleading estimates if the policy's effect is heterogeneous between groups. Moreover, in multi-period DID designs, the treatment effect may vary not only across different groups but also over time within the same group, meaning that the effect of the treatment may change as time progresses. In such cases, the assumption of parallel trends, where the treatment and control groups would have followed the same trend in the absence of treatment, is often failed. To address the issue of heterogeneous effects, alternative estimators that are robust to such effects have been recently proposed in the literature. These estimators aim to provide more accurate estimates by accounting for the heterogeneity in treatment effects across groups and dynamic treatment effects over time.

[De Chaisemartin & D'Haultfoeuille (2023)](https://doi.org/10.1093/ectj/utac017) describe in detail the issue of heterogeneous treatment effects in linear regressions with period and group fixed effects and review alternative estimators robust to these effects proposed in the recent literature.

Among the estimators allowing dynamic treatment effects (which is relevant in our case with the effects of increased retirement ages on mental health susceptible to change over time) they mainly describe the ones proposed by Callaway and Sant'Anna (2021), Borusyak et al. (2021), and De Chaisemartin and D'Haultfoeuille (2021), which consider different covariate specifications and treatment initiation timing to estimate treatment effects under the assumption of parallel trends.

Callaway and Sant'Anna (2021) propose estimators for binary and staggered treatments under the assumption of parallel trends. They define their parameters of interest based on the timing of treatment initiation and the duration of treatment. They suggest that, in a staggered adoption design, groups can be aggregated into cohorts that start receiving the treatment at the same period. They then impose a weaker parallel trends assumption, where a cohort must be on the same trend as the never-treated / not-yet-treated groups from a certain period onwards, but before that cohort may have been on a different trend. Their assumption is conditional on the design and is not testable. They consider time-invariant covariates and assume that trends are parallel once conditioned on them. In STATA, their estimator corresponds to the package *csdid*.

Borusyak et al. (2021) also propose estimators for binary and staggered treatments accounting for the heterogeneity of effects between groups and periods. Their approach involves estimating the treatment effects using a nonparametric method based on local linear regression. They impose parallel trends for every group and between every pair of consecutive time periods. Their estimators can be obtained by running a TWFE regression of the outcome on group and time fixed effects, and fixed effects for every treated cell. This regression is then used to predict the counterfactual outcome of treated observations, and the treatment effect is obtained by subtracting the counterfactual from the actual outcome. Their design readily generalizes to more complicated specifications, such as triple-differences, or models allowing for group-specific linear trends. In contrast to Callaway and Sant'Anna (2021), they define their parameters of interest using only the never-treated group and time-varying variables as controls. Moreover, including group fixed effects allows for better precision of estimates and, consequently, lower standard errors. In STATA, their estimator corresponds to the package *did\_imputation*.

De Chaisemartin and D’Haultfœuille (2021) propose treatment effect estimators robust to heterogeneous and dynamic treatment effects and that can be used even if the treatment is not binary or the design is not staggered. Their estimators are equivalent to those proposed by Callaway and Sant'Anna (2021) when no covariates are included. However, when covariates are included, their estimators differ and consider time-varying covariates. This allows them to include group-specific linear trends in the estimation. Their estimators can also be used with a binary treatment switching on and off, with a discrete treatment, or with a continuous and staggered treatment (groups start getting treated at different dates, with differing intensities, but once a group gets treated its treatment intensity never changes). However, when allowing for dynamic effects, fewer units can be used as controls in their methodology. Without dynamic effects, at period t, any unit whose treatment has not changed between t − 1 and t can be used as a valid control. With dynamic effects, only units whose treatments have not changed from period 1 to t can be used as valid controls. Therefore, the need for “stayers” is very strong: many units need to keep the same value of the treatment for a large number of time periods. In STATA, their estimator corresponds to the package *did\_multiplegt*.

When specifying the treatment as binary and staggered, the estimator of Borusyak et al. (2021) seems to be the most appropriate in our case. In comparison to Callaway and Sant'Anna (2021)’s estimator, it offers lower standard errors thus being preferable to use. And, in comparison to De Chaisemartin and D’Haultfœuille (2021), it does not present the limitation in the need for “stayers”.

Below are the results of the Borusyak et al. (2021)’s estimator performed with the *did\_imputation* package in STATA. The following model is implemented to estimate the effect of reforms increasing retirement ages (defined as 1 at the moment when an individual faces an increase for the first time) on mental health (EuroD 0-12) with a repeated cross-section specification, time-varying controls, and cell-by-year fixed effects. The standard errors are clustered by cell. The data used is based on 4-digit ISCO codes.

*did\_imputation eurod mergeid year first\_treated [aw=cciw], fe(cell year) controls(age nb\_children nb\_grandchildren partnerinhh thinclog life\_insurance sphus chronic) autosample cluster(cell)*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | **All individuals** | **Males** | **Females** |
| **Full** | | 0.004  (0.206) | -0.093  (0.264) | 0.041  (0.237) |
| **JQI physical environment** | **<= 25th percentile**  **>= 75th percentile**  *p-value for equality* | 0.079  (0.310)  **-0.697\*\*\***  **(0.252)**  ***0.052*** | 0.008  (0.468)  **-0.443\*\*\***  **(0.147)**  *0.358* | 0.090  (0.429)  **-1.205\*\*\***  **(0.313)**  ***0.015*** |
| **JQI social environment** | **<= 25th percentile**  **>= 75th percentile**  *p-value for equality* | -0.357  (0.256)  -0.272  (0.221)  *0.802* | -0.337  (0.409)  -0.080  (0.453)  *0.674* | -0.078  (0.293)  **-0.345\*\***  **(0.154)**  *0.420* |
| **JQI skills and discretion** | **<= 25th percentile**  **>= 75th percentile**  *p-value for equality* | 0.066  (0.430)  **0.582\***  **(0.298)**  *0.324* | -0.365  (0.555)  0.211  (0.225)  *0.336* | -0.054  (0.435)  **0.709\***  **(0.382)**  *0.188* |
| **JQI working time quality** | **<= 25th percentile**  **>= 75th percentile**  *p-value for equality* | **0.894\*\*\***  **(0.245)**  -0.134  (0.342)  ***0.015*** | 0.725  (0.472)  **-0.643\*\*\***  **(0.077)**  ***0.004*** | 0.707  (0.438)  0.114  (0.288)  *0.258* |
| **JQI intensity** | **<= 25th percentile**  **>= 75th percentile**  *p-value for equality* | **0.519\***  **(0.293)**  **-0.412\*\***  **(0.199)**  ***0.009*** | -0.030  (0.209)  -0.304  (0.408)  *0.550* | **0.936\*\*\***  **(0.319)**  **-0.519\*\*\***  **(0.189)**  ***0.000*** |
| **JQI prospects** | **<= 25th percentile**  **>= 75th percentile**  *p-value for equality* | 0.357  (0.305)  0.291  (0.303)  *0.878* | 0.559  (0.468)  **-0.505\***  **(0.289)**  ***0.053*** | -0.056  (0.258)  **0.514\***  **(0.266)**  *0.124* |
| **JQI overall** | **<= 25th percentile**  **>= 75th percentile**  *p-value for equality* | 0.070  (0.278)  -0.123  (0.326)  *0.675* | -0.257  (0.337)  0.082  (0.439)  *0.540* | **0.589\*\***  **(0.237)**  -0.074  (0.398)  *0.152* |

\* p<0.1, \*\* p < 0.05, \*\*\*p < 0.01

Potentially, applying the estimator of De Chaisemartin and D’Haultfœuille (2021) could also be interesting in case of exploration of the continuous or categoric and not staggered specification of treatment (groups start getting treated at different dates, with differing intensities, which can change over time – this can potentially allow us to account for different intensity of pension reforms).