Web appendix of two-way fixed effects estimators with heterogeneous treatment effects

Clément de Chaisemartin* Xavier D'Haultfœuille†

April 18, 2019

Abstract

In this web appendix, we first discuss whether common trends necessarily implies homogenous treatment effect. Second, we consider various extensions of the decomposition results (Theorems 1 and 6) in the paper. Third, we discuss our alternative estimand when the design is sharp but the treatment is stochastic. Fourth, we discuss inference with fuzzy designs. The fifth section discusses briefly all the papers that are included in our literature review (see Section 7 of the paper). Finally, the last section gathers the proofs of all the additional results in this Web Appendix.

1 Can common trends hold with heterogeneous treatment effects?

Throughout the paper, we assume that groups experience common trends, but that the effect of the treatment may be heterogeneous between groups and / or over time. We now discuss two examples where this may happen. We then argue that the mechanisms behind these examples are fairly general. Thus treatment effects are often likely to be heterogeneous, even when common trends are plausible.

First, assume one wants to learn the effect of the minimum wage on the employment levels of some US counties. For simplicity, let us assume that the minimum wage can only take two values, a low and a high value. Also, let us assume that there are only two periods, the 90s and the 2000s. Between these two periods, the amount of competition from China for the US industry increased substantially. Thus, for the common trends assumption to hold for counties A and B, the effect of that increase in competition should be the same in those two counties, in the counterfactual state of the world where A and B have a low minimum wage at both dates.

*University of California at Santa Barbara, clementdechaisemartin@ucsb.edu

†CREST, xavier.dhaultfoeuille@ensae.fr

For that to be true, the economy of those two counties should be pretty similar. For instance, if A has a very service-oriented economy, while B has a very industry-oriented economy, it is unlikely that their employment levels will react similarly to Chinese competition.

Now, if the economies of A and B are similar, they should also have similar effects of the minimum wage on employment, thus implying that the treatment effect is homogenous between groups. On the other hand, the treatment effect may vary over time. For instance, the drop in the employment levels of A and B due to Chinese competition will probably be higher if their minimum wage is high than if their minimum wage is low. This is equivalent to saying that the effect of the minimum wage on employment diminishes from the first to the second period: due to Chinese competition in the second period, the minimum wage may have a more negative effect on employment then.

Second, assume one wants to learn the effect of a job training program implemented in some US counties on participants' wages. Let us suppose that individuals self-select into the training according to a Roy model:

$$D_{i,g,t} = 1\{Y_{i,g,t}(1) - Y_{i,g,t}(0) > c_{g,t}\}, \tag{38}$$

where $c_{g,t}$ represents the cost of the training for individuals in county g and period t. Here, the common trends condition requires that average wages without the training follow the same evolution in all counties. As above, for this to hold counties used in the analysis should have similar economies, so let us assume that those counties are actually identical copies of each other: at each period, their distribution of wages without and with the training is the same. Therefore, $(g,t) \mapsto E(Y_{g,t}(1) - Y_{g,t}(0))$ is constant. However, $c_{g,t}$ may vary across counties and over time: some counties may subsidize the training more than others, and some counties may change their subsidies over time. Then, $(g,t) \mapsto \Delta^{TR}_{g,t} = E(Y_{i,g,t}(1) - Y_{i,g,t}(0)|Y_{i,g,t}(1) - Y_{i,g,t}(0) > c_{g,t})$ will not be constant, despite the fact that all counties in the sample have similar economies and experience similar trends on their wages

Overall, when the treatment is assigned at the group $\times$ period level as in the minimum wage example, the economic restrictions underlying the common trends assumption may also imply homogeneous treatment effect between groups. However, those restrictions usually do not imply that the treatment effect is constant over time. Moreover, when the treatment is assigned at the individual level, as in the job training example, the economic restrictions underlying the common trends assumption neither imply homogeneous treatment effects between groups, nor homogeneous treatment effects over time.

2 Extensions of the decomposition results

We consider hereafter several extensions of our decompositions of β_{fe} and β_{fd} in the paper. First, we consider decompositions of β_{fe} and β_{fd} under the common trends assumption, and under the assumption that the treatment effect is stable over time. Second, we extend our Theorems 1 and 6 to ordered treatments. Third, we investigate the effect of including covariates in the regression. Fourth, we study two-way fixed effects 2SLS regressions. Except when considering 2SLS regressions, we focus hereafter on sharp designs with a non-stochastic treatment, to ease the exposition. Nevertheless, all the results generalize outside of that special case.

2.1 β_{fe} and β_{fd} as weighted sums of ATEs of switching groups

We first show that under an additional condition, β_{fe} and β_{fd} can be written as a weighted sum of the ATEs of switching groups.

Assumption S1 *(Stable treatment effect) For all g and $t \geq 1$, $\Delta_{g,t}D_{g,t-1} = \Delta_{g,t-1}D_{g,t-1}$.*

The stable treatment effect assumption requires that the ATE of every group treated in $t-1$ does not change from $t-1$ to t. By iteration, the ATE of a group treated for instance from period t_0 to $\overline{t}$ is unrestricted before t_0 but should be constant from t_0 to $\overline{t}$. Assumption S1 rules out dynamic effects, and it rules out the possibility that the treatment effect changes over time. Therefore, it may be implausible and should be carefully discussed.

We now show that under the common trend and stable treatment effects assumptions, β_{fe} and β_{fd} may identify weighted averages of ATEs. Let $N_S = \sum_{(g,t):D_{g,t}\neq D_{g,t-1}} N_{g,t}$ and, for all g and $t \geq 1$,

$$w^S_{g,t} = \frac{(D_{g,t} - D_{g,t-1})\sum_{t'\geq t}\frac{N_{g,t'}}{N_{g,t}}\varepsilon_{g,t'}}{\sum_{(g,,t):t\geq 1}\left[\frac{N_{g,t}}{N_S}(D_{g,t} - D_{g,t-1})\sum_{t'\geq t}\frac{N_{g,t'}}{N_{g,t}}\varepsilon_{g,t'}\right]},$$

$$w^S_{fd,g,t} = \frac{(D_{g,t} - D_{g,t-1})\varepsilon_{fd,g,t}}{\sum_{(g,,t):t\geq 1}\frac{N_{g,t}}{N_S}(D_{g,t} - D_{g,t-1})\varepsilon_{fd,g,t}}.$$

Theorem S1 *Suppose that Assumptions 1-3 and S1 hold. Then,*

$$\beta_{fe} = \sum_{(g,t):D_{g,t}\neq D_{g,t-1},t\geq 1}\frac{N_{g,t}}{N_S}w^S_{g,t}\Delta_{g,t},$$

$$\beta_{fd} = \sum_{(g,t):D_{g,t}\neq D_{g,t-1},t\geq 1}\frac{N_{g,t}}{N_S}w^S_{fd,g,t}\Delta_{g,t}.$$

Moreover, $w^S_{fd,g,t} \geq 0$ for all g and $t \geq 1$. If Assumption 4 holds and $N_{g,t}/N_{g,t-1}$ does not vary across g for all $t \geq 1$, $w^S_{g,t} \geq 0$ for all g and $t \geq 1$.

Theorem S1 shows that in sharp designs, under the common trends and stable treatment effect assumptions, β_{fe} and β_{fd} identify weighted sums of ATEs of switching cells. The weights differ from those in Theorems 1 and 6. Now the weights attached to β_{fe} are all positive in staggered adoption designs, while the weights attached to β_{fd} are all positive in all sharp designs. Therefore, in staggered adoption (resp. sharp) designs, β_{fe} (resp. β_{fd}) relies on the assumption that the treatment effect is stable over time, but it does not require that treatment effects be homogeneous between groups.

2.2 Non-binary, ordered treatment

We now consider the case where the treatment takes a finite number of ordered values, $D \in \{0, 1, ..., \overline{d}\}$. We keep here the same notation as before but simply define potential outcomes for all the possible treatment values. For instance $Y_{i,g,t}(\overline{d})$ is the counterfactual outcome of i in (g, t) if she receives treatment value $\overline{d}$. In this set-up, Regressions 1 and 2 and Assumptions 1-3 remain identical.

On the other hand, we need to modify the treatment effect parameters we consider. In lieu of Δ^{TR}, we consider the average causal response (ACR) on the treated,

$$\text{ACR} = E\left(\frac{1}{N_1}\sum_{i,g,t} Y_{i,g,t}(D_{g,t}) - Y_{i,g,t}(0)\right).$$

Similarly, for all (g, t) such that $D_{g,t} \neq 0$, we consider, instead of $\Delta_{g,t}$,

$$\text{ACR}_{g,t} = E\left(\frac{1}{N_{g,t}D_{g,t}}\sum_{i=1}^{N_{g,t}} Y_{i,g,t}(D_{g,t}) - Y_{i,g,t}(0)\right).$$

Then, similarly to (2), the following decomposition holds:[1]

$$\text{ACR} = \sum_{(g,t):D_{g,t}\neq 0} \frac{N_{g,t}D_{g,t}}{N_1}\text{ACR}_{g,t}.$$

Theorem S2 below shows that with a non binary treatment, a similar decomposition as in Theorems 1 and 6 holds for β_{fe} and β_{fd}, with weights that generalize $w_{g,t}$ and $w_{fd,g,t}$ to the non-binary treatment case. Namely,

$$w^O_{g,t} = \frac{\varepsilon_{g,t}}{\sum_{g,t}\frac{N_{g,t}D_{g,t}}{N_1}\varepsilon_{g,t}},$$

$$w^O_{fd,g,t} = \frac{\varepsilon_{fd,g,t} - \frac{N_{g,t+1}}{N_{g,t}}\varepsilon_{fd,g,t+1}}{\sum_{g,t}\frac{N_{g,t}D_{g,t}}{N_1}\left(\varepsilon_{fd,g,t} - \frac{N_{g,t+1}}{N_{g,t}}\varepsilon_{fd,g,t+1}\right)}.$$

[1]Any equation with a numbering lower than (38) refers to an equation in the paper.

Note that if the treatment is binary, $w^O_{g,t} = w_{g,t}$ and $w^O_{fd,g,t} = w_{fd,g,t}$.

Theorem S2 *Suppose that Assumptions 1-3 hold and $D \in \{0, ..., \overline{d}\}$. Then*

$$\beta_{fe} = \sum_{(g,t):D_{g,t}\neq 0} \frac{N_{g,t}D_{g,t}}{N_1} w^O_{g,t} ACR_{g,t},$$
$$\beta_{fd} = \sum_{(g,t):D_{g,t}\neq 0} \frac{N_{g,t}D_{g,t}}{N_1} w^O_{fd,g,t} ACR_{g,t}.$$

Theorem S2 shows that under Assumption 3, when the treatment is not binary β_{fe} and β_{fd} identify weighted sums of the ACRs in each (g,t) cell that is not untreated. Because of this, and since the proof of Corollary 1 does not rely on the nature of the treatment, Corollary 1 directly applies to ordered treatments as well, by just replacing $w_{g,t}$ and $N_{g,t}$ by $w^O_{g,t}$ and $N_{g,t}D_{g,t}$, respectively. Corollary 2 extends as well to this set-up, by simply modifying the no-correlation condition appropriately.

2.3 Including covariates

Often times, researchers also include a vector of covariates $X_{g,t}$ as control variables in their regression. We show in this section that our results can be extended to this case. We start by redefining the two regressions we consider in this context.

Regression 1X *(Fixed-effects regression with covariates)*

Let $\widehat{\beta}^X_{fe}$ and $\widehat{\gamma}_{fe}$ denote the coefficients of $D_{g,t}$ and $X_{g,t}$ in an OLS regression of $Y_{i,g,t}$ on group and period fixed effects, $D_{g,t}$ and $X_{g,t}$. Let $\beta^X_{fe} = E\left(\widehat{\beta}^X_{fe}\right)$ and $\gamma_{fe}(\boldsymbol{X}) = E\left(\widehat{\gamma}_{fe}|\boldsymbol{X}\right)$.

Regression 2X *(First-difference regression with covariates)*

Let $\widehat{\beta}^X_{fd}$ and and $\widehat{\gamma}_{fd}$ denote the coefficients of $D_{g,t}-D_{g,t-1}$ and $X_{g,t}-X_{g,t-1}$ in an OLS regression of $Y_{g,t} - Y_{g,t-1}$ on period fixed effects, $D_{g,t} - D_{g,t-1}$ and $X_{g,t} - X_{g,t-1}$, among observations for which $t \geq 1$. Let $\beta^X_{fd} = E\left(\widehat{\beta}^X_{fd}\right)$ and $\gamma_{fd}(\boldsymbol{X}) = E\left(\widehat{\gamma}_{fd}|\boldsymbol{X}\right)$.

Then, we need to modify the common trends assumption as follows.

Assumption S2 *(Common trends for β^X_{fe}) $E\left[(Y_{g,t}(0) - X'_{g,t}\gamma_{fe}(\boldsymbol{X})) - (Y_{g,t-1}(0) - X'_{g,t-1}\gamma_{fe}(\boldsymbol{X}))\right]$ does not vary across g for all $t \in \{1, ..., \overline{t}\}$.*

Assumption S3 *(Common trends for β^X_{fd}) $E\left[(Y_{g,t}(0) - X'_{g,t}\gamma_{fd}(\boldsymbol{X})) - (Y_{g,t-1}(0) - X'_{g,t-1}\gamma_{fd}(\boldsymbol{X}))\right]$ does not vary across g for all $t \in \{1, ..., \overline{t}\}$.*

Assumptions S2 and S3 are implied by the linear and constant treatment effect models that are often invoked to justify the use of FE and FD regressions with covariates. For instance, the use of Regression 1X is often justified by the following model:

$$Y_{g,t} = \gamma_g + \lambda_t + \theta D_{g,t} + X'_{g,t}\gamma_{fe} + \eta_{g,t}, \quad E(\eta_{g,t}|D_{g,t}, X_{g,t}) = 0. \tag{39}$$

Equation (39) implies Assumption S2, but it does not imply Assumption 3. This is why we consider the former common trends assumption instead of the latter in this subsection. Similarly, the linear and constant treatment effect model rationalizing the use of Regression 2X implies Assumption S3 but not Assumption 3.

Assumption S2 requires that once netted out from the partial linear correlation between the outcome and X, $Y(0)$ satisfies the common trends assumption. It may be more plausible than Assumption 3, for instance if there are group-specific trends affecting the outcome but if those group-specific trends can be captured by a linear model in X. A similar interpretation applies to Assumption S3.

Theorem S3 below generalizes Theorems 1 and 6 to the case where there are covariates in the regression.

Theorem S3 *Suppose that Assumptions 1-2 hold. Then:*

1. *If Assumption S2 holds as well,* $\beta_{fe}^X = \sum_{(g,t):D_{g,t}=1} \frac{N_{g,t}}{N_1} w_{g,t}\Delta_{g,t}$.
2. *If Assumption S3 holds as well,* $\beta_{fd}^X = \sum_{(g,t):D_{g,t}=1} \frac{N_{g,t}}{N_1} w_{fd,g,t}\Delta_{g,t}$.

Theorem S3 shows that under a modified version of the common trends assumption accounting for the covariates, β_{fe}^X identifies the same weighted sum of the $\Delta_{g,t}^{TR}$s as β_{fe} in Theorem 1. The same holds for β_{fd}^X. Therefore, if the regression without covariates has some negative weights and one worries that treatment effect heterogeneity may be systematically related to the weights, the addition of covariates will not alleviate that concern. But adding covariates may still have some benefits. For instance, the common trends assumption underlying β_{fe}^X may be more credible than that underlying β_{fe}.

Finally, Theorem 2 can easily be extended to the case with covariates. Under versions of Assumptions 6 and 7 written conditional on X, a conditional version of the Wald-TC estimand identifies Δ^S under the common support condition $\text{Supp}(X_{d,g,t}) = \text{Supp}(X)$. We refer to de Chaisemartin and D'Haultfœuille (2018) for further details.

2.4 2SLS regressions

Researchers have sometimes estimated 2SLS versions of Regressions 1 and 2. Our main conclusions also apply to those regressions. Let $\widehat{\beta}_{fe}^{2SLS}$ denote the coefficient of $D_{i,g,t}$ in a 2SLS regression of $Y_{g,t}$ on group and period fixed effects and $D_{i,g,t}$, using a variable $Z_{g,t}$ constant within each group $\times$ period as the instrument for $D_{i,g,t}$. $Z_{g,t}$ typically represents an incentive for treatment allocated at the group $\times$ period level. For instance, Duflo (2001) studies the effect of years of schooling on wages in Indonesia, using a primary school construction program as an instrument. Specifically, she estimates a 2SLS regression of wages on cohort and district of birth fixed effects and years of schooling, using the interaction of belonging to a cohort entering primary school after the program was completed and the number of schools constructed in one's district of birth as the instrument for years of schooling.

Remark that $\widehat{\beta}_{fe}^{2SLS} = \widehat{\beta}_{fe}^{Y}/\widehat{\beta}_{fe}^{D}$, where $\widehat{\beta}_{fe}^{Y}$ (resp. $\widehat{\beta}_{fe}^{D}$) is the coefficient of $Z_{g,t}$ in the reduced-form regression of $Y_{g,t}$ (resp. the first-stage regression of $D_{g,t}$) on group and period fixed effects and $Z_{g,t}$. Then let $\beta_{fe}^{2SLS} = E[\widehat{\beta}_{fe}^{Y}]/E[\widehat{\beta}_{fe}^{D}]$.[2] Following Imbens and Angrist (1994), for any $z \in \text{Supp}(Z)$ let $D_{i,g,t}(z)$ denote the potential treatment of unit i in (g,t) if $Z_{i,g,t} = z$. It follows from Theorem 1 that under a common trends assumption on $D_{i,g,t}(0)$, $E[\widehat{\beta}_{fe}^{D}]$ is equal to a weighted sum of the average effects of the instrument on the treatment in each group and time period, with potentially many negative weights. Similarly, under a common trends assumption on $Y_{i,g,t}(D_{i,g,t}(0))$ instead of $Y_{i,g,t}(0)$, $E[\widehat{\beta}_{fe}^{Y}]$ is equal to a weighted sum of the average effects of the instrument on the outcome, again with potentially many negative weights. For instance, in Duflo (2001), under a common trends assumption on $D_{i,g,t}(0)$, the number of years of schooling individuals would complete if zero new schools were constructed in their district, the first stage coefficient identifies a weighted sum of the effect of one new school on years of schooling in every district, with many negative weights.[3]

Hence, it is only if the average effects of $Z_{g,t}$ on $Y_{i,g,t}$ and $D_{i,g,t}$ are constant across groups and periods, or if the weights are uncorrelated to treatment effects as in Assumption 5, that the reduced-form and first-stage coefficients respectively identify the average effect of $Z_{i,g,t}$ on $Y_{i,g,t}$ and $D_{i,g,t}$. Then, this implies that β_{fe}^{2SLS} identifies, under the conditions in Imbens and Angrist (1994), the LATE of $D_{i,g,t}$ on $Y_{i,g,t}$ among units that comply with the instrument.[4]

[2] We do not consider here $E[\widehat{\beta}_{fe}^{2SLS}]$, as the 2SLS estimator has actually no expectation. Moreover, under conditions similar to those imposed in Section 6 of the paper, β_{fe}^{2SLS} is the probability limit of $\widehat{\beta}_{fe}^{2SLS}$, which makes β_{fe}^{2SLS} the proper estimand here.

[3] New schools were constructed in every district, so this application falls into the heterogeneous adoption case.

[4] In the special case with two groups and two periods, a binary incentive for treatment, and where only group 1 in period 1 receives the incentive, de Chaisemartin (2010) and Hudson et al. (2015) show that in a 2SLS regression of $Y_{i,g,t}$ on $1\{g = 1\}$, $1\{t = 1\}$ and $D_{i,g,t}$, using $Z_{g,t} = 1\{g = 1\}1\{t = 1\}$ as the instrument, the coefficient of $D_{i,g,t}$ identifies a LATE under common trends assumptions on $Y_{i,g,t}(D_{i,g,t}(0))$ and $D_{i,g,t}(0)$. However, the

3 Additional results on the alternative estimand

3.1 Sharp design and stochastic treatment

In this section, we assume that the first point of Assumption 2 in the paper holds, but the second point fails: the design is sharp, but the treatment status of each (g,t) cell can be stochastic. Then, the Wald-TC estimand introduced in Section 3.3 of the paper is not an estimand anymore, as it depends on the realization of the $D_{g,t}$s. The Wald-TC estimand introduced in Section 4.2 is also not suitable, because its definition depends on $E[D_{g,t}]$, and this object cannot be estimated consistently with a sharp design but a stochastic treatment.[5] We therefore need to redefine a Wald-TC estimand in such cases. For all $t \in \{1,...,\overline{t}\}$ and for all $(d,d') \in \{0,1\}^2$, recall that $N_{d,d',t} = \sum_{g:D_{g,t}=d,D_{g,t-1}=d'} N_{g,t}$. Then, for all $t \in \{1,...,\overline{t}\}$ and as long as $\min(E[N_{1,0,t}],E[N_{0,0,t}]) > 0$ or $\min(E[N_{1,1,t}],E[N_{0,1,t}]) > 0$, let

$$\begin{aligned}
DID_{+,t} =& \sum_{g=0}^{\overline{g}} \frac{N_{g,t}P(D_{g,t}=1,D_{g,t-1}=0)}{E[N_{1,0,t}]} E(Y_{g,t}-Y_{g,t-1}|D_{g,t}=1,D_{g,t-1}=0) \\
&- \sum_{g=0}^{\overline{g}} \frac{N_{g,t}P(D_{g,t}=0,D_{g,t-1}=0)}{E[N_{0,0,t}]} E(Y_{g,t}-Y_{g,t-1}|D_{g,t}=0,D_{g,t-1}=0) \\
DID_{-,t} =& \sum_{g=0}^{\overline{g}} \frac{N_{g,t}P(D_{g,t}=1,D_{g,t-1}=1)}{E[N_{1,1,t}]} E(Y_{g,t}-Y_{g,t-1}|D_{g,t}=1,D_{g,t-1}=1) \\
&- \sum_{g=0}^{\overline{g}} \frac{N_{g,t}P(D_{g,t}=0,D_{g,t-1}=1)}{E[N_{0,1,t}]} E(Y_{g,t}-Y_{g,t-1}|D_{g,t}=0,D_{g,t-1}=1).
\end{aligned}$$

When $\min(E[N_{1,0,t}],E[N_{0,0,t}]) = 0$ (resp. $\min(E[N_{1,1,t}],E[N_{0,1,t}]) = 0$), we simply let $DID_{+,t} = 0$ (resp. $DID_{-,t} = 0$). Finally, using again the convention that $0/0 = 0$, let

$$W_{TC} = \frac{\sum_{t=1}^{\overline{t}} E[N_{1,0,t}]DID_{+,t} + E[N_{0,1,t}]DID_{-,t}}{\sum_{t=1}^{\overline{t}} E[N_{1,0,t}] + E[N_{0,1,t}]}.$$

The causal effect identified by our new Wald-TC estimand is

$$\begin{aligned}
\Delta^S =& \frac{\sum_{(g,t),t\geq 1} N_{g,t}P(D_{g,t}=1,D_{g,t-1}=0)E(Y_{g,t}(1)-Y_{g,t}(0)|D_{g,t}=1,D_{g,t-1}=0)}{\sum_{(g,t),t\geq 1} N_{g,t}(P(D_{g,t}=1,D_{g,t-1}=0)+P(D_{g,t}=0,D_{g,t-1}=1))} \\
&+ \frac{\sum_{(g,t),t\geq 1} N_{g,t}P(D_{g,t}=0,D_{g,t-1}=1)E(Y_{g,t}(1)-Y_{g,t}(0)|D_{g,t}=0,D_{g,t-1}=1)}{\sum_{(g,t),t\geq 1} N_{g,t}(P(D_{g,t}=1,D_{g,t-1}=0)+P(D_{g,t}=0,D_{g,t-1}=1))}. \qquad (40)
\end{aligned}$$

discussion above shows that this result does not generalize to applications with multiple groups and periods, a non-binary instrument, or a fuzzy design where all groups receive heterogeneous levels of the incentive, as in Duflo (2001).

[5] When the treatment varies at the individual level, on the other hand, $E[D_{g,t}]$ can be consistently estimated if for instance the $D_{i,g,t}$ are i.i.d. within each cell (g,t).

Each cell (g,t) may switch from untreated to treated or from treated to treated from $t-1$ to t. Δ^S is a weighted average of the ATEs in each (g,t) cell conditional on those two events, where each conditional ATE receives a weight proportional to the number of units in the cell multiplied by the probability of the event. Notice that if Point 2 of Assumption 2 in the paper holds, the objects $N_{d,d',t}$, $DID_{+,t}$, $DID_{-,t}$, W_{TC} and Δ^S defined above are equal to the objects with the same name in Section 3.3 of the paper, which is why we do not change their names.

We show below that Δ^S is identified by W_{TC} under Assumptions S4 and S5. These conditions replace Assumptions 6 and 7 now that we assume that the treatment is stochastic.

Assumption S4 *(Common trends for groups with the same treatment at $t-1$)*

1. *For all $t \geq 1$, $E(Y_{g,t}(0) - Y_{g,t-1}(0)|D_{g,t-1} = 0)$ does not vary across g and $E(Y_{g,t}(0) - Y_{g,t-1}(0)|D_{g,t}, D_{g,t-1} = 0)$ does not depend on $D_{g,t}$.*
2. *For all $t \geq 1$, $E(Y_{g,t}(1) - Y_{g,t-1}(1)|D_{g,t-1} = 1)$ does not vary across g, and $E(Y_{g,t}(1) - Y_{g,t-1}(1)|D_{g,t}, D_{g,t-1} = 1)$ does not depend on $D_{g,t}$.*

Point 1 requires that conditional on being untreated at $t-1$, all groups experience the same evolution from $t-1$ to t of their expected $Y(0)$. It also requires that conditional on being untreated at $t-1$, the evolution from $t-1$ to t of group's g expected $Y(0)$ does not depend on her period-t treatment. This second requirement is not specific to our Wald-TC estimand: it arises due to the stochastic nature of the treatment. Consider a simple example with two groups two periods, and where no group is treated at period 0 and one of the two groups is randomly selected for treatment at period 1. Assume also that group 1 has a probability $p_1 > 1/2$ of being treated at period 1. Then, if Point 1 of Assumption S4 holds, the DID estimator $\widehat{DID} = \sum_{g=0}^{1}(Y_{g,1} - Y_{g,0})(2D_{g,1} - 1)$ is unbiased for $p_1\Delta_{1,1} + (1-p_1)\Delta_{0,1}$. On the other hand, if the evolutions from period 0 to 1 of groups' 0 and 1 expected $Y(0)$ do depend on their period-1 treatment, $\widehat{DID}$ is biased. Point 2 of Assumption 4 is similar to Point 1, except that it concerns the evolution of $Y(1)$. Finally, if Point 2 of Assumption 2 holds, Assumption S4 reduces to Assumption 6.

Assumption S5 *(Existence of "stable" groups) For all $t \in \{1, ..., \overline{t}\}$,*

1. *If there exists g such that $P(D_{g,t} = 1, D_{g,t-1} = 0) > 0$, then there exists $g' \neq g$ such that $P(D_{g,t} = 0, D_{g,t-1} = 0) > 0$.*
2. *If there exists g such that $P(D_{g,t} = 0, D_{g,t-1} = 1) > 0$, then there exists $g' \neq g$ such that $P(D_{g,t} = 1, D_{g,t-1} = 1) > 0$.*

Assumption S5 generalizes Assumption 7 when the treatment is stochastic. It ensures that if at least one group has a positive probability to switch from treatment value 0 to 1 (resp. 1 to 0), then at least another group has a positive probability to remain untreated (resp. treated).

Theorem S4 *If Assumption 1, Point 1 of Assumption 2, and Assumptions S4-S5 hold, then* $W_{TC} = \Delta^S$.

Overall, Theorem S4 shows that in sharp designs with a stochastic treatment, an identification result similar to that in sharp designs with a non stochastic treatment can be obtained, though the estimand, the causal parameter and the identifying assumptions have to be modified. Importantly however, the estimator we proposed with a non stochastic treatment remains consistent if the treatment is stochastic. We refer to Section 4.2.1 below for more details.

3.2 Non-binary, ordered treatment

Theorem 2 can be extended to the case where the treatment is not binary but takes values in $\mathcal{D} = \{0, ..., \overline{d}\}$. For any (g,t) such that $t \geq 1$ and $D_{g,t} \neq D_{g,t-1}$, let

$$\text{ACR}^S_{g,t} = E\left[\frac{1}{N_{g,t}|D_{g,t} - D_{g,t-1}|}\sum_{(i=1}^{N_{g,t}} Y_{i,g,t}\left(\max(D_{g,t}, D_{g,t-1})\right) - Y_{i,g,t}\left(\min(D_{g,t}, D_{g,t-1})\right)\right]$$

be the average effect of going from $\min(D_{g,t}, D_{g,t-1})$ to $\max(D_{g,t}, D_{g,t-1})$ units of treatment, normalized by $|D_{g,t} - D_{g,t-1}|$, in all the (g,t) cells whose treatment changes from $t-1$ to t. The causal effect we consider hereafter is

$$\text{ACR}^S = \sum_{(g,t):D_{g,t}\neq D_{g,t-1}, t\geq 1} \frac{N_{g,t}|D_{g,t} - D_{g,t-1}|}{N_{D,S}} ACR^S_{g,t},$$

where $N_{D,S} = \sum_{(g,t):t\geq 1} N_{g,t}|D_{g,t} - D_{g,t-1}|$. Note that $\text{ACR}^S = \Delta^S$ when $D_{g,t}$ is binary.

We identify ACR^S under the following two conditions, which generalize Assumptions 6-7 to non-binary treatments.

Assumption S6 *(Common trends for groups with the same treatment at* $t-1$*) For every* d *and for all* (g,t) *such that* $D_{g,t-1} = d$*,* $E(Y_{g,t}(d)) - E(Y_{g,t-1}(d))$ *does not vary across* g*.*

Assumption S7 *(Existence of "stable" groups) For all* $t \in \{1, ..., \overline{t}\}$*: For all* $(d, d') \in \mathcal{D}^2$*,* $d \neq d'$*, if there is at least one* $g \in \{0, ..., \overline{g}\}$ *such that* $D_{g,t-1} = d$ *and* $D_{g,t} = d'$*, then there exists at least one* $g' \neq g, g' \in \{0, ..., \overline{g}\}$ *such that* $D_{g',t-1} = D_{g',t} = d$*.*

When the treatment takes a large number of values, Assumption S7 may be violated. A solution, then, is to consider a modified treatment variable $\widetilde{D}_{g,t} = h(D_{g,t})$ that groups together several values of $D_{g,t}$, to ensure that Assumption S7 holds for $\widetilde{D}_{g,t}$. For instance, if the treatment can be equal to 0, 1, 2, or 3, and there is a group whose treatment switches from 2 to 3 between period 0 and 1, but no group whose treatment remains equal to 2 between those two dates, one may define $\widetilde{D}_{g,t} = \min(D_{g,t}, 2)$ if there is a group whose treatment is equal to 3 at periods 0 and 1. Then, Theorem S5 below still holds, after replacing $D_{g,t}$ by $\widetilde{D}_{g,t}$ in the numerator of the Wald-TC estimand defined below, and if Assumption S6 is replaced by the requirement that $E(Y_{g,t}(d)) - E(Y_{g,t-1}(d))$ only depends on t and $h(d)$.

In order to define W_{TC} in this context, let us introduce, for all $(d, d', t) \in \mathcal{D}^2 \times \{1, ..., \bar{t}\}$,

$$DID_{d,d',t} = [1\{d < d'\} - 1\{d' < d\}] \Bigg[\sum_{(g,t):D_{g,t}=d', D_{g,t-1}=d, t\geq 1} \frac{N_{g,t}}{N_{d,d',t}} [E(Y_{g,t}) - E(Y_{g,t-1})] - \sum_{(g,t):D_{g,t}=D_{g,t-1}=d, t\geq 1} \frac{N_{g,t}}{N_{d,d,t}} [E(Y_{g,t}) - E(Y_{g,t-1})] \Bigg],$$

where $N_{d,d',t}$ is defined as in (3) for any $(d, d') \in \mathcal{D}^2$. Then

$$W_{TC} = \sum_{t=1}^{\bar{t}} \sum_{(d,d')\in\mathcal{D}^2, d\neq d'} \frac{N_{d,d',t}}{N_{D,S}} DID_{d,d',t}.$$

If the treatment is binary, the Wald-TC estimand defined above is equal to that defined in Section 3.3 of the paper.

Theorem S5 *Suppose that Assumptions 1, 2, S6 and S7 hold, and $D_{i,g,t} \in \mathcal{D}$. Then $W_{TC} = ACR^S$.*

Theorem S5 generalizes Theorem 2 to non-binary treatments. We can also extend Theorem 3 in the same way to construct placebo tests of Assumption S6.

4 Estimation and inference with stochastic treatments

4.1 Sensitivity measures

With stochastic treatments, the sensitivity measure defined in Corollary 1 of the paper slightly changes, and the interpretation of that measure also slightly changes. However, we still have that $\widehat{\underline{\sigma}}_{fe}$, the estimator of that measure defined in Section 6 of the paper, converges towards a limiting sensitivity measure when the number of groups tends to infinity, as in Theorem 7 of the paper. Let

$$\Delta^{TR}(\boldsymbol{D}) = \sum_{g,t} \frac{N_{g,t}D_{g,t}}{N_1}\Delta^{TR}_{g,t}(\boldsymbol{D}),$$

$$\sigma(\boldsymbol{\Delta^{TR}}(\boldsymbol{D})) = \left(\sum_{g,t} \frac{N_{g,t}D_{g,t}}{N_1}\left(\Delta_{g,t}(\boldsymbol{D}) - \Delta^{TR}(\boldsymbol{D})\right)^2\right)^{1/2},$$

$$\sigma(\boldsymbol{w^{TR}}) = \left(\sum_{g,t} \frac{N_{g,t}D_{g,t}}{N_1}\left(w^{TR}_{g,t} - 1\right)^2\right)^{1/2}.$$

Under Assumption 9, one can show that

$$E\left(\widehat{\beta}_{fe}|\boldsymbol{D}\right) = \sum_{g,t} \frac{N_{g,t}D_{g,t}}{N_1} w^{TR}_{g,t}\Delta^{TR}_{g,t}(\boldsymbol{D}).$$

Let $\widetilde{\beta}_{fe} = E(\widehat{\beta}_{fe}|\boldsymbol{D})$. Now, applying Corollary 1 to $\sigma(\boldsymbol{\Delta^{TR}}(\boldsymbol{D}))$, $\widetilde{\beta}_{fe}$ and $\Delta^{TR}(\boldsymbol{D})$, we obtain that the minimal value of $\sigma(\boldsymbol{\Delta^{TR}}(\boldsymbol{D}))$ compatible with $\widetilde{\beta}_{fe}$ and $\Delta^{TR}(\boldsymbol{D}) = 0$ is

$$\underline{\widetilde{\sigma}}^{TR}_{fe} = \frac{|\widetilde{\beta}_{fe}|}{\sigma(\boldsymbol{w^{TR}})}.$$

Essentially, Corollary 1 still holds conditional on $\boldsymbol{D}$, and the sensitivity measure is now a random variable.

Now, assume that the vectors $(Y_{g,0}, ..., Y_{g,\bar{t}}, D_{g,0}, ..., D_{g,\bar{t}})_{g\geq 0}$ are mutually independent, the analogue of Assumption 17 in the paper with stochastic treatments. Then, by the law of large numbers applied to triangular arrays, $\underline{\widetilde{\sigma}}^{TR}_{fe}$ will converge, under appropriate moment conditions, to a fixed quantity $\underline{\sigma}^{TR}_{fe}$. Now, using similar arguments, $\widehat{\beta}_{fe} - \widetilde{\beta}_{fe}$, converges to 0 in probability, so

$$\underline{\widehat{\sigma}}_{fe} = \frac{|\widehat{\beta}_{fe}|}{\sigma(\boldsymbol{w^{TR}})},$$

the estimator of the sensitivity measure defined in Section 6 of the paper, converges in probability to $\underline{\sigma}^{TR}_{fe}$ as $\overline{g}$ tends to infinity.

4.2 Average treatment effect on switchers

4.2.1 Sharp, stochastic designs

In this case, the estimator $\widehat{W}_{TC}$ defined in Section 6 of the paper is a consistent estimator of Δ^S defined in (40), provided some conditions hold. These conditions include Assumption 1, Point 1 of Assumption 2, Assumptions S4-S5 and the mutual independence of the vectors $(Y_{g,0}, ..., Y_{g,\bar{t}}, D_{g,0}, ..., D_{g,\bar{t}})_{g\geq 0}$. As in Theorem 8 in the paper, consistency can be proven by combining Theorem S4 above with the weak law of large numbers for triangular arrays. Also, inference can be conducted in the same way as in Section 6 of the paper.

4.2.2 *Fuzzy designs*

We now suppose that observations' treatment can vary within each cell (g, t). We then adopt the framework of Section 4.2 of the paper. To estimate $\tilde{\Delta}_S$, we can use an estimator of W_{TC} in Equation (4) where $E(D_{g,t})$ and $E(Y_{g,t})$ are respectively replaced by $D_{g,t}$ and $Y_{g,t}$. Estimating W_{TC} also requires forming estimators of three sets of groups at each period: $\mathcal{G}_{+,t} = \{g : E(D_{g,t}) > E(D_{g,t-1})\}$, the set of groups whose treatment increases between $t-1$ and t; $\mathcal{G}_{-,t} = \{g : E(D_{g,t}) < E(D_{g,t-1})\}$, the set of groups whose treatment decreases between $t-1$ and t; and $\mathcal{G}_{=,t} = \{g : E(D_{g,t}) = E(D_{g,t-1})\}$, the set of groups whose treatment remains stable between $t-1$ and t. We estimate these sets by

$$\begin{aligned}
\widehat{\mathcal{G}}_{+,t} &= \{g : D_{g,t} - D_{g,t-1} > c_{N_{g,t}}\},\\
\widehat{\mathcal{G}}_{-,t} &= \{g : D_{g,t} - D_{g,t-1} < -c_{N_{g,t}}\},\\
\widehat{\mathcal{G}}_{=,t} &= \{g : D_{g,t} - D_{g,t-1} \in [-c_{N_{g,t}}, c_{N_{g,t}}]\}.
\end{aligned}$$

for some sequence $c_{N_{g,t}} > 0$ converging towards 0 when $N_{g,t}$ goes to infinity, but such that $\sqrt{N_{g,t}}c_{N_{g,t}}$ converges towards $+\infty$.

We now discuss informally the properties of our estimator, leaving formal results for future research. If the $(D_{i,g,t})_{i=1\dots N_{g,t}}$ are assumed to be i.i.d. in each (g, t) cell and if $N_{g,t} \to +\infty$ for all (g, t), $D_{g,t}$ will converge to $E(D_{g,t})$. We can thus expect $\widehat{W}_{TC}$ to be consistent and asymptotically normal in such a case, because the probability of incorrectly classifying a (g, t) cell as, e.g., belonging to $\mathcal{G}_{+,t}$ while it actually belongs to $\mathcal{G}_{=,t}$ will go to 0 given the requirements we imposed on the sequence $c_{N_{g,t}}$. de Chaisemartin and D'Haultfœuille (2018) show a result similar to the result we are hypothesizing here but with a fixed number of groups $\overline{g}$ and assuming also that the $(Y_{i,g,t})_{i=1\dots N_{g,t}}$ are i.i.d. (see Theorem S6 in their Web Appendix). We conjecture that their result would still hold, in an asymptotic framework where $N_{g,t}$ goes to infinity for each (g, t), and $\overline{g}$ also goes to infinity.

5 Detailed literature review

We now review the 33 papers that use two-way fixed effects or closely related regressions that we found in our literature review. For each paper, we use the following presentation:
Authors (year), Title. *Where the two-way fixed effects estimator is used in the paper.*
Description of the two-way fixed effects estimator used in the paper, and how it relates to Regression 1 or 2. Assessment of whether the stable groups assumption holds in this paper. Assessment of whether the research design is sharp or fuzzy.

1. **Chandra et al. (2010), Patient Cost-Sharing and Hospitalization Offsets in the Elderly.** *First line of Tables 2 and 3.*

In the regressions in the first line of Tables 2 and 3, the outcomes (e.g. a measure of utilization for plan p in month t) are regressed on plan fixed effects, month fixed effects, and an indicator of whether plan p had increased copayments in month t (see regression equation at the bottom of page 198). This regression corresponds to Regression 1. The period analyzed runs from January 2000 to September 2003. The stable groups assumption is satisfied until January 2002, when the HMO plans also become treated. This is a sharp design.

2. **Duggan and Morton (2010), The Effect of Medicare Part D on Pharmaceutical Prices and Utilization.** *Tables 2 and 3.*
   In regression Equation (1), the dependent variable is the change in the price of drug j between 2003 and 2006, the explanatory variables are the Medicare market share for drug j in 2003, and some control variables. This regression corresponds to Regression 2, with some control variables. The stable groups assumption is presumably not satisfied: it seems unlikely that there are drugs whose Medicare market share in 2003 is equal to 0. This is a sharp design.

3. **Aizer (2010), The Gender Wage Gap and Domestic Violence.** *Table 2.*
   In regression Equation (2), the dependent variable is the log of female assaults among females of race r in county c and year t, and the explanatory variables are race, year, county, race $\times$ year, race $\times$ county, and county $\times$ year fixed effects, as well as the gender wage gap in county c, year t, and race r, and some control variables. This regression is a "three-way fixed effects" version of Regression 1, with some control variables. The stable groups assumption is presumably satisfied: it seems likely that between each pair of consecutive years, there are counties where the gender wage gap does not change. This is a fuzzy design: the treatment of interest is the gender wage gap in a couple (see the bargaining model in Appendix 1), which varies within (year,county) cells.

4. **Algan and Cahuc (2010), Inherited Trust and Growth.** *Figure 4.*
   Figure 4 presents a regression of changes in income per capita from 1935 to 2000 on changes in inherited trust over the same period and a constant. This regression corresponds to Regression 2. The stable groups assumption is satisfied: there are countries where inherited trust does not change from 1935 to 2000. This is a sharp design.

5. **Ellul et al. (2010), Inheritance Law and Investment in Family Firms.** *Table 7.*
   In the regressions presented in Table 7, the dependent variable is the capital expenditure of firm j in year t, and the explanatory variables are firm fixed effects, an indicator for whether year t is a succession period for firm j, some controls, and three treatment variables: the interaction of the succession indicator with the level of investor protection in the country where firm j is located, the interaction of the succession indicator with the level of

inheritance laws permissiveness in the country where firm j is located, and the interaction of the succession indicator with the level of inheritance laws permissiveness and the level of investor protection in the country where firm j is located. This regression is similar to Regression 1 with controls, except that it has three treatment variables. The stable groups assumption is presumably not satisfied: for instance, it seems unlikely that there are countries with no investor protection at all. This is a sharp design.

6. **Bustos (2011), Trade Liberalization, Exports, and Technology Upgrading: Evidence on the Impact of MERCOSUR on Argentinean Firms.** *Tables 3 to 12.* In regression Equation (11), the dependent variable is the change in exporting status of firm i in sector j between 1992 and 1996, and the explanatory variables are the change in trade tariffs in Brasil for products in sector j over the same period, and some control variables. This regression corresponds to Regression 2, with some controls. The stable groups assumption is presumably satisfied: it seems likely that there are sectors where trade tariffs in Brasil did not change between 1992 and 1996. This is a sharp design.

7. **Anderson and Sallee (2011), Using Loopholes to Reveal the Marginal Cost of Regulation: The Case of Fuel-Economy Standards.** *Table 5 Column 2.* In the regression in Table 5 Column (2), the dependent variable is an indicator for whether a car sold is a flexible fuel vehicle, and the explanatory variables are state and month fixed effects, the percent of gas stations that have ethanol fuel in each month $\times$ state, and some controls. This regression corresponds to Regression 1. The stable groups assumption is presumably satisfied: it seems likely that between each pair of consecutive months, there are states where the percent ethanol availability does not change. This a fuzzy design: the treatment of interest is whether a car buyer has access to ethanol fuel, which varies within (month,state) cells.

8. **Bagwell and Staiger (2011), What Do Trade Negotiators Negotiate About? Empirical Evidence from the World Trade Organization.** *Table 3, OLS columns.* In regression equations (15a) and (15b), the dependent variable is the ad valorem tariff level bound by country c on product g, while the explanatory variables are country and product fixed effects, and two treatment variables which vary at the country $\times$ product level. These regressions are similar to Regression 1, except that they have two treatment variables. The stable groups assumption is not applicable here, as none of the two sets of fixed effects included in the regression correspond to an ordered variable. This is a sharp design.

9. **Zhang and Zhu (2011), Group Size and Incentives to Contribute: A Natural Experiment at Chinese Wikipedia.** *Tables 3 and 4, Columns 4-6.* In the regression in, say, Table 3 Column (4), the dependent variable is the total number

of contributions to Wikipedia by individual i at period t, regressed on individual fixed effects, an indicator for whether period t is after the Wikipedia block, the interaction of this indicator and a measure of social participation by individual i, and some controls. This regression corresponds to Regression 1 with some controls. The stable groups assumption is satisfied: there are individuals with a social participation measure equal to 0. This is a sharp design.

10. **Hotz and Xiao (2011), The Impact of Regulations on the Supply and Quality of Care in Child Care Markets.** *Table 7, Columns 4 and 5.*
    In Regression Equation (1), the dependent variable is the outcome for market m in state s and year t, and the explanatory variables are state and year fixed effects, various measures of regulations in state s in year t, and some controls. This regression corresponds to Regression 1 with several treatment variables and with some controls. The stable groups assumption is presumably satisfied: between each pair of consecutive years, it is likely that there are states whose regulations do not change. This is a sharp design.

11. **Mian and Sufi (2011), House Prices, Home Equity-Based Borrowing, and the US Household Leverage Crisis.** *Tables 2 and 3.*
    In Regression Equation (1), the dependent variable is the change in homeowner leverage from 2002 to 2006 for individual i living in zip code z in MSA m, and the dependent variable is the change in the house price for that individual, instrumented by MSA-level housing supply elasticity. This regression is the 2SLS version of Regression 2, with some controls. The stable groups assumption is presumably not satisfied: it is unlikely that some MSAs have an housing supply elasticity equal to 0. This is a sharp design.

12. **Wang (2011), State Misallocation and Housing Prices: Theory and Evidence from China.** *Table 5, Panel A.*
    In regression Equation (15), the dependent variable is the quantity of housing services in household i's residence in year t, while the explanatory variables are an indicator for period t being after the reform, a measure of mismatch in household i, the interaction of the measure of mismatch and the time indicator, and some controls. This regression is similar to Regression 1 with some controls, except that it has a measure of mismatch in household i instead of household fixed effects. The stable groups assumption is presumably satisfied: it is likely that some households have a mismatch equal to 0. This is a sharp design.

13. **Duranton and Turner (2011), The Fundamental Law of Road Congestion: Evidence from US Cities.** *Table 5.*
    In the regressions presented in, say, the first column of Table 5, the dependent variable is the change in vehicle kilometers traveled in MSA s between decades t and t-1, and the

explanatory variables are the change in kilometers of roads in MSA s between decades t and t-1, and decade effects. This regression corresponds to Regression 2. The stable groups assumption is presumably satisfied: it is likely that between each pair of consecutive decades, there are some MSAs where the kilometers of roads do not change. This is a sharp design.

14. **Acemoglu et al. (2011), The Consequences of Radical Reform: The French Revolution.** *Table 3.*
    In regression Equation (1), the dependent variable is urbanization in polity j at time t, while the explanatory variables are time and polity fixed effects, and the number of years of French presence in polity j interacted with the time effects. This regression corresponds to Regression 1. The stable groups assumption is satisfied as there are several polities that did not experience any year of French presence. This is a sharp design.

15. **Baum-Snow and Lutz (2011), School Desegregation, School Choice, and Changes in Residential Location Patterns by Race.** *Tables 2 to 6.*
    In regression Equation (1), the dependent variable is, say, whites public school enrolment in MSA j in year t, while the explanatory variables are MSA and region $\times$ time fixed effects, and an indicator for whether MSA j is desegregated. This regression corresponds to Regression 1 with controls. The stable groups assumption is satisfied: between each pair of consecutive years, there are MSAs whose desegregation status does not change. This is a sharp design.

16. **Dinkelman (2011), The Effects of Rural Electrification on Employment: New Evidence from South Africa.** *Tables 4 and 5 Columns 5-8, Table 8 Columns 3-4, Table 9 Column 2, and Table 10 Columns 2, 4, and 6.*
    In regression Equation (3), the dependent variable is, say, the first difference of the female employment rate for community j between periods 0 and 1, and the explanatory variables are district fixed effects, the change of electrification status of community j between periods 0 and 1, and some statistical controls. The land gradient in community j is used as an instrument for the change in electrification. This regression corresponds to the 2SLS version of Regression 2 with some controls. The stable groups assumption is presumably satisfied: it is likely that there are communities whose land gradient is 0. This is a sharp design.

17. **Enikolopov et al. (2011), Media and Political Persuasion: Evidence from Russia.** *Table 3.*
    In regression Equation (5), the dependent variable is the share of votes for party j in election-year t and subregion s, and the explanatory variables are subregion and election fixed effects, and the share of people having access to NTV in subregion s in election-year t. This regression corresponds to Regression 1. The stable groups assumption is not satisfied:

the share of people having access to NTV strictly increases in all regions between 1995 and 1999, the two elections used in the analysis. This a fuzzy design: the treatment of interest is whether a person has access to NTV, which varies within (subregion,year) cells.

18. **Fang and Gavazza (2011), Dynamic Inefficiencies in an Employment-Based Health Insurance System: Theory and Evidence.** *Tables 2, 3, 5, and 6, Column 3.*
    In regression Equation (7), the dependent variable is the health expenditures of individual j working in industry i in period t and region r, and the explanatory variables are individual effects, region specific time effects, and the job tenure of individual j. The death rate of establishments in industry i in period t and region r is used as an instrument for the job tenure of individual j. This regression is the 2SLS version of Regression 2 with controls. The stable groups assumption is presumably satisfied: between each pair of consecutive years, it is likely that there are some industry $\times$ region pairs where the death rate of establishments does not change. This a fuzzy design: the instrument of interest is whether a person's former employee closed down over the current year, which varies within (industry,year) cells.

19. **Gentzkow et al. (2011), The Effect of Newspaper Entry and Exit on Electoral Politics.** *Tables 2 and 3.*
    In regression Equation (2), the dependent variable is the change in voter turnout in county c between elections year t and t-1, and the explanatory variables are state $\times$ year effects, and the change in the number of newspapers in county c between t and t-1. This regression corresponds to Regression 2 with controls. The stable groups assumption is satisfied: between each pair of consecutive years, there are some counties where the number of newspapers does not change. This is a sharp design.

20. **Bloom et al. (2012), Americans Do IT Better: US Multinationals and the Productivity Miracle.** *Table 2, Columns 6-8.*
    In the regression in, say, Column 6 of Table 2, the dependent variable is the log of output per worker in firm i in period t, while the explanatory variables are firms and time fixed effects, the log of the amount of IT capital per employee $\ln(C/L)$, the interaction of $\ln(C/L)$ and an indicator for whether the firm is owned by a US multinational, the interaction of $\ln(C/L)$ and an indicator for whether the firm is owned by a non-US multinational, and some controls. This regression is similar to Regression 1 with some controls, except that it has three treatment variables. The stable groups assumption is presumably satisfied: between each pair of consecutive years, it is likely that there are some firms where the amount of IT capital per employee $\ln(C/L)$ does not change. This is a sharp design.

21. **Simcoe (2012), Standard Setting Committees: Consensus Governance for Shared Technology Platforms.** *Table 4, Columns 1-3.*

In regression Equation (5), the dependent variable is a measure of time to consensus for project i submitted to committee j, while the explanatory variables are an indicator for projects submitted to the standards track, a measure of distributional conflict, the interaction of the standards track and distributional conflict, and some controls variables. This regression is similar to Regression 1 with some controls, except that it has a measure of distributional conflict instead of committee fixed effects. The stable groups assumption is presumably not satisfied: it is unlikely that there is any committee where the measure of distributional conflict is equal to 0. This is a sharp design.

22. **Moser and Voena (2012), Compulsory Licensing: Evidence from the Trading with the Enemy Act.** *Table 2.*
In the regression equation in the beginning of Section III, the dependent variable is the number of patents by US inventors in patent class c at period t, and the explanatory variables are patent class and time fixed effects, the interaction of period t being after the trading with the enemy act and the number of licensed patents in class c, and some control variables. This regression corresponds to Regression 1 with some controls. The stable groups assumption is satisfied: there are patent classes where no patent was licensed. This is a sharp design.

23. **Forman et al. (2012), The Internet and Local Wages: A Puzzle.** *Tables 2 and 4.*
In regression Equation (1), the dependent variable is the difference between log wages in 2000 and 1995 in county i, and the explanatory variables are the proportion of businesses using Internet in county i in 2000, and control variables. This regression corresponds to Regression 2 with some controls. The stable groups assumption is satisfied: there are counties with no Internet investment in 2000. This a fuzzy design: the treatment of interest is whether a business uses Internet, which varies within (county,year) cells.

24. **Besley and Mueller (2012), Estimating the Peace Dividend: The Impact of Violence on House Prices in Northern Ireland.** *Table 1, Columns 3 and 5-7.*
In regression Equation (1), the dependent variable is the price of houses in region r at time t, while the explanatory variables are region and time fixed effects, and the number of people killed because of the civil war in region r at time t-1. This regression corresponds to Regression 1. The stable groups assumption is presumably satisfied: between each pair of consecutive years, it is likely that there are some regions where the number of people killed because of the civil war does not change. This is a sharp design.

25. **Dafny et al. (2012), Paying a Premium on Your Premium? Consolidation in the US Health Insurance Industry.** *Table 3.*
In regression Equation (3), the dependent variable is the the concentration of the hospital industry in market m and year t, and explanatory variables are time fixed effects, market

fixed effects, and the change in concentration in market m induced by a merger interacted with an indicator for t being after the merger. This regression corresponds to Regression 1. The stable groups assumption is satisfied: there are many markets where the merger did not change concentration. This is a sharp design.

26. **Hornbeck (2012), The Enduring Impact of the American Dust Bowl: Short- and Long-Run Adjustments to Environmental Catastrophe.** *Table 2.* In regression Equation (1), the dependent variable is, say, the change in log land value in county c between period t and 1930, and the explanatory variables are state × year fixed effects, the share of county c in high erosion regions, the share of county c in medium erosion regions, and some control variables. This regression is similar to Regression 1 with controls, except that it has two treatment variables. The stable groups assumption is satisfied: many counties have 0% of their land situated in medium or high erosion regions. This a fuzzy design: the treatments of interest are whether a piece of land is in high or in medium erosion regions, which varies within (county,year) cells.

27. **Bajari et al. (2012), A Rational Expectations Approach to Hedonic Price Regressions with Time-Varying Unobserved Product Attributes: The Price of Pollution.** *Table 5.*
In, say, the first regression equation in the bottom of page 1915, the dependent variable is the change in the price of house j between sales 2 and 3, and the explanatory variables are the change in various pollutants in the area around house j between sales 2 and 3, and some controls. This regression is similar to Regression 2 with controls, except that it has several treatment variables. The stable groups assumption is presumably satisfied: it is likely that for each pair of consecutive sales, there are houses where the level of each pollutant does not change. This is a sharp design.

28. **Dahl and Lochner (2012), The Impact of Family Income on Child Achievement: Evidence from the Earned Income Tax Credit.** *Table 3.*
In regression Equation (4), the dependent variable is the change in test scores for child i between years a and a-1, while the explanatory variables are the change in the EITC income of her family and some controls, and the change in the expected EITC income of her family based on her family income in year a-1 is used to instrument for the actual change of her family's EITC income. This regression is a 2SLS version of Regression 2 with controls, except that it does not have years fixed effects. The stable groups assumption is presumably satisfied: it is likely that for each pair of consecutive years, there are children whose family's expected EITC income does not change. This is a sharp design.

29. **Imberman et al. (2012), Katrina's Children: Evidence on the Structure of Peer Effects from Hurricane Evacuees.** *Tables 3-6.*

In regression Equation (1), the dependent variable is the test score of student i in school j in grade g and year t, and the explanatory variables are school and grade × year fixed effects, the fraction of Katrina evacuee students received by school j in grade g and year t, and some controls. This regression is a three-way fixed effects version of Regression 1. The stable groups assumption is satisfied: there are schools that did not receive any Katrina evacuee. This a fuzzy design: the treatment of interest is the proportion of evacuees in one's class, which varies within (school,grade,year) cells.

30. **Chaney et al. (2012), The Collateral Channel: How Real Estate Shocks Affect Corporate Investment.** *Table 5.*
In regression Equation (1), the dependent variable is the value of investment in firm i and year t divided by the lagged book value of properties, plants, and equipments (PPE), and the explanatory variables are firm and time fixed effects and the market value of firm i in year t divided by its lagged PPE, and some controls. This regression corresponds to Regression 1, with some controls. The stable groups assumption is presumably satisfied: it is likely that between each pair of consecutive years, there are firms whose market value divided by their lagged PPE does not change. This is a sharp design.

31. **Aaronson et al. (2012), The Spending and Debt Response to Minimum Wage Hikes.** *Tables 1, 2, and 5.*
In regression Equation (1), the outcome variable is, say, income of household i at period t, and the explanatory variables are household and time fixed effects, and the minimum wage in the state where household i lives in period t. This regression corresponds to Regression 1. The stable groups assumption is satisfied: between each pair of consecutive periods, there are states where the minimum wage does not change. This is a sharp design.

32. **Brambilla et al. (2012), Exports, Export Destinations, and Skills.** *Table 5.*
In the regression in, say, the first column of Table 2, the dependent variable is a measure of skills in the labor force employed by firm i in industry j at period t, and the explanatory variables are firm and industry × period fixed effects, the ratio of exports to sales in firm i at period t, and some controls. This regression corresponds to Regression 1, with some controls. The stable groups assumption is presumably satisfied: it is likely that between each pair of consecutive periods, there are firms whose ratio of exports to sales does not change. This is a sharp design.

33. **Faye and Niehaus (2012), Political Aid Cycles.** *Table 3, Columns 4 and 5, and Tables 4 and 5.*
In regression Equation (2), the dependent variable is the amount of donations received by receiver r from donor d in year t, and the explanatory variables are donor × receiver fixed effects, an indicator for whether there is an election in country r in year t, a measure of

alignment between the ruling political parties in countries r and d at t, and the interaction of the election indicator and the measure of alignment. This regression corresponds to Regression 1. The stable groups assumption is presumably not satisfied: it is unlikely that there are donor-receiver pairs that are perfectly unaligned. This is a sharp design.

6 Proofs

Two useful lemmas

The first lemma below generalizes Lemma 1 to ordered treatments. The second lemma proposes a decomposition of difference-in-differences different from that in Lemma 1, which relies on Assumption S1.

Lemma S1 *If Assumptions 1-3 hold and $D_{i,g,t} \in \{0, , ..., \overline{d}\}$.*

$$\begin{aligned} E(Y_{g,t}) - E(Y_{g,t'}) - (E(Y_{g',t}) - E(Y_{g',t'})) =& D_{g,t}ACR_{g,t} - D_{g,t'}ACR_{g,t'} \\ & - (D_{g',t}ACR_{g',t} - D_{g',t'}ACR_{g',t'}) \,. \end{aligned}$$

Lemma S2 *If Assumptions 1-3 and S1 hold (and $D_{i,g,t}$ is binary),*

$$E(Y_{g,t}) - E(Y_{g,t-1}) - (E(Y_{g',t}) - E(Y_{g',t-1})) = (D_{g,t} - D_{g,t-1})\Delta_{g,t} - (D_{g',t} - D_{g',t-1})\Delta_{g',t}.$$

Proof of Lemma S1

We have $E(Y_{g,t}) = E(Y_{g,t}(0)) + E[Y_{g,t}(D) - Y_{g,t}(0)]$. The result follows by decomposing similarly the three other terms $E(Y_{g,t'})$, $E(Y_{g',t})$ and $E(Y_{g',t'})$, using Assumption 3, and finally using the definition of $\text{ACR}_{g,t}$.

Proof of Lemma S2

By Lemma 1 and Assumption 1,

$$\begin{aligned} & E(Y_{g,t}) - E(Y_{g,t-1}) - (E(Y_{g',t}) - E(Y_{g',t-1})) \\ =& D_{g,t}\Delta_{g,t} - D_{g,t-1}\Delta_{g,t-1} - D_{g',t}\Delta_{g',t} + D_{g',t-1}\Delta_{g',t-1} \\ =& (D_{g,t} - D_{g,t-1})\Delta_{g,t} - (D_{g',t} - D_{g',t-1})\Delta_{g',t}. \end{aligned}$$

6.1 Proof of Theorem S1

Proof of the decomposition for the fixed-effect regression

First, we have

$$\begin{aligned}\sum_{g,t} N_{g,t}\varepsilon_{g,t}E(Y_{g,t}) =& \sum_{g,t} N_{g,t}\varepsilon_{g,t}[E(Y_{g,t}) - E(Y_{0,t})] \\ =& \sum_{g}\sum_{t=1}^{\bar{t}} \left[\sum_{t'\geq t} N_{g,t'}\varepsilon_{g,t'}\right] [E(Y_{g,t}) - E(Y_{g,t-1}) - (E(Y_{0,t}) - E(Y_{0,t-1}))] \\ =& \sum_{g}\sum_{t=1}^{\bar{t}} \left[\sum_{t'\geq t} N_{g,t'}\varepsilon_{g,t'}\right] [(D_{g,t} - D_{g,t-1})\Delta_{g,t} - (D_{0,t} - D_{0,t-1})\Delta_{0,t}] \\ =& \sum_{(g,t):D_{g,t}\neq D_{g,t-1},t\geq 1} \left[N_{g,t}(D_{g,t} - D_{g,t-1})\sum_{t'\geq t}\frac{N_{g,t'}}{N_{g,t}}\varepsilon_{g,t'}\right]\Delta_{g,t}. \end{aligned} \tag{41}$$

The first equality follows by (7). The second equality follows from summation by part and (7). The third equality uses Lemma 2. The fourth equality stems from the fact that by (7), the terms with $g = 0$ vanish.

Similarly,

$$\begin{aligned}\sum_{g,t} N_{g,t}\varepsilon_{g,t}D_{g,t} =& \sum_{g}\sum_{t=1}^{\bar{t}} \left[\sum_{t'\geq t} N_{g,t'}\varepsilon_{g,t'}\right] [D_{g,t} - D_{g,t-1}] \\ =& \sum_{(g,t):D_{g,t}\neq D_{g,t-1},t\geq 1} N_{g,t}(D_{g,t} - D_{g,t-1})\sum_{t'\geq t}\frac{N_{g,t'}}{N_{g,t}}\varepsilon_{g,t'}. \end{aligned} \tag{42}$$

The first result follows by combining (5), (41) and (42).

Proof of the decomposition for the first-difference regression

First, we have

$$\begin{aligned}& \sum_{(g,t):t\geq 1} N_{g,t}\varepsilon_{fd,g,t}\left(E(Y_{g,t}) - E(Y_{g,t-1})\right) \\ =& \sum_{(g,t):t\geq 1} N_{g,t}\varepsilon_{fd,g,t}\left(E(Y_{g,t}) - E(Y_{g,t-1}) - (E(Y_{0,t}) - E(Y_{0,t-1}))\right) \\ =& \sum_{(g,t):t\geq 1} N_{g,t}\varepsilon_{fd,g,t}\left[(D_{g,t} - D_{g,t-1})\Delta_{g,t} - (D_{0,t} - D_{0,t-1})\Delta_{0,t}\right] \\ =& \sum_{(g,t):t\geq 1} N_{g,t}\varepsilon_{fd,g,t}(D_{g,t} - D_{g,t-1})\Delta_{g,t}. \end{aligned}$$

The first equality follows from (7). The second equality follows from Lemma S2. The third equality follows from (7) again. The result follows by combining (25) with the last display.

Proof that $w^S_{g,t} \geq 0$ under Assumption 4 and if $N_{g,t}/N_{g,t-1}$ does not depend on g

Under Assumption 4, one has that $D_{g,t} = 1\{t \geq a_g\}$, with $a_g \in \{0, ..., \bar{t}+1\}$. Therefore, given the form of $w^S_{g,t}$, we just have to prove that for all g,

$$\sum_{t \geq a_g} N_{g,t}\varepsilon_{g,t} \geq 0. \tag{43}$$

Because $N_{g,t}/N_{g,t-1}$ does not vary across g for all $t \geq 1$, we have $N_{g,t} = N_{g,0}\gamma_t$ for some $\gamma_t \geq 0$. Moreover, $\varepsilon_{g,t} = D_{g,t} - D_{g,.} - D_{.,t} + D_{.,.}$. Let $\widetilde{\gamma}_t = \gamma_t / \sum_{t \geq 0} \gamma_t$, then $D_{g,.} = \sum_{t \geq a_g} \widetilde{\gamma}_t$, and $D_{.,.} = \sum_{t \geq 0} \widetilde{\gamma}_t D_{.,t}$. Hence,

$$\begin{aligned}
\frac{1}{N_{g,0}\sum_t \gamma_t} \sum_{t \geq a_g} N_{g,t}\varepsilon_{g,t} &= D_{g,.}\left(1 - D_{g,.} + D_{.,.}\right) - \sum_{t \geq a_g} \widetilde{\gamma}_t D_{.,t} \\
&= D_{g,.}\left(1 - D_{g,.} + \sum_{t < a_g} \widetilde{\gamma}_t D_{.,t}\right) - \left(\sum_{t \geq a_g} \widetilde{\gamma}_t D_{.,t}\right)(1 - D_{g,.}).
\end{aligned} \tag{44}$$

Now, because $D_{.,t} \leq 1$,

$$\sum_{t \geq a_g} \widetilde{\gamma}_t D_{.,t} \leq \sum_{t \geq a_g} \widetilde{\gamma}_t = D_{g,.}.$$

Hence, in view of (44),

$$\frac{1}{N_{g,0}\sum_t \gamma_t} \sum_{t \geq a_g} N_{g,t}\varepsilon_{g,t} \geq D_{g,.} \sum_{t < a_g} \widetilde{\gamma}_t D_{.,t} \geq 0.$$

Therefore, (43) and the result follows.

Proof that $w^S_{fd,g,t} \geq 0$

We just have to focus on the cases where $D_{g,t} \neq D_{g,t-1}$. Note that $\varepsilon_{fd,g,t} = D_{g,t} - D_{g,t-1} - (D_{.,t} - D_{.,t-1})$. Then, if $D_{g,t} - D_{g,t-1} = 1$, the numerator of $w^S_{fd,g,t}$ has the same sign as $1 - (D_{.,t} - D_{.,t-1})$, which is positive. if $D_{g,t} - D_{g,t-1} = -1$, the numerator of $w^S_{fd,g,t}$ has the same sign as $1 + (D_{.,t} - D_{.,t-1})$, which is also positive. Because the denominator sums terms that are always positive, it is positive as well. The result follows.

6.2 Proof of Theorem S2

The reasoning is exactly the same as in Theorem 1, except that we rely on Lemma S1 instead of Lemma 1.

6.3 Proof of Theorem S3

Proof for β_{fe}

First, remark that $\widehat{\beta}^X_{fe}$ is the coefficient of $D_{g,t}$ in the regression of $Y_{g,t} - X'_{g,t}\widehat{\gamma}_{fe}$ on group and time fixed effects and $D_{g,t}$. Therefore, by the Frisch-Waugh theorem,

$$\widehat{\beta}^X_{fe} = \frac{\sum_{g,t} N_{g,t}\varepsilon_{g,t}\left(Y_{g,t} - X'_{g,t}\widehat{\gamma}_{fe}\right)}{\sum_{g,t} N_{g,t}\varepsilon_{g,t}D_{g,t}}.$$

Let $Y^*_{g,t} = Y_{g,t} - X'_{g,t}\gamma_{fe}(\boldsymbol{X})$. Then, by definition of β^X_{fe}, $\gamma_{fe}(\boldsymbol{X})$ and the law of iterated expectations,

$$\beta^X_{fe} = \frac{\sum_{g,t} N_{g,t}\varepsilon_{g,t}E(Y^*_{g,t})}{\sum_{g,t} N_{g,t}\varepsilon_{g,t}D_{g,t}}.$$

Then, reasoning as in the proof of Theorem 1, we obtain

$$\sum_{g,t} N_{g,t}\varepsilon_{g,t}E(Y^*_{g,t}) = \sum_{g,t} N_{g,t}\varepsilon_{g,t}\left[E(Y^*_{g,t} - E(Y^*_{g,0}) - E(Y^*_{0,t}) + E(Y^*_{0,0})\right].$$

Now, under Assumptions S2, we can follow the same steps as those used to establish the first point of Lemma 1 to show that

$$\begin{aligned}&E(Y^*_{g,t} - E(Y^*_{g,0}) - E(Y^*_{0,t}) + E(Y^*_{0,0})\\ =&D_{g,t}\Delta^{TR}_{g,t} - D_{g,0}\Delta^{TR}_{g,0} - \left(D_{0,t}\Delta^{TR}_{0,t} - D_{0,0}\Delta^{TR}_{0,0}\right).\end{aligned}$$

The rest of the proof follows exactly that of Theorem 1.

Proof for β_{fd}

As above, we have, by the Frisch-Waugh theorem,

$$\widehat{\beta}^X_{fd} = \frac{\sum_{(g,t):t\geq 1} N_{g,t}\varepsilon_{fd,g,t}\left(Y_{g,t} - X'_{g,t}\widehat{\gamma}_{fd} - (Y_{g,t-1} - X'_{g,t-1}\widehat{\gamma}_{fd})\right)}{\sum_{(g,t):t\geq 1} N_{g,t}\varepsilon_{fd,g,t}\left[D_{g,t} - D_{g,t-1}\right]}.$$

Then, letting $Y^{**}_{g,t} = Y_{g,t} - X'_{g,t}\gamma_{fd}(\boldsymbol{X})$, we have:

$$\beta^X_{fd} = \frac{\sum_{g,t} N_{g,t}\varepsilon_{g,t}E(Y^{**}_{g,t})}{\sum_{g,t} N_{g,t}\varepsilon_{g,t}D_{g,t}}.$$

Then, reasoning as in the proof of Theorem 6, we obtain

$$\sum_{(g,t):t\geq 1} N_{g,t}\varepsilon_{fd,g,t}\left(E[Y^{**}_{g,t}] - E[Y^{**}_{g,t-1}]\right) = \sum_{(g,t):t\geq 1} N_{g,t}\varepsilon_{fd,g,t}\left[E[Y^{**}_{g,t}] - E[Y^{**}_{g,t-1}] - \left(E[Y^{**}_{0,t}] - E[Y^{**}_{0,t-1}]\right)\right].$$

As in the proof for β_{fe} above, we obtain, under Assumption S3,

$$E[Y^{**}_{g,t}] - E[Y^{**}_{g,t-1}] - \left(E[Y^{**}_{0,t}] - E[Y^{**}_{0,t-1}]\right) = D_{g,t}\Delta^{TR}_{g,t} - D_{g,0}\Delta^{TR}_{g,0} - \left(D_{0,t}\Delta^{TR}_{0,t} - D_{0,0}\Delta^{TR}_{0,0}\right).$$

The rest of the proof follows exactly that of Theorem 6.

6.4 Proof of Theorem S4

Fix t and let us suppose first that $E[N_{1,0,t}] > 0$. Then, by Assumption 4, we also have $E[N_{0,0,t}] > 0$. Moreover, using the convention that $P(A)E(U|A) = 0$ for any r.v. U if $P(A) = 0$, we have

$$\begin{aligned}
E[N_{1,0,t}]DID_{+,t} &= \sum_{g=0}^{\overline{g}} N_{g,t}P(D_{g,t} = 1, D_{g,t-1} = 0)E(Y_{g,t} - Y_{g,t-1}|D_{g,t} = 1, D_{g,t-1} = 0) \\
&- \frac{E[N_{1,0,t}]}{E[N_{0,0,t}]} \sum_{g=0}^{\overline{g}} N_{g,t}P(D_{g,t} = 0, D_{g,t-1} = 0)E(Y_{g,t} - Y_{g,t-1}|D_{g,t} = 0, D_{g,t-1} = 0) \\
&= \sum_{g=0}^{\overline{g}} N_{g,t}P(D_{g,t} = 1, D_{g,t-1} = 0)E(Y_{g,t}(1) - Y_{g,t}(0)|D_{g,t} = 1, D_{g,t-1} = 0) \\
&+ \sum_{g=0}^{\overline{g}} N_{g,t}P(D_{g,t} = 1, D_{g,t-1} = 0)E(Y_{g,t}(0) - Y_{g,t-1}(0)|D_{g,t} = 1, D_{g,t-1} = 0) \\
&- \frac{E[N_{1,0,t}]}{E[N_{0,0,t}]} \sum_{g=0}^{\overline{g}} N_{g,t}P(D_{g,t} = 0, D_{g,t-1} = 0)E(Y_{g,t}(0) - Y_{g,t-1}(0)|D_{g,t} = 0, D_{g,t-1} = 0) \\
&= \sum_{g=0}^{\overline{g}} N_{g,t}P(D_{g,t} = 1, D_{g,t-1} = 0)E(Y_{g,t}(1) - Y_{g,t}(0)|D_{g,t} = 1, D_{g,t-1} = 0) \\
&+ \sum_{g=0}^{\overline{g}} N_{g,t}P(D_{g,t} = 1, D_{g,t-1} = 0)E(Y_{g,t}(0) - Y_{g,t-1}(0)|D_{g,t-1} = 0) \\
&- \frac{E[N_{1,0,t}]}{E[N_{0,0,t}]} \sum_{g=0}^{\overline{g}} N_{g,t}P(D_{g,t} = 0, D_{g,t-1} = 0)E(Y_{g,t}(0) - Y_{g,t-1}(0)|D_{g,t-1} = 0) \\
&= \sum_{g=0}^{\overline{g}} N_{g,t}P(D_{g,t} = 1, D_{g,t-1} = 0)E(Y_{g,t}(1) - Y_{g,t}(0)|D_{g,t} = 1, D_{g,t-1} = 0). \qquad (45)
\end{aligned}$$

The third and fourth equalities follow from Assumption S4. Remark that (45) still holds if $E[N_{1,0,t}] = 0$.

Similarly, one can show that irrespective of whether $E[N_{0,1,t}] > 0$ or not,

$$E[N_{0,1,t}]DID_{-,t} = \sum_{g=0}^{\overline{g}} N_{g,t}P(D_{g,t} = 0, D_{g,t-1} = 1)E(Y_{g,t}(1) - Y_{g,t}(0)|D_{g,t} = 0, D_{g,t-1} = 1).$$

The result follows given the definition of W_{TC} and Δ^S.

6.5 Proof of Theorem S5

First, notice that

$$\text{ACR}^S = E\left[\frac{1}{N_{D,S}}\sum_{(i,g,t):t\geq 1} Y_{i,g,t}\left(\max(D_{g,t},D_{g,t-1})\right) - Y_{i,g,t}\left(\min(D_{g,t},D_{g,t-1})\right)\right]. \tag{46}$$

Reasoning as in the proof of Theorem 2, we get that for all $t \geq 1$,

$$\begin{aligned} N_{d,d',t}DID_{d,d',t} &= \sum_{g:D_{g,t}=d',D_{g,t-1}=d} N_{g,t}E[Y_{g,t}(\max(d,d')) - Y_{g,t}(\min(d,d'))] \\ &= \sum_{(i,g):D_{g,t}=d',D_{g,t-1}=d} E\left[Y_{i,g,t}(\max(D_{g,t},D_{g,t-1})) - Y_{i,g,t}(\min(D_{g,t},D_{g,t-1}))\right]. \end{aligned}$$

For all (g,t), there exists one $(d,d') \in \mathcal{D}^2$ such that $D_{g,t} = d'$ and $D_{g,t-1} = d$. Hence,

$$\begin{aligned} \sum_{t=1}^{\bar{t}} \sum_{(d,d')\in\mathcal{D}^2:d\neq d'} N_{d,d',t}DID_{d,d',t} &= \sum_{t=1}^{\bar{t}} \sum_{(d,d')\in\mathcal{D}^2} N_{d,d',t}DID_{d,d',t} \\ &= \sum_{(i,g,t):t\geq 1} E\left[Y_{i,g,t}(\max(D_{g,t},D_{g,t-1})) - Y_{i,g,t}(\min(D_{g,t},D_{g,t-1}))\right]. \end{aligned}$$

The result follows by definition of W_{TC} and Equation (46).

References

Aaronson, D., Agarwal, S. and French, E. (2012), 'The spending and debt response to minimum wage hikes', *American Economic Review* **102**(7), 3111–3139.

Acemoglu, D., Cantoni, D., Johnson, S. and Robinson, J. A. (2011), 'The consequences of radical reform: The french revolution', *American Economic Review* **101**(7), 3286–3307.

Aizer, A. (2010), 'The gender wage gap and domestic violence', *American Economic Review* **100**(4), 1847–1859.

Algan, Y. and Cahuc, P. (2010), 'Inherited trust and growth', *American Economic Review* **100**(5), 2060–2092.

Anderson, S. T. and Sallee, J. M. (2011), 'Using loopholes to reveal the marginal cost of regulation: The case of fuel-economy standards', *American Economic Review* **101**(4), 1375–1409.

Bagwell, K. and Staiger, R. W. (2011), 'What do trade negotiators negotiate about? empirical evidence from the world trade organization', *American Economic Review* **101**(4), 1238–1273.

Bajari, P., Fruehwirth, J. C., Kim, K. I. and Timmins, C. (2012), 'A rational expectations approach to hedonic price regressions with time-varying unobserved product attributes: The price of pollution', *American Economic Review* **102**(5), 1898–1926.

Baum-Snow, N. and Lutz, B. F. (2011), 'School desegregation, school choice, and changes in residential location patterns by race', *American Economic Review* **101**(7), 3019–3046.

Besley, T. and Mueller, H. (2012), 'Estimating the peace dividend: The impact of violence on house prices in northern ireland', *American Economic Review* **102**(2), 810–833.

Bloom, N., Sadun, R. and Van Reenen, J. (2012), 'Americans do it better: Us multinationals and the productivity miracle', *American Economic Review* **102**(1), 167–201.

Brambilla, I., Lederman, D. and Porto, G. (2012), 'Exports, export destinations, and skills', *American Economic Review* **102**(7), 3406–3438.

Bustos, P. (2011), 'Trade liberalization, exports, and technology upgrading: Evidence on the impact of mercosur on argentinian firms', *American Economic Review* **101**(1), 304–340.

Chandra, A., Gruber, J. and McKnight, R. (2010), 'Patient cost-sharing and hospitalization offsets in the elderly', *American Economic Review* **100**(1), 193–213.

Chaney, T., Sraer, D. and Thesmar, D. (2012), 'The collateral channel: How real estate shocks affect corporate investment', *American Economic Review* **102**(6), 2381–2409.

Dafny, L., Duggan, M. and Ramanarayanan, S. (2012), 'Paying a premium on your premium? consolidation in the us health insurance industry', *American Economic Review* **102**(2), 1161–1185.

Dahl, G. B. and Lochner, L. (2012), 'The impact of family income on child achievement: Evidence from the earned income tax credit', *American Economic Review* **102**(5), 1927–1956.

de Chaisemartin, C. (2010), A note on instrumented difference in differences. Working Paper.

de Chaisemartin, C. and D'Haultfœuille, X. (2018), 'Fuzzy differences-in-differences', *The Review of Economic Studies* **85**(2), 999–1028.

Dinkelman, T. (2011), 'The effects of rural electrification on employment: New evidence from south africa', *American Economic Review* **101**(7), 3078–3108.

Duflo, E. (2001), 'Schooling and labor market consequences of school construction in indonesia: Evidence from an unusual policy experiment', *American Economic Review* **91**(4), 795–813.

Duggan, M. and Morton, F. S. (2010), 'The effect of medicare part d on pharmaceutical prices and utilization', *American Economic Review* **100**(1), 590–607.

Duranton, G. and Turner, M. A. (2011), 'The fundamental law of road congestion: Evidence from us cities', *American Economic Review* **101**(6), 2616–2652.

Ellul, A., Pagano, M. and Panunzi, F. (2010), 'Inheritance law and investment in family firms', *American Economic Review* **100**(5), 2414–2450.

Enikolopov, R., Petrova, M. and Zhuravskaya, E. (2011), 'Media and political persuasion: Evidence from russia', *American Economic Review* **101**(7), 3253–3285.

Fang, H. and Gavazza, A. (2011), 'Dynamic inefficiencies in an employment-based health insurance system: Theory and evidence', *American Economic Review* **101**(7), 3047–77.

Faye, M. and Niehaus, P. (2012), 'Political aid cycles', *American Economic Review* **102**(7), 3516–3530.

Forman, C., Goldfarb, A. and Greenstein, S. (2012), 'The internet and local wages: A puzzle', *American Economic Review* **102**(1), 556–575.

Gentzkow, M., Shapiro, J. M. and Sinkinson, M. (2011), 'The effect of newspaper entry and exit on electoral politics', *American Economic Review* **101**(7), 2980–3018.

Hornbeck, R. (2012), 'The enduring impact of the american dust bowl: Short-and long-run adjustments to environmental catastrophe', *American Economic Review* **102**(4), 1477–1507.

Hotz, V. J. and Xiao, M. (2011), 'The impact of regulations on the supply and quality of care in child care markets', *American Economic Review* **101**(5), 1775–1805.

Hudson, S., Hull, P. and Liebersohn, C. (2015), Interpreting instrumented difference-in-differences, Technical report, Working Paper (available upon request).

Imbens, G. W. and Angrist, J. D. (1994), 'Identification and estimation of local average treatment effects', *Econometrica* **62**(2), 467–475.

Imberman, S. A., Kugler, A. D. and Sacerdote, B. I. (2012), 'Katrina's children: Evidence on the structure of peer effects from hurricane evacuees', *American Economic Review* **102**(5), 2048–2082.

Mian, A. and Sufi, A. (2011), 'House prices, home equity-based borrowing, and the us household leverage crisis', *American Economic Review* **101**(5), 2132–2156.

Moser, P. and Voena, A. (2012), 'Compulsory licensing: Evidence from the trading with the enemy act', *American Economic Review* **102**(1), 396–427.

Simcoe, T. (2012), 'Standard setting committees: Consensus governance for shared technology platforms', *American Economic Review* **102**(1), 305–336.

Wang, S.-Y. (2011), 'State misallocation and housing prices: Theory and evidence from china', *American Economic Review* **101**(5), 2081–2107.

Zhang, X. M. and Zhu, F. (2011), 'Group size and incentives to contribute: A natural experiment at chinese wikipedia', *American Economic Review* **101**(4), 1601–1605.