WIFO WORKING PAPERS
689/2024

Private Enforcement of Procurement Rules: The Heterogeneous Effect of the EU Remedies Directive

Eshien Chong
Michael Klien
Stéphane Saussier

Private Enforcement of Procurement Rules: The Heterogeneous Effect of the EU Remedies Directive

Eshien Chong (French Competition Authority), Michael Klien (WIFO), Stéphane Saussier (Panthéon Sorbonne)

WIFO Working Papers 689/2024
November 2024

Abstract

We examine the impact of the EU Remedies Directive on the use of direct awards. Theoretically, strengthening bid protest mechanisms should reduce the use of direct awards as aggrieved bidders are more likely to challenge such awards. By utilising the staggered implementation of the Remedies Directive across EU countries, we are able to test this hypothesis using a large database of public procurement contracts. While we do not find an effect on average, we observe a systematic effect for public buyers with a high initial propensity to use direct awards. In this group, the estimates suggest a reduction in the use of direct awards of more than 50 percentage points two years after implementation.

E-Mail: michael.klien@wifo.ac.at

2024/1/W/23128

Media owner (publisher), producer: Austrian Institute of Economic Research
1030 Vienna, Arsenal, Objekt 20 | Tel. (43 1) 798 26 01 0 | https://www.wifo.ac.at
Place of publishing and production: Vienna

Free download: https://www.wifo.ac.at/publication/pid/55574518

Private enforcement of procurement rules: The heterogeneous effect of the EU Remedies Directive

Eshien Chong* Michael Klien† Stéphane Saussier‡

November 18, 2024

Abstract

We examine the impact of the EU Remedies Directive on the use of direct awards. Theoretically, strengthening bid protest mechanisms should reduce the use of direct awards as aggrieved bidders are more likely to challenge such awards. By utilizing the staggered implementation of the Remedies Directive across EU countries, we are able to test this hypothesis using a large database of public procurement contracts. While we do not find an effect on average, we observe a systematic effect for public buyers with a high initial propensity to use direct awards. In this group, the estimates suggest a reduction in the use of direct awards of more than 50 percentage points two years after implementation.

Keywords: Direct awards, public procurement, bid protests, remedies, EU

JEL Codes: L14, L51, D73, D82

*French Competition Authority and University of Paris I Sorbonne. <chong.iae@univ-paris1.fr>
†WIFO - Austrian Institute of Economic Research <Michael.Klien@wifo.ac.at>
‡University of Paris I Sorbonne. <saussier@univ-paris1.fr>

1 Introduction

European public procurement rules govern how contracts worth more than 2,400 billion EUR (16% of the GDP in EU member countries) are awarded each year. There is considerable debate in the area of public procurement regarding the use of direct awards. While many practitioners view them as cost-effective and flexible, direct awards are also often seen as a discretionary way to circumvent competition, which could potentially invite corruption and favoritism. The European Court of Justice (ECJ) took a particularly skeptical stance, calling the inappropriate use of direct awards the *"most serious breach of Community law in the field of public procurement"*[1].

To reduce the perceived too frequent use of direct awards, the EU amended its procurement rules in 2007 by adopting the EU Remedies Directive (2007/66/EC). At its core, the Remedies Directive seeks to strengthen private enforcement of procurement rules. The idea was to provide an incentive to aggrieved bidders who had not been awarded a public contract, to bring an illegal contract award to court. The directive aimed to increasing the reward for appealing bidders, as under the new system successful challenges would imply that a new tender for the contract would have to be launched.

There is little empirical evidence to indicate the extent to which private enforcement can enhance compliance with procurement rules. In a study conducted 10 years after the adoption of the Directive, an evaluation by the European Commission concluded that the effectiveness of the remedies system was challenging to ascertain: *"The measurement of the effectiveness of the remedies system in this aspect is however very difficult as it consists of preventive impacts: i.e. the mere existence of the Remedies Directives avoids breaches of EU public procurement law before they occur. The practical effect of the deterrence role of the Remedies Directives is that fewer illicit practices can be observed and hence fewer review decisions are requested. However, the importance of the absence of complaints (i.e. the*

[1]Case C-26/03, Judgment of the Court (First Chamber) of 11 January 2005. Stadt Halle and RPL Recyclingpark Lochau GmbH v Arbeitsgemeinschaft Thermische Restabfall- und Energieverwertungsanlage TREA Leuna.

number of complaints that would have been lodged in the absence of the Remedies Directives) cannot be reasonably estimated."European Commission (2017, p.44)

Consequently, the European commission's evaluation was based exclusively on information from surveys and public consultations (perceptions of impact) and the fact that there were more than 50,000 first-instance remedy decision in the period 2009-2012.

Despite these complications, the primary objective of our study is to estimate the deterrence effect of the regulation. We build upon the theoretical framework by Marshall et al. (1994), hereafter referred to as MMR94, which generates clear-cut and testable predictions regarding the impact of strengthening the remedies system on the behavior of public buyers. In our empirical analysis, we leverage the staggered implementation of the Remedies Directive across EU countries to construct a counterfactual, which should allow us to obtain a reasonable estimate of procurement practices in the absence of the Remedies Directive. Unlike the evaluation by European Commission (2017), we do not attempt to analyze the effect of the directive on the number of complaints but on the use of direct awards, the most prominent and potentially illicit practice. While the number of complaints or remedy decisions may be a mediating factor that theoretically could fall or rise due to the Directive, the expectation for the use of direct awards is theoretically straightforward (as discussed in the theoretical section).

The paper is related to several strands of existing literature. Firstly, it complements the general theoretical literature on private enforcement, including works such as (Kovacic, 1998), McAfee et al. (2008), and Kaplow (2017). As this literature takes the effectiveness of private enforcement for granted, our study supplies empirical evidence to support these theories. It is noteworthy that the rich theoretical literature on private enforcement since Becker and Stigler (1974) is not matched by studies working on its empirical foundations. This is not only true for procurement, but also for other areas such as antitrust enforcement. As Peyer (2012) observes, despite the increased reliance of EU community law on private enforcement there is a paucity of systematic empirical evidence. Secondly, as far as we know the present paper

is the first causal empirical test of the hypothesis laid down in Marshall et al. (1994), which posits that strengthening bid protest mechanisms will reduce the number of direct awards used by public buyers. As the Remedies Directive was specifically designed to enhance the bid protest scheme, our empirical design allows us to quantify the associated deterrent effect. Finally, the paper is related to the broader literature analyzing mechanisms to increase compliance with procurement rules such as Di Tella and Schargrodsky (2003) or Olken (2007). Previous studies have analyzed the potential of increased government monitoring or higher public sector wages, whereas our analysis confirms that private enforcement is a potential avenue for anti-corruption reforms.

The following section presents a theoretical analysis of the potential impact of a strengthened bid protest mechanism on the behavior of procuring parties and the likelihood of procurement officers utilizing direct awards. Section 3 provides a detailed examination of the regulatory framework and the EU Remedies Directive. Section 4 elucidates the data and empirical strategy employed in the study, while Section 5 presents and discusses the findings. Section 6 offers a concluding summary.

2 Theory: Bid protests, private enforcement and the choice of award mechanism

To understand the mechanism through which the EU Remedies Directive should affect the way public contracts are awarded, we revisit the core features of the model developed in Marshall et al. (1994) (henceforth MMR94), which analyzes how the introduction of a bid-protest scheme affects the use of direct awards. Several aspects of their model lend themselves to study the Remedies Directive. In their model, there are two award mechanisms, namely competitive procurement and sole-sourcing, i.e. direct awards. While the former has strictly lower direct cost, there are cases when (large) indirect cost make sole-sourcing more appropriate. Indirect cost of competitive procurement range from administrative cost of holding

a tender to switching cost related to technological lock-ins. Moreover, rents from direct awards might also induce high effort in firms supplying products to the government[2]. The optimal choice between competitive procurement and direct award is therefore dependent on the project-specific direct and indirect cost of the two award mechanisms.

Indeed, in recent years a considerable number of contributions have analyzed this trade-off underlying the model of MMR94 in more detail. Starting with the very strong result in Bulow and Klemperer (1996), implying that competitive bidding is always preferable to other award mechanisms,[3] subsequent research has stressed that competitive procurement might not always be optimal. For instance, Manelli and Vincent (1995) show that open competitive tenders may be inefficient when the procurer cares not only about the price, but also about the quality of a good. In their setting, open competitive tenders may lead a procurer to end up with a trading partner offering a low quality good.

More generally, when one takes into account that there may be difficulties in specifying fully the scope of cooperation between contracting partners, and therefore the project to be procured, competitive procedures may not always be the most efficient way of awarding a contract. As suggested by Goldberg (1977) and more recently in Bajari et al. (2009), negotiations and direct-awards may outperform open competitive tenders because contracting parties can exchange pre-contract information. While competitive tenders stimulates competition it also stifles communication. In a project, sellers may have important information about suitable specifications of a project, which the procurer may find useful when drafting a contract. Such pre-contract information can help in avoiding costly *ex post* haggling and adaptation. Hence, when projects are complex, and when *ex post* adaptations can be important and contracts are difficult to specify *ex ante*, negotiations may yield benefits that outweigh competitive tenders, by facilitating communication and coordination.

[2]This argument has been developed further in Calzolari and Spagnolo (2009).

[3]In their seminal contribution, Bulow and Klemperer (1996) formally show that the benefits of an extra bidder *always* outweigh the benefits for an auctioneer using the optimal auction design without the extra participant. This result suggests that open competitive tenders unambiguously performs better than any other mechanisms as a means to select a trading partner.

Finally, it is also worthwhile to note that open competitive tenders are far from being the norm in private sector procurement practices (Leffler et al., 2003; Bajari et al., 2009). Leffler et al. (2003) found that there is about 50/50 split between competitive tenders and negotiation in private sector sales of timber in North Caroline and Mississippi. For private sector non residential construction projects in North California, unrestricted open competitive tenders are used only in 18% of the projects procured (Bajari et al., 2009). Kelman (1990) also note that private firms do not use open competitive tenders as much as in public procurement. They also tend to leave higher margins to suppliers, change suppliers less often, and are more satisfied with the quality of goods and services.

Taken together, the economic literature has identified a number of compelling cases for using direct awards instead of competitive bidding. This is well embodied in the model in MMR94, where an optimizing procurement strategy will consider the sum of direct and indirect cost when choosing an award mechanism. However, in the setup of MMR94, the award mechanism is chosen by a procurement official whose objectives differ from the government. This agency problem manifests in a bias towards direct awards, which has several potential origins. On the one hand, favoritism for a specific firm or the features of a certain product are more easily pursued using direct awards. Alternatively, procurement officials might undervalue the direct cost of procurement compared to indirect costs such as switching the supplier or administrating a competitive tender. In either case, this bias generates instances of unjustified direct awards.

This important feature of the model in MMR94, the potential abuse of direct awards, has received ample attention in procurement research. Particularly with respect to corruption and favoritism, the abuse of discretionary power has been a long-standing subject. From this point of view, direct awards can be regarded as a way to manipulate market attribution. For instance, in a survey of 82 Norwegian exporters, the main purpose of bribe paying firms is to obtain a public contract through direct award and without any call for tenders (Søreide,

2006).[4] In the same vein, Tran (2011) offers some compelling empirical evidence on how a larger scope of discretionary power granted to a procurement official can be abused. Using internal bribery records of a firm participating in public procurement of a medical device in Vietnam, he found that procurement procedures matter. The study found open auctions less prone to corruption compared to direct awards.

How can the risk of granting discretionary power to public procurers be mitigated? The literature on economics of corruption (Tirole, 1992; Laffont, 2001) suggests that the incidence of corruption will depend on the wage rate of the bureaucrat (efficiency wage), the monitoring system and legal remedies.[5] Empirically, cross-country comparison by van Rijckeghem and Weder (2001) suggests that corruption may be lower in the public sector when relative wages are higher. Di Tella and Schargrodsky (2003) and Olken (2007) provide some empirical support to the idea that more government audits and monitoring can help to reduce corruption.

The focal point of this paper is better enforcement of procurement rules through bid protest mechanisms. The introduction or strengthening of a bid protest system can also be understood as a way to fight corruption and favoritism. Instead of scrutiny through administrative controls, the remedy system generates 'decentralized enforcement' of public procurement rules (Kovacic, 1998).

The availability of remedy procedures for excluded bidders, who can benefit from the detection and cancellation of unjustified direct awards, is also the main mechanism in the model of MMR94. In a situation without the possibility of a bid protest, the procurement official would choose the award mechanism based on the direct and indirect cost, weighted depending on the intensity of the bias. After the introduction of a bid protest scheme, the procurement official has to factor in the possibility that excluded bidders protest against their

[4]Note, however, that open competitive tenders (auctions) are not immune to corruption in itself, as shown by anecdotal evidence and in theory (Celentani and Ganuza, 2002; Burguet and Che, 2004; Compte et al., 2005). This can happen because even when open competitive tenders are used to select a contractor, the procurer may still have some discretionary power. This is the case when other dimensions of a good or service to be procured matter in addition to the price (e.g. quality).

[5]See Aidt (2003) for a review.

exclusion. However, for the excluded bidder it is uncertain whether the exclusion was justified (high indirect cost) or not. In this setting, MMR94 identify a unique perfect Bayesian equilibrium in which protests are an effective, albeit imperfect, deterrent of unjustified direct-awards.[6]

The hypothesis based on the theoretical framework by MMR94 is that the strengthened bid protest scheme will reduce the propensity to use direct awards as award mechanism. The details on how the Remedies Directive has affected the incentives for the involved parties is explained in the next section.

3 The EU Remedies Directive

Already before the Remedies Directive, EU procurement rules specified that direct awards should be used only in a limited set of cases. For instance if a previous open tender received no offers, or if for technical or legal reasons (e.g. intellectual property rights) only a certain supplier can offer the required service.[7] To tackle the perceived too frequent use of non-competitive direct awards, the EU Remedies Directive (directive 2007/66/EC) was adopted in 2007. The directive should have been transposed by the end of 2009, but a considerable number of countries missed the implementation deadline. The actual date of implementation by country is shown in Table 1.

The directive contained two main features. Firstly, a standstill period of at least 10 days between choosing a winning offer and the signature of the contract. By giving loosing bidders time to examine the decision and file a bid protest before the contract is concluded, the directive strengthens the legal position of aggrieved bidders because any protest before the signature results in a automatic suspension of the procurement process until the review body takes its decision. Disregard of the standstill period obliges national review bodies to

[6]It comes at the cost of over-deterrence, as high protest costs for procurement officials (e.g. project delays) lead to situations where competitive bidding is chosen albeit a direct-award would be optimal. In the paper, MMR94 also take into account other features of the protest process such as settlements. This is, however, beyond the analysis of this paper.

[7]For the full set of exemptions see article 32 of the procurement directive 2014/24/EU.

Table 1: Implementation Dates of Directive

Country	**Day**	**Month**	**Year**
AT	16	03	2010
BE	16	02	2010
BG	12	07	2010
CY	22	11	2010
CZ	01	01	2010
DE	24	04	2009
DK	01	06	2010
EE	01	07	2010
ES	05	08	2010
FI	01	06	2010
FR	07	05	2009
GR	30	09	2010
HU	01	01	2010
IE	25	03	2010
IT	20	10	2010
LT	02	03	2010
LU	01	12	2010
LV	15	06	2010
MT	01	06	2010
NL	19	02	2010
PL	01	01	2010
PT	14	01	2011
RO	12	03	2009
SE	15	07	2010
SI	03	07	2011
SK	01	01	2010
UK	20	12	2009

Source: European Commission (2017)

render the signed contract void.

Secondly, the directive adopted more stringent rules against direct awards that did not live up to minimum requirements of transparency. Typical cases of this sort are when a contracting authority directly awards a public contract to a contractor, without the required prior notice publication in the Official Journal of the European Union (OJEU). Under the new rules national courts are able to render contracts ineffective if they have been illegally awarded without transparency. Before the Remedies Directive, excluded bidders were only able to claim damages, which is notoriously difficult to prove in court.

We expect that the Remedies Directive worked very similar to the bid protest mechanism discussed in MMR94. Compared to the previous situation, the directive increased the expected benefit of a protest for excluded bidders. The reason is that now the firms have a renewed chance to win the contract in a new tender. Given the higher propensity to challenge a contract award, procurement officials will reduce the number of direct awards. Hence, a very straightforward prediction from MMR94 is that the Remedies Directive will lead to fewer direct awards. The extent to which this has been the case is subject to the next section.

Interestingly, in contrast to the prediction for the use of direct awards, it is theoretically ambiguous whether the directive will lead to an actual increase in bid protests. If public buyers expects considerably more bid protests, the expectation of being challenged could lead to a strong reduction in direct awards, which might eventually lead to fewer bid protests overall. This also motivates why the most clear and direct evidence for the effect of the directive should be observable in the use of direct awards.

4 Data and empirical strategy

4.1 Data and descriptive statistics

To compare the use of direct awards before and after the Remedies Directive, our main source of data is the Official Journal of the European Union (OJEU), which contains contract award notifications from all EU member states.[8] All public procurement contracts above certain value thresholds[9] fall under EU-procurement rules and have to be notified with the OJEU, which are then publicly available in Tenders Electronic Daily (TED). The individual award notices contain information on the public buyer, the procured good or service, the value of the contract as well as the winner of the contract. The resulting micro dataset contains contract awards from all 27 member states that were full members of the EU in 2008. Although the data would be available from 2006 to 2020, we restrict the analysis to the period 2006 to 2015 because the EU procurement system was subject to important changes from 2016 onwards.[10] As the implementations of the Remedies Directive vary between March 2009 and July 2011, this leaves us with at least 2 complete year of contract awards before and after the directive.[11]

During our observational period, there were roughly 4,000,000 awards in the TED database, thereof 118,000 (3%) direct awards. Although the use of direct awards is considered a general problem across the EU, the large variation in the use of direct awards suggests that the rules were applied very differently across public buyers and across member states (see first column in Table 2). In some countries direct awards accounted for more than 10% of all observed tenders, whereas in others it was used very rarely (<1%). Moreover, also within the member

[8]This data base is also used by the European Commission and Eurostat to measure public procurement activity in the EU.

[9]The thresholds are adjusted to GDP growth and updated biannually. In 2011, the thresholds for services and supplies was at 200K. For public buyers in certain sectors such as water, energy, transport or telecom, higher thresholds applied. In the case of public works, the threshold was at 5000K in 2011.

[10]Most importantly, the updated procurement directives 2014/24/EU, 2014/25/EU, and 2014/23/EU, which for instance made the use of e-procurement mandatory, went into effect in 2016.

[11]Most of the empirical analysis is restricted to comparisons between countries that have already implemented the directive to countries that have not yet implemented the directive, thus discarding all data after the last country has implemented the directive.

states, the use of direct awards varied strongly across public buyers. As the right part of the tables shows, direct awards were used as the standard award mechanism by some public buyers (in more than 20% of all their projects) whereas the bulk of public buyers used it very infrequently (in less than 5% of all projects). Although the high shares do not necessarily imply illegal practices, there are obviously systematic differences regarding the application of EU procurement rules.

If we combine the data from the OJEU with the information on the implementation date, the effect of the directive on the use of direct awards appears sizable, but only clearly visible for the sub-population that used them frequently (see Figure 1). Public buyers with a high initial share of direct awards exhibit a considerable drop from 50 to 20 percentage points after the implementation of the directive. Public buyers with an intermediate initial share (15% direct awards before the directive) also registered a decrease that coincides with the start of the directive, but with a much smaller magnitude of several percentage points. The group of public buyers with very few direct awards prior to the directive appears to experiences a small increase. Even if there are several problems associated with this simple comparisons, we interpret the striking changes corresponding with the introduction of the directive as a first indication to its effect.

4.2 Empirical strategy

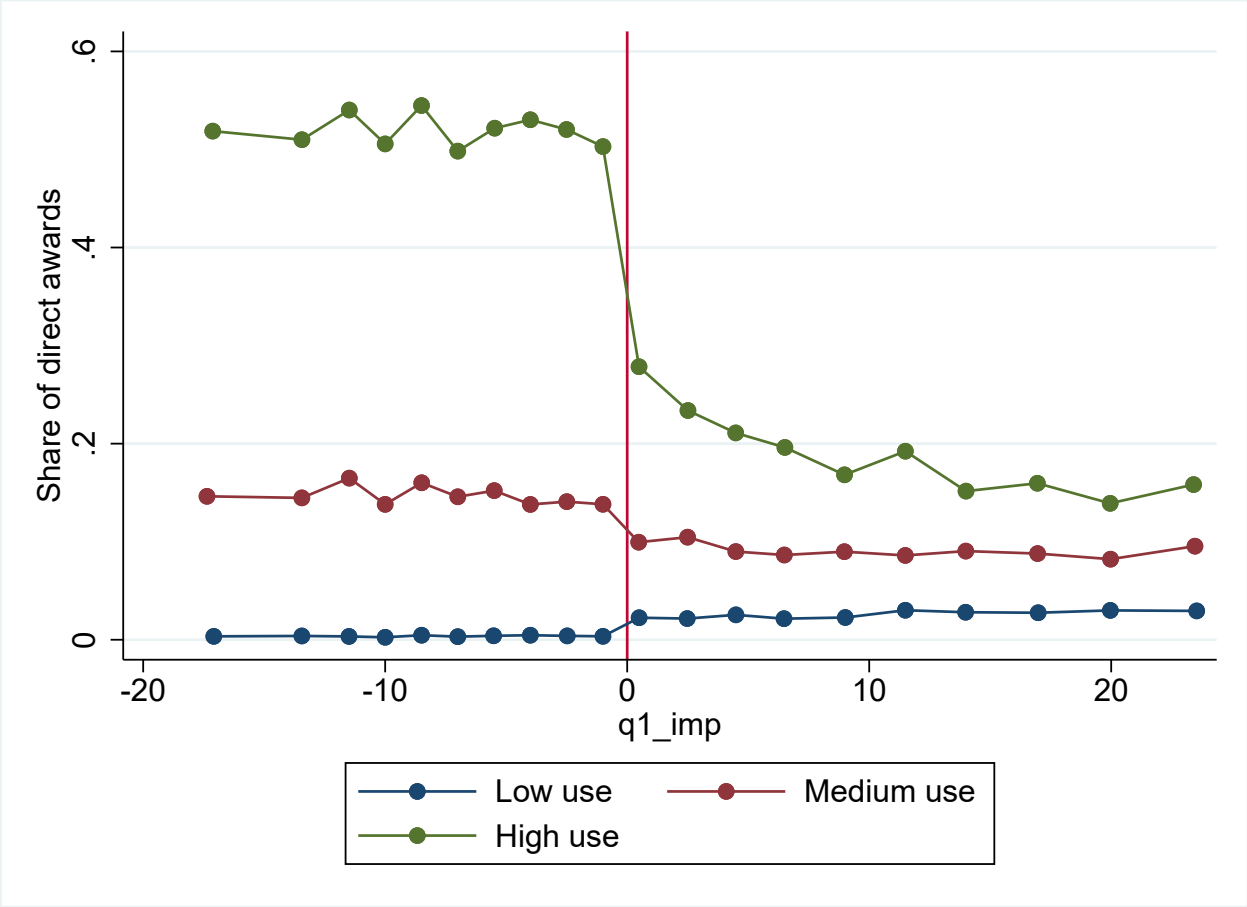
To analyze its effect we exploit the staggered implementation of the Remedies Directive and employ state-of-the art event study estimators. In recent years, there has been a proliferation in research analyzing empirical designs, in which units are treated at different points in time, and researchers seek to leverage this differential timing. The analysis is more complicated than in a standard Difference-in-Difference (DiD) case because all units are treated at some point, leaving no never-treated control-group. An important problem identified by the literature is that the canonical two-way fixed effects model (TWFE) – a specification with treatment dummy and unit and time fixed effects – can lead to biased treatment estimates.

Table 2: Share of direct awards

Country	**% Direct awards**	**Within country distribution**		
		High	**Medium**	**Low**
AT	2	3	4	93
BE	2	2	3	94
BG	4	2	21	76
CY	0	0	1	99
CZ	14	25	22	53
DE	5	8	9	83
DK	1	1	2	97
EE	11	17	36	47
ES	7	9	13	78
FI	0	0	0	100
FR	1	1	4	95
GR	3	6	4	90
HU	14	20	20	60
IE	0	0	0	100
IT	5	9	7	84
LT	15	32	31	37
LU	1	1	4	95
LV	4	5	17	79
MT	0	0	0	100
NL	2	3	6	91
PL	6	8	8	84
PT	1	2	1	97
RO	11	17	20	63
SE	1	1	2	98
SI	4	6	18	76
SK	18	36	11	53
UK	1	1	2	97
Total	4	6	8	86

High refers to share of public buyers with more than 20% direct awards, Medium for cases above 5% but below 20%, and Low for cases below 5%.

Figure 1: Use of direct awards before and after the directive for different initial levels of direct awards



As units switch between treatment and control group, the dynamics of the treatment effect become important. The resulting problem is outlined in Borusyak and Jaravel (2017), who demonstrate that the TWFE model generates 'forbidden' comparisons between not-yet and already treated units. In the presence of differences between short and long-term effects, the aggregate effect (averaged over time and cohorts) has no causal interpretation because the weights of individual treatment effects can be negative. While a fully dynamic model – with fixed effects for units, time and relative treatment time – avoids this problem, Borusyak and Jaravel (2017) show that it is essentially unidentified.

Several papers have proposed ways to address these problems, and we follow Borusyak et al. (2021) who introduce both a robust but also efficient imputation estimator that builds on the two standard DiD assumptions: A parallel trend assumption that potential outcomes without treatment are characterized by parallel trends and the assumption that there are no anticipatory effects. We will analyze both assumptions and run tests to identify possible violations in the empirical results section.

The estimator proceeds in three steps. First, the estimator runs a model for the non-treated potential outcomes ($=Y_{it}(0)$) using only the not-yet-treated observations. In our setup, Y_{it} measures the share of direct awards by public buyer i in quarter t. Public buyers are identified in the data using the name of the procuring entity.[12]

In the second step this estimate is then extrapolated to impute the counterfactual for the treated units, which allows us to obtain the individual level treatment effect $\tau_{it} = Y_{it} - Y_{it}(0)$, where Y_{it} is the observed outcome for the unit i at time t. In the third step, the individual estimates are averaged to get an estimate of the overall treatment effect of interest. We consider both static treatment effects but also a dynamic version that shows how the effect evolves with increasing time. In the dynamic version, we estimate separate treatment effects for each quarter after the treatment until the 11th quarter. After the 11th quarter, all units are treated. Regarding pre-trends, we use up to four quarters before the implementation.

[12]We use a deduplication approach to identify the same units even if spelled slightly differently that is based on cosine string-similarity with a term frequency–inverse document frequency (tf-idf) weighting.

While longer pre-trends would be numerically feasible, the power of the joint tests would be lower thereby defeating its purpose (see (Borusyak et al., 2021)).

As the descriptive statistics have indicated that the effect of the treatment seems to vary greatly by initial use of direct awards (before the directive), we estimate separate regressions for three groups of public buyers with different initial levels: $[0, 0.05), [0.05, 0.20), [0.20, 1.00)$. The choice of the thresholds is clearly arbitrary but alternative estimations with different thresholds deliver a very similar picture.

There are several reasons why we chose this estimator for our baseline regressions. Firstly, estimating individual level treatment effects through imputation is intuitive and also flexible. It is quite straightforward to calculate aggregated treatment effects through weighting therefore avoiding the pitfall of negative weights. Secondly, the flexibility is a large advantage in unbalanced datasets where imputation also allows comparisons for units even if they are not observed in every quarter. This is the case in the underlying dataset when we analyze procurement practices on the level of public buyers. As many institutions do not procure goods in every quarter, they would otherwise be discarded.[13] Thirdly, the efficiency of the estimator combines a robust approach without unnecessarily discarding statistical power. Borusyak et al. (2021) show much smaller variances compared to other frequently used estimators such as Callaway and Sant'Anna (2021) or Sun and Abraham (2021). Fourth, the higher efficiency also lends more credibility to testing the parallel trends assumption, where low-powered estimators could potentially shroud evidence suggesting a violation of the assumption. Finally, despite the fact that we use the imputation estimator as our main approach, we will also apply multiple alternative estimators developed in Callaway and Sant'Anna (2021), Sun and Abraham (2021), and Wooldridge (2021) to gauge the robustness of our baseline results. Due to the different sizes of public buyers, with some having only a single procurement per year whereas others procure multiple items each quarter, we opt to weight the regressions

[13] This is also the reason why we do not use the estimator proposed by De Chaisemartin and d'Haultfoeuille (2020) which uses only public buyers which procure in every quarter and as a result drops a very large part of the data.

with the number of contracts that a public buyer awards in each quarter. This should help to ensure that the estimates are less affected by noise from small authorities.

To account for the fact that the treatment varies on the country level, we cluster the standard errors on the country level. As this leaves us with only 27 clusters, we also present robustness tests that our baseline results carry over to other standard error assumptions.

5 Empirical results

5.1 Baseline estimates

We start by presenting the results from our baseline model where we estimate a static and a dynamic treatment effect for the three groups of public buyers (see Table 3). For the group with less than 5% direct awards prior to the directive, which is by far the largest group, we find a very small and statistically insignificant effect of the directive. Also in the dynamic version, the initially positive and significant effect dissipates over time and switches to a negative sign 8 quarters after the treatment, before reverting back. All estimated effects remain below a 1% point change in the use of direct awards, and are thus overall very small. There is also no evidence of a significant differential pre-trend.

Regarding the medium users, with an initial share of direct awards between 5 and 20%, the average effect is negative but statistically insignificant. The average effect is estimated at 4.3% points reduction. In the dynamic version, the treatment effect is negative but statistically significant only for a few time horizons. Moreover, the pre-trends in the two quarters prior to implementation of the Remedies Directive are statistically significant, questioning the assumption of no pre-trends. The pre-trends are negative, suggesting a reduction in the use of direct-awards already before the implementation of the directive. This could be interpreted as an anticipation effect because public buyers appear to have used fewer direct awards already two quarters before the implementation date.

The largest effect of the directive, as already suggested by the descriptive statistics, can

be found for those public buyers which frequently used direct awards (in more than 20% of all their awards before the directive). On average, the effect of the directive amounted to a reduction of 40 percentage points, which is a very large effect (see column 5). The effect size also has a clear dynamic path in that it increases continuously over the quarters following the implementation of the directive and reaches values above 40 percentage points after 3 quarters. The effect after 11 quarters, the maximum time frame until all units are treated, reaches a value of 67%, which is considerably larger than the short-term effect. Importantly for the event study approach is the fact there there is no evidence of a statistically significant differential pre-trend. The estimated effects in the four quarters before the treatment are not only statistically insignificant, but also their size is very small compared to the estimated treatment effects and varies around zero.

5.2 Alternative estimators

As outlined in empirical strategy section, the imputation estimator by Borusyak et al. (2021) is only one of several recently suggested estimators to analyze event study designs. To gauge the sensitivity of our results with respect to the chosen estimator, we also present results using the alternatives developed in Sun and Abraham (2021), Callaway and Sant'Anna (2021), Wooldridge (2021), as well as a standard TWFE-Modell. The obtained dynamic model effects for public buyers with a high-initial share of direct awards is presented graphically in Figure 2, the table with the results for all three subgroups are show in 6.

The results from the different estimators are qualitatively similar and suggest a statistically significant and sizeable reduction in the use of direct awards for high-use public buyers. There are, however, noticeable differences in the estimates effect size and a growing spread in the long run. Several patterns emerge: Firstly, the TWFE results yields the lowest effect size of the Remedies Directive, suggesting a reduction of 30 to 40% after 11 quarters. The highest effect sizes are obtained by the estimator from Callaway and Sant'Anna (2021), indicating a reduction in direct awards of more than 70% after 10 quarters. Secondly, some estima-

Table 3: Baseline results

	Low use		Medium use		High use	
	(1)	(2)	(3)	(4)	(5)	(6)
tau	0.00314		-0.0433		-0.419***	
	(1.87)		(-1.71)		(-9.55)	
tau0		0.00295***		-0.0375		-0.319***
		(4.02)		(-1.77)		(-3.98)
tau1		0.00733***		-0.0161		-0.220***
		(7.24)		(-0.85)		(-3.50)
tau2		0.00453***		-0.0295*		-0.364***
		(4.16)		(-2.10)		(-8.97)
tau3		0.00446*		-0.0366		-0.479***
		(2.43)		(-1.23)		(-9.64)
tau4		0.00256		-0.0516		-0.535***
		(0.74)		(-1.29)		(-7.12)
tau5		0.00408		-0.0609		-0.449***
		(1.40)		(-1.95)		(-8.88)
tau6		0.000841		-0.0593*		-0.424***
		(0.40)		(-2.38)		(-9.83)
tau7		-0.00721*		-0.0366		-0.446***
		(-2.08)		(-0.74)		(-6.44)
tau8		0.000680		-0.0929*		-0.464***
		(0.15)		(-2.55)		(-8.09)
tau9		0.00606***		-0.0503*		-0.545***
		(4.06)		(-2.16)		(-19.76)
tau10		0.00612***		-0.0206		-0.670***
		(4.95)		(-1.10)		(-21.79)
pre1	-0.00113	-0.00113	-0.0283*	-0.0283*	0.0283	0.0283
	(-0.68)	(-0.68)	(-2.14)	(-2.14)	(0.38)	(0.38)
pre2	0.0000296	0.0000296	-0.0333***	-0.0333***	-0.0311	-0.0311
	(0.02)	(0.02)	(-3.60)	(-3.60)	(-0.83)	(-0.83)
pre3	-0.000215	-0.000215	0.0101	0.0101	0.0667	0.0667
	(-0.17)	(-0.17)	(0.55)	(0.55)	(1.21)	(1.21)
pre4	-0.000495	-0.000495	0.00686	0.00686	0.0679	0.0679
	(-0.79)	(-0.79)	(0.26)	(0.26)	(0.99)	(0.99)
N	148928	148928	15940	15940	14975	14975

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

tors like Callaway and Sant'Anna (2021) and Borusyak et al. (2021) show strong increases in effect size over time, whereas the other estimators like Wooldridge (2021) and Sun and Abraham (2021) remain more stable after five quarters. Both these estimators yield an effect size of roughly -50% after 11 quarters. Third, although the approaches are quite different, and the calculation and interpretation of the pre-trends varies across the cases, none of the estimators finds evidence for a pre-trend prior to treatment[14]. Taken together, the different estimators corroborate the previous finding that the directive had a strong deterrent effect on the use of direct awards. However, the magnitude of the estimated effect of the directive varies considerably.

Looking at the results for the other two types of public buyers, those with medium to low-initial levels of direct awards, we also find qualitatively similar results to the baseline findings. As shown in Figures 3 and 4, in the group of public buyers with a medium level of direct awards before the implementation of the directive, the estimated effect size tends to range between 0 and -10 percentage points. As the 95%-confidence-intervals show, however, the effect is rarely statistically significant. There are also some noticeable differences in the estimated effect, with the estimators from Abraham and Sun (2018) and Wooldridge (2021) producing larger effects than Callaway and Sant'Anna (2021), Borusyak et al. (2021) or a simple TWFE estimator. The results for the group of low-users of direct awards is comparatively clear, with estimates bunched slightly on the positive side around +2% but by and large not statistically different from zero. To conclude, while the strong and continued effect of the directive on public buyers with a high initial level of direct awards is corroborated by the different estimators, the case is less clear-cut for medium users. The estimators would suggest a smaller but mostly negative effect which is, however, not precise enough and seldom statistically significant. Regarding the group of public buyers with low use of direct awards, the directive seems to have had not statistically significant effect on their behavior.

[14]The stata package implementing the estimator of Wooldridge (2021) does not produce pre-trend tests

Figure 2: Effect of the directive for high-users, different event study estimators

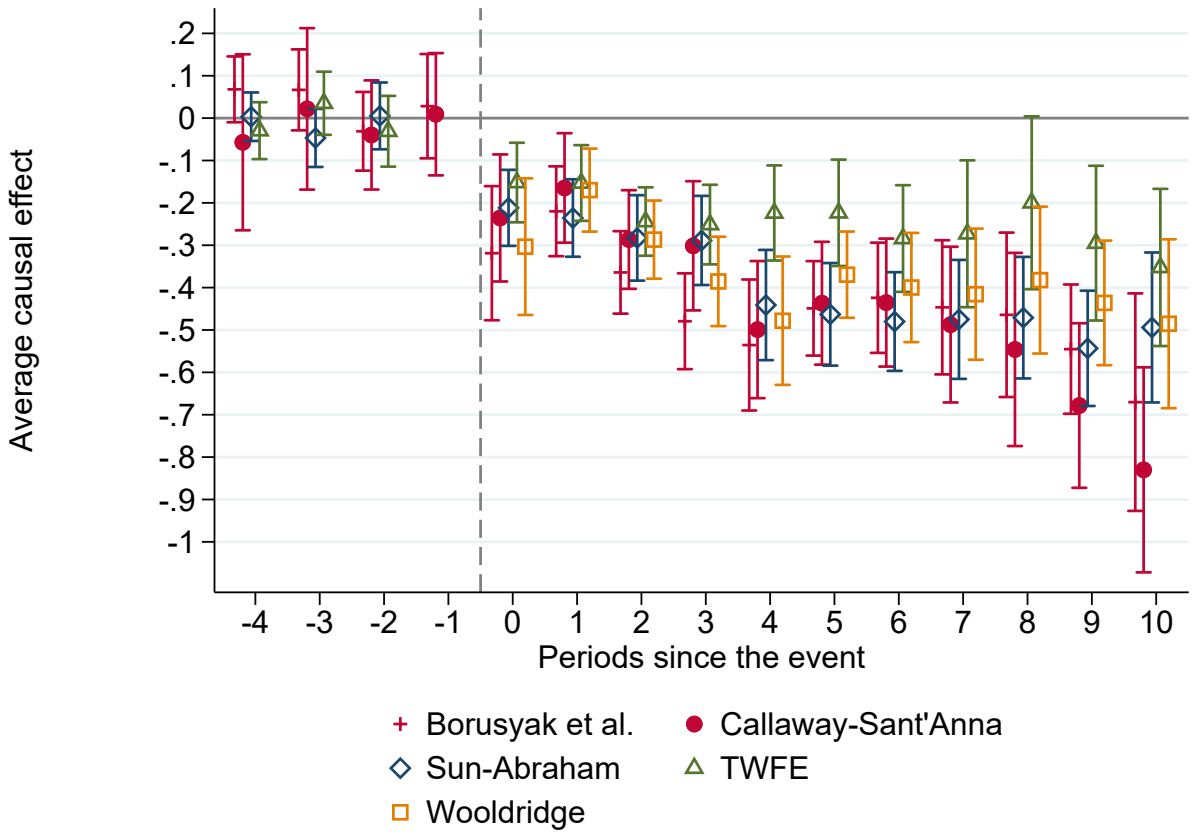


Figure 3: Effect of the directive for medium-users, different event study estimators

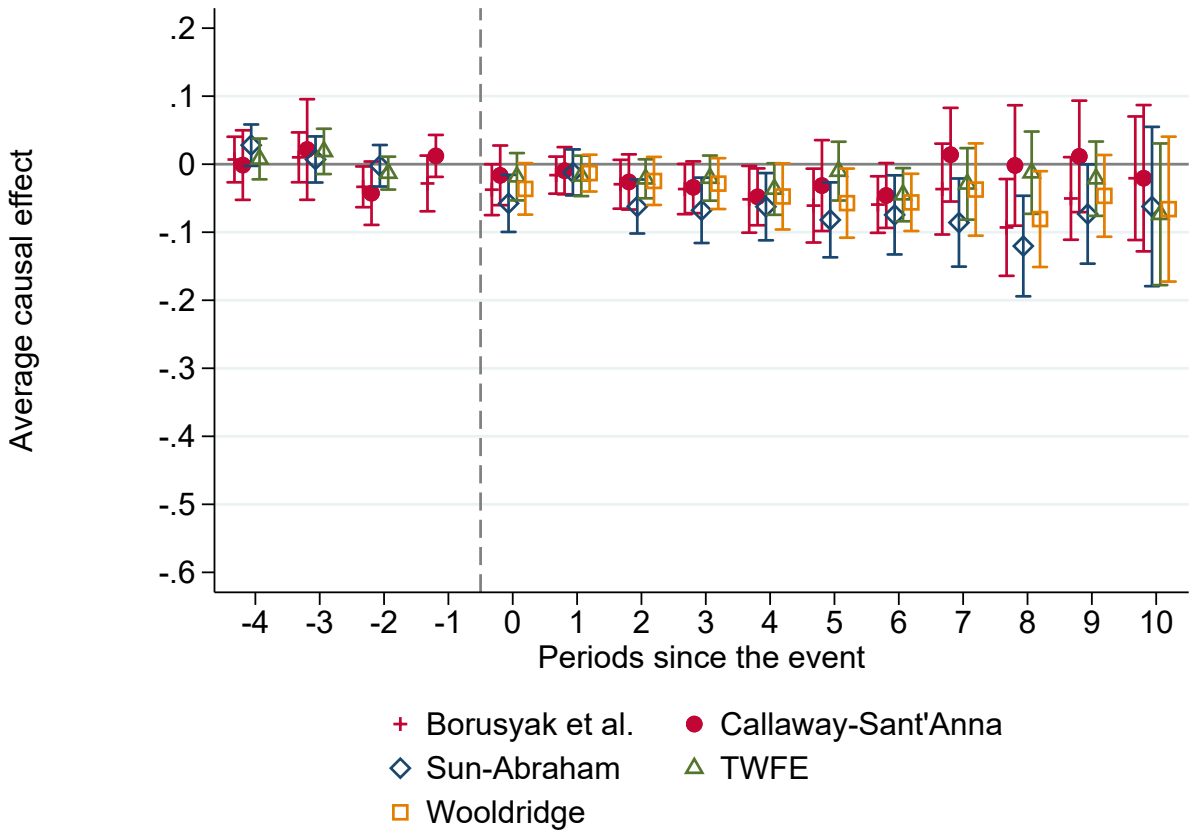
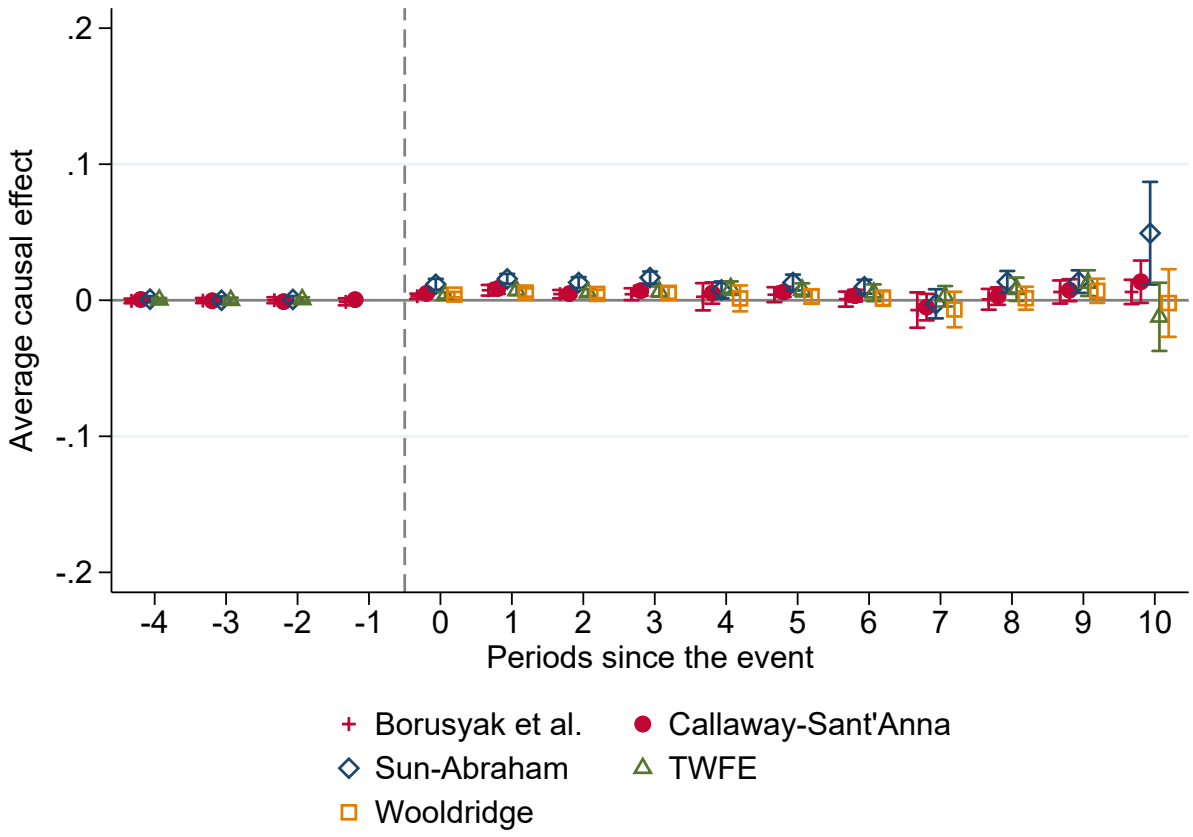


Figure 4: Effect of the directive for low-users, different event study estimators



5.3 Robustness tests

5.3.1 Country level estimates

A potential drawback of our approach is that we leverage sub-national variation on the level of public-buyers whereas the treatment is on the country level. Although this modeling approach is closest to the theoretical framework of MMR94, we can also analyze the effect on the level of countries. To this end, we aggregate the data on the country-level instead of the public-buyer level and re-estimate the previous regressions. To differentiate between countries where we would expect a larger effect, i.e. because direct awards were more frequently used before the directive, we distinguish countries according to the same intervals as before: $[0, 0.05), [0.05, 0.20), [0.20, 1.00)$. The respective results are shown in Table 4.

The estimates qualitatively confirm the results from the public-buyer level estimates, but there are some noticeable differences. Firstly, the effect size for the high-use group decreases for 27% in the static version (column 5) and dynamic estimates peak at -50% after 11 quarters, which is also somewhat lower than previously. Secondly, on the country-level the effect for the group of countries with a low initial share are positive and statistically significant. However, as the dynamics (column 2) show, this result becomes weaker over time and insignificant or even with a negative effect after 11 quarters. Thirdly, the effects for the medium group are closer to zero than in the previous regressions, with only weak statistical significance. Taken together, the country-level regressions corroborate the strong and persistent effect of the directive on public buyers with a high initial share, here countries with a high initial share. Conversely, the qualitative results for low and medium use groups differ somewhat between the two levels of analysis, yielding no conclusive picture.

5.3.2 Country-specific linear trend

A potential pitfall of our chosen modelling approach is that we might confound a country-specific trend with the measured treatment effect. Countries with a high initial level might

Table 4: Country-level results

	Low use		Medium use		High use	
	(1)	(2)	(3)	(4)	(5)	(6)
tau	0.0137***		-0.0372		-0.269***	
	(3.60)		(-1.88)		(-5.90)	
tau0		0.00939***		-0.0469		-0.128*
		(3.89)		(-1.49)		(-2.09)
tau1		0.0151***		0.00605		-0.213**
		(5.87)		(0.29)		(-3.13)
tau2		0.00896***		-0.00996		-0.261***
		(3.44)		(-0.27)		(-3.62)
tau3		0.0113*		-0.0698*		-0.306***
		(2.49)		(-2.35)		(-7.07)
tau4		0.0329*		-0.0450		-0.293***
		(2.14)		(-1.34)		(-4.77)
tau5		0.0144**		-0.0435		-0.356***
		(3.17)		(-1.79)		(-8.49)
tau6		0.00744		-0.0566***		-0.352***
		(1.55)		(-3.33)		(-10.02)
tau7		0.000311		-0.0443		-0.316***
		(0.10)		(-1.33)		(-5.36)
tau8		0.00965*		-0.0269		-0.325***
		(2.10)		(-0.92)		(-6.11)
tau9		0.0106***		-0.0574***		-0.419***
		(7.33)		(-4.14)		(-19.44)
tau10		-0.00356***		-0.0839***		-0.494***
		(-6.84)		(-7.60)		(-25.94)
pre1	-0.00264	-0.00264	-0.0626	-0.0626	-0.0397	-0.0397
	(-1.43)	(-1.43)	(-1.74)	(-1.74)	(-0.59)	(-0.59)
pre2	-0.00119	-0.00119	-0.0181	-0.0181	-0.0293	-0.0293
	(-0.78)	(-0.78)	(-0.55)	(-0.55)	(-0.45)	(-0.45)
pre3	-0.00131	-0.00131	-0.00935	-0.00935	0.104*	0.104*
	(-0.94)	(-0.94)	(-0.24)	(-0.24)	(2.01)	(2.01)
pre4	-0.000454	-0.000454	-0.0296	-0.0296	-0.0397	-0.0397
	(-0.45)	(-0.45)	(-1.47)	(-1.47)	(-0.64)	(-0.64)
N	614	614	535	535	530	530

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

already be on a declining path in their use of direct awards. We can partially address this problem by estimating a model that includes a country specific linear trend. The results shown in Table 5 suggest that the effects are negligible although the terms enter the regressions highly statistically significant (not shown here). The measured treatment effects are very similar to the baseline results particularly what concerns the large effect for the public buyers with a high initial level. Somewhat different results arise for the low use group, which more consistently exhibits a positive effect of the directive. Conversely, the previously mildly negative effect for the medium group disappears after controlling for a linear country trend in the use of direct awards.

5.3.3 Weighting and clustering

Two additional modelling choice reviewed here are the use of weights and the level of clustering. In the baseline regressions we used the number of contracts per public buyer as a weight to increase the precision of the estimates. Alternatively, one might prefer to give equal weight to each public buyer, which will tend to increase the weight of smaller units compared to previous estimates. The associated results are shown in Table 6. As can be seen in the table, removing weights has a slight effect on the effect size, which tends to decrease slightly, but does not alter the qualitative picture. Interestingly, without weights the pre-trends for high-users become negative and two quarters before the implementation even statistically significant. As the comparison with the (weighted) baseline results shows, the weights might therefore help to ensure that the pre-trends assumptions holds, at least conditionally. When weights are used, the pre-trend coefficients tend to be positive and/or closer to zero.

In addition to using weights, the previous regressions unanimously clustered standard errors at the country level, due to the fact that the treatment was on the country level. As can be seen in Table 7, switching from country to public-buyer level clustering does not change the results with respect to statistical significance.

Table 5: Country-level linear trend

	Low use		Medium use		High use	
	(1)	(2)	(3)	(4)	(5)	(6)
tau	0.00409***		-0.00285		-0.409***	
	(3.93)		(-0.16)		(-11.12)	
tau0		0.00301***		-0.0184		-0.243***
		(3.58)		(-1.00)		(-6.26)
tau1		0.00749***		0.00269		-0.203***
		(7.05)		(0.23)		(-3.40)
tau2		0.00461***		-0.00389		-0.392***
		(4.56)		(-0.41)		(-11.55)
tau3		0.00566***		-0.00604		-0.510***
		(3.88)		(-0.33)		(-13.66)
tau4		0.00379		-0.0181		-0.507***
		(1.39)		(-0.72)		(-7.84)
tau5		0.00528***		-0.0286		-0.428***
		(3.39)		(-0.98)		(-8.81)
tau6		0.00352*		0.00564		-0.355***
		(2.46)		(0.28)		(-8.78)
tau7		-0.00648**		0.0299		-0.492***
		(-2.87)		(0.86)		(-5.52)
tau8		-0.00109		-0.0467		-0.510***
		(-0.34)		(-1.62)		(-4.54)
tau9		0.00707***		0.0437		-0.582***
		(3.38)		(1.43)		(-10.63)
tau10		0.0294***		0.271***		-0.727***
		(14.58)		(9.77)		(-16.44)
pre1	-0.00418***	-0.00418***	0.0220	0.0220	-0.00621	-0.00621
	(-4.41)	(-4.41)	(0.71)	(0.71)	(-0.06)	(-0.06)
pre2	-0.00257*	-0.00257*	-0.00159	-0.00159	-0.0717	-0.0717
	(-2.50)	(-2.50)	(-0.08)	(-0.08)	(-1.03)	(-1.03)
pre3	-0.00208*	-0.00208*	0.0333	0.0333	0.0503	0.0503
	(-2.40)	(-2.40)	(1.12)	(1.12)	(0.67)	(0.67)
pre4	-0.00138*	-0.00138*	0.0268	0.0268	0.0666	0.0666
	(-2.20)	(-2.20)	(0.78)	(0.78)	(0.91)	(0.91)
N	148928	148928	15940	15940	14975	14975

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6: No weights

	Low use		Medium use		High use	
	(1)	(2)	(3)	(4)	(5)	(6)
tau	0.0124***		-0.0639		-0.365***	
	(4.99)		(-1.95)		(-14.57)	
tau0		0.0121***		-0.0514		-0.214***
		(3.92)		(-1.76)		(-7.10)
tau1		0.0160***		-0.0266		-0.258***
		(4.14)		(-1.25)		(-6.30)
tau2		0.0127***		-0.0595*		-0.288***
		(4.97)		(-2.40)		(-8.30)
tau3		0.0166***		-0.0688*		-0.330***
		(6.92)		(-2.32)		(-12.20)
tau4		0.00787*		-0.0532		-0.417***
		(2.41)		(-1.22)		(-6.68)
tau5		0.0138***		-0.0851		-0.481***
		(4.83)		(-1.52)		(-13.14)
tau6		0.0118***		-0.0809**		-0.499***
		(3.47)		(-3.05)		(-18.63)
tau7		-0.00310***		-0.0672		-0.467***
		(-4.13)		(-1.55)		(-9.49)
tau8		0.0132*		-0.133*		-0.481***
		(2.29)		(-1.96)		(-15.69)
tau9		0.0136**		-0.0782*		-0.588***
		(3.15)		(-2.34)		(-34.57)
tau10		0.0614***		-0.0766***		-0.520***
		(134.04)		(-3.63)		(-25.92)
pre1	-0.00363*	-0.00363*	-0.0296	-0.0296	-0.0633	-0.0633
	(-2.52)	(-2.52)	(-1.07)	(-1.07)	(-1.34)	(-1.34)
pre2	0.000104	0.000104	-0.0190	-0.0190	-0.101***	-0.101***
	(0.10)	(0.10)	(-0.88)	(-0.88)	(-5.48)	(-5.48)
pre3	-0.00153	-0.00153	-0.00545	-0.00545	-0.0327	-0.0327
	(-0.95)	(-0.95)	(-0.31)	(-0.31)	(-1.25)	(-1.25)
pre4	-0.000104	-0.000104	0.0248	0.0248	-0.0150	-0.0150
	(-0.09)	(-0.09)	(0.77)	(0.77)	(-0.33)	(-0.33)
N	148928	148928	15940	15940	14975	14975

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 7: Clustering at public-buyer level

	Low use		Medium use		High use	
	(1)	(2)	(3)	(4)	(5)	(6)
tau	0.00314		-0.0433**		-0.419***	
	(1.61)		(-2.78)		(-9.29)	
tau0		0.00295**		-0.0375*		-0.319***
		(3.06)		(-1.96)		(-3.95)
tau1		0.00733***		-0.0161		-0.220***
		(3.68)		(-1.16)		(-4.07)
tau2		0.00453**		-0.0295		-0.364***
		(2.92)		(-1.61)		(-7.33)
tau3		0.00446*		-0.0366		-0.479***
		(2.00)		(-1.94)		(-8.31)
tau4		0.00256		-0.0516*		-0.535***
		(0.50)		(-2.06)		(-6.79)
tau5		0.00408		-0.0609*		-0.449***
		(1.45)		(-2.20)		(-7.89)
tau6		0.000841		-0.0593**		-0.424***
		(0.30)		(-2.80)		(-6.39)
tau7		-0.00721		-0.0366		-0.446***
		(-1.09)		(-1.07)		(-5.53)
tau8		0.000680		-0.0929*		-0.464***
		(0.17)		(-2.56)		(-4.69)
tau9		0.00606		-0.0503		-0.545***
		(1.40)		(-1.63)		(-7.00)
tau10		0.00612		-0.0206		-0.670***
		(1.34)		(-0.44)		(-5.12)
pre1	-0.00113	-0.00113	-0.0283	-0.0283	0.0283	0.0283
	(-0.87)	(-0.87)	(-1.35)	(-1.35)	(0.45)	(0.45)
pre2	0.0000296	0.0000296	-0.0333*	-0.0333*	-0.0311	-0.0311
	(0.03)	(0.03)	(-2.18)	(-2.18)	(-0.66)	(-0.66)
pre3	-0.000215	-0.000215	0.0101	0.0101	0.0667	0.0667
	(-0.23)	(-0.23)	(0.54)	(0.54)	(1.37)	(1.37)
pre4	-0.000495	-0.000495	0.00686	0.00686	0.0679	0.0679
	(-0.57)	(-0.57)	(0.40)	(0.40)	(1.71)	(1.71)
N	148928	148928	15940	15940	14975	14975

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

5.3.4 Fuzzy record linkage of public buyers

Finally, a big drawback of the TED dataset is that public-buyers are not identified by a unique id. We have to identify public buyers by their names, subject to a frequently used deduplication approach. This approach uses a cosine string-similarity measure with a term frequency–inverse document frequency (tf-idf) weighting to identify the same units even if spelled slightly differently. Due to its computational intensity we implement this appraoch using the python library 'sparse_dot_topn'[15] and use a 95% threshold measure in conjunction with 3-grams. To understand if this somewhat arbitrary deduplication affects the estimation results, we use an exact matching as an alternative. Hence, public buyers are identified as the same unit only if their names are identical across contract awards. This increases the number of disginguished public buyers from 292K to 333K.

The result of this exercise is shown in Table 8, and largely corroborates the previous baseline results. The coefficient estimates for the high use group decrease slightly from 42% (baseline) to 38%, but show the very similar dynamics over time. Also for the medium and low use groups, the results are in line with those from the baseline regressions, yielding a mixed and inconclusive result for these two groups. To sum, identifying public buyers more narrowly does not significantly alter the estimated effects of the directive.

6 Discussion and conclusion

This study tries to fill the empirical gap on the effects of private enforcement. Using the EU Remedies Directive as a case that improved the prospects of private enforcement, we are able to test the hypothesis from Marshall et al. (1994) that bid protest mechanisms can help to reduce illicit practices. We can show empirically that strengthening the bid protest mechanism had a deterrent effect on the use of direct awards. However, the effect of the directive was not homogeneous across all public buyers. While we find no effect on

[15] `https://github.com/ing-bank/sparse_dot_topn`

Table 8: Exact match for public buyer names

	Low use		Medium use		High use	
	(1)	(2)	(3)	(4)	(5)	(6)
tau	0.00219		-0.0462		-0.382***	
	(1.29)		(-1.77)		(-8.54)	
tau0		0.00286***		-0.0375		-0.321***
		(3.74)		(-1.88)		(-3.95)
tau1		0.00681***		-0.0246		-0.210***
		(6.72)		(-1.27)		(-3.43)
tau2		0.00399***		-0.0338*		-0.322***
		(3.40)		(-2.18)		(-7.23)
tau3		0.00410*		-0.0362		-0.372***
		(2.19)		(-1.25)		(-8.57)
tau4		0.000168		-0.0530		-0.490***
		(0.06)		(-1.27)		(-6.27)
tau5		0.00254		-0.0561		-0.420***
		(0.86)		(-1.67)		(-7.87)
tau6		-0.000267		-0.0643**		-0.402***
		(-0.12)		(-2.64)		(-8.46)
tau7		-0.00791*		-0.0373		-0.441***
		(-2.28)		(-0.71)		(-5.84)
tau8		-0.00172		-0.0981*		-0.424***
		(-0.36)		(-2.30)		(-8.86)
tau9		0.00488**		-0.0753**		-0.571***
		(2.95)		(-2.73)		(-13.14)
tau10		0.00544***		-0.000713		-0.666***
		(4.34)		(-0.04)		(-22.02)
pre1	-0.00106	-0.00106	-0.0195	-0.0195	0.0144	0.0144
	(-0.56)	(-0.56)	(-1.16)	(-1.16)	(0.20)	(0.20)
pre2	0.000154	0.000154	-0.0339***	-0.0339***	-0.0492	-0.0492
	(0.11)	(0.11)	(-3.88)	(-3.88)	(-1.31)	(-1.31)
pre3	-0.000314	-0.000314	0.0150	0.0150	0.0607	0.0607
	(-0.22)	(-0.22)	(0.77)	(0.77)	(1.13)	(1.13)
pre4	-0.000551	-0.000551	0.0107	0.0107	0.0632	0.0632
	(-0.71)	(-0.71)	(0.39)	(0.39)	(0.94)	(0.94)
N	144238	144238	16190	16190	14738	14738

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

average, the directive had a very strong effect on public buyers that used direct awards as their standard procurement mechanism. This differential effect of the directive might also be interpreted as an indication that the directive was well targeted. Although the results for the group of buyers with low or medium shares of direct awards were inconclusive, finding no evidence of a reduction might suggest that there are no signs of over-deterrence for public buyers which used direct awards only occasionally. In contrast, frequent users had to adapt their behavior as a result of the EU-directive.

Although this study is able to show that private enforcement can help to improve compliance with procurement rules there are several limitations. First, our empirical results hinge on the assumption that the treatment date was randomly assigned. While we have no reason to believe this was not the case, we are basically unable to test this assumption. The largely insignificant effects on the estimated pre-trends at least give no empirical ground for such concerns. Second, we do not know if public buyers tried to evade the new regulations by switching from above to below threshold contracts. As there is not systematic data on below threshold procurement, knowledge about above-threshold procurement contracts might not deliver the full picture. Unfortunately, the TED database does not provide much data to further analyze how the directive affected the behavior of public bidders, and maybe even more importantly, the associated outcomes in terms of price.

References

Abraham, S. and Sun, L. (2018). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *SSRN Electronic Journal*.

Aidt, T. S. (2003). Economic analysis of corruption: A survey. *The Economic Journal*, 113(491):F632–F652.

Bajari, P., MacMillan, R. S., and Tadelis, S. (2009). Auctions versus negotiations in procurement: an empirical analysis. *Journal of Law, Economics, and Organization*, 25(2):372–399.

Becker, G. S. and Stigler, G. J. (1974). Law enforcement, malfeasance, and compensation of enforcers. *The Journal of Legal Studies*, 3(1):1–18.

Borusyak, K. and Jaravel, X. (2017). Revisiting event study designs. *SSRN 2826228*.

Borusyak, K., Jaravel, X., and Spiess, J. (2021). Revisiting Event Study Designs: Robust and Efficient Estimation. Papers 2108.12419, arXiv.org.

Bulow, J. and Klemperer, P. (1996). Auctions versus negotiations. *American Economic Review*, 86(1):pp. 180–194.

Burguet, R. and Che, Y.-K. (2004). Competitive procurement with corruption. *RAND Journal of Economics*, 35(1):50–68.

Callaway, B. and Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230.

Calzolari, G. and Spagnolo, G. (2009). Relational contracts and competitive screening. CEPR Discussion Paper No. 7434.

Celentani, M. and Ganuza, J.-J. (2002). Corruption and competition in procurement. *European Economic Review*, 46(7):1273–1303.

Compte, O., Lambert-Mogiliansky, A., and Verdier, T. A. (2005). Corruption and competition in procurement auctions. *RAND Journal of Economics*, 36(1):1–15.

De Chaisemartin, C. and d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–96.

Di Tella, R. and Schargrodsky, E. (2003). The role of wages and auditing during a crackdown on corruption in the city of buenos aires*. *Journal of law and economics*, 46(1):269–292.

European Commission (2017). Commission staff working document evaluation of the modifications introduced by directive 2007/66/ec to directives 89/665/eec and 92/13/eec concerning the european framework for remedies in the area of public procurement/ refit evaluation accompanying the document report from the commission to the european parliament and the council on the effectiveness of directive 89/665/eec and directive 92/13/eec, as modified by directive 2007/66/ec, concerning review procedures in the area of public procurement. SWD/2017/013 final.

Goldberg, V. P. (1977). Competitive bidding and the production of precontract information. *Bell Journal of Economics*, 8(1):250–261.

Kaplow, L. (2017). Optimal design of private litigation. *Journal of Public Economics*, 155:64–73.

Kelman, S. (1990). *Procurement and Public Management: The Fear of Discretion and the Quality of Government Performance*. AEI Press.

Kovacic, W. E. (1998). Law, economics, and the reinvention of public administration: Using relational agreements to reduce the cost of procurement regulation and other forms of government intervention in the economy. *Admin. L. Rev.*, 50:141.

Laffont, J.-J. (2001). *Incentives and Political Economy*. Oxford University Press, Oxford, UK.

Leffler, K. B., Rucker, R. R., and Munn, I. A. (2003). Transaction costs and the collection of information: presale measurement on private timber sales. Working paper, Department of Econonomics, University of Washington.

Manelli, A. M. and Vincent, D. R. (1995). Optimal procurement mechanisms. *Econometrica*, 63(3):591–620.

Marshall, R. C., Meurer, M. J., and Richard, J.-F. (1994). Curbing agency problems in the procurement process by protest oversight. *The Rand Journal of Economics*, pages 297–318.

McAfee, R. P., Mialon, H. M., and Mialon, S. H. (2008). Private v. public antitrust enforcement: A strategic analysis. *Journal of Public Economics*, 92(10-11):1863–1875.

Olken, B. A. (2007). Monitoring corruption: Evidence from a field experiment in Indonesia. *Journal of Political Economy*, 115(2):200–249.

Peyer, S. (2012). Private antitrust litigation in germany from 2005 to 2007: Empirical evidence. *Journal of Competition Law and Economics*, 8(2):331–359.

Søreide, T. (2006). Corruption in international business transactions: The perspective of Norwegian firms. In Rose-Ackerman, S., editor, *International Handbook on the Economics of Corruption*, chapter 13, pages 381–417. Edward Elgar, Cheltenham,UK/Northampton, Mass.

Sun, L. and Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2):175–199.

Tirole, J. (1992). Collusion and the theory of organization. In Laffont, J.-J., editor, *Advances in Economic Theory*, volume 2 (Sixth World Congress). Cambridge University Press, Cambridge.

Tran, A. (2011). Which regulations reduce corruption? Evidence from the internal records of a bribe-paying firm. Working Paper.

van Rijckeghem, C. and Weder, B. (2001). Bureaucratic corruption and the rate of temptation: Do wages in the civil service affect corruption, and by how much? *Journal of development economics*, 65(2):307–331.

Wooldridge, J. M. (2021). Two-way fixed effects, the two-way mundlak regression, and difference-in-differences estimators. *Available at SSRN 3906345*.

Appendix A: Regressions

Table 9: Five Estimators: High users

	BJS	CS	AS	W	TWFE
	(1)	(2)	(3)	(4)	(5)
tau0	-0.319***	-0.236**	-0.212***	-0.303***	-0.152**
	(-3.95)	(-3.08)	(-4.63)	(-3.69)	(-3.17)
tau1	-0.220***	-0.165*	-0.236***	-0.170***	-0.153***
	(-4.07)	(-2.50)	(-5.05)	(-3.40)	(-3.36)
tau2	-0.364***	-0.286***	-0.283***	-0.287***	-0.244***
	(-7.33)	(-4.82)	(-5.49)	(-6.09)	(-5.93)
tau3	-0.479***	-0.301***	-0.289***	-0.385***	-0.251***
	(-8.31)	(-3.87)	(-5.38)	(-7.16)	(-5.24)
tau4	-0.535***	-0.499***	-0.441***	-0.478***	-0.224***
	(-6.79)	(-6.04)	(-6.65)	(-6.19)	(-3.91)
tau5	-0.449***	-0.437***	-0.463***	-0.369***	-0.223***
	(-7.89)	(-5.91)	(-7.50)	(-7.11)	(-3.49)
tau6	-0.424***	-0.435***	-0.480***	-0.400***	-0.284***
	(-6.39)	(-5.65)	(-8.08)	(-6.08)	(-4.42)
tau7	-0.446***	-0.487***	-0.475***	-0.415***	-0.273**
	(-5.53)	(-5.20)	(-6.63)	(-5.27)	(-3.09)
tau8	-0.464***	-0.546***	-0.471***	-0.382***	-0.200
	(-4.69)	(-4.69)	(-6.44)	(-4.32)	(-1.92)
tau9	-0.545***	-0.678***	-0.543***	-0.436***	-0.295**
	(-7.00)	(-6.85)	(-7.83)	(-5.82)	(-3.17)
pre1	0.0283	0.00918			
	(0.45)	(0.12)			
pre2	-0.0311	-0.0397	0.00528		-0.0310
	(-0.66)	(-0.60)	(0.13)		(-0.73)
pre3	0.0667	0.0219	-0.0468		0.0351
	(1.37)	(0.23)	(-1.34)		(0.92)
pre4	0.0679	-0.0570	0.00336		-0.0295
	(1.71)	(-0.54)	(0.12)		(-0.86)
N	14975	14975	14807	14975	82628

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

BJS = Borusyak et al. (2021); CS = Callaway and Sant'Anna (2021), AS = Abraham and Sun (2018), W = Wooldridge (2021)

Table 10: Five Estimators: Medium users

	BJS	CS	AS	W	TWFE
	(1)	(2)	(3)	(4)	(5)
tau0	-0.0375*	-0.0164	-0.0576**	-0.0363*	-0.0185
	(-1.96)	(-0.73)	(-2.69)	(-1.88)	(-1.04)
tau1	-0.0161	-0.00958	-0.0121	-0.0131	-0.0172
	(-1.16)	(-0.54)	(-0.70)	(-0.95)	(-1.14)
tau2	-0.0295	-0.0261	-0.0622**	-0.0246	-0.0215
	(-1.61)	(-1.26)	(-3.07)	(-1.37)	(-1.47)
tau3	-0.0366	-0.0339	-0.0678**	-0.0286	-0.0205
	(-1.94)	(-1.75)	(-2.76)	(-1.50)	(-1.21)
tau4	-0.0516*	-0.0480*	-0.0625*	-0.0474	-0.0365
	(-2.06)	(-2.25)	(-2.48)	(-1.91)	(-1.89)
tau5	-0.0609*	-0.0314	-0.0818**	-0.0572*	-0.0102
	(-2.20)	(-0.92)	(-2.91)	(-2.21)	(-0.46)
tau6	-0.0593**	-0.0460	-0.0745*	-0.0561**	-0.0449*
	(-2.80)	(-1.89)	(-2.51)	(-2.61)	(-2.25)
tau7	-0.0366	0.0139	-0.0858**	-0.0373	-0.0290
	(-1.07)	(0.40)	(-2.60)	(-1.08)	(-1.08)
tau8	-0.0929*	-0.00188	-0.120**	-0.0807*	-0.0125
	(-2.56)	(-0.04)	(-3.19)	(-2.25)	(-0.41)
tau9	-0.0503	0.0115	-0.0732*	-0.0466	-0.0213
	(-1.63)	(0.27)	(-1.97)	(-1.52)	(-0.77)
pre1	-0.0283	0.0122			
	(-1.35)	(0.77)			
pre2	-0.0333*	-0.0427	-0.00233		-0.0131
	(-2.18)	(-1.80)	(-0.15)		(-1.06)
pre3	0.0101	0.0216	0.00694		0.0188
	(0.54)	(0.57)	(0.40)		(1.11)
pre4	0.00686	-0.00129	0.0279		0.00770
	(0.40)	(-0.05)	(1.80)		(0.50)
N	15940	15940	15934	15940	124210

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

BJS = Borusyak et al. (2021); CS = Callaway and Sant'Anna (2021), AS = Abraham and Sun (2018), W = Wooldridge (2021)

Table 11: Five Estimators: Low users

	BJS	CS	AS	W	TWFE
	(1)	(2)	(3)	(4)	(5)
tau0	0.00295**	0.00509***	0.0117***	0.00413***	0.00355**
	(3.06)	(5.85)	(5.74)	(3.98)	(3.03)
tau1	0.00733***	0.00865***	0.0157***	0.00566***	0.00714***
	(3.68)	(4.27)	(8.46)	(3.63)	(4.28)
tau2	0.00453**	0.00490***	0.0130***	0.00465**	0.00556***
	(2.92)	(4.29)	(6.44)	(2.96)	(4.29)
tau3	0.00446*	0.00700***	0.0166***	0.00546*	0.00636***
	(2.00)	(3.72)	(7.34)	(2.42)	(3.32)
tau4	0.00256	0.00524	0.00757*	0.00136	0.00886***
	(0.50)	(1.31)	(2.39)	(0.28)	(3.49)
tau5	0.00408	0.00613**	0.0130***	0.00275	0.00761**
	(1.45)	(3.14)	(4.39)	(1.01)	(3.10)
tau6	0.000841	0.00337	0.00996***	0.00155	0.00550
	(0.30)	(1.53)	(3.81)	(0.55)	(1.75)
tau7	-0.00721	-0.00506	-0.00262	-0.00680	0.00291
	(-1.09)	(-1.02)	(-0.48)	(-1.02)	(0.76)
tau8	0.000680	0.00304	0.0135***	0.00149	0.00808
	(0.17)	(0.93)	(3.33)	(0.35)	(1.84)
tau9	0.00606	0.00740	0.0137**	0.00693	0.0126**
	(1.40)	(1.80)	(3.24)	(1.53)	(2.61)
pre1	-0.00113	0.000431			
	(-0.87)	(0.56)			
pre2	0.0000296	-0.00104	0.000614		0.000567
	(0.03)	(-1.36)	(0.49)		(0.68)
pre3	-0.000215	-0.000406	-0.00000116		0.00000674
	(-0.23)	(-0.56)	(-0.00)		(0.01)
pre4	-0.000495	0.000549	0.000773		0.000352
	(-0.57)	(0.48)	(0.67)		(0.51)
N	148928	148928	144960	148928	1216220

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

BJS = Borusyak et al. (2021); CS = Callaway and Sant'Anna (2021), AS = Abraham and Sun (2018), W = Wooldridge (2021)