Universität St.Gallen

An empirical evaluation of environmental Alternative Dispute Resolution methods

Petyo Bonev, Shigeru Matsumoto

August 2022 Discussion Paper no. 2022-08

School of Economics and Political Science,
Department of Economics

University of St.Gallen

Editor: Mirela Keuschnigg
University of St.Gallen
School of Economics and Political Science
Department of Economics
Müller-Friedberg-Strasse 6/8
CH-9000 St.Gallen
Email seps@unisg.ch
Publisher: School of Economics and Political Science
Department of Economics
University of St.Gallen
Müller-Friedberg-Strasse 6/8
CH-9000 St.Gallen

Electronic Publication: http://www.seps.unisg.ch

An empirical evaluation of environmental Alternative Dispute Resolution methods[1]

Petyo Bonev, Shigeru Matsumoto

Author's address: Petyo Bonev (Assistant Professor of Econometrics)
Email: Petyo.Bonev@unisg.ch
Swiss Institute for Empirical Economic Research
University of St. Gallen
Varnbüelstrasse 14
9000/St. Gallen
Phone +41 71 224 26 94

Shigeru Matsumoto (Professor of Economics)
Email: hmatsumoto@aoyamagakuin.jp
Aoyama Gakuin University

[1] We thank seminar participants at the University of St. Gallen and the University of Reading for their helpful and critical comments. We also thank Stefan Bühler, Stephen Kastoryano, Michael Knaus and Michael Lechner for their ideas and improvement suggestions.

Abstract

This paper empirically evaluates different Alternative Dispute Resolution methods. Using a novel dataset on environmental disputes from Japan, we show that consensus-based approaches such as mediation lead on average to shorter duration and higher satisfaction than top-down approaches such as arbitration. Moreover, our findings suggest that the benefits depend on the transaction cost of resolving a dispute: while disputes with high transaction costs tend to benefit more from top-down approaches, disputes with lower costs benefit more from consensual resolution methods.

Keywords

environmental policy, environmental disputes, Alternative Dispute Resolution, Coase Theorem

JEL Classification

C21, C41, C78, D04, D74, D83, Q34, Q53, Q58.

1 Introduction

Alternative Dispute Resolution (ADR) methods are generally considered to be superior to court settlement because of their lower costs, reduced complexity and shorter duration, Shavell (1995).[1] Because of these advantages, ADR methods have become a widely used approach to settle disputes. As an example, in a survey of corporate counsel in Fortune 1000 companies, Stipanowich and Lamare (2014) found that at least 42 % of companies commonly use the ADR method mediation in consumer disputes, 50% of the companies use it in in employment disputes, and 47% in commercial disputes, see also Fanning (2021). Other ADR approaches such as arbitration seem to be similarly popular, Balzer and Schneider (2021).

Accordingly large is the interest in ADR in the law and economics literature. Coben and Stienstra (2021) provide a literature overview that contains over 600 ADR-related empirical studies published between 2013 and 2020 alone. One of the central open questions is which is the optimal dispute resolution method, where optimality can be defined with regard to a variety of characteristics such as fairness, cost and effectiveness, and party satisfaction, see e.g. Menkel-Meadow (2006). Recent theoretical studies have therefore focused on the comparative performance of different resolution methods, see e.g. Salamanca (2021) for a comparison of mediation and cheap talk, Meirowitz et al. (2019) for third party mediation vs. unmediated negotiation, Olszewski (2011) for for conventional vs final offer arbitration, Hörner et al. (2015) for meditaion vs. arbitration, and Goltsman et al. (2009) for unmediated negotiation vs. mediation vs. arbitration.

Despite the importance of comparing the performance different ADR methods, the empirical literature on this topic is scarce. The main challenge is that parties involved into a dispute select into different ADR methods based on unobserved determinants of

[1] "ADR" is a broad term and often used for any method that is "alternative to court", i.e. any resolution method in which a dispute is resolved without a trial (The New York State Unified Court System, http://ww2.nycourts.gov/adr). However, scholars typically assume that ADR methods involve the participation of a third, neutral, party.

the outcomes. Important examples for such unobserved confounders are risk preferences, Ashenfelter et al. (1992), preference for control over the dispute process, Shapiro and Brett (1993) and the subjective judgment of the expected outcome of the two parties, Shavell (1995), among others. As a result, a comparison of different ADR methods yields potentially biased results.

In this paper, we empirically evaluate different alternative methods for resolving environmental disputes in Japan. The environmental ADR system in Japan is characterized by two unique features. First, each environmental conflict is assigned by a neutral third party, called environmental counselor, to one of three possible dispute resolution methods: (1) administrative guidance by the counselor, a method close in spirit to conventional arbitration; (2) mediated discussion between parties; (3) negotiation between the environmental counselor and the plaintiff, a method in which a solution is reached only if the plaintiff agrees to the proposal by the counselor. Second, prior to assigning a case to a given method, the environmental counselor collects an institutionally predetermined set of case characteristics which are recorded in a standardized form. Examples for such characteristics are the type and source of pollution, the type of damage, characteristics of the area, as well as of the involved parties. Based on these characteristics and on the own legal discretion, the environmental counselor then assigns the case to one of the three ADR methods.

To causally evaluate the relative performance of each ADR method, we assembled a novel dataset that consists of all environmental disputes in Japan that occurred between 2009 and 2018. Our dataset consists of more than a quarter of million observations and for each case, we observe a rich set of characteristics. Importantly, we have access to *all case characteristics used by the environmental counselor in the treatment assignment process.* This setup gives rise to a quasi-experimental research design. In particular, we adopt a selection-on-observables econometric approach and use the idiosyncratic variation in the treatment assignment that is caused by differences in counselors' preferences and experience. Using this exogenous treatment variation, we identify the effect of each ADR method on two outcomes. The first one is the duration of a dispute. The second one is the satisfaction of the plaintiff with the resolution of the dispute. Information on the latter is obtained as part of the official handling of each case through a survey.

Our study delivers several important findings. First, the two ADR methods that require the consent of at least of one of the two parties perform on average better than the top-down (also referred to as "commanded", Menkel-Meadow (2006)) arbitration approach. In particular, consensus-based approaches reduce the average duration of a dispute by 11 days, which amounts to roughly 20% of the average duration of an environmental dispute. Similar findings are obtained for a variety of alternative treatment effects defined for the duration variable such as an additive effect on the survival function and a multiplicative effect on the hazard function. Moreover, consensus-based approaches increase the share of plaintiffs satisfied with the dispute resolution outcome by more than 6%. This finding is somewhat surprising given the theoretical predictions of the literature that emphasizes the advantage of arbitration over mediation to enforce dispute solutions, Goltsman et al. (2009).[2] One possible explanation is that parties are better able to find a satisfactory solution to a problem than a third neutral because they better know their preferences and sets of possible strategies. As the seminal paper of Shavell (1995) puts it, "... because there is a limit to the degree to which the legal process can be made sensitive to the particular situations of disputants. This limit is due principally to difficulties that would be faced by courts in determining the detailed characteristics of different situations. But the parties themselves know much about their own situations; typically, they know their situations far better than the courts could." Since conventional arbitration is very similar to a court procedure, this observation could be extended to a general comparison of commanded and consensual approaches.

Next, to get a better understanding of the underlying causal mechanisms, we study

[2]The paper by Hörner et al. (2015) shows that the lack of enforcement powers need not be of a disadvantage.

treatment effect heterogeneity with state-of-the-art doubly robust machine learning estimation techniques (see e.g. Semenova and Chernozhukov (2021)). We find that the benefits of consensus depend on the transaction costs of the agreement. While disputes that involve lower number of involved parties are resolved quicker and have a higher plaintiff satisfaction under a consensus-based approach, disputes with numerous involved parties and or complex (infrastructure related) dispute subjects tend to do better under the arbitration resolution method. Thus, our results are in remarkable alignment with the Coase theorem, Coase (1960), which states that in the case of externalities, courts should only intervene if there are substantial transaction costs - otherwise, bargaining parties will reach themselves the most efficient solution. Our insights can be directly used to formulate an optimal policy design.

Finally, we also find that environmental counselors are conservative in the sense that the assignment decision is driven largely by legal rather than welfare and efficiency considerations. In particular, the probability to assign a case to the arbitration-type procedure is (1) generally high and (2) particularly high when the initial report has found that a local environmental regulation has been violated. Our analysis suggests that if the duration of a dispute and the satisfaction of the plaintiff are used as objectives, environmental counselors misspecify the large majority of cases.

Our paper contributes in several ways to the existing literature. First, we contribute to the small empirical literature on comparing the performance of different ADR processes (the so-called "comparative process" studies, Menkel-Meadow (2006)). To the best of our knowledge, our study is the first to causally compare consensual and commanded ADR methods with actual field data and an empirical strategy that accounts for potential exogeneity. Related papers either provide only qualitative evidence, Rosenberg and Folberg (1993), descriptive evidence, Delikat and Kleiner (2003), Farber and White (1994), or use lab experiments, Eisenkopf and Bächtiger (2013), Wilkenfeld et al. (2003), or use parametric regression techniques that do not explicitly deal with the selection problem, Beardsley et al. (2006), FEY and RAMSAY (2010) and Speight and Thomas (1997). The only field experiment to our knowledge is by McGillicuddy et al. (1987), but the sample size is only 36, so that no proper statistical analysis is feasible. Two important comparative process studies that use field data and adopt a quasi-experimental approach are Collins and Urban (2014) and Backus et al. (2020). The first one compares mediated and non-mediated bargaining, while the second one evaluates the effect of direct communication (via text messages) between parties on the bargaining success in a nonmediated price bargaining setting (Ebay transactions).

Second, we contribute to the yet very small empirical ADR literature on environmental disputes. Typically, this literature does not compare the performance of different ADR methods, see e.g. Matsumoto (2011) and the references therein.

Third, our paper contributes to the literature that tries to understand the benefits and pitfalls of compulsory assignment to different ADR processes, Lysaught (2009). Our findings echo some of the experimental findings on arbitrator behavior, see e.g. the important study by Ashenfelter and Bloom (1984). In particular, like arbitrators in binding arbitration, environmental counselors display "fairness" in the sense that their assignment decisions are driven by regulatory (and thus "fair") motives. At the same time, we find that the choice of the ADR method is insensitive to other factors commonly thought to impact arbitrators' behavior such as the relative bargaining power of the parties.

Fourth, our paper can be interpreted to be an extended empirical "test" of the Coase theorem. This literature has traditionally used either lab experiments, see e.g. the early work of Coursey et al. (1987), or case studies, see e.g. Hanley and Summer (1995) and Farnsworth (1999), with Bleakley and Ferrie (2014) being one of the few studies using quasi-experimental evidence with field data, see Medema (2020) for a comprehensive survey of the literature and Deryugina et al. (2021) for a review of the applications of the Coase theorem in environmental context. The major insight of our paper is to demonstrate a Coase-like result in a setting with private environmental goods, in which there are neither (competetive) markets, nor functioning price systems.

Finally, we also contribute to the literature on treatment effect evaluation in duration models. In particular, we provide identification results for the nonparametric identification of additive and multiplicative treatment effects on the hazard function, with the latter being a generalization of a result in Abbring and Berg (2005).

The paper is structured as follows. In section 2, we describe the institutional setup and the data. In section 3, we set out the empirical framework. In section 4, we present our empirical results. Section 5 concludes. Additional results are provided in an appendix.

2 Institutional setup and data

2.1 Institutional setup

A historical background of environmental pollution resolution in Japan. During its rapid economic development after the World War II, Japan experienced a steep rise in pollution problems. These problems included serious health hazards caused by the water contamination by heavy metals, as well as by air pollution caused by sulfur dioxide. These pollution problems occurred mainly due to the lack of proper environmental regulations and were further aggravated by the lack of a proper pollution management. In particular, the central government prioritized economic development over pollution management[3] and it ignored requests from local governments facing pollution problems, see e.g. Kitamura (2018). Many of the pollution problems were disputed in civil courts and were settled after the government or the polluting firms compensated the victims. However, it took a very long time until settlement.

[3] Environmental regulations laws at that time had a so-called "prastabilierte harmonie clause" that expressed the idea that environmental regulations needed to be implemented in a way that did not have a significant adverse effect on the business activities of firms.

As a response to growing public discontent, the Japanese government enacted "Environmental Pollution Prevention Act" in 1967. However, many deficiencies in the law were pointed out soon after the enactment, and the law was modified in the 1970 Diet substantially.[4] In addition, the government also enacted "The Act on the Settlement of Environmental Pollution Disputes", whose purpose is to resolve pollution problems "promptly" and "appropriately." This act serves as the contemporary legal basis for settling environmental disputes.

Based on the above experience, the central government has promoted the decentralization of environmental administration. It has recommended the residents and companies facing pollution problems to contact to the pollution complaint counseling desk of nearby municipalities or prefectures (local office) initially. Consequently, the majority of the pollution problems are now resolved with the assistance of municipal pollution complaint counselors without being disputed in court.

Current institutional setup When an environmental problem arises, affected residents and firms can either go to court or file a complaint with their local government. When the case contains a criminal offense, the complaint is transferred to the police. When the complaint either involves inter-prefecturial issues, or is of grave matter, or has nationwide implications, it is transferred to the Environmental Dispute Coordination Commission (EDCC), a administrative commission established as an external agency of the Prime Minister's Office. This commission provides a variety of ADR services such as conciliation, mediation and arbitration. If a case is not of a nationwide importance but has prefecture-wide implications, it is transferred to a Prefectural Pollution Examination Commission (PPEC), that assists the negotiation between the parties in the dispute. If none of the above conditions apply, the case is handled by a local Environment Pollution Complaint Counselor (for short, environmental counselor). Importantly, a complaint cannot be filed at other local governments, i.e. the plaintiff cannot choose

[4]The 1970 Diet is called "Pollution Diet" because the main agenda of the diet was to pollution management.

where to file a complaint. Thus, a Tiebout-type sorting is not possible.

When a pollution complaint is handled by a local counselor, a brief field investigation of the cause is conducted. Upon this investigation, the counselor enters a comprehensive description of the pollution case in a standardized sheet which is provided by the EDCC. This description includes information about the type of pollution, its cause and source, the type of damage, type of area, land use zone, type and of claimants, related local regulations and information on whether these regulations are violated. We provide a thorough description of this information below. In addition, the counselor also records the exact date of the complaint arrival and enters a unique id number.

In a second step, the counselors has to choose a method for resolving the pollution dispute. The assignment decision is made on the basis of the recorded information. The three main options are (1) administrative guidance to the polluter, which, given the asymmetric nature of an environmental dispute, is close in spirit to conventional arbitration; (2) mediated discussions between the parties, which is close in spirit to mediated negotiations; and (3) persuasion to the complainant, a form of dispute resolution, in which the case is resolved only if the claimant agrees to a solution proposed by the counselor.[5] When deciding on a method, the responsible counselor consults with other environmental counselors in the department. Overall, the environmental counselor and in particularly the counsel office have substantial discretion in the assignment process and the assignment decision is practically binding for the parties.[6]

After a complaint is resolved, the counselor selects the degree of satisfaction of the complainant from four levels based on the reactions of the complainant during and after the treatment processing. Once a record is complete, it is sent back to the EDCC. Recording of all of the above mentioned characteristics is obligatory and is used by EDCC

[5] An additional method is "investigation of the pollution cause". Since the pollution complaints that are assigned to this fourth method are quite different from other pollution complaints in terms of the processing purpose, we focus only on the the first three ADR methods.

[6] The decision can in principle be contested by the involved parties. However, since the environmental counsel is part of the local municipality, and since the municipality has a substantial leverage over the involved parties (i.e. through issuing licenses and many other activities), contest rarely takes place.

to monitor the quality of the assignment process and the dispute resolution activities of the environmental counselors. The information is also used for learning purposes and definition of best practices.[7]

The local environmental counselors. The precise role, objectives and career path of the local environmental counselors are regulated by the "The Act on the Settlement of Environmental Pollution Disputes" from 1970. The counselors are elected by the local governments. Upon election, they are tenured. Although the EDCC monitors closely their work and might intervene if claimants are predominantly dissatisfied with the dispute resolution at a given local government, there are no predefined career incentives. In particular, promotion depends solely on seniority and not on specified goals. The objectives for the environmental counselors are loosely specified by the aforementioned act as "rapid and appropriate resolution of the cases". In practice, however, statistics on the average duration of resolving a case are never used by EDCC to incentivize the counselors. A given counselor is assigned to a case based on the availability/work load and is not a choice by the involved parties. This precludes the possibility for strategic behavior of the parties at the stage of the treatment (i.e. ADR method) assignment.

2.2 Data and descriptive analysis

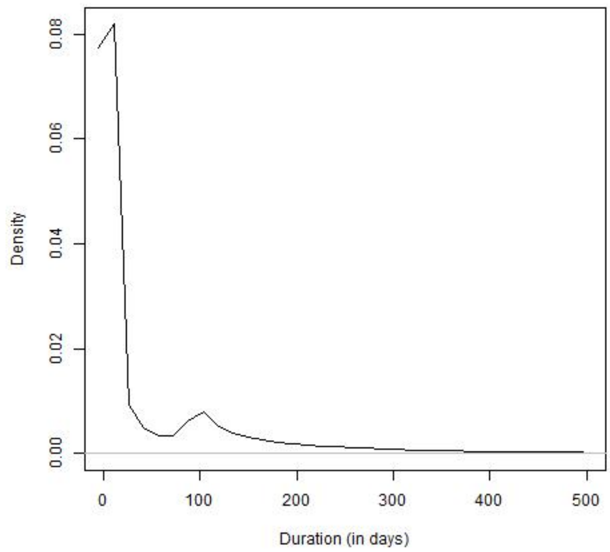
We have collected a unique dataset that contains all environmental conflict cases from all 47 prefectures in Japan for the period 2009-2018. We obtained the data from the EDCC through an information disclosure request procedure. From the available observations, we exclude all cases that are not handled by the local environmental counselors. These disputes are transferred to the police or to one of the central environmental agencies. As explained above, these cases are typically more severe in terms of the caused damage and or the extent of law violation. We exclude them from the analysis because the resolution method is not at the discretion of the local environmental counselors and because we do

[7] These best practices are discussed during the annual EDCC meetings and eventually published in the EDCC magazine "Chosei".

not observe their outcomes. With this choice, our sample contains 279230 observations.

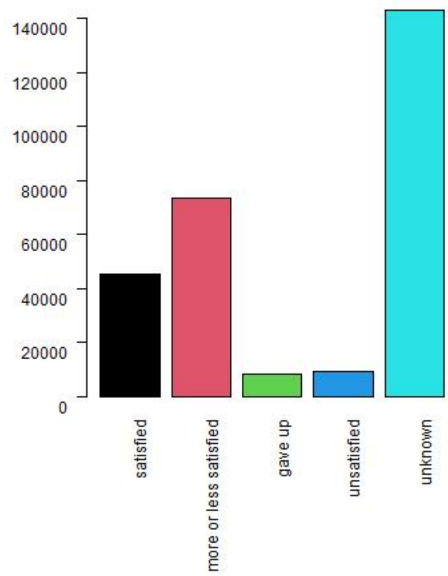
Outcome variables There are two outcome variables. The first one is the duration of a dispute and is defined as the time between the day on which the plaintiff files the complaint at the local environmental agency and the day on which the dispute is officially resolved. The mean duration is 55 days and the standard deviation is 195 days, indicating a substantial heterogeneity. Figure 1 shows a nonparametric estimate of the duration density. The graph is truncated at 500 days because there are only few cases ($< 0.1\%$) that continue longer. The bulk of the observations (77%) are resolved for less than 50 days. There appear to be a bunching at day 100. There is no regulatory reason for this bunching and it appears to be entirely driven by parties trying to resolve a conflict within 100 days.

Figure 1: Nonparametric estimate of the duration density.



The second outcome variable is the satisfaction of the plaintiff. At the end of each dispute case, the environmental counselor selects the degree of satisfaction of the complainant from four levels based on the reactions of the complainant during and after the treatment processing. The counselor can choose one from the following options: "satisfied" (1), "more or less satisfied" (2), "gave up" (3) , "unsatisfied" (4), and "unknown" (5). The absolute frequencies of these values are displayed as a barplot in figure 2. In 51% of the cases, the dispute is not rated (category 5), in 16% the plaintiff is satisfied, and in another 26% the plaintiff is more or less satisfied. Each of the categories 3 and 4 constitutes roughly 3 % of the cases. Thus, in the majority of the cases, plaintiffs are satisfied with the resolution or do not rate it. There is a small negative correlation between the satisfaction of the plaintiff and the duration of the resolution process.[8]

Figure 2: Distribution of satisfaction. "Satisfied" (1), "more or less satisfied" (2), "gave up" (3) , "unsatisfied" (4), and "unknown" (5)



In addition, we observe precisely how each conflict was resolved. A conflict may be

[8]When we exclude the non-rated cases, and assign numerical values 1-4 to first four categories, the correlation with the duration variable is 3.9%, indicating that higher dissatisfaction is associated with longer duration.

resolved because the cause for the conflict has disappeared by itself (33.5% of all cases), or because the polluter has taken measures to abolish it (48.3%), or because the two parties were able to reach a compromise solution (0.8%), or due to reasons that are not observed (17.4%). In the case that the polluter has taken measures to abolish the cause of the conflict, we observe the exact type of the measures. Examples include relocation or improvement of machinery and facilities, improvements in the work methods, removal of causative substances and suspension of operation. Figure 11 in appendix A displays the distribution of these measures.

Finally, we observe whether preventive measures were taken by the polluter, i.e. measures that ensure that the cause of the conflict would not reappear. When no such measures were taken (13.1% of all cases), we observe why this was the case. The main reasons are lack of money (2.1 % of all cases with no preventive measures), technological constraints (6.5%) or constraints imposed through other laws and regulations (1%). In all other cases, no measures were necessary because of an agreement between the parties (47%), or we do not observe the precise reason (43.4%).

Treatment An environmental conflict can be assigned to one of three conflict resolution methods by the local environmental counselor. The first one consists of administrative guidance for the polluter (91% of all cases). In this method, a solution is elaborated by the counselor and imposed on the parties. Thus, this method resembles conventional arbitration. One difference, however, is that the method is asymmetric in its consequences in the sense that the measures for changing the status quo are taken by the polluter only.

The second method consists of discussions between the parties (3%). The role of the counselor is to serve as a channel of communication, which is called facilitator mediation in the literature, see Wilkenfeld et al. (2003).[9]

[9] It is also known as “process communication” or “third-party consultation”.

In the last method, called "persuasion of complainant", the counselor negotiates with the plaintiff trying to persuade her to accept the status quo. A case is resolved only if the plaintiff agrees to that. This method is assigned to roughly 6% of all disputes.

From these three categories we construct a binary treatment variable $D_i \in \{0, 1\}$ in the following way. We define D_i to be equal to 1 whenever the solution method assigned to case i is either the second or the third one (and else D_i is equal to 0).

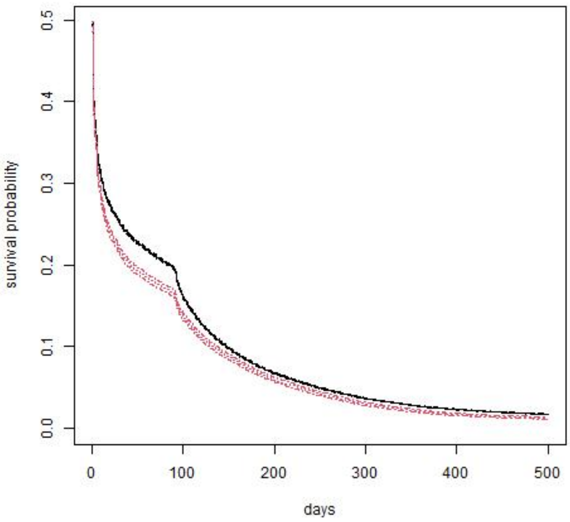
The main advantages of this treatment definition are the following. First, it allows to compare the performance of ADR methods that differ on an important dimension: the degree of control over the process and outcome. In particular, it amounts to comparing ADR methods, in which consensus of at least one of the parties is necessary ($D = 1$) to ADR methods, in which the solution is imposed in a top-down way. The literature refers to the former as "consensual" and to the latter as "commanded" (also referred to as mandated) ADR methods, see Menkel-Meadow (2015). Second, it allows to pool more observations into the treatment group, which improves the statistical analysis, and in particular it makes it possible to estimate Group Average Treatment Effects, as we do in section 4.2.2 below. As we demonstrate in the appendix, this pooling does not lead to a lost of information: the differences in performance between the two consensual methods are not significant.

Figure 2a displays the nonparametric Kaplan-Meier estimates of the survival probabilities for treated $D = 1$ (dotted line) and nontreated $D = 0$ (solid line) with the corresponding 95% confidence bounds (dashed lines).[10] The estimated survival distribution of the nontreated stochastically dominates the survival distribution of the treated, with the largest difference occurring around day 50 and being roughly 0.025. Thus, the conflicts solved by a method based on consensus have in general a shorter duration. Figure 2b shows the distributions of satisfaction for treated (black) and nontreated (grey).

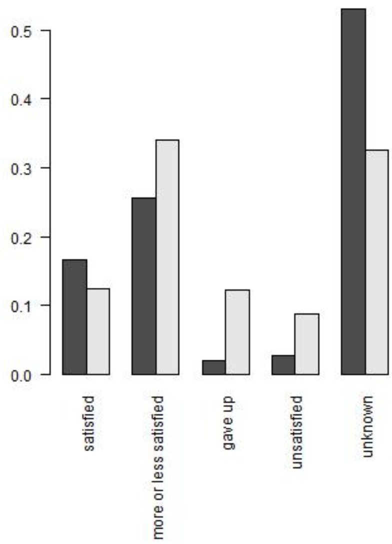
[10]A survival probability is defined as the probability that the duration of a conflict T will exceed a cerain number of days t: $P\{T > t\}$. The figure plots the conditional probabilities $P\{T > t|D = 1\}$ and $P\{T > t|D = 0\}$ for t between 0 and 500 days. The confidence intervals are calculated with the Greenwood formula.

Figure 3: Distributions of the two outcome variables for treated and nontreated. Left: distributions of the duration variable. Right: distribution of the satisfaction outcome.

(a) Kaplan-Meier estimates of the survival probabilities for treated (dotted line) and nontreated (solid line) with the corresponding 95% confidence bounds (dashed lines).



(b) Distribution of satisfaction for treated (black) and nontreated (grey).



A higher proportion among the treated rates the process as satisfactory. With respect to all other categories, those unsatisfied have a higher share among nontreated, and those who choose not to answer have a larger share among treated.

Pre-treatment conflict characteristics We a observe a rich set of conflict characteristics. First, we observe in which city the complaint has been filed. In addition, we observe the type of area (commercial, semi-commercial, residential, industrial, semi-industrial, special industrial, controlled urbanization, urban planning, and others) in which the reason for conflict has occurred. By far the most frequent area type is the residential (44%), followed by the controlled urbanization (14.5%) and the semi-industrial type (9%), see figure 12 in appendix A. Second, we observe the precise type(s) of pollution that caused the conflict. The most common types of pollution are air pollution(36%), offensive odor (17 %) and water pollution (7.5%), see figure 13 in appendix A. When the pollution is waste, we also observe the precise type of waste (household (34.4% of all waste-related cases), industry (28.1%), agriculture (23.1%) and construction waste (14.3 %)). Furthermore, we also observe the corresponding reason for each pollution case. We are able to distinguish between 17 different reasons/sources of pollution such as burning fields (28%), construction work (22.1 %) and industrial waste water (2.7%), see figure 14 in appendix A. In addition, we observe both the frequency as well as the precise day time in which the pollution occurs (figures 15 and 16 in appendix A, respectively).

On the plaintiff's side, we observe the type of person who filed the complaint (the plaintiff herself (85.6%), on behalf of several individuals (5%), mediated by a public institution (3.6%), mediated by a third party (< 1%), and others), her/his residential code, how many households were affected (see figure 17), and the type of damage incurred. The latter can be physical health damage (8.6%), property damage (1.5%), animal or plant damage (< 1%), sensory or psychological (83.4%), or others.

Finally, we observe regulatory characteristics of each case. In particular, we observe which type of environmental regulation is related to the particular case (air pollution control act (19%), water pollution control act (6%), "soil contamination countermeasures" law (¡0.5%), noise regulation law (20%), vibration regulation law (3%), offensive odor prevention law (11%), local government pollution law (26%) and others), whether the regulation has actually been violated, and if yes, what type of violation has occurred (standard vs. license violation). Like all other pretreatment characteristics, these regulatory details are filled by the local environmental counselor upon receiving the complaint and prior to assigning it to a resolution method.

3 Empirical framework

3.1 Treatment effects of interest

3.1.1 Average Treatment Effects

We study treatment effects within the Rubin Causal Model, Imbens and Rubin (2015). Let $Y_i(d)$ be the potential outcome for the hypothetical case that individual i assigned to treatment d, where $i \in \{1, \ldots, n\}$ and $d \in \{0, 1\}$, while Y_i is the observed outcome. Furthermore, let X_i represent the random vector that consists of all observed pretreatment characteristics of case i. With this notation, we can define the effects of interest.

Treatment effects on a duration outcome. The main effect we consider in this paper is defined as $\Delta_1 = \{E\}[Y(1) - Y(0)]$, where $Y(d)$ is the potential duration of a dispute. Thus Δ_1 measures the average effect from resolving a dispute with a consensus-based approach ($d = 1$) relative to a top-down approach ($d = 0$) on the duration of that dispute. The reason we focus on this effect is that it is easiest to interpret, identify and estimate. Moreover, studying heterogeneity of treatment effects is also easiest when one focuses on average additive effects.

In addition, we also study the average additive treatment effect on the survival function, $\Delta_2(t) = \{E\}[Y_t(1) - Y_t(0)]$, where for a given elapsed duration t, $Y_t(d)$ indicates whether an individual survives longer than t: $Y_t(d) = \mathbb{1}\{Y(d) > t\}$.[11] Estimating $\Delta_2(t)$ for a variety of t delivers insights where in the duration distribution the losses or gains from the treatment are realized. For each t, $\Delta_2(t)$ can be identified and estimated in an analogous way as Δ_1.

Finally, we also study treatment effects on the hazard rate of the duration of a dispute $\theta_{Y(d)}$. The hazard rate represents the "exit rate" out of a conflict for those conflicts that have continued at least up to t. The hazard is helpful because it provides insights on the gains of the treatment for those conflicts that typically last longer. We consider the average multiplicative effect $\Delta_3(t) = \mathbb{E}\left(\frac{\theta_{Y(1)}(t)}{\theta_{Y(0)}(t)}\right)$ on the hazard, as well as the average additive effect $\Delta_4(t) = \mathbb{E}[\theta_{Y(1)}(t) - \theta_{Y(0)}(t)]$.

The treatment effects $\Delta_1, \Delta_2, \Delta_3$ and Δ_4 represent a comprehensive set of treatment effects that can be considered when the outcome is a duration variable, Abbring and Berg (2005).

Treatment effects on the satisfaction. The outcome "satisfaction" is a categorical, partially ordered variable with 5 categories ("1 satisfied", "2 more or less satisfied", "3 unsatisfied", "4 plaintiff gave up" and "5 no answer"). One strategy is therefore to estimate a separate average treatment effect for each of the 5 categories. Denote the corresponding treatment effects by Δ_k^S, $k = 1, \ldots, 5$. In addition, we construct an aggregated binary outcome variable by grouping together favorable outcomes (categories 1 and 2), $Y = 1$, and unfavorable outcomes (categories 3-5), $Y = 0$. An implicit assumption here is that the environmental counselor would prefer to know the level of satisfaction, so that no answer is considered as an unfavorable outcome. We denote this aggregated treatment effect by Δ_S.

[11] $\Delta_2(t)$ can be also written as $P\{Y(1) > t\} - P\{Y(0) > t\}$.

3.1.2 Group Average Treatment Effects

We also consider Group Average Treatment Effects (GATEs). A GATE is defined as $\mathbb{E}[Y(1) - Y(0)|Z = z]$, where Z is a (possibly proper) subset of all observed case pre-treatment characteristics X. Thus, a GATE represents an average treatment effect for a subgroup of cases that shares given characteristics. Studying GATES helps understand who are the "winners and losers" from using a given ADR method (treatment effect heterogeneity). Because of their computational intensity, we study GATES only for Δ_1 and Δ_S.[12]

3.2 Identification of treatment effects

Identification of $\Delta_1, \Delta_2, \Delta_k^S, \Delta_S$. Our main assumption is the so called *Conditional Independence Assumption (CIA)* (or "Selection on observables"), which postulates that conditional on observed covariates X, the treatment is independent from potential outcomes. Formally, $D \perp\!\!\!\perp (Y(1), Y(0))|X$. This assumption is motivated by the institutional setup. In particular, we observe all dispute characteristics that are known to the environmental counselors at the point in time in which they make an assignment decision. Thus, any remaining randomness in the assignment process must be caused either by differences in the preferences across counselors or by uncertainty in the counselors' decision making process. The latter can be driven by e.g. learning effects or by true idiosyncratic intrinsic uncertainty. [13] Thus, since we observe in which environmental constituency a given case is solved, the CIA assumption amounts to treating the intrinsic randomness of environmental counselors as independent of the potential outcomes.

Common support assumption. Our second assumption is that which states that no

[12] $\Delta_2, \Delta_3, \Delta_4$ depend on t and are estimated for a large number of t, it is not feasible to estimate GATES.

[13] Experimental evidence for unpredictability in arbitrators' decision making process is provided by Ashenfelter and Bloom (1984), see also the discussion in Ashenfelter (1987). Individual intrinsic uncertainty commonly used in doubly stochastic models such as duration models, see Chesher (2002), Heckman (1991) and Bonev (2020) for a discussion.

set of case characteristics perfectly determines the treatment status. Formally,

$$0 < P\{D = 1|X = x\} < 1 \qquad \text{for all } x. \tag{1}$$

We study empirically its validity in section 4.1.2.

SUTVA. Our third assumption is the Stable Unit Treatment Assumption (SUTVA), which states that there are no equilibrium effects. Since each case is solved "for itself", i.e. in isolation from all other cases, this assumption is satisfied per design of the system.

A common result in the treatment evaluation literature is that under CIA, Common Support and SUTVA, the treatment effects $\Delta_1, \Delta_2, \Delta_k^S, \Delta_S$ are identified.

Identification of Δ_3, Δ_4. Identification of the two treatment effects on the hazard is more involved because of dynamic selection. In particular, at a point in time t, the distributions of the unobservable factors of the outcome will differ among treated and nontreated even conditional on unobservables. This holds because the unobservables interact with the treatment status and over time, the exit process differ among the two groups, see Van den Berg (2001). Therefore, the approach in the biostatistics literature (see e.g. Higbee et al. (2020) and Burnett et al. (2018)) to estimate the observed hazard ratio $\frac{\theta_Y(t|D=1,X)}{\theta_Y(t|D=0,X)}$ does not lead to estimands of the treatment effect Δ_3. Similar complications arise with Δ_4.

We solve this problem by deriving a novel identification result for Δ_3. In Lemma B.1 in the appendix, we show that under multiplicative separability of the treatment effect and under a generalized separability of the unobserved heterogeneity, Δ_3 is identified and equal to $\lim_{t\to 0} \frac{f_Y(t|D=1)}{f_Y(t|D=0)}$, where f_Y denotes the density of the duration variable Y. Our results generalizes the identification result presented by Abbring and Berg (2005). Both additional assumptions are common in the literature on the Mixed Proportional Hazard model, Van den Berg (2001). We also show that under the conditions of Lemma

B.1, $\ln \Delta_3$ and Δ_4 have the same signs. Estimating Δ_3 is therefore sufficient to identify the sign of the additive effect on the hazard.

3.3 Estimation

Estimation of ATEs. Our main approach for estimating $\Delta_1, \Delta_2, \Delta_k^S, \Delta_S$ is the recently developed Doubly robust Machine Learning (DML) approach von Chernozhukov et al. (2018). It is a fully nonparametric estimator that relies on the so-called score. We provide a brief explanation in appendix C.1. A detailed and intuitive introduction into estimation with DML can be found in Knaus (2020). We estimate all nuisance parameters (conditional expectations of the outcome and the propensity score) with flexible random forests estimators.

To estimate $\Delta_3 = \lim_{t \to 0} \frac{f_Y(t|D=1)}{f_Y(t|D=0)}$, in a first step we estimate the densities $f_Y(t|D = 1), f_Y(t|D = 0)$ at a vicinity of 0 with nonparametric kernel estimators. In a second step, we plot these density to graphically inspect the limit. The reason for this ad hoc procedure is that estimating a limit is a challenging task which requires many observations very close to 0.

We do not estimate Δ_4 for two reasons. First, its sign is identified from the sign of Δ_3. Second, estimation of Δ_4 requires substantial parametric assumptions.

Estimation of GATEs. To estimate the GATEs, we follow the newly developed doubly robust machine learning approach described in Semenova and Chernozhukov (2021). In particular, we regress the predicted differences in scores on observed covariates Z, see appendix C.2 for a brief explanation.

4 Empirical results

4.1 Analysis of the treatment assignment

The main objectives of this section are twofold. First, we study the determinants of the treatment assignment: How do local environmental counselors decide which type of resolution method to assign to a given case? Are there certain cases which (almost surely) determine the resolution method? Is there heterogeneity in the way environmental counselors use their assignment discretion? Second, we study the validity of the common support assumption.

4.1.1 Determinants of the treatment assignment

Our quantitative analysis consists of two parts. Both parts are based on a regression analysis, in which the propensity score $P\{D = 1|X\}$ (i.e. the probability to be treated) is estimated as a function of observed case characteristics X. $P\{D = 1|X\}$ is estimated in two different ways. Our baseline model is a flexible Logit regression model with time and regional dummies, see appendix D.1.1 for the precise model specification. The second one is a flexible nonparametric random forest approach. The results from these approaches are helpful in assessing the role of the parametric specifications. In appendix D.2, we show that these two approaches lead to nearly identical out-of-sample prediction. Thus, we use the simple Logit to interpret the estimation results, while in subsequent estimation procedures that potentially involve different sample definitions, we use the nonparametric random forest approach.

Part 1. In part 1, we assess the *relative* importance of the covariates in the assignment process. This can be inferred from the signs of the coefficient of the logit model. The estimated coefficients are displayed in table 2 in appendix D. The following conclusions can be made. First, the year fixed effects are not significant, implying that there is not much change in the assignment process in 10 years of observation. Second, the majority of the pollution type dummies have a positive and significant estimates. Since the omitted pollution type dummy is air pollution, these results indicate that air pollution is more likely to be assigned to the top-down approach ($D = 0$). Consistent with this finding, the estimates of the pollution source/reason dummies indicate that incinerator pollution (the omitted category) and burning fields increase the probability for $D = 0$. Third, cases that are registered by public institutions on behalf of private complainants are less likely to end up in a consensus-based approach than cases registered by the complainants themselves. Fourth, complaints from residential areas (the omitted category) are more likely to be assigned to $D = 1$. Fifth, the coefficients of the number of affected households dummies show no monotonic relationship, indicating that counselors do not pay attention to how many households are affected. In particular, more affected households do not necessarily have a larger (in magnitude) relative importance. Sixth, the coefficients of the different regulation types that are related to the case are predominantly insignificant and are with small coefficients. One exception is the coefficient of the indicator that a local government regulation is affected, which is negative and highly significant. Thus, such cases are more likely to end up with $D = 0$. Finally, the dummies for non-violation of an environmental or non-environmental regulation have large positive and significant coefficients, which implies that violations increase the probability that the environmental counselors assign these cases to administrative measures for the polluter.

One drawback of this analysis is that, with all attributes being represented by mutually exclusive dummy variables, the coefficients display the *relative* importance of a certain value of a given covariate compared to other values of the same covariate. In particular, it is hard to infer the predictive importance of a given covariate or a value based on the estimates alone. To mitigate this problem, we perform two further sets of analysis, each based on the estimated propensity score.

Part 2. In part 2, we evaluate which case characteristics are associated with very high or very low predicted propensity score is associated with certain case characteristics.

Cases characterized by such characteristics will be particularly likely to be assigned to either the consensual or the top-down approach. However, it is technically very difficult to estimate $P(\text{high propensity score}|X)$. Therefore, we estimate instead

$$P\{X_j = x_j|e \leq 0.01\} - P\{X_j = x_j|e \geq 0.01\} \qquad \text{and} \tag{2}$$

$$P\{X_j = x_j|e \geq 0.9\} - P\{X_j = x_j|e \leq 0.9\}, \tag{3}$$

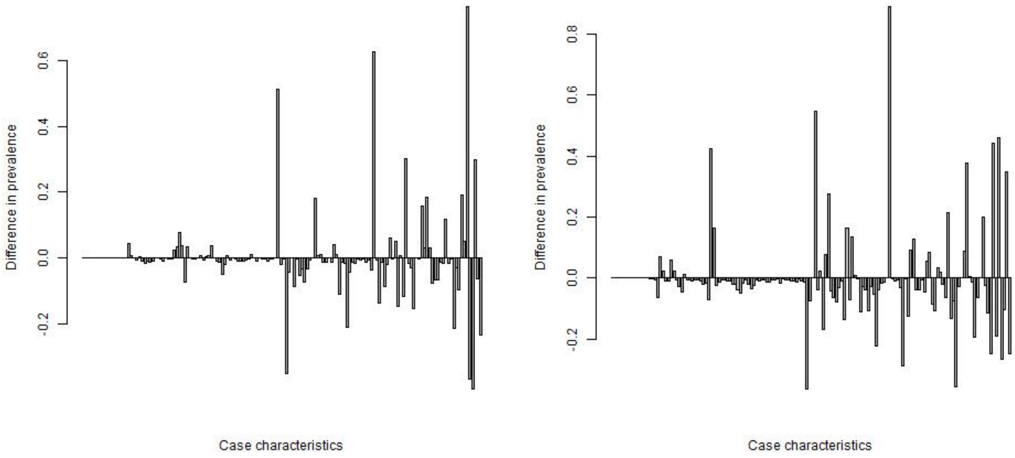
where $\hat{e}$ denotes the probability $P\{D = 1\}$, X_j is a given observed characteristics of a conflict, and x_j is any value from the support of X_j.[14] A high difference (2) (or (3)) for a value x_j would indicate that this value has a high prevalence among characteristics with (low) high treatment probability, so that it predicts a propensity score close to 0 (or 1).[15] The choices of the thresholds 0.01 and 0.9 reflect availability of observations near both extremes. Figures (4a) and (4b) display the estimates of (2) and (3) respectively. On the x-axis of each figure, each bar represents a given value of a given variable. The y-axis represents the differences (2) and (3). The highest value of (2) is realized for cases in which an environmental regulation is violated. The share of such cases among those observations with $\hat{e} \leq 0.01$ is 86% compared to 9.2% among all other observations, implying a difference (2) of 0.76. Thus, whenever a regulatory rule is violated, environmental counselors tend to assign administrative measures to the polluter (a non-consensus based approach). However, violation of non-environmental regulations appears to be less important, with a prevalence of 36% among the $\hat{e} \leq 0.01$-observations and 6.5% among all others. Further important reasons for assigning a case to $D = 0$ are when the pollution source is burning fields (a difference (2) equal to 0.62) and when the type of pollution is air pollution (a difference (2) equal to 0.51). Interestingly, these two

[14]Note that this is possible because all observed independent variables are discrete (nominally-scaled) characteristics.

[15]To see this, note that due to the Bayes law, it holds $P\{X|\hat{e}\} = \frac{P\{\hat{e}|X\}P\{X\}}{P\{\hat{e}\}}$. Since we are interested in $P\{\hat{e}|X\}$ and not in $P\{X|\hat{e}\}$, we have to take into account the weight of the characteristics X (i.e. $P\{X\}$) in the population. The differences (2) and (3) are an ad hoc way to do this.

Figure 4: Comparison of prevalence of characteristics for observations with very low (left) and very high (right) predicted propensity score.

(a) Estimates of equation (2). (b) Estimates of equation (3).



characteristics are not highly correlated with the number of households affected (less than 10% of cases with burning fields affect more than 5 households), or with the type of damage. In general, the type of damage seems not to be an important predictor of the treatment arm.

On the other hand, cases with pollution source being an aircraft or the type of pollution being noise pollution are particularly likely to be assigned to a consensus-based resolution approach (an estimated (3) of 0.89 and 0.54, respectively).

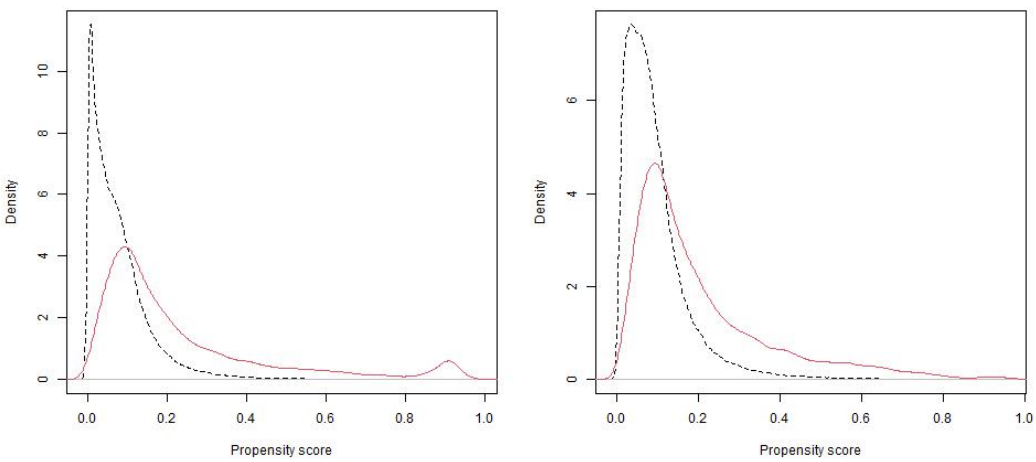
Summary of the evidence. The following conclusions can be drawn based on our results. First, a violation of a regulation leads with a high probability to a top-down resolution approach in which environmental counselors decide which measures must be taken by the polluter. Because counselors are not compelled by law to decide in this way, the tendency reflects their beliefs about the chances that such cases are properly resolved by the involved parties themselves. In this sense, counselors act as if they are risk averse. Second, the type and source of pollution are the two other major predictors of the assigned treatment arm. The type of damage, on the other hand, plays no important role as a predictor. One potential explanation is that cases with sufficiently high damages have been considered as criminal offenses and have been directed to the police, which "levels the ground" for all types of damage. And last but not least, the type of complainant and the number of affected households are also not important predictors of the treatment. This implies that the relative bargaining power of the complainant plays only a negligible role in the assignment decision.

4.1.2 The common support assumption

In this section, we study the validity of the common support assumption. To this end, we estimate the densities of the propensity score separately for treated and non-treated. The two densities are shown in figure 5a. The solid line represents the estimated density for the treated, and the dashed line for the nontreated.

Figure 5: Estimated densities of the distributions of predicted propenstity score for treated (solid line) and nontreated (dashed line).

(a) Estimation with the full sample. (b) Estimation excluding strong predictors.



The figure reveals that the common support assumption is violated for observations with both very low and very high propensity score. In particular, the support of the density of the nontreated ends before 0.6, while for the nontreated it extends almost until 1 with an bump-like increase between 0.8 and 1. Similarly, a large fraction of the nontreated have estimated propensity score between 0 and 0.01, which is not matched by density of the treated.

We deal with this problem in several ways. Our main approach is to "trim" the sample by simply excluding observations with very high or low propensity score. While this makes the estimation of treatment effects tractable, these effects are local since they are restricted to that particular group of observations. The obvious drawback is that this group is not straightforward to interpret, as a propensity score is not a direct characteristics of a case.

To account for this drawback, we complement the trimming approach with two further approaches. First, we exclude observations with characteristics that make a case particular likely to be assigned to either $D = 1$ or $D = 0$. In particular, based on our above analysis of the prediction power of covariates, we exclude observations that either (i) violate an environmental regulation or (ii) have a pollution type that is aircraft. The new sample consists of 224783 observations. To see how this exclusion impacts the supports, we re-estimate the densities of the propensity score for treated and nontreated on this restricted sample. These densities are shown in figure 5b. The common support is substantially increased both on the left and on the right, even though for propensity scores larger than 0.4 there is still lack of support.

Our final approach to deal with the lack of common support is regression based-extrapolation, which we describe in the robustness checks in appendix E.

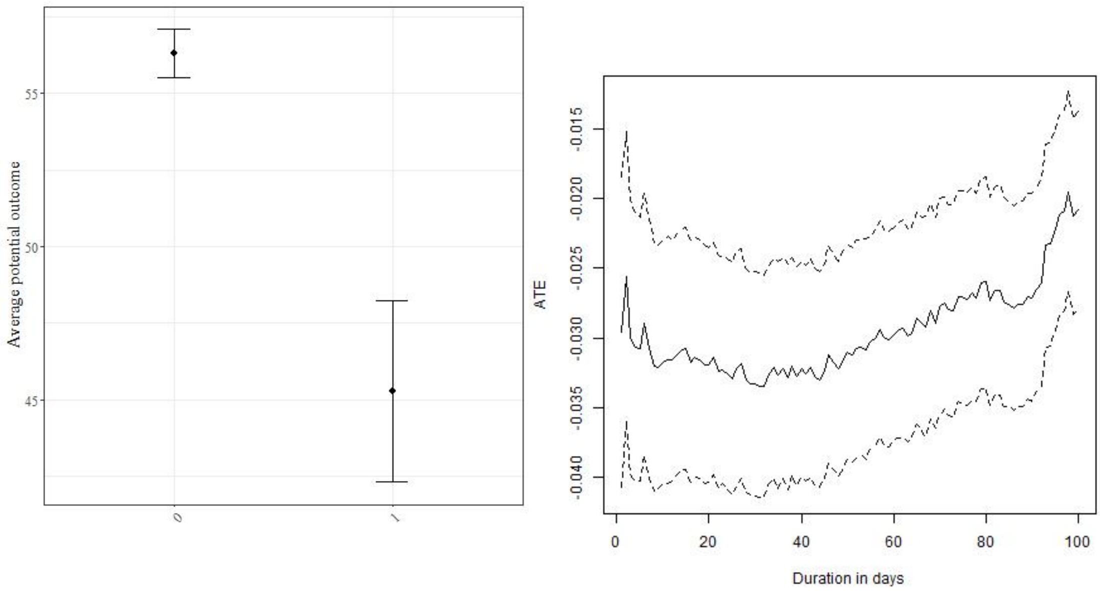
4.2 Empirical comparison of consensual and top-down ADR methods

4.2.1 Average treatment effects

ATE on duration outcomes. Consider first the average treatment effect Δ_1 on the duration of a dispute. Figure 6a displays the estimated average potential durations under $D = 0$ and $D = 1$, that is $\hat{E}[Y(0)]$ and $\hat{E}[Y(1)]$, as well as the corresponding 95% confidence bounds.[16] The difference $\hat{\Delta}_1 = \hat{E}[Y(1)] - \hat{E}[Y(0)]$ is negative and equal to -11.045 days. A 95% confidence interval is equal to $[-14.08, -8.00]$, which implies that the estimate is statistically significant. To put it into perspective, $\hat{\Delta}_1$ equals 20% of the average duration and about 6% of the standard deviation in the sample. This suggests that the effect is also economically significant.

Figure 6: Average treatment effects on duration outcomes.

(a) Average potential durations. (b) $\hat{\Delta}_2(t)$ for different survival times t.



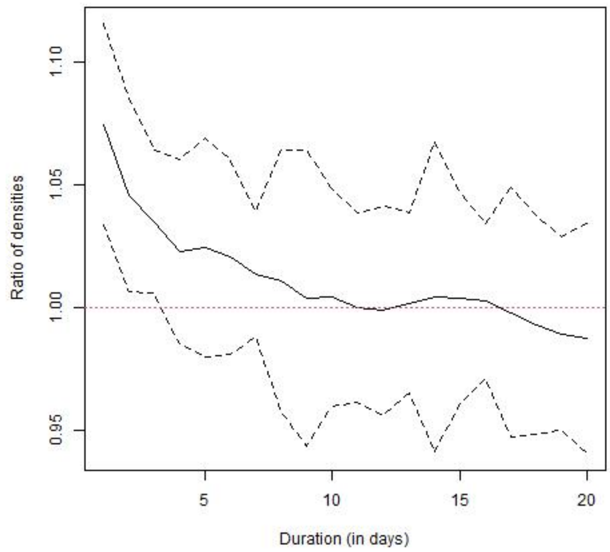
Next, we consider the average treatment effect on the survival function $\Delta_2(t)$. The estimated effects for different survival times t are plotted in figure 6b. For any t, the

[16]Here, $Y(d), d = 0, 1$ denotes potential durations.

estimate $\hat{\Delta}_2(t)$ is negative and statistically significant, and ranges between -2.0% and -3.5%.

Finally, figure 7 displays the ratio of the estimated densities $\frac{f_Y(t|D=1)}{f_Y(t|D=0)}\lim_{t\to 0}\frac{f_Y(t|D=1)}{f_Y(t|D=0)}$ for a variety of survival times t. As shown in Lemma B.1 in the appendix, the limit of this ratio at $t = 0$ can be interpreted as an estimate of Δ_3. The ratio is smooth near 0 and appears to converge to 1.075. Thus, the consensual ADR method increases the exit rate out of a conflict by 7.5% relative to the top-down approach. The limit is also significant at the 5% level, which is revealed by the confidence bounds (the dashed lines) near $t = 0$.

Figure 7: A ratio of estimated densities. The limit at 0 provides an estimate of the multiplicative ATE Δ_3.



Thus, all three treatment effects suggest that a consensual ADR approach leads on average to a shorter dispute duration. The gains are realized for all survival times, both as an ex ante probability, as well as an exit rate for conflicts that have "survived" up to a given duration.

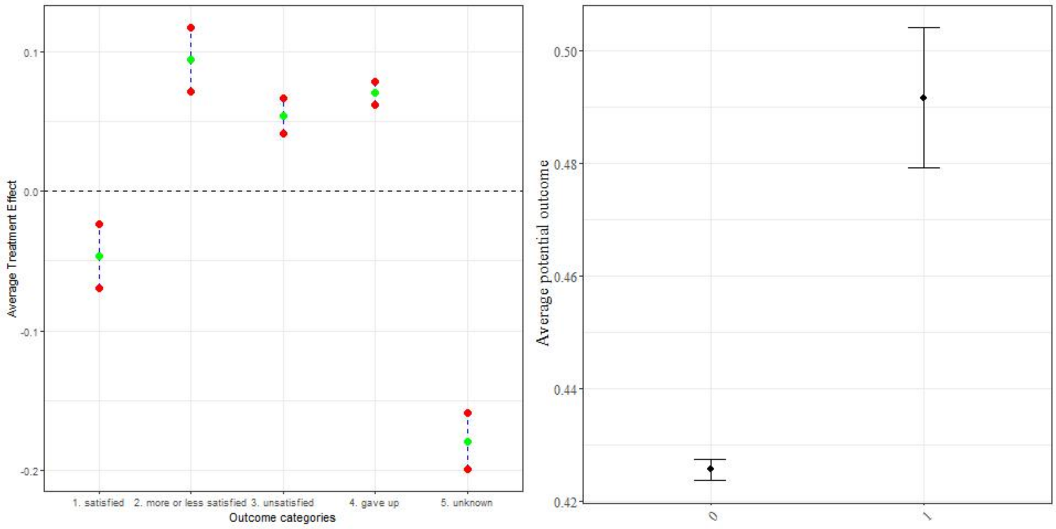
ATE on satisfaction outcomes Consider next the satisfaction outcome. Figure 8a shows the estimated treatment effects Δ_k^S, $k = 1, 2, \ldots, 5$, for all five satisfaction categories separately. The consensus-based approaches lowers the probability that the plaintiff is satisfied with a case by 4.6% and the effect is statistically significant. However, the effect on the likelihood of being more or less satisfied is positive and equal to 9.5%, and it is also statistical significant. The other three effects are 7.03% (with a standard error of 0.012), 5.37% (0.0065), and -17.91% (0.01). This provides an exact characterization of the changes in the distribution of satisfaction due to the treatment. This characterization, however, is difficult to interpret because of ex ante differences in the shares of the different outcomes. As an example, while the relative change in category 3 is larger than the change in category 1, the latter has a larger ex ante share (see figure 2b). We therefore study the estimated effect on the dichotomized outcome. Figure 8b presents the estimated average potential outcomes (APOs) $\hat{E}[Y(1)], \hat{E}[Y(0)]$ (where now $Y(d)$ denotes potential satisfaction). The difference corresponds to the ATE Δ_S. This difference is positive and equal to 6.5%. The corresponding 95%-confidence interval is equal to $[0.053, 0.078]$.

Thus, our results suggest that on average, a switch from a top-down to a consensus-based conflict resolution approach leads also to an increased satisfaction of the plaintiff.

Robustness checks. Note that the above results were obtained on the full sample. To account for a possible violation of the common support assumption, we conduct a variety of robustness checks. First, we reestimate the ATEs using a trimmed sample which excludes observations with very small (≤ 0.01) or high (≥ 0.4) estimated propensity score. Second, reestimate the ATEs we exclude cases that violate environmental regulations, or result from burning fields or aircraft as pollution source. Finally, we estimate the ATEs using regression-based approaches that extrapolate on the regions of no overlap. The results and approaches are described in appendix E. All three approaches yield results consistent with the main results above. In all sections below, we use a

Figure 8: Average treatment effects on satisfaction outcomes.

(a) ATE on all five satisfaction categories. (b) APO, dichotomized satisfaction.



trimmed sample approach.

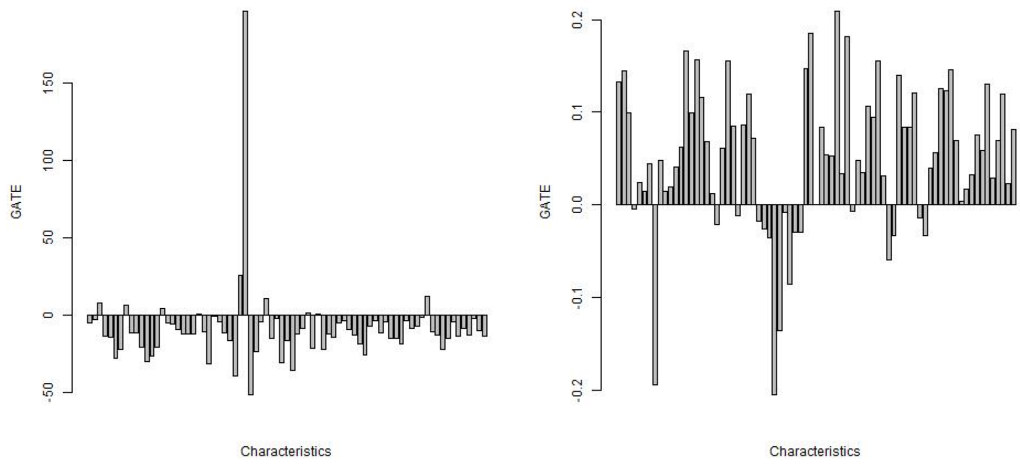
4.2.2 Treatment effect heterogeneity

In this section, we study two questions. First, which cases benefit from a consensual and which from a top-down approach? Here benefit is defined as shorter duration and higher plaintiff satisfaction. Second, if short duration and high satisfaction were the objectives of the environmental counselors, is the way the treatment assignment is practiced on average compatible with these objectives?

Treatment effect heterogeneity. To answer the first question, we estimate GATEs (i.e. $\mathbb{E}[Y(1) - Y(0)|Z = z]$) for each possible value z of each pretreatment characteristic. The results are displayed in figure 9a for the duration and in figure 9b for the satisfaction outcome. The x-axis in both figures displays all possible values of all pretreatment characteristics. The y-axis measures the corresponding GATEs. The figures show that most duration GATEs are negative and most satisfaction GATEs are positive, which indicates gains for the specific groups of disputes from obtaining the treatment.

Figure 9: Group Average Treatment Effects (GATEs).

(a) Duration as outcome. (b) Satisfaction as outcome.



In particular, the majority of cases would have between 5 and 20 days shorter duration (on average) under the consensual approach (with the maximum benefit being 50 days reduction). Similarly, for most cases, the consensual approach would lead to an increase of plaintiff satisfaction between 5% and 20%.

Of particular interest are the GATEs for groups of cases with characteristics associated with particularly high or low propensity score. In section 4.1, we showed that violating a environmental regulation as well as having a dispute that results from burning fields is particularly likely to lead to a top-down resolution approach (i.e. a propensity score close to 0). The duration GATEs for these groups are -13.20 days and -2.33 days (p-values: 0.05 and 0.0003) , respectively, so that the duration of such disputes would on average be reduced by a consensus based resolution approach. Similarly, the satisfaction GATEs for these groups are 0.13 and 0.15 and both p-values are smaller than 10^{-6}. This suggests group-specific gains in satisfaction from a consensus-based approach.

On the other hand, consider now dispute cases associated with noise pollution or with pollution that is related to an aircraft source. We showed that such cases are particularly likely to be assigned to a consensus-based approach (high propensity score). Due to violation of the Common Support Assumption for $e(x) > 0.4$, however, most of these observations are excluded when estimating treatment effects, so that no reliable statement about their GATEs is possible. When we ignore the violation of the assumption and estimate GATEs using the full sample, we obtain duration GATEs for these two groups of -13.55 days and -51.88 days, respectively, which implies that the two groups would benefit in terms of duration from being assigned to a consensus-based approach. The satisfaction results are less clear, with GATES being small but negative (-0.0046 and -0.0013).

Thus, our results so far provide a first indication that the current practice of the treatment assignment is not compatible with the above stated objectives. In particular, while the majority of single-characteristic-GATEs indicate a benefit associated with a consensual approach, the majority of cases are assigned to the top-down approach. This mismatched is particularly relevant for characteristics associated with high probability to be assigned to the top-down approach.

Finally, we study which characteristics are associated with gains from arbitration $D = 0$. These characteristics can be grouped in three groups. The first one contains cases characterized by the pollution sources "car", "train" and "aircraft". The pollution type in these categories is mainly noise pollution and the damage is psychological. The second group contains disputes in which more than one household are affected. The third one contains cases in which the registering person is either representing a large group of complainants, or a public institution that acts on behalf of many complainants, or a third party acting on behalf of a large number of complainants.[17] We come back to the interpretation of this observation below.

Does treatment assignment lead to optimal outcomes (on average)? To

[17]These characteristics are associated with gains for both outcomes duration and satisfaction. Additional properties only for satisfaction are regular frequency of the pollution, leakage as a pollution source, soil contamination as pollution type and controlled urbanization area as a pollution location.

answer this question, one has to compare (1) the likelihood of a given case to be assigned to the treatment with (2) its gains from the treatment. The former corresponds to the propensity score of a case with particular observed characteristics x, $e(x) = P(D = 1|X = x)$. The latter corresponds to the conditional average treatment effect of this particular $X = x$, $\mathbb{E}[Y(1) - Y(0)|X = x]$. An optimal treatment assignment process would yield treatment effects that depend monotonically on the propensity score: the higher the propensity score, the more favorable the effect of receiving the treatment.

To study the relationship between $e(X)$ and $\mathbb{E}[Y(1) - Y(0)|X = x]$, we adopt the following econometric approach. In a first step, using random forests and a cross-validation approach, we predict for each observation x_i its propensity score $\tilde{e}(x_i)$.[18] In a second step, we regress the outcomes Y (Duration or Satisfaction) on the predicted propensity score for treated and nontreated separately.[19] Because (i) the propensity score $e(X)$ is a balancing score,Rosenbaum and Rubin (1983), and because (ii) $\hat{e}(x)$ is predicted on a training sample, this procedure yields consistent estimates of the average potential outcomes, that is, $\hat{E}[Y|\hat{e}(x), D = 1] \xrightarrow{p} E[Y(1)|e(x)]$ and $\hat{E}[Y|\hat{e}(x), D = 0] \xrightarrow{p} E[Y(0)|e(x)]$.

The estimated conditional expectations are plotted against the estimated propensity score in figures 10a and 10b (duration and satisfaction, respectively). In both figures, the dashed black line denote the conditional expectation $\hat{E}[Y|\hat{e}(x), D = 1]$, i.e. the regression for the subsample of the treated, while the red continuous line denotes the conditional expectation for the nontreated ($D = 0$). For a given argument $e(x)$ on the x-axis, the difference between the two lines can be interpreted as the average treatment effect for all individuals with that propensity score value $e(x)$, i.e. $E[Y(1) - Y(0)|e(x)]$, while integrating over all $e(x)$ yields the ATE. Note that the x-axis is truncated at 0.4, because beyond this point the Common Support assumption is violated.

Figure 10a reveals that for very small values of the propensity score ($e(x) < 0.03$),

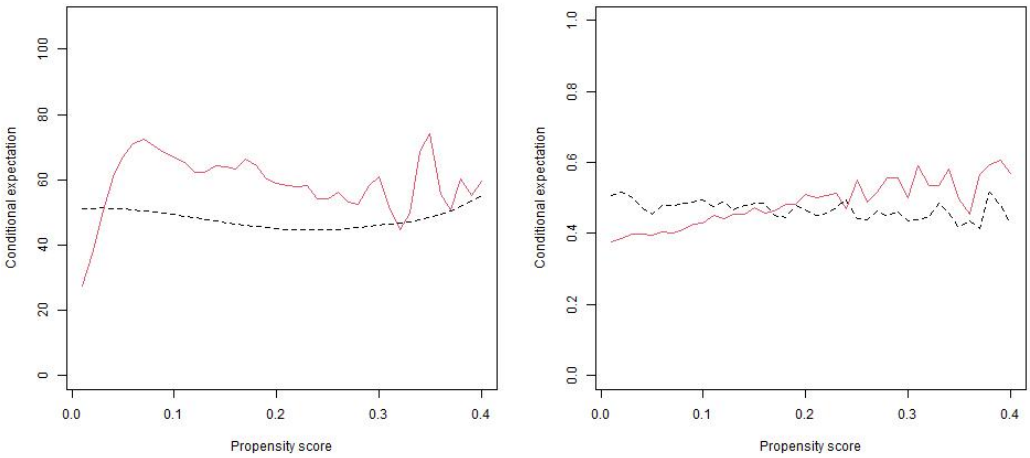
[18]We thank Michael Knaus and Michael Lechner from the University of St. Gallen for suggesting this procedure to us.

[19]We use a simple Nadaraya-Watson kernel regression for the duration variable and a neural network for the binary satisfaction variable.

Figure 10: Nonparametric regressions of outcomes on the propensity score for treated (dashed black line) and nontreated (continuous red line).

(a) Duration as outcome.

(b) Satisfaction as outcome.



the average duration of the treated exceeds the average duration of the nontreated. This indicates that these observations have been optimally assigned. For all other values, however, counselors assign the cases to a resolution method that leads to a negative treatment effect. The number of these cases is 198806 or approximately 71% of all observations. Figure 10b displays a similar misspecification, with most of the observations assigned to a treatment arm that yields a negative treatment effect and the assignment process following a monotonic pattern that reverses the optimal one.

These estimates reveal that the treatment assignment process is not optimal with respect to the duration of a conflict and the satisfaction of the plaintiff.

4.2.3 Intepretation of empirical results in light of economic theory

Our empirical findings have several important implications. First, the consensual ADR approach appears to perform on average better relative to the top-down approach with respect to the available outcome measures. This result is somewhat contrary to existing theoretical predictions as well as to empirical results from lab experiments. As an example, Wilkenfeld et al. (2003) finds no significant difference between passive forms of mediation (which correspond in spirit to our consensual methods) and ADR methods in which the third party endorses a particular solution (close in spirit to our commanded ADR treatment). Similarly, Eisenkopf and Bächtiger (2013) find that a third party with enforcement powers is more likely to reach an optimal outcome than a mediator who cannot threaten punishment for noncooperation.

One possible explanation for these differences is that the literature has predominantly focused on the information flow between parties and on the role of the third party as an informational gate keeper, see e.g. Balzer and Schneider (2021) for a recent theoretical contribution. In a such a setting, the third party in top-down approaches has the advantage of being able to enforce the solution. Accordingly, lab experiments have been set up with this role in mind. However, while governing the information flow plays a key role in every ADR method, there are further dimensions that contribute to finding a solution. This holds particularly for mediation. As the paper of Fanning (2021) puts it (p. 2449), real-world mediation "has many other reputed benefits[...]. These include the mediator's skill at trading off bargainers' relative preferences on multiple issues, her acknowledgment of each side's grievances, her ability to create a less confrontational atmosphere for negotiation, and her ability to establish commonly accepted facts". Therefore, one possible explanation for our results are that these advantages of consensual approaches outweigh the benefits of top-down approaches related to enforcing the solution.

One further explanation for the superiority of consensus-based approaches lies in the nature of solution finding. In particular, in real-world settings, parties may explore a very large number of strategies to find a solution. Importantly, parties are informed about their own preferences and situations in a much better way than the third party. This information leaves potential room to find creative solutions and common interests. This is in contrast to a top-down approach, which "tends to produce only winners and losers - not solutions to joint problems", as formulated by the influential study of Carver and Vondra (1994). In lab experiments, on the contrary, there is always a limited number of possible strategies for each party. In addition, lab games are predominantly set up as zero-sum games in which a pie has to be divided. While such a design might be a reasonable approximation of real-world settings in some cases such as employer-employee wage bargaining, Ashenfelter and Bloom (1984), in environmental context, the damage is hard to be assessed in a monetary way and solutions may involve complex non-monetary compromises. This complexity is difficult to reproduce in a lab experiment.

Our treatment heterogeneity analysis sheds light on the nature of the advantages of consensual and top-down approaches. In particular, our results point out that consensual methods perform better when the transaction costs related to the solution finding process are relatively small. To see this, consider again the three groups of characteristics associated with gains from a top-down approach: (1) noise pollution from cars, trains and aircraft, (2) cases with numerous households affected and (3) cases in which an agent (a third party) represents the plaintiff. A common property of all three aforementioned groups is the complexity associated with solving a dispute. The first group primarily concerns large infrastructural issues related to the location of a street, rails or a flight corridor. Such issues typically involve multiple regional authorities and interest groups, which arguably involves increased administrative and bargaining costs. The second group represent cases that potentially require high coordination efforts. Finally, when involved parties hire agents, inefficiencies might arise because the communication gets more complex, see e.g. Schotter et al. (2000) for evidence from a lab experiment.

We support the claim that the above groups of cases are associated with transaction costs with descriptive evidence. In particular, table 1 in appendix A shows that these groups of disputes have on average longer duration and higher uncertainty (measured by the standard deviation of the duration). To avoid conflation of treatment selection and actual complexity, the results are presented separately for treated and nontreated.

From the point of view of economic theory, our findings align remarkably well with the predictions of the so-called Coase Theorem. Formulated first in Coase (1960), it states that under certain conditions (functioning price system and competitive markets), bargaining parties will find the efficient distribution of goods as long as (1) there are well defined property rights and (2) there are little to no transaction costs. In such cases, the efficient solution is reached regardless of how the property rights are distributed and there is no need for the planner to intervene. However, whenever transaction costs are high, different distributions of property rights will lead to different distributions of goods, and not all of the latter will be Pareto efficient. This gives rise to potential benefits from regulation. Our treatment effect heterogeneity findings reflect these theoretical statements. Thus, our results can be considered as an empirical test of the Coase Theorem. However, it must be pointed out that not all of the premises of the Coase Theorem hold in our setup. Most importantly, the goods and/or the related externalities are "traded" outside of a market, and there is no functioning price system, at least not directly (what is the price for having clean ground water?). Instead, the source of efficiency in the bargaining process comes entirely from people knowing how to deal with their problems better than a third party. In this sense, our results can be interpreted as evidence for an extended Coase-type result.

5 Concluding remarks

In this paper, we empirically compare consensus-based and commanded ADR approaches. We find that on average, consensus-based approaches shorten the duration of a dispute and increase the satisfaction of the plaintiff. Furthermore, we empirically establish a Coase-Theorem-type result: a third party top-down intervention may improve outcomes only when the transaction costs are high.

These findings have two important implications for policy design. First, our results point out the need to have a very good ex ante assessment of the transaction costs associated with solving a dispute. This is particularly important in ADR systems with compulsory assignment to an ADR method. Second, agents responsible for the assignment process should be incentivized to consider in their decisions the transaction costs rather than follow entirely the regulatory norms.

Finally, there are several open questions for future research. First and most importantly, the link between the measured outcomes (such as duration and satisfaction of the plaintiff) and welfare needs to be better understood. Particularly in environmental context, this is a challenging task because there is rarely an accurate monetary assessment of the damage. Moreover, research needs better understand the resulting satisfaction of the parties: is parties' stated satisfaction entirely driven by the outcome or is it also impacted by the way the outcome has been achieved? Given the importance and prevalence of surveys in the empirical ADR literature as a way to measure welfare, this distinction has not received its deserved attention, and theories of procedural utility (see e.g. Frey and Stutzer (2005)) might suggest a path to go forward.

A Further descriptive characteristics

Figure 11: Types of measures taken by the defendant to abolish the cause of the conflict. (1): relocation of offices, (2): relocation of machines and facilities, (3): improvement of machines and facilities, (4): repair and recovery of failures, (5): improvement of work method and usage, (6): change or shortening of sales and operation hours, (7): suspension of operations, (8): removal or recovery of causative substances, (9): preventive measures for complaint's buildings, etc, (10): others.

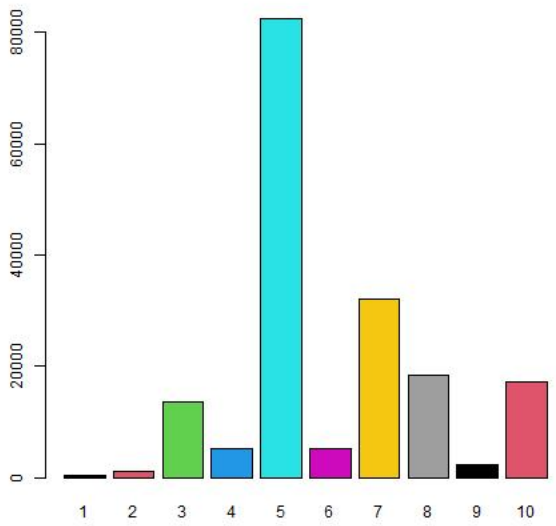


Figure 12: Type of location in which the reason for the conflict has occured.

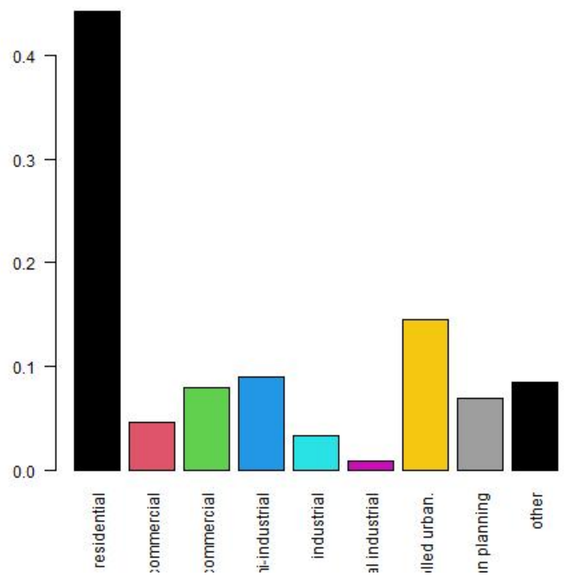


Figure 13: Type of pollution.

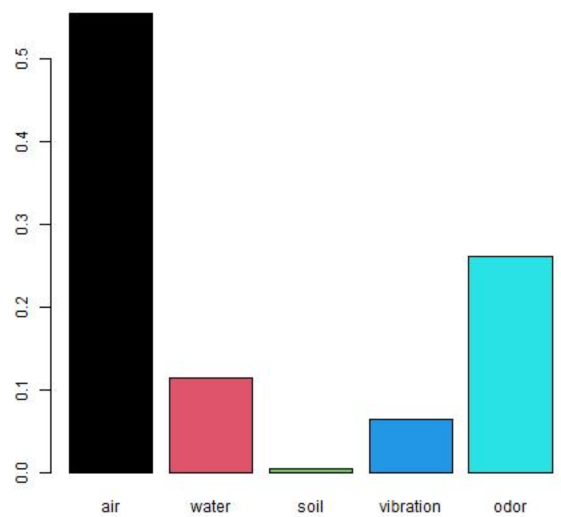


Figure 14: Most common reasons/sources of pollution.

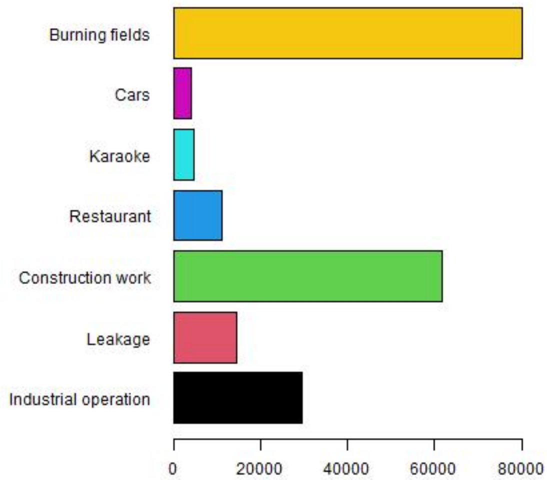


Figure 15: Frequency of problem/pollution reason.

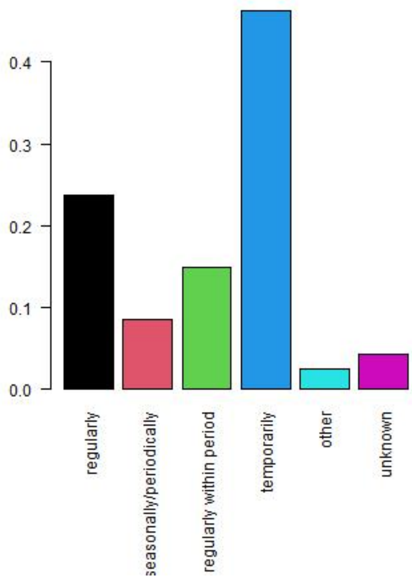


Figure 16: Time of the occurrence of problem/pollution reason.

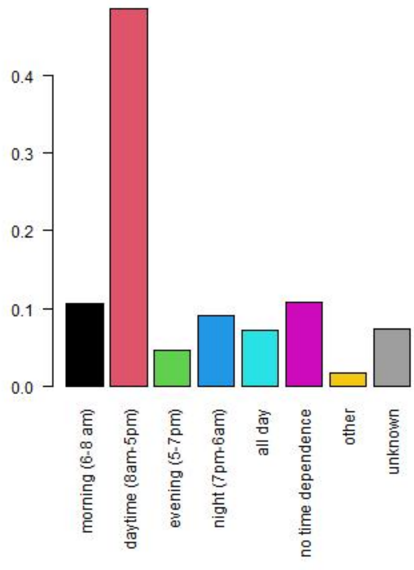


Figure 17: Number of affected households.

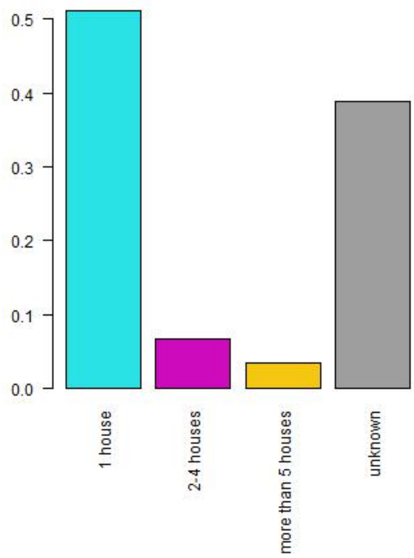


Table 1: Mean and standard deviation of duration for different characteristics

PANEL A: According to the source of pollution						
	Average Duration			Standard deviation		
	Transport*	All others	Difference (1-2)	Transport*	All others	Difference (1-2)
Treated	68.94	45.51	23.43	273.66	137.58	136.08
Nontreated	69.66	55.55	14.11	216.73	198.74	17.99
PANEL B: According to number of affected households						
	Average Duration			Standard deviation		
	1	> 1	Difference (2-1)	1	> 1	Difference (2-1)
Treated	53.52	61.10	7.58	162.73	168.31	5.58
Nontreated	69.03	77.33	8.3	207.50	263.43	55.93
PANEL C: According to whether there is an external agent involved						
	Average Duration			Standard deviation		
	NO	YES	Difference (2-1)	NO	YES	Difference (2-1)
Treated	48.39	45.51	-2.88	156.30	141.93	-14.37
Nontreated	59.53	40.64	-18.89	205.65	167.60	-38.05

B Identification of $\Delta_3(t)$.

To study identification of the effect on the hazard, we need some additional notation. Let V be a one-dimensional unobserved non-negative random variable that represents unobserved heterogeneity. Define $\theta_{Y(d)}(t|X,V)$ and $S_{Y(d)}(t|X,V)$ to be the hazard rate and the survival function of the **individual** potential duration $Y(d)$.[20] Define the following average treatment effects (ATEs):

$$\Delta_3(t|X) = \mathbb{E}_{V|Y\geq t}[\frac{\theta_{Y(1)}(t|X,V)}{\theta_{Y(0)}(t|X,V)}]$$
$$\Delta_4(t|X) = \mathbb{E}_{V|Y\geq t}[\theta_{Y(1)}(t|X,V) - \theta_{Y(0)}(t|X,V)]$$

Note that Δ_3, Δ_4 build the expectation over the distribution of the unobserved heterogeneity among survivors at a given elapsed duration, $V|Y \geq t$ while Δ_1 and $\Delta_2(t)$ over the total population ($Y \geq 0$).

We now turn to the identification of Δ_3. Abbring and Berg (2005) show that under a MPH structure of the nontreated potential hazard and a multiplicative treatment effect, the effect is identified under CIA, SUTVA and CSA. We now show that identification can be achieved in a much more general model. We need the following additional assumptions:

Separability of genuine duration dependence and unobserved heterogeneity (Separability for short): The nontreated hazard satisfies

$$\theta_{Y(0)}(t|x,v) = \psi_{Y(0)}(t,x)r(x,v), \tag{4}$$

where x, v are arbitrary values in the domains of X and V, respectively, and ψ and r

[20]The conditional hazard rate and survival function of a random duration $\tilde{Y}$ given $\tilde{X}$ are defined as

$$\theta_{\tilde{Y}}(t|\tilde{x}) := \lim_{dt\to 0} \frac{P\{\tilde{Y} \in [t, t+dt)|\tilde{Y} \geq t, \tilde{x}\}}{dt}$$

and $S_{\tilde{Y}}(t|\tilde{x}) = P\{\tilde{Y} \geq t|\tilde{x}\}$, respectively. The unconditional counterparts are defined analogously.

are positive real-valued functions such that $\int_0^t \psi_{Y(0)}(u,x)du$ is well defined and for each x in the domain of X, $\mathbb{E}[r(x,V)]$ is finite.

This is a very minimal requirement. In particular, unlike in the MPH model, observed and unobserved covariates are allowed to interact arbitrarily. Furthermore, the genuine duration dependence is allowed to depend on the value of the observed covariates. The multiplicativity of ψ and r ensures that dynamic selection (i.e. change of the distribution of risks over time with the share of bad risks increasing) can be distinguished from the genuine duration dependence. This assumption is the main assumption of the (fully separable) MPH model. Existence of $\int_0^t \psi(u,x)du$ and finiteness of $\mathbb{E}[r(x,V)]$ are also generalizations of MPH assumptions.

Nonparametric identification of nonseparable models of the type (4) are considered in Bonev (2020) in the context of non-treatment effect models. Identification of ψ and r is shown under additional assumptions such as access to multiple spells and monotonicity of r w.r.t its second component. We show that such assumptions are not necessary when the objective is to identify the treatment effect of D. Notably, we will not require independence of X and V, and assumption that is required by the standard MPH model.

In addition to the separability assumption above, we also need the following assumption:

Separability of the treatment effect (STE): It holds

$$\theta_{Y(1)}(t|x,v) = \beta_D(x)\theta_{Y(0)}(t|x,v), \tag{5}$$

where $\beta_D(x)$ is a real positive deterministic function of x.

Under STE, $\Delta_Y(t|x) = \beta_D(x)$. In the empirical literature, the STE assumption is always assumed through the specification $\psi(t,x) = \exp\{X\beta\}$. We allow the treatment effect to depend arbitrarily on observed covariates, which makes the model very general. We can now state the following result.

Lemma B.1. *Let CSA hold.*

1. *Then* $\Delta_3(t|x) - 1$ *and* $\Delta_4(t|x)$ *have the same sign.*
2. *If in addition SUTVA, CIA, Common Support and STE hold, then* Δ_3 *is identified from the data.*

Proof 1 (Proof of Lemma B.1)**.** *We first proof 2. Define* $V_x = r(x, V)$ *and let* G_x *be the distribution of* $V|X = x$*. Furthermore, denote by* L_F *the Laplace transform corresponding to a given distribution function* F*,*

$$L_F(s) = \int_0^\infty e^{-sy} F(dy).$$

We also define $\psi_{Y(1)}(t, x) = \beta_D(x)\psi_{Y(0)}(t, x)$ *and* $\Psi_{Y(d)}(t, x) = \int_0^t \psi_{Y(d)}(u, x)du$*. Using a well known result, it holds*

$$S_{Y(d)}(t|x) = \mathbb{E}[S_{Y(d)}(t|x, V)] = L_{G_x}(\Psi_{Y(d)}). \tag{6}$$

Furthermore, it holds

$$\frac{\partial S_{Y(1)}(t|x)/\partial t}{\partial S_{Y(0)}(t|x)/\partial t} = \frac{\beta_D(x)\psi_{Y(0)}(t, x)L'_{G_x}(\beta_D(x)\Psi_{Y(0)}(t, x))}{\psi_{Y(0)}(t, x)L'_{G_x}(\Psi_{Y(0)}(t, x))} \tag{7}$$

$$= \frac{\beta_D(x)L'_{G_x}(\beta_D(x)\Psi_{Y(0)}(t, x))}{L'_{G_x}(\Psi_{Y(0)}(t, x))}, \tag{8}$$

where the finite mean assumption is used to ensure that L_{G_x} *is finite. The last equality implies that*

$$\lim_{t\to 0}\Big(\frac{\partial S_{Y(1)}(t|x)/\partial t}{\partial S_{Y(0)}(t|x)/\partial t}\Big) = \beta_D(x) = \Delta_3(t|x). \tag{9}$$

To proof 1, let Ω_V be the support of the distribution function F_V of V. We can write

$$\Delta_4 = \int_{\Omega_V} \theta_{Y(1)}(t|x,V) - \theta_{Y(0)}(t|x,V) dF_{V|Y(d)\geq t, X=x} \tag{10}$$

$$\overset{CSA}{=} \psi_{Y(0)}(t,x)(\Delta_3(t,x)-1)\int_{\Omega_V} r(x,v) dF_{V|Y(d)\geq t, X=x}. \tag{11}$$

The proof follows because $\psi_{Y(0)}(.)$ and $r(.)$ are positive functions.

Several remarks are to be made. The main reason for the generality of this result is that no averaging over survivors is necessary. The proportionality assumption STE ensure that the individual treatment effect does not depend on observed heterogeneity and that thus no averaging over different distributions of survivors must be done. The separability assumption makes it possible to identify the effect without independence of X and V, as long as selection into treatment arms is random conditionally on observables (CIA).

C Estimation

C.1 Doubly robust estimation

The DML estimator relies on the equality

$$\mathbb{E}[Y(d)] = \mathbb{E}\Big[\mu(d,X) + \frac{\mathbb{1}\{D=d\}(Y-\mu(d,X))}{1\{D=1\}p(X) - 1\{D=0\}(1-p(X))}\Big] =: \mathbb{E}[G(d,X)], \tag{12}$$

where $\mu(d,X) = \mathbb{E}[Y|D=d,X]$ is the regression of the outcome on observed covariates (one regression for each group $d=1,0$) and $G_{d,i}$ is referred to as a "score".

The DML estimator is based on (estimated and) predicted values of $G(d,X_i)$ for each individual, $\hat{G}_{d,i}$ and a subsequent average:

$$\hat{\Delta} = \frac{1}{n}\sum_{i=1}^{n} \hat{G}_{1,i} - \hat{G}_{0,i} \tag{13}$$

C.2 Estimation of GATEs

To estimate GATEs, we follow the approach described in detail in Semenova and Chernozhukov (2021). In a first steps, we estimate the nuisance parameters using a leave-one-out approach. On the training sample, the scores are predicted. In a second step, using a simple simple OLS regression, the predicted nuisance parameter $\hat{\Gamma}_i = G_{1,i} - G_{0,i}$ (i.e. the difference in the predicted scores for each individual) is regressed on the covariates Z_i,

$$G\hat{A}TE(Z) = \text{argmin} \sum_{i=1}^{n} \left(\hat{\Gamma}_i - Z_i\beta\right). \tag{14}$$

D Analysis of the treatment assignment

D.1 Appendix to section 4.1.1 (Determinants of the treatment assignment)

D.1.1 Model specifications

With these specifications, our baseline logit model is

$$\ln \frac{P\{D_{ijt} = 1\}}{P\{D_{ijt} = 0\}} = \beta_0 + X_{ijt}\beta + \gamma_j \mathbb{1}\{prefecture = j\} + \eta_t \mathbb{1}\{year = t\} + \varepsilon_{ijt}, \tag{15}$$

where $\beta_0, \beta, , \gamma_j, \eta_t$ are coefficients, X_{ijt} are observed case characteristics of case i in region j in period t, $\mathbb{1}\{prefecture = j\}$ are region dummies, and $\mathbb{1}\{year = t\}$ are period dummies. To make the model computationally tractable, we define the dummies in the following way. First, the time dummies are yearly dummies. Second, the region fixed effects are on prefecture level. This decision is driven by the availability of the data - for over 90 cities, there is only one observation. Fixed effects on a city level would thus make it impossible to distinguish between the treatment effect and the time-constant impact of the region.

D.2 Alternative estimation methods: Random forests

We now assess the role of the parametric assumption in predicting the outcome. In particular, we estimate a Random forests regression. In a cross-validation procedure, we compare the out-of-sample predicted values of the treatment of the two regressions.[21] The two regressions agree on all but 352 cases, meaning that prediction based on the two approaches is equivalent.

[21] We use the optimal Bayes assignment rule: if a predicted propensity score exceeds 0.5, the predicted value is 1, otherwise 0.

D.3 Further tables and figures

Table 2: Baseline Logit Regression

	Coefficients	SE	t-stat	p-values
(Intercept)	-13.236	196.97	-0.067	0.947
year 1991	-0.290	240.28	-0.001	0.999
year 1993	0.746	278.55	0.003	0.998
year 1995	0.533	238.51	0.002	0.998
year 1996	0.565	278.55	0.002	0.998
year 1997	0.361	219.12	0.002	0.998
year 1998	9.912	196.97	0.050	0.960
year 1999	-0.974	208.96	-0.005	0.996
year 2000	0.053	202.34	0.000	1.000
year 2001	9.176	196.97	0.047	0.963
year 2002	8.839	196.97	0.045	0.964
year 2003	8.195	196.97	0.042	0.966
year 2004	7.713	196.97	0.039	0.969
year 2005	7.684	196.97	0.039	0.969
year 2006	8.683	196.97	0.044	0.965
year 2007	8.544	196.97	0.043	0.966
year 2008	8.663	196.97	0.044	0.965
year 2009	9.041	196.97	0.046	0.963
year 2010	9.049	196.97	0.046	0.963
year 2011	9.032	196.97	0.046	0.963
year 2012	9.041	196.97	0.046	0.963
year 2013	9.004	196.97	0.046	0.963
year 2014	9.051	196.97	0.046	0.963

year 2015	9.131	196.97	0.046	0.963
year 2016	9.024	196.97	0.046	0.963
year 2017	9.021	196.97	0.046	0.963
year 2018	9.162	196.97	0.047	0.963
prefecture 2	-0.110	0.11	-1.000	0.317
prefecture 3	0.123	0.10	1.230	0.219
prefecture 4	-0.094	0.08	-1.175	0.240
prefecture 5	-0.058	0.10	-0.580	0.562
prefecture 6	-0.230	0.10	-2.300	0.021
prefecture 7	-0.315	0.09	-3.500	0.000
prefecture 9	-0.065	0.07	-0.929	0.353
prefecture 10	-0.050	0.07	-0.714	0.475
prefecture 11	-0.280	0.06	-4.667	0.000
prefecture 12	-0.118	0.06	-1.967	0.049
prefecture 13	-0.136	0.05	-2.720	0.007
prefecture 14	-0.386	0.06	-6.433	0.000
prefecture 15	-0.321	0.07	-4.586	0.000
prefecture 16	0.098	0.11	0.891	0.373
prefecture 17	-0.420	0.11	-3.818	0.000
prefecture 18	0.109	0.09	1.211	0.226
prefecture 19	-0.496	0.10	-4.960	0.000
prefecture 20	0.338	0.07	4.829	0.000
prefecture 21	-0.214	0.08	-2.675	0.007
prefecture 22	-0.089	0.06	-1.483	0.138
prefecture 23	-1.127	0.07	-16.100	0.000
prefecture 24	-0.053	0.07	-0.757	0.449
prefecture 25	0.583	0.07	8.329	0.000

prefecture 26	-0.014	0.06	-0.233	0.816
prefecture 27	-0.041	0.06	-0.683	0.495
prefecture 28	-0.106	0.06	-1.767	0.077
prefecture 29	-0.182	0.09	-2.022	0.043
prefecture 30	0.246	0.09	2.733	0.006
prefecture 31	0.235	0.11	2.136	0.033
prefecture 32	0.056	0.12	0.467	0.640
prefecture 33	0.165	0.07	2.357	0.018
prefecture 34	0.455	0.07	6.500	0.000
prefecture 35	0.278	0.08	3.475	0.001
prefecture 36	0.250	0.10	2.500	0.012
prefecture 37	-0.077	0.10	-0.770	0.441
prefecture 38	-0.254	0.08	-3.175	0.001
prefecture 39	0.644	0.11	5.855	0.000
prefecture 40	0.536	0.06	8.933	0.000
prefecture 41	0.043	0.11	0.391	0.696
prefecture 42	0.116	0.08	1.450	0.147
prefecture 43	0.024	0.08	0.300	0.764
prefecture 44	0.096	0.08	1.200	0.230
prefecture 45	0.153	0.08	1.912	0.056
prefecture 46	-0.050	0.08	-0.625	0.532
prefecture 47	-0.242	0.08	-3.025	0.002
typepollution1 A02	-0.046	0.07	-0.657	0.511
typepollution1 A03	0.476	0.17	2.800	0.005
typepollution1 A04	0.307	0.04	7.675	0.000
typepollution1 A05	0.476	0.06	7.933	0.000
typepollution1 A06	2.084	0.27	7.719	0.000

typepollution1 A07	0.119	0.04	2.975	0.003
typepollution1 A041	1.859	0.11	16.900	0.000
location A2	-0.033	0.03	-1.100	0.271
location A3	-0.068	0.03	-2.267	0.023
location A4	-0.231	0.03	-7.700	0.000
location A5	-0.288	0.04	-7.200	0.000
location A6	-0.297	0.08	-3.712	0.000
location A7	-0.155	0.03	-5.167	0.000
location A8	0.106	0.03	3.533	0.000
location B1	-0.009	0.03	-0.300	0.764
typedamage 2	0.383	0.05	7.660	0.000
typedamage 3	-0.119	0.09	-1.322	0.186
typedamage 4	-0.489	0.02	-24.450	0.000
typedamage 5	-0.471	0.04	-11.775	0.000
pollutionreason B	0.471	0.05	9.420	0.000
pollutionreason C	-0.037	0.07	-0.529	0.597
pollutionreason D	0.041	0.06	0.683	0.495
pollutionreason E	0.256	0.05	5.120	0.000
pollutionreason F	0.450	0.06	7.500	0.000
pollutionreason G	-0.078	0.08	-0.975	0.330
pollutionreason H01	1.301	0.06	21.683	0.000
pollutionreason H02	2.282	0.11	20.745	0.000
pollutionreason H03	4.169	0.10	41.690	0.000
pollutionreason I	0.169	0.13	1.300	0.194
pollutionreason J01	1.535	0.06	25.583	0.000
pollutionreason J02	0.140	0.09	1.556	0.120
pollutionreason J03	1.182	0.06	19.700	0.000

pollutionreason K	-0.529	0.05	-10.580	0.000
pollutionreason L	1.970	0.08	24.625	0.000
pollutionreason M	0.494	0.05	9.880	0.000
pollutionreason N	2.666	0.06	44.433	0.000
typeperson 2	0.234	0.03	7.800	0.000
typeperson 3	-0.563	0.05	-11.260	0.000
typeperson 4	-0.007	0.07	-0.100	0.920
typeperson 5	-0.267	0.05	-5.340	0.000
frequency 2	0.040	0.03	1.333	0.183
frequency 3	-0.329	0.03	-10.967	0.000
frequency 4	-0.379	0.02	-18.950	0.000
frequency 5	0.362	0.04	9.050	0.000
frequency 6	0.264	0.04	6.600	0.000
numberaffected 2	-0.370	0.03	-12.333	0.000
numberaffected 3	-0.239	0.04	-5.975	0.000
numberaffected 4	-0.421	0.02	-21.050	0.000
time 2	-0.070	0.03	-2.333	0.020
time 3	0.062	0.05	1.240	0.215
time 4	0.084	0.03	2.800	0.005
time 5	0.189	0.03	6.300	0.000
time 6	0.310	0.03	10.333	0.000
time 7	0.359	0.05	7.180	0.000
time 8	0.158	0.04	3.950	0.000
regulation 2	0.023	0.08	0.288	0.773
regulation 3	0.346	0.21	1.648	0.099
regulation 4	-0.014	0.04	-0.350	0.726
regulation 5	0.103	0.07	1.471	0.141

regulation 6	-0.028	0.05	-0.560	0.575
regulation 7	-0.291	0.04	-7.275	0.000
regulation 8	0.258	0.04	6.450	0.000
violation B	1.719	0.04	42.975	0.000
violation C	1.288	0.04	32.200	0.000
violationother B	0.889	0.04	22.225	0.000
violationother C	0.929	0.04	23.225	0.000

Table 2: Baseline Logit regression

E Estimation of average treatment effects: robustness checks

We address the violation of the common support assumption in several ways. First, we reestimate the average treatment effects on duration and satisfaction Δ_1 and Δ_S using a trimmed sample. In particular, we exclude all observations that have an estimated propensity score either below 0.01 or above 0.4. Below the lower threshold, there are only 125 treated observations (and roughly 28016 nontreated). Above the upper threshold, there are only 503 nontreated above this threshold (and roughly 3232 treated). The resulting estimates are -12.246 and 0.053, both significant at all conventional levels, and thus very similar to the estimates on the full sample.

Second, we reestimate Δ_1 and Δ_S on a sample, where we exclude observations that violate local environmental regulations or have burning fields or aircraft as pollution origin. As shown in treatment assignment analysis, such observations are particularly likely to be assigned to either nontreatment or treatment. The estimated effects are now -12.103 and 0.055 and also significant.

Third, we estimate Δ_1 and Δ_S using a regression based approach that interpolates on the regions where the common support assumption is violated. In particular, we use the equality

$$ATE = \mathbb{E}[\mathbb{E}[Y|D=1,X]] - \mathbb{E}[\mathbb{E}[Y|D=0,X]] \tag{16}$$

and estimate the conditional expectations $\mathbb{E}[Y|D = 1, X]$ and $\mathbb{E}[Y|D = 0, X]$ using a linear (in the case of Δ_1) or Probit (in the case of Δ_S) regression. The estimates are -4.39 for Δ_1 and 0.077 for Δ_S. Thus, both estimates have the same sign as the corresponding estimates obtained with all previous approaches. The former estimate is of somewhat smaller magnitude, while the latter is somewhat of higher magnitude.

References

Abbring, J. and Berg, G. V. D. (2005). Social experiments and instrumental variables with duration outcomes. IFS Working Papers W05/19, Institute for Fiscal Studies.

Ashenfelter, O. (1987). Arbitrator behavior. *The American Economic Review*, 77(2):342–346.

Ashenfelter, O. and Bloom, D. E. (1984). Models of arbitrator behavior: Theory and evidence. *The American Economic Review*, 74(1):111–124.

Ashenfelter, O., Currie, J., Farber, H. S., and Spiegel, M. (1992). An experimental comparison of dispute rates in alternative arbitration systems. *Econometrica*, 60(6):1407–1433.

Backus, M., Blake, T., Pettus, J., and Tadelis, S. (2020). Communication and bargaining breakdown: An empirical analysis. Technical report, National Bureau of Economic Research.

Balzer, B. and Schneider, J. (2021). Managing a conflict: optimal alternative dispute resolution. *The RAND Journal of Economics*, 52(2):415–445.

Beardsley, K. C., Quinn, D. M., Biswas, B., and Wilkenfeld, J. (2006). Mediation style and crisis outcomes. *The Journal of Conflict Resolution*, 50(1):58–86.

Bleakley, H. and Ferrie, J. (2014). Land openings on the georgia frontier and the coase theorem in the short-and long-run. *Unpublished paper, University of Chicago.[1064].*

Bonev, P. (2020). Nonparametric identification in nonseparable duration models with unobserved heterogeneity. Working Papers 2020-05, University of St. Gallen.

Burnett, R., Chen, H., Szyszkowicz, M., Fann, N., Hubbell, B., Pope, C. A., Apte, J. S., Brauer, M., Cohen, A., Weichenthal, S., Coggins, J., Di, Q., Brunekreef, B., Frostad, J., Lim, S. S., Kan, H., Walker, K. D., Thurston, G. D., Hayes, R. B., Lim, C. C., Turner, M. C., Jerrett, M., Krewski, D., Gapstur, S. M., Diver, W. R., Ostro, B., Goldberg, D., Crouse, D. L., Martin, R. V., Peters, P., Pinault, L., Tjepkema, M., van Donkelaar, A., Villeneuve, P. J., Miller, A. B., Yin, P., Zhou, M., Wang, L., Janssen, N. A. H., Marra, M., Atkinson, R. W., Tsang, H., Quoc Thach, T., Cannon, J. B., Allen, R. T., Hart, J. E., Laden, F., Cesaroni, G., Forastiere, F., Weinmayr, G., Jaensch, A., Nagel, G., Concin, H., and Spadaro, J. V. (2018). Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter. *Proceedings of the National Academy of Sciences*, 115(38):9592–9597.

Carver, T. B. and Vondra, A. A. (1994). Alternative dispute resolution: Why it doesn't work and why it does. *Harvard Business Review*, 72:120–120.

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., and Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, 21(1):C1–C68.

Chesher, A. (2002). Semiparametric identification in duration models. CeMMAP working paper CWP20/02, Centre for Microdata Methods and Practice, London, UK.

Coase, R. H. (1960). The problem of social cost. *The journal of Law and Economics*, 56(4):837–877.

Coben, J. and Stienstra, D. (2021). Adr empirical research studies (summer 2013-fall 2020).

Collins, J. M. and Urban, C. (2014). Mandatory Mediation and the Renegotiation of Mortgage Contracts. *The Economic Journal*, 125(589):1734–1763.

Coursey, D. L., Hoffman, E., and Spitzer, M. L. (1987). Fear and loathing in the coase theorem: Experimental tests involving physical discomfort. *The Journal of Legal Studies*, 16(1):217–248.

Delikat, M. and Kleiner, M. M. (2003). An empirical study of dispute resolution mechanisms: Where do plaintiffs better vindicate their rights? *Dispute Resolution Journal*, 58(4):56.

Deryugina, T., Moore, F., and Tol, R. S. (2021). Environmental applications of the coase theorem. *Environmental Science & Policy*, 120:81–88.

Eisenkopf, G. and Bächtiger, A. (2013). Mediation and conflict prevention. *Journal of Conflict Resolution*, 57(4):570–597.

Fanning, J. (2021). Mediation in reputational bargaining. *American Economic Review*, 111(8):2444–72.

Farber, H. S. and White, M. J. (1994). A comparison of formal and informal dispute resolution in medical malpractice. *The Journal of Legal Studies*, 23(2):777–806.

Farnsworth, W. (1999). Do parties to nuisance cases bargain after judgment? a glimpse inside the cathedral. *The University of Chicago Law Review*, 66(2):373–436.

FEY, M. and RAMSAY, K. W. (2010). When is shuttle diplomacy worth the commute? information sharing through mediation. *World Politics*, 62(4):529–560.

Frey, B. S. and Stutzer, A. (2005). Beyond outcomes: measuring procedural utility. *Oxford Economic Papers*, 57(1):90–111.

Goltsman, M., Hörner, J., Pavlov, G., and Squintani, F. (2009). Mediation, arbitration and negotiation. *Journal of Economic Theory*, 144(4):1397–1420.

Hanley, N. and Summer, C. (1995). Bargaining over common property resources: Applying the coase theorem to red deer in the scottish highlands. *Journal of Environmental Management*, 43(1):87–95.

Heckman, J. J. (1991). Identifying the hand of the past: distinguishing state dependence from heterogeneity. *The American Economic Review*, 81:75–79.

Higbee, J. D., Lefler, J. S., Burnett, R. T., Ezzati, M., Marshall, J. D., Kim, S.-Y., Bechle, M., Robinson, A. L., and Pope III, C. A. (2020). Estimating long-term pollution exposure effects through inverse probability weighting methods with cox proportional hazards models. *Environmental Epidemiology (Philadelphia, Pa.)*, 4(2).

Hörner, J., Morelli, M., and Squintani, F. (2015). Mediation and peace. *The Review of Economic Studies*, 82(4):1483–1501.

Imbens, G. W. and Rubin, D. B. (2015). *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction.* Cambridge University Press.

Kitamura, Y. (2018). *Municipal Environmental Administration Law.* Daiichi Hoki. Tokyo, Japan. In Japanese.

Knaus, M. C. (2020). Double machine learning based program evaluation under unconfoundedness. *arXiv preprint arXiv:2003.03191.*

Lysaught, G. (2009). Searle civil justice institute.

Matsumoto, S. (2011). A duration analysis of environmental alternative dispute resolution in japan. *Ecological Economics*, 70(4):659 – 666.

McGillicuddy, N. B., Welton, G. L., and Pruitt, D. G. (1987). Third-party intervention: A field experiment comparing three different models. *Journal of Personality and Social Psychology*, 53(1):104.

Medema, S. G. (2020). The coase theorem at sixty. *Journal of Economic Literature*, 58(4):1045–1128.

Meirowitz, A., Morelli, M., Ramsay, K. W., and Squintani, F. (2019). Dispute resolution institutions and strategic militarization. *Journal of Political Economy*, 127(1):378–418.

Menkel-Meadow, C. (2006). Empirical studies of adr: The baseline problem of what adr is and what it is compared to.

Menkel-Meadow, C. (2015). Mediation, arbitration, and alternative dispute resolution (adr). *International Encyclopedia of the Social and Behavioral Sciences, Elsevier Ltd.*

Olszewski, W. (2011). A welfare analysis of arbitration. *American Economic Journal: Microeconomics*, 3(1):174–213.

Rosenbaum, P. R. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55.

Rosenberg, J. D. and Folberg, H. J. (1993). Alternative dispute resolution: An empirical analysis. *Stan. L. Rev.*, 46:1487.

Salamanca, A. (2021). The value of mediated communication. *Journal of Economic Theory*, 192:105191.

Schotter, A., Zheng, W., and Snyder, B. (2000). Bargaining through agents: An experimental study of delegation and commitment. *Games and Economic Behavior*, 30(2):248–292.

Semenova, V. and Chernozhukov, V. (2021). Debiased machine learning of conditional average treatment effects and other causal functions. *The Econometrics Journal*, 24(2):264–289.

Shapiro, D. L. and Brett, J. M. (1993). Comparing three processes underlying judgments of procedural justice: A field study of mediation and arbitration. *Journal of personality and social psychology*, 65(6):1167.

Shavell, S. (1995). Alternative dispute resolution: an economic analysis. *The Journal of Legal Studies*, 24(1):1–28.

Speight, A. and Thomas, D. (1997). Arbitrator decision-making in the transfer market: an empirical analysis. *Scottish Journal of Political Economy*, 44(2):198–215.

Stipanowich, T. J. and Lamare, J. R. (2014). Living with adr: Evolving perceptions and use of mediation, arbitration, and conflict management in fortune 1000 corporations. *Harv. Negot. L. Rev.*, 19:1.

Van den Berg, G. J. (2001). Duration models: specification, identification, and multiple durations. In Heckman, J. and Leamer, E., editors, *Handbook of Econometrics*, volume 5, chapter 55, pages 3381–3460. Elsevier.

Wilkenfeld, J., Young, K., Asal, V., and Quinn, D. (2003). Mediating international crises: Cross-national and experimental perspectives. *The Journal of Conflict Resolution*, 47(3):279–301.