

Consequences of Inconvenient Truth: Evidence from Sentencing Disparity*

Michal Šoltés[†]

Preliminary and Incomplete Version: Comments Welcome

The most recent version

Tuesday 9th February, 2021

Abstract

TODO

JEL Codes: C90, D02, D83, K40.

Keywords: information disclosure, institutional trust, sentencing disparity, dispute resolutions

*I thank Jakub Drápal, Ole Jann, Filip Matějka, and Nikolas Mittag. This study was supported by Charles University, GAUK project No. 218218 and by a grant from the CERGE-EI Foundation's Teaching Fellows program. This project also received funding from the European Research Council under the European Union's Horizon 2020 research and innovation programme (grant agreement No 678081). All opinions expressed are those of the author(s) and have not been endorsed by the CERGE-EI Foundation. The experiment can be found in the AEA RCT Registry (AEARCTR-0005374).

[†]Email address: msoltes@cerge-ei.cz; CERGE-EI, a joint workplace of Charles University and the Economics Institute of the Czech Academy of Sciences, Politických vězňů 7, 111 21 Prague, Czech Republic.; Charles University, Faculty of Law, nám. Curieových 901/7, 116 40 Prague, Czech Republic.

1 Introduction

Public administrations have been repeatedly found reluctant to disclose an inconvenient truth even though not publishing the information may jeopardize public health and safety and undermine principles of modern democracies. Censorship of information after the Chernobyl catastrophe by Soviet propaganda and repeated effort to cover-up a spread of diseases in China are two prominent examples. In these examples as and many other cases, the aversion to disclose information was supposedly motivated by concerns that it might reveal incompetence and systematic failure of responsible authorities. And consequently lead to the public distrust in the system and its institutions.¹

In many cases, the concerns may be valid. In general, information about performance and competence of public administration to deliver on their responsibilities (e.g., public health ensured by health officers) affects public trust in institutions. Which, in turn, is likely to shape citizens' behavior with important economic and social consequences. For example, Acemoglu et al. (2020) documented that providing the general public with a positive information about performance of state courts (reduced delays) in Pakistan, changed citizens' attitude and increased likelihood of using state courts instead of relying on alternative dispute resolution. Naturally, one may wonder whether disclosing negative – and for the public administration inconvenient – information would lead to public distrust and avoidance of such institutions. Or, as others would argue to promote accountability, increase public interest in addressing the inconvenient situation, and lead to an improvement in performance.

In this paper, I study how citizens response to information about sentencing disparity among judges in the Czech republic. The fact that judges' decisions systematically differ (e.g. propensity to incarcerate) depending on a personality of judge is arguably one of the most worrying signals of competence of the judicial systems. It challenges formal rules of equality before the law and clear, stable, and predictable application of the law. And the information about sentencing disparity classifies as inconvenient truth for the judicial system to disclose. Inconsistency in sentencing decisions is not unique to the Czech judicial system. In fact, sentencing disparity has been documented in many other judicial systems worldwide and been extensively discussed by both scholars and media.²

¹For example, the Associated Press (2020) reports on the COVID-19 outbreak in China as follows: *"In Wuhan, local leaders were accused of telling doctors in December not to publicize the spreading virus in order to avoid casting a shadow over the annual meeting of a local legislative body. As the virus spread, doctors were ordered to delete posts on social media that appealed for donations of medical supplies. That prompted complaints authorities were more worried about image than public safety."*

²For example, Reuters (2017) wrote about disparities in judges' decisions in the USA the following: *"[Reuters] found broad variations among judges and courts in how frequently immigrants are deported. Houston, Charlotte and Atlanta have been labeled "asylum free zones" by immigrant advocates who say those courts disproportionately deny claims for protection. Within specific courts, too, rulings vary widely*

I conduct a survey experiment with 2410 respondents who were provided with varying yet not deceptive information about sentencing disparity among judges at regional courts in the Czech republic. The source of the variation comes from the different levels of sentencing disparities (i.e. shares of cases sentenced to community service) among judges within a court across different courts; while at one court, the propensity to sentence to community service among judges is virtually the same, at a different court the sentencing pattern shows significant differences. The former court serves as information for the active control group, while the latter for the treatment group. The information provided consists of shares of cases in which convicted offenders were sentenced to community services as opposed to different types of punishments (e.g. incarceration) for one of the most frequent crimes - *failure to pay alimony*. In a between subjects design, I measure the effect of the provided information on: (i) declared institutional trust; (ii) courts related behavior such as willingness to apply to courts and demand for alternative dispute resolutions; and (iii) policy preferences regarding the judicial system, including willingness to actively engage in addressing the underlying issue.

The results suggest that provided information about sentencing disparity does not have adverse consequences. Information about sentencing disparity did *not* undermine public trust in the judicial system neither it motivated the public to avoid the formal judicial system and seek for substitutions. Moreover, I found suggestive evidence that the information increases citizens' engagement in addressing sentencing inconsistency. In particular, respondents became more interested in signing a petition that invites politicians to suggest particularizing sentencing principles that would assist judges in their sentencing decisions. Additionally, subjects exposed to the treatment information found fairness of the judicial system as more important policy goal than other subjects. The effect is driven by two groups of respondents: (i)

[TO BE ADDED]

2 Design of Survey Experiment

To conduct the survey experiment, I partnered with Behavio, a private company administering a panel of regular respondents. Respondents were invited by email to take part in an online survey about courts and justice. 2410 respondents were asked completed the survey. Additional to the data collected in the experiment, I have basic demographic characteristics of respondents collected in previous surveys. With the exception of the final task, the survey experiment was run on a platform of Behavio which is familiar to the respondents.

from judge to judge."

The experiment consists of four stages. First, upon starting the survey respondents were asked three questions regarding their prior attitude to, and previous experience with the judicial system. Second, respondents were randomized into the treatment and the active control groups and were presented with the corresponding information. Third, after the treatment, respondents were asked to complete five tasks (questions). Finally, the final task consists of reading and signing a petition that was posted on a different website and thus respondents interested in the petition had to leave the Behavior platform. The order of tasks in the first and the third stages were randomized at the individual level. The full script is available in the online appendix.

2.1 Attitude

The aim of the first stage was to understand the respondents' initial attitude towards the judicial system. To do so, respondents were asked to what extent on a four-level scale they agree with two statements: (i) "Depending on a judge, similar cases can be sentenced differently"; (ii) "Overall, the judicial system in the Czech republic works well". Additionally, I asked how experienced with the judicial system they and/or people close to them are.

2.2 Treatment

I provided respondents with varying yet not deceptive information about sentencing disparity among judges within a regional court; the treatment group was informed about a court with high sentencing disparity among judges, while the active control group about a court with a negligible sentencing disparity among judges. The source of the variation comes from the fact that courts differ in their sentencing disparities among judges; in some regional courts judges vary in their sentencing patterns significantly, while in others judges exhibit indistinguishable sentencing patterns. The information relies on variation within a given court, rather than between courts or among all judges across all courts, as some of the regional disparities in sentencing are justifiable and do not represent the intended variation.³

Data about sentencing decisions are complicated and multidimensional⁴ which makes it complicated to convey an understandable message. I rely on one of the most common

³For example, losing a driving license/a driving disqualification in a city with functional public transportation is arguably less lenient punishment in terms of economic and social consequences than losing a driving license in regions at the foothills of mountains with limited public transportation. These and similar considerations may lead to some desired sentencing disparity across regions.

⁴One has to consider different offenses and offense subsections, different types and extent of punishment, and combination of more types of punishment.

crimes in the Czech republic – *failure to pay alimony*⁵ – and present shares of cases in which a judge sentences convicted person to community services as the main punishment. The offense of *failure to pay alimony* satisfies four important criteria needed for the treatment to be based on a credible information: (i) there are enough observations, so I can provide aggregate statistics based on at least 80 cases per judge (over a span of three years of 2016-2018); (ii) compared to other offenses, in the objective elements of a crime *failure to pay alimony* is a homogenous crime; (iii) while it is not part of the information provided in the treatment, the differences highlighted in the treatment are statistically significant; (iv) since it is a general type of crime, cases can be assigned to judges at random and importantly, according to the courts' schedule, they indeed are.

The treatment and the control slides present the information in a style news would do. The core of the slide is a simple self-explanatory bar graph accompanied by a few additional pieces of information that provides interpretation of the graph. In particular, the treatment slide consists of a bar graph showing shares of cases in which the convicted criminals were sentenced to community service by different judges (22 %, 18 %, 29 %, 7 %, and 8 %), the headline says: "Judges sentence differently." Next, the slide explains that the judge C (29 %) sentenced to community service almost a third of the convicted offenders, whereas for some it is less than 10 % and instead they apply difference types of punishments. Finally, the slide highlights that cases are assigned at random, and that being assigned to the judge C implies up to threefold higher probability of being sentenced to community service. In the control group, the slide shows a bar graph with shares of cases that were sentenced to community service by different judges (17 %, 14 %, 16 %, 17 %). The headline says: "Judges sentenced very similarly." The control slide further explains that regardless the judge assigned, a convicted offender has very similar probability of being sentenced to community service.⁶⁷

2.3 Experimental Outcomes

The outcomes collected are classified in three main categories: (i) declared institutional trust; (ii) reliance on the judicial system; and (iii) policy preferences.

Institutional Trust To measure institutional trust, I adopt standard survey questions of declared institutional trust similar to those used by international institutions such as the WVS and the Eurostat. In particular, respondents were asked to indicate their trust

⁵Formally, the crime is called Section 196 Negligence of Mandatory Support.

⁶The screens of the control and the treatment slides can be found in the online appendix.

⁷Once respondents finished the experiment, the Behavior sent them debriefing letter that informs them that the information presented represents only one regional court and the situation may differ in different courts. The debriefing letter can be found in the online appendix.

level on a scale of: *a great deal*; *quite a lot*; *not very much*, and *none at all* towards four different institutions. One of the institutions was the judicial system. The choice of the others was led by connection to the judicial system. The closest institution to the judicial system is the police, as police officers often cooperate on criminal cases. Next institution is the government which is responsible for functional judicial system⁸; and finally the public broadcasting, that can be viewed responsible for a lack of information about sentencing disparity.

Reliance on Judicial System Next, I propose two measures to answer whether information about sentencing disparity reduce respondents' willingness to apply to the court and (if so,) are the respondents more likely to search for alternatives to the formal judicial system? Since these questions ask about actual (intended) behavior, they provide more convincing measures of real life consequences of the treatment information than the declared level of trust.

To understand whether providing information about sentencing disparity reduces willingness to apply to the court, I cooperate with a NGO (*vasevyzivne.cz*), that assists single-parents to sue out for alimony. In the experiment, I briefly explain a problem of a typical client of the NGO, i.e. a single mother who consider applying to the court to sue for alimony or not. Applying is potentially beneficial, but it also may lead to high cost, both in terms of money and time. Then, I asked the respondents whether they would recommend her to apply to the court or not and informed them that their advice may be used by the NGO as a material in similar cases. Presumably, the belief that their responses will potentially serve as guideline for other people in actual problems increases the cost of ill-concerned answer.

Should information about sentencing disparity discourage respondents from applying to the judicial system, it seems reasonable to expect that they may be interested in substitute to the judicial system. In similar vein, Acemoglu et al. (2020) document substitutability between formal and informal courts motivated by perception of a poor performance of the formal courts. As the next task, I explained that in some cases, it is possible to rely on alternative dispute resolution instead of the judicial system and then I offered a free booklet with basic information about the alternative dispute resolutions. They first provided an indicate answer of their interest and if it was affirmative they were asked to provide their email addresses to have the booklet sent. The two-steps procedure was done for two reasons. First, it evaded legal concerns about using their email addressed for different purposes than inviting them to conduct the survey and second, it imposed small but positive cost on the action.

⁸For example, the government (the Minister of Justice) plays a role in appointing new judges.

Policy Preferences Previous literature documented that new information about how public institutions work causes changes in policy preferences and may evoke public reaction (e.g. petitions and political protests). To measure this effect, I collect two outcomes regarding the policy preferences. First, I asked respondents to imagine that they were the prime minister of the Czech Republic and then gave them a list of four policy issues that were supposed to be ranked according to the perceived priority. The acutest issue was supposed to be ranked as the top priority, the second most acutest as the second priority and so on. The four policy issues were: (i) fairness of the judicial system; (ii) sufficient highway infrastructure; (iii) high-quality teachers in educational systems; and (iv) safety situation in the Czech Republic.

Second, I elicited respondents' willingness to become actively involved in policy debate. Subjects were presented with an extract of petition inviting politicians (members of the Committee on Constitutional and Legal Affairs, Chamber of Deputies, Parliament of the Czech Republic) to suggest particularizing sentencing principles. Respondents were asked to answer whether they were interested in reading the full text of the petition and signing it. If interested, the respondents were referred to a google forms website with the full text of the petition. The text highlights the importance and far-reaching consequences of sentencing decisions in one's life and suggests that it may be beneficial to have a manual that would lead to more consistent sentencing. Importantly, the petition was explicit that the manual would be designed to assist judges in their sentencing decisions, but would not in anyway undermine their independence and discretion. If interested in signing the petition, the respondents could have left an email address to have the signature sheet sent.⁹ I collected individual declaratory answers regarding their interest in reading and signing the petition. Once respondents left the Behavio website and opened the petition I could not observe responses at the individual level. However, since individuals in the treatment and the active control groups were referred to different forms of the identical petition, I observe the number of email addressed left for the control and the treatment group separately.

2.4 Randomization

The groups are balanced on both observed demographic characteristics and their prior attitudes towards the judicial system. Roughly 19 % (21 % in the control group) of the respondents in the treatment group reported to be single, 18 % (17 %) live in cohabitation, 43 % (44 %) are married. Additional 17 % (15 %) reported they are currently divorced and only 4% (3%) are widowed. The reported marital status reflects their status at the

⁹The respondents were informed that the petition will be sent once there are at least 1000 signatures.

time of the survey experiment but not their history; for example, respondents classified as married could have experienced a divorce. Slightly more than 70 % of respondents have at least one child and the average number of children is 1.44 in the control group and 1.49 in the treatment group. In both groups, there are fewer male than female respondents (46.6 % in the treatment group and 48.5% in the control group). For more details, see table 4 in the appendix.

3 Results

3.1 Prior Attitude Towards Judicial System

A majority of respondents (91.4 %) agree that sentencing decisions are sensitive to a personality of a judge and that depending on a judge assigned, similar cases can be sentenced differently. The measure of the general approval of the judicial system is less unequivocal, as 52.5 % agree that the judicial system works well, while 47.6 % are less optimistic and disagree. The immediate implication is that many respondents (46.7 %) tend to approve of the judicial system, despite the perceived sentencing disparity. Figure 6 in the appendix shows the aggregate levels of responses.

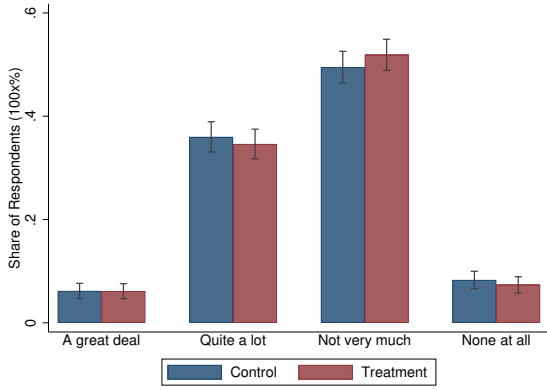
Perception of the judicial system varies with the level of experience with the judicial system. More experienced respondents tend to hold more negative prior beliefs. Consider the 10 % of the most experienced respondents, the majority of them (60 %) strongly agree that depending on judge assigned, similar cases can be sentenced differently and do not consider the judicial system working well. The less experienced respondents consider the judicial system working rather well and do not view the sentencing disparity as that extreme, yet they still admit it may exist. Figure 7 in Appendix shows a frequency of different combinations of prior approval (x-axis) and perception of sentencing heterogeneity (y-axis) by groups that report different experience with the judicial system. Formally, Wilcoxon rank-sum test rejects the equality of attitudes ($p\text{-value} = 0.000$) between the most experience and the less experienced respondents.

3.2 Experimental Evidence

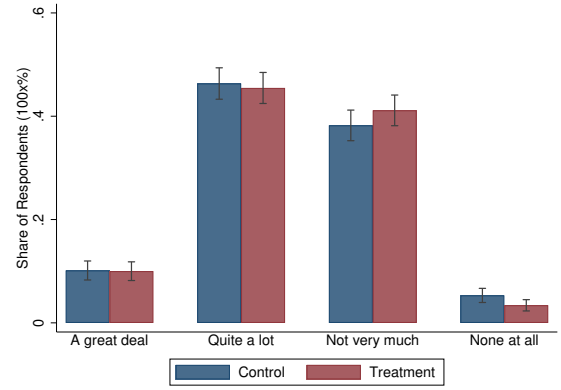
3.2.1 Institutional Trust

The most trusted institution is the police (56 % of respondents report one of the two highest level of trust), followed by the judicial system (40 %), and the public broadcasting (23 %). The national government enjoys the lowest level of trust (22 %). The results are consistent with other surveys. For example, according to Eurobarometer (2018), similar

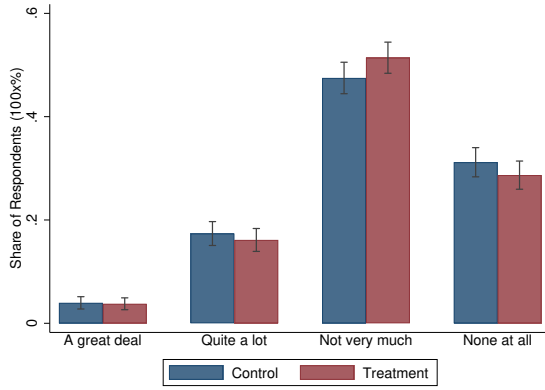
Figure 1: Levels of Institutional Trust



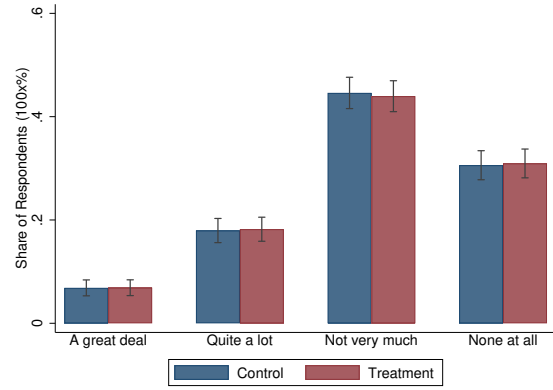
(a) Trust in Judicial System



(b) Trust in Police



(c) Trust in Government



(d) Trust in Public Broadcasting

Notes: Declared level of institutional trust by treatment status: (a) in the judicial system; (b) in the police; (c) in the government; and (d) in the public broadcasting. For each institution, respondents were asked to choose from four levels of trust: (i) a great deal; (ii) quite a lot; (iii) not very much; (iv) none at all.

shares of Czech respondents tend to trust in the police (63 %), the judicial system (43 %), and the national government (28 %).¹⁰ Overall, my results, while less optimistic, tend to resemble those from Eurobarometer (2018).

To test the effect of information about sentencing disparity on institutional trust, I conduct four exercises that all suggest that the treatment does not have an effect. I start by collapsing the first two possible answers: *a great deal* of trust; and *quite of a lot* of trust into one category denoted as *High Trust*. For each institution, I: (i) estimate

¹⁰Eurobarometer (2018) did not ask about the institutional trust in the public broadcasting.

univariate OLS regressions (1)

$$High\ Trust = \alpha + \beta\ Treatment + \varepsilon \quad (1)$$

(ii) estimate OLS regressions with a set of dummies capturing individual prior attitude towards the judicial system and basic demographic characteristics (income, age, level of education, gender, and the number of kids); (iii) simulate the exact p-value for the sharp null hypothesis derived from the potential outcome framework (Athey and Imbens, 2017) and then test

$$High\ Trust_i(0) = High\ Trust_i(1) \quad \forall i = 1, \dots, N. \quad (2)$$

Finally, since the levels of institutional trust represent ordinal scale, I assign a rank (1,2,3,4) to these categories and apply the Wilcoxon rank-sum test.

To provide results that are robust against limited attention of respondents, I run the OLS regressions for two different samples: (i) a *baseline* sample that contains all observations at the level of randomization; and (ii) a sample that drops observations for 10 % of respondents who spent the least time on the treatment and control slides, respectively, That limits the number of observations to 2,168. Additionally, I also dispose 10 % of respondents who spent the least time on the slide with the institutional trust task, reaching 2,008 observations (83 % of the initial dataset). I refer to this sample as the *restricted* sample and argue that no effect in the *restricted* sample, is credible evidence that the lack of effect is not caused by the limited attention of respondents.

Table 1 summarizes the results. Panel A shows results from the OLS regression run in the *baseline* sample and panel B adds results from the *restricted* sample. The first column for each institution provides strong evidence that the average treatment effect is economically and statistically insignificant. For all four institutions the point estimates are, in absolute value, safely less than 2 percentage points. Providing information about sentencing disparities among judges thus did not lead to an institutional mistrust. Similarly, the exact p-value test and the Wilcoxon rank-sum test presented in panel C for the full sample indicate null average treatment effect.

The null average treatment effect masks heterogeneity. Among mothers¹¹, who are arguably more sensitive to inconvenience information regarding *failure to pay alimony*, the treatment tends to increase distrust. Second columns for each institution in panel A and B report results with focus on *Treatment x Mother*, a product of treatment and mother dummy variables, controlling for prior beliefs and attitude towards judicial system and demographic characteristics, such as age, gender, level of education, income

¹¹A female respondent is characterized as a mother if she has at least one child.

Table 1: Treatment Effect on Declared Institutional Trust

Panel A: Baseline Sample								
	Judicial System		Police		Government		Broadcasting	
Treatment	-0.013 (0.020)	0.010 (0.027)	-0.013 (0.020)	0.009 (0.025)	-0.008 (0.017)	-0.001 (0.021)	-0.004 (0.018)	-0.009 (0.022)
Treat. x Mother		-0.060* (0.034)		-0.061 (0.039)		-0.030 (0.033)		0.018 (0.035)
Prior Attitude	No	Yes	No	Yes	No	Yes	No	Yes
Demo. Char.	No	Yes	No	Yes	No	Yes	No	Yes
N	2 410	2 407	2 410	2 407	2 410	2 407	2 410	2 407

Panel B: Restricted Sample								
	Judicial System		Police		Government		Broadcasting	
Treatment	-0.010 (0.022)	0.020 (0.029)	-0.013 (0.022)	0.050 (0.033)	-0.013 (0.028)	0.000 (0.018)	0.003 (0.019)	-0.021 (0.029)
Treat. x Mother		-0.056 (0.037)		-0.045 (0.042)		-0.050 (0.036)		0.026 (0.038)
Prior Attitude	No	Yes	No	Yes	No	Yes	No	Yes
Demo. Char.	No	Yes	No	Yes	No	Yes	No	Yes
N	2 008	2 005	2 008	2 005	2 008	2 005	2 008	2 005

Standard errors in parentheses.
* (p<0.10), ** (p<0.05), *** (p<0.01)

Panel C: Exact p-value and p-value of the Wilcoxon test (Baseline sample)				
	Judicial System	Police	Government	Broadcasting
Exact p-value	0.510	0.518	0.625	0.810
Wilcoxon test				
Full sample	0.652	0.760	0.509	0.702
Only mothers	0.062	0.094	0.915	0.515

Panel A shows results from univariate and multivariate OLS regressions of *Treatment* on a dummy for high level (great deal of trust and quite a lot) of institutional trust.

Panel B shows results from univariate and multivariate OLS regressions of *Treatment* on a dummy for high level (great deal of trust and quite a lot) of institutional trust using a sample without 25 % of respondents who spent the least time on the task.

Panel C presents p-values of two alternative measures; the exact p-value (Athey and Imbens, 2017) derived from 20,000 simulations and the p-value of the Wilcoxon rank-sum test that tests for the same rank of institutional trust.

and whether a respondent is a mother. Despite only a marginal statistical significance, the panel A and the Wilcoxon test in panel C suggest treatment heterogeneity. Taking the point estimates at their face values, share of trusting respondents among mothers, who are arguably more sensitive to information about *failure to pay alimony*, decreases by 6 percentage points more than among non-mothers.

3.2.2 Reliance on Judicial System

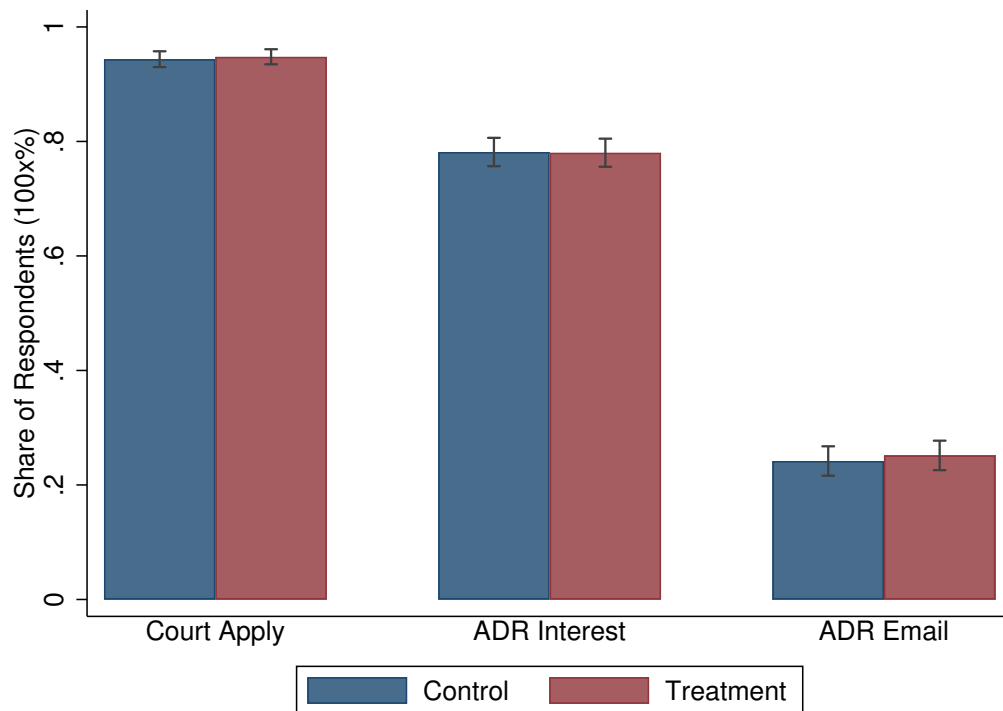
Second group of outcomes consists of a measure of respondents' intention to apply to a court, *Court Apply*, and two measures of their willingness to learn about alternative dispute resolution – a possible substitute to the judicial system – *ADR Interest* and *ADR Mail*. In general, respondents exhibited a strong will to apply to the judicial system, as almost 94 % of respondents recommended to a single mother to apply to the court in a situation in which she hesitates. Additionally, if offered, a majority (76 %) would be interested in receiving a booklet about alternative dispute resolutions. However, when the respondents were asked to provide their email addresses, so the booklet can be sent, only 24 % of all respondents and 31 % of those who declared their interest did so. See figure 2.

Information about sentencing disparities affected neither the (intended) reliance on the judicial systems, nor the demand for alternative dispute resolution. Table 2 reports results from a univariate OLS regression, a regression with focus on treatment effect on mothers, and exact p-value. First columns for each measures show that the average treatment effects are statistically insignificant and economically negligible. In terms of the magnitude of the point estimates, the effects are even smaller than the effect on institutional trust. All point estimates of the average treatment effects are bounded between -1 and 1 percentage points. Panel B shows results estimated on a sample of respondents that spend enough time on the treatment and control slides and on the slide with the corresponding task and confirms the results.¹²

Looking at the heterogeneity of the treatment effect, it seems that not even more sensitive groups of respondents change their behavior. In particular, mothers tend to respond to the inconvenient information about sentencing disparity in the same manner other respondents did – not at all. Panel C shows exact p-value and provides additional evidence of strong null results for all three measures.

¹²Since the overlap of the dropped samples differ across tasks, the number of observations does too.

Figure 2: Reliance on Judicial System



Notes: Shares of affirmative responds by the treatment status. *CourtApply* captures whether respondents recommend to a single mother to apply to the court when she hesitates. *ADRInterest* measures shares of respondents who indicate their interest in receiving information about alternative dispute resolution (ADR). *ADRMail* shows share of respondents who provide their email addresses to have the information about ADR sent.

Table 2: Treatment Effect on Reliance on Judicial System

Panel A: Baseline Sample						
	Court Apply		ADR Interest		ADR Mail	
Treatment	0.000 (0.010)	0.006 (0.013)	-0.009 (0.017)	-0.023 (0.022)	-0.002 (0.017)	-0.013 (0.022)
Treat. x Mother		-0.016 (0.020)		0.043 (0.035)		0.035 (0.035)
Prior Attitude	No	Yes	No	Yes	No	Yes
Demo. Char.	No	Yes	No	Yes	No	Yes
N	2 410	2 407	2 410	2 407	2 410	2 407

Panel B: Restricted Sample						
	Court Apply		ADR Interest		ADR Mail	
Treatment	0.002 (0.010)	0.001 (0.013)	-0.007 (0.018)	-0.022 (0.024)	0.006 (0.019)	0.002 (0.025)
Treat. x Mother		0.002 (0.020)		0.045 (0.037)		0.016 (0.038)
Prior Attitude	No	Yes	No	Yes	No	Yes
Demo. Char.	No	Yes	No	Yes	No	Yes
N	2 020	2 017	2 023	2 020	2 023	2 020

Standard errors in parentheses

* (p<0.10), ** (p<0.05), *** (p<0.01)

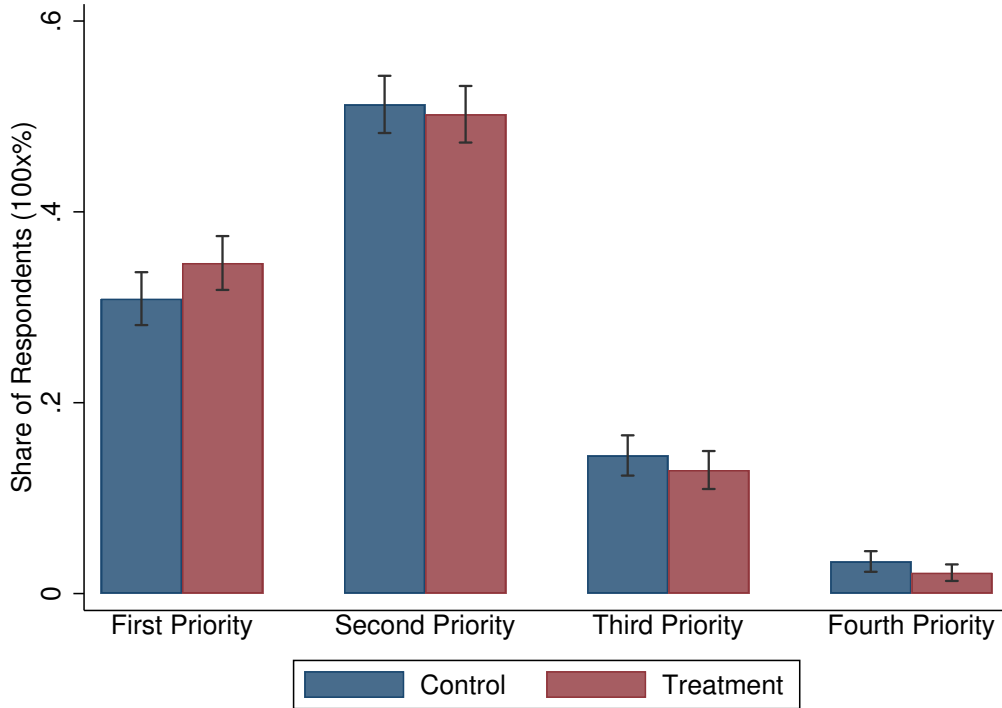
Panel C: Exact p-value (Baseline sample)			
	Court Apply	ADR interest	ADR Mail
Exact p-value	0.993	0.620	0.898

Panel A shows results from univariate and multivariate OLS regressions of *Treatment* on a dummy variable for affirmative response from three tasks.

Panel B shows results from univariate and multivariate OLS regressions of *Treatment* on a dummy variable for affirmative response from three tasks using a sample without 25 % of respondents who spent the least time on the corresponding task.

Panel C shows the exact p-value (Athey and Imbens, 2017) derived from 20,000 simulations.

Figure 3: Rank of Fairness of Judicial System as Priority



Notes: Share of respondents who rank fairness of the judicial system as the first, second, third, and the fourth priority by the control and treatment groups. The remaining issues to be ranked to were sufficient highway infrastructure, safety in the Czech republic, high-quality teachers in the education system.

3.2.3 Policy preferences

Information about sentencing disparity evokes response in both measures of policy preferences.

The respondents view fairness of the judicial system is a relevant policy issue. A third (32.6 %) of them ranked fairness of the judicial system as the top priority and additional 50 % as the second most important priority. While the perception of fairness of the judicial system is likely affected by the environment of the survey experiment itself (e.g. through experimenter demand effect) and thus it is barely generalizable, it is a good signal of a relevance of the issue. Figure 3 shows shares of respondents who ranked fairness of the judicial system as the first, the second, the third priority by both the treatment and the control groups. The figure suggests that the ranking among the treated respondents is slightly shifted towards the higher priority compared to the control group.

The share of respondents who would address the fairness of the judicial system as the top priority in the treatment group is 34 %, while in the control group it is 31.2 %. The difference is not statistically significant. It becomes marginally significant, once one

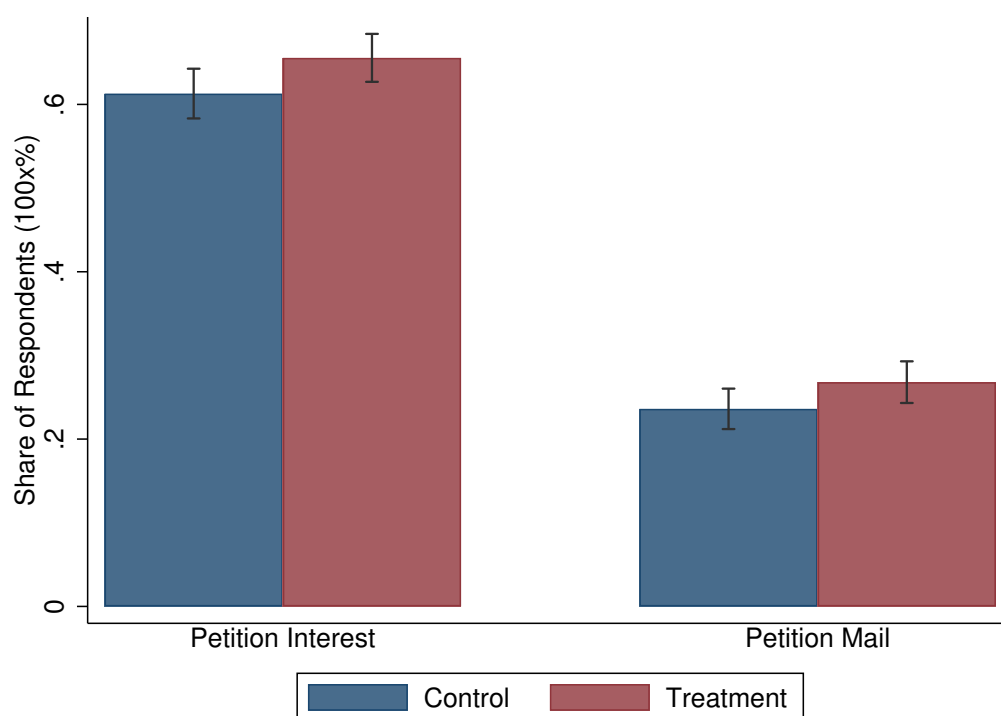
looks at the *restricted* sample of more attentive respondents. See table 3. To formally test the differences in the rankings of the policy issues, I rely on the Wilcoxon rank-sum test. Panel D in table 3 shows that the ranks marginally differ and especially once one focuses on the restricted sample. That confirms the apparent shift of ranking of the judicial system among the treatment group. There seems to be no heterogeneity effect with respect to mother-status.

I collected two measures of respondents' reaction regarding the petition. At the individual level. I measured respondents' interest in reading an signing petition, as a dummy indicator that can be matched with other individual characteristics. At the level of the control and the treatment groups, I collected the number of email addresses provided. After declaring one's interest to read and sign petition, the respondents were referred read the text of the petition elsewhere. Once they left the platform, it is impossible to track their decisions (providing email address) at the individual level and match it to their characteristics. However, since the control and the treatment group was referred to different petitions (with identical text), I can measure a number of email addresses left by respondents in the control and the treatment group.

Figure 4 shows shares of respondents interested in reading and signing petition and those who provided their email addresses by the control and the treatment groups. More than 60 % of respondents showed their interest in reading and signing petition. The share was higher in the treatment group by 3.4 percentage points (5.8 %). Once the respondents were asked to provide me with an email addresses, the share of affirmative action fell to 25 %. The gap between the treatment and the control groups remains almost constant at 3.2 percentage points (13.6 %).

Formal tests suggest that the average treatment effects on respondents' interest in: (i) reading and signing petition; and (ii) providing an email address are marginally significant. The former effect becomes slightly larger once I drop observations of least attentive respondents. See Panel B in table 3. The average treatment effect, however, masks sizeable heterogeneity. The second column of panel A shows that controlling for additional individuals characteristics and including a product of treatment status and mother status leads to null treatment effect. Instead, it seems that the whole effect is driven by mothers who act on that information. In particular, the inconvenient information about sentencing disparity in cases of failure to pay alimony increases the likelihood that a mother will be interested in reading and signing petition by 8 percentage points. The effect is even larger, once I focus on the *restricted* sample.

Figure 4: Interest in Reading and Signing Petition.



Notes: Share of respondents who demonstrate their interest in signing petition. *Petition Interest* measures shares of respondents of the *baseline* sample who indicate their interest in reading the petition. *Petition Mail* measures shares of respondents of all respondents invited in the survey who provided me with their email addresses to have the petition sent.

Table 3: Treatment Effect on Political Preferences

Panel A: Baseline Sample					
	Petition		Petition Mail	Top Priority JS	
Treatment	0.033*	-0.000	0.032*	0.027	0.038
	(0.019)	(0.025)	(0.017)	(0.019)	(0.025)
Treat. x Mother		0.082**			-0.019
		(0.040)			(0.038)
Prior Attitude	No	Yes	No	No	Yes
Demo. Char.	No	Yes	No	No	Yes
N	2 410	2 407	2 410	2 410	2 407

Panel B: Restricted Sample					
	Petition		Petition Mail	Top Priority JS	
Treatment	0.047**	0.006		0.034*	0.042
	(0.021)	(0.027)		(0.020)	(0.027)
Treat. x Mother		0.103**			-0.015
		(0.043)			(0.042)
Prior Attitude	No	Yes		No	Yes
Demo. Char.	No	Yes		No	Yes
N	2 037	2 034		2 018	2 015

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Panel C: Exact p-value (Baseline sample)			
	Petition	Petition Mail	Top Priority JS
Exact p-value	0.086*	0.079*	0.155

Panel D: Rank Judicial System - Wilcoxon test		
	Baseline sample	Restricted sample
Wilcoxon test	0.060*	0.041**

Panel A and B show results from OLS regressions of *Treatment* on dummy variables that measure whether one (i) is interested in reading a petition; (ii) provide email address to have the petition sent; (iii) ranks fairness of judicial system as the top or second priority.

Panel B is estimated on a sample without the least attentive respondents.

Panel C shows the exact p-value (Athey and Imbens, 2017) derived from 20,000 simulations.

4 Discussion

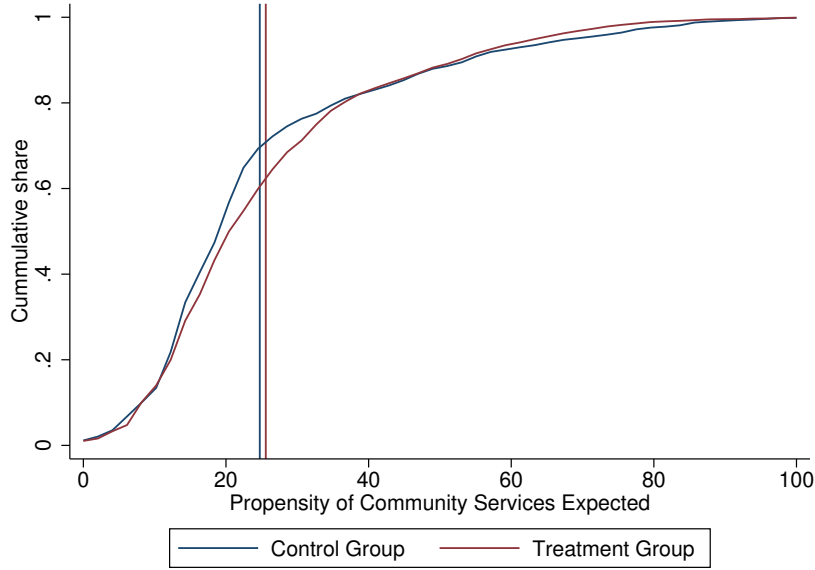
I next discuss the interpretation of the treatment and add thoughts on external validity and generalizability of the results for cases in which the information about sentencing disparity would become widespread news.

4.1 Interpretation of the treatment effect

An important concern in studying trust and trustworthiness based on information treatment is whether respondents trust the information provided. At the beginning of the experiment, all respondents were notified that the information provided to them are truthful and based on data from the Ministry of Justice. This may evoke tension between implicitly asking respondents to trust the information based on data provided by the Ministry of Justice and, at the same time, ask them whether they trust in the judicial system. To understand the degree of potential risk of mistrusting the information, I elicited the perceived credibility of the information. To limit effects of asking the question regarding the credibility of the information provided on respondents' actions and decisions in the experiment, the question came at the very end of the experiment. 87 % of respondents declared that they view the data credible and only less than 1 % of them selected an extreme choice of *definitely not credible*. The shares are virtually identical in both groups. This suggests that the results are not systematically affected by respondents' mistrust in the information treatment. See figure 8 in Appendix.

For the proper interpretation of the results, it is important how respondents understand and interpret the information and what difference in perception is caused by the treatment. The core of the treatment is to provide information about varying levels of sentencing disparity, i.e. the second moment of the distribution of sentencing decisions. This is a novel feature, as most of the information and survey experiments existing in the literature exogenously vary beliefs about the first moment of a relevant distribution, e.g. a probability of audit, a share of high-skilled immigrants. Varying signals about sentencing disparity is challenging, as it may be more complicated for the general public to understand the information and interpret it correctly. Additionally, one varies other properties (e.g. maximum, minimum) of the data used to derive the information treatment. Conveniently, in the two courts used in the control and the treatment groups, respectively, the (unweighted) average shares of convicted offenders sentenced to community services for *failure to pay alimony*, i.e. the first moment of the distribution, were numerically identical (16 %). Nevertheless, it is still plausible that the information provided affects respondents' perception of propensity to sentence to community services systematically differently in the control and the treatment group.

Figure 5: Expected Share of Cases Sentenced to Community Service



Notes: The figure shows empirical cumulative distribution function of respondents' estimates of share of cases that are typical sentenced to community service at the national level for both the control and the treatment group.

To understand the threat, I elicited the respondents' expectation regarding the average propensity to sentenced to community services, i.e. the first moment of the distribution. In particular, the respondents were asked in what percentage of cases in the Czech Republic of *failure to pay alimony* is a convicted offender sentenced to the community service. On average, respondents in both groups overestimate the true shares. While the national average corresponds to the one of the presented cases, i.e. 16 %, the respondents in the control group expected 24.7 % and in the treatment groups is 25.6 %. The difference between the groups is not statistically significant (p-value = 0.255). Figure 5 shows empirical cumulative distribution functions of expected share of cases sentenced to the community service and suggests that most of their estimates, in both groups, are concentrated between 10 % and 30 %. The fact that the cumulative distributions functions resemble each other and the averages are not statistically different suggest that the information provided does not affect the perception of the propensity of sentences to community services systematically differently in the control and the treatment groups.

4.2 Implications

[TO BE ADDED]

5 Concluding Remark

[TO BE ADDED]

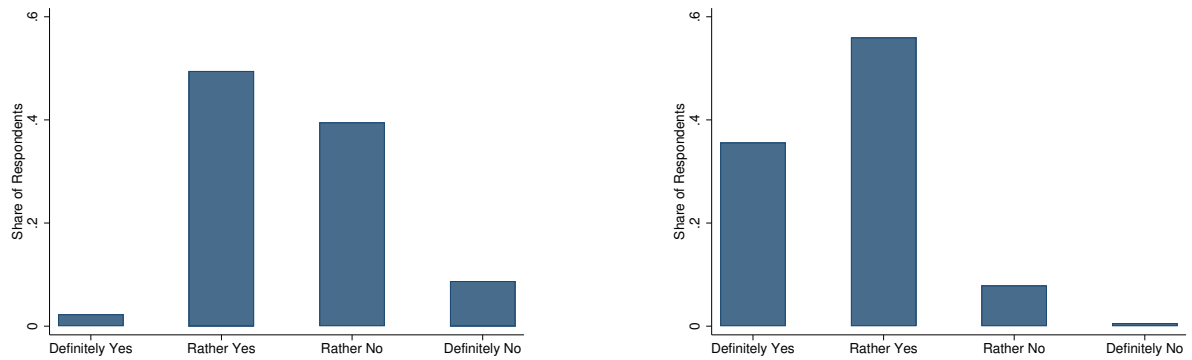
References

- Acemoglu, D., Cheema, A., Khwaja, A. I., and Robinson, J. A. (2020). Trust in state and nonstate actors: Evidence from dispute resolution in pakistan. *Journal of Political Economy*, 128(8):3090–3147.
- Athey, S. and Imbens, G. W. (2017). The econometrics of randomized experiments. In *Handbook of economic field experiments*, volume 1, pages 73–140. Elsevier.
- Eurobarometer, S. (2018). Public opinion in the european union.

Appendix

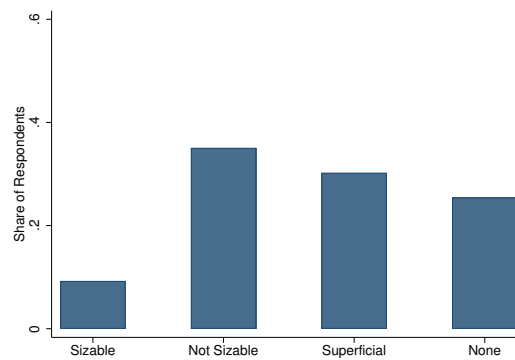
Figures

Figure 6: Prior Attitude Towards Judicial System



(a) Prior Approval of the JS

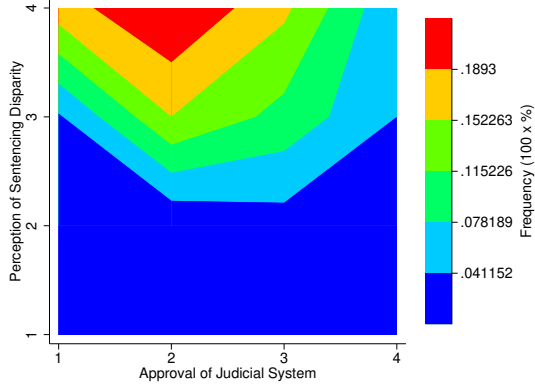
(b) Perception of Sentencing Disparity



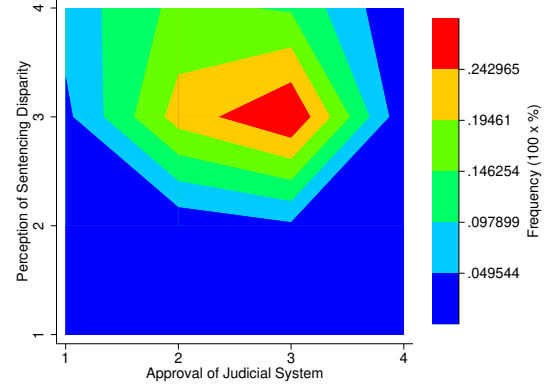
(c) Prior Experience with the JS

(a) The judicial system in the Czech Republic works well. (b) Judges regularly differ in sentencing decisions in similar cases. (c) Considering how often you or people you know well come into contact with judicial system, how experienced you think you are?

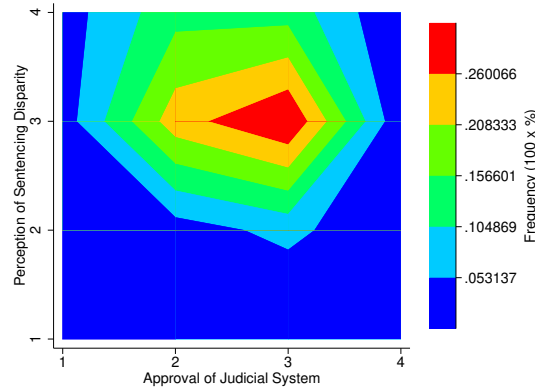
Figure 7: Attitude Towards Judicial System by Experience



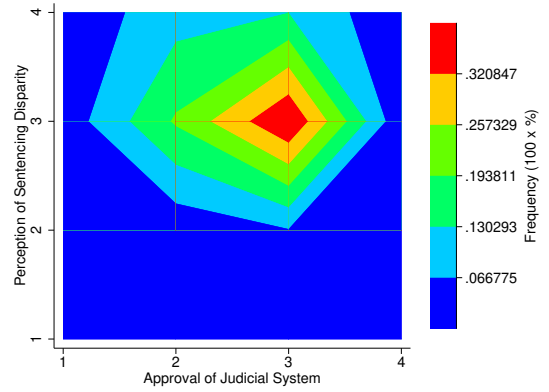
(a) Sizable Experience With the JS



(b) Not Sizeable Experience with the JS



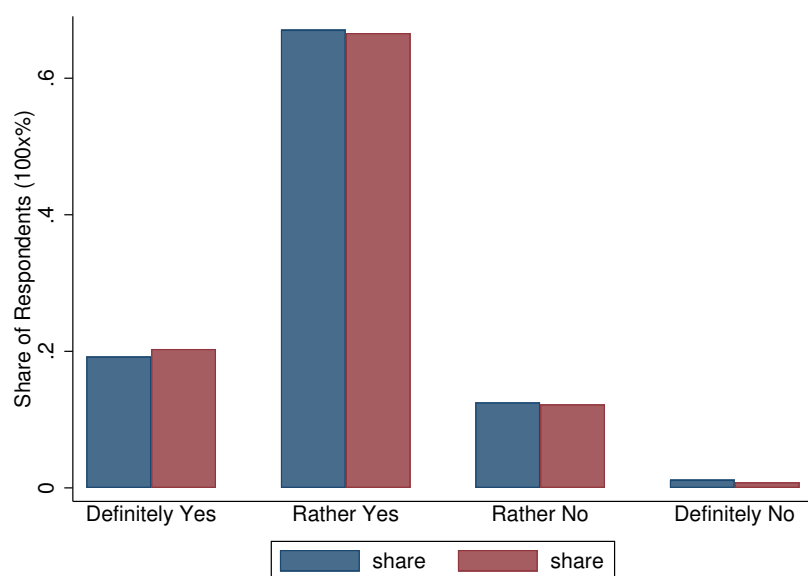
(c) Superficial Experience with the JS



(d) None Experience with the JS

Notes: Level of approval of the judicial system (x-axis) and perception of sentencing disparity (y-axis) by groups of respondents with different levels of experience with the judicial system.

Figure 8: Perception of Credibility of Information Provided



Notes: The figure shows shares of respondents classified by how credible they perceive information provided by the experimenter.

Tables

Table 4: Mean Characteristics of Treatment and Control Groups

Variable	Mean Control	Mean Treatment	t-test (p-value)
Measures of Prior Attitude			
Approval of Judicial System			
Definitely yes	0.026	0.020	0.340
Rather yes	0.488	0.501	0.567
Rather no	0.392	0.398	0.771
Definitely no	0.094	0.081	0.311
Perception of Sentencing Disparity			
Definitely yes	0.372	0.341	0.138
Rather yes	0.547	0.572	0.261
Rather no	0.075	0.082	0.564
Definitely no	0.006	0.006	0.968
Experience with the Judicial System			
Sizable	0.087	0.098	0.375
Not sizeable	0.354	0.347	0.734
Superficial	0.304	0.301	0.877
None	0.255	0.254	0.956
Demographic Characteristics			
Education			
University	0.249	0.255	0.765
Highschool	0.697	0.692	0.825
Elementary	0.054	0.053	0.901
Marital Status			
Single	0.210	0.187	0.175
Cohabitation	0.171	0.182	0.466
Married	0.442	0.423	0.369
Divorced	0.150	0.167	0.277
Widowed	0.027	0.041	0.087
Male	0.485	0.466	0.400
Age	44.71	44.90	0.760
At least 1 child	0.717	0.722	0.818
Number of children	1.435	1.490	0.295
N	1,036	1,060	

Online Appendix

Script

Throughout the experiment, the respondents were shown 14 or 15 slides, depending on their choices. Furthermore, at the end of the session, the respondents could choose to go to a website with petition to sign. The survey experiment consists of three blocks: (i) prior-treatment questions, (ii) treatment slides, and (iii) post-treatment questions.

Prior-treatment Questions

Prior the treatment, respondents are asked three questions to elicit their prior beliefs regarding the sentencing disparity, their experience with the judicial system, and their approval of the judicial system.

Slide 1. *You are about to participate in a survey about courts in the Czech republic. All information provided are truthful and based on data from the Ministry of Justice.*

Slide 2. *We are going to show you three statements and ask you how much you agree with these statements.*

Slide 3. *Sentencing decisions depend on the particular judge assigned to the case. Judges regularly differ in sentencing decisions in similar cases.*

- *Definitely yes*
- *Rather yes*
- *Rather no*
- *Definitely no*

Slide 4. *The judicial system in the Czech Republic works well.*

- *Definitely yes*
- *Rather yes*
- *Rather no*
- *Definitely no*

Slide 5. *Considering how often you or people you know well come into contact with judicial system, how experienced you think you are?*

- *Sizable*
- *Not sizeable*
- *Superficial*
- *None*

Treatment

In this section of the survey, I show different information to the control group and to the treatment group. Both the control and the treatment slides consist of a figure of judges showing shares of cases sentenced to community services and a brief explanation of what the figure represents. Figure 9 (10) shows the control (treatment) slide.

Slide 6. *One of the most frequent crimes in the Czech republic is failure to pay alimony. In last 3 years, courts in the Czech Republic sentenced more than 13,000 cases. The punishment can be suspended sentence, incarceration, community service, and fine. Now, we will show you sentencing decisions of judges at a regional court in the Czech Republic.*

Slide (Control Group). *At this regional court, 16 % of the convicted are sentenced to the community services as their main punishment.*

Judges sentence very similarly.

Regardless which judge is assigned to the case, the offender has very similar probability of being sentenced to community services.

Slide (Treatment Group). *At this regional court, 16 % of the convicted are sentenced to the community services as their main punishment.*

Judges sentence differently.

Judge C sentence to the community service a third of his/her cases.

Other judges sentence to the community service less than 10 % cases, instead they choose different types of punishments.

Cases are assigned to judges at random: an offender assigned to the judge C has three time higher probability of being sentenced to community services compared to a situation in which he was assigned to a different judge.

Post-treatment Questions

Slide 7. *The figure showed the situation from one regional court. Based on your opinion, what is the situation in the Czech republic? In what % of cases are offenders sentenced to the community services as the main punishment for failure to pay alimony?*

Slide 8. *How much do you trust in the following institutions ... { the judicial system; the police; the government; public broadcasting }*

- *A great deal*
- *Quite a lot*
- *Not very much*

- *None at all*

Slide 9. *Suppose you are the prime minister of the Czech Republic. Rank the following issues according to the priorities you would approach them.*

- *Fair judicial system*
- *Sufficient highway infrastructure*
- *High-quality teachers in education system*
- *Safety situation in the Czech Republic*

Slide 10. *In some cases, it is possible to substitute the formal judicial system by alternative dispute resolution (the arbitration) that has several advantages compared to the juridical system.*

- *Want to know more*
- *Not interested*

Slide (only if Slide 10: want to know more). *If you interested, we can send you a booklet with information about the arbitration. What is your email address:*

Slide 11. *We would like your advice. Your response can be used by a NGO (vasevyzivne.cz) as information for its clients. Please read the following story: Jane has two kids and their father does not pay alimony. She hesitates to apply to the court, because she is not sure whether a court could help or it would be only a waste of time and energy.*

What would you recommend to her?

- *She should apply to the court*
- *There is no good reason to apply to the court*

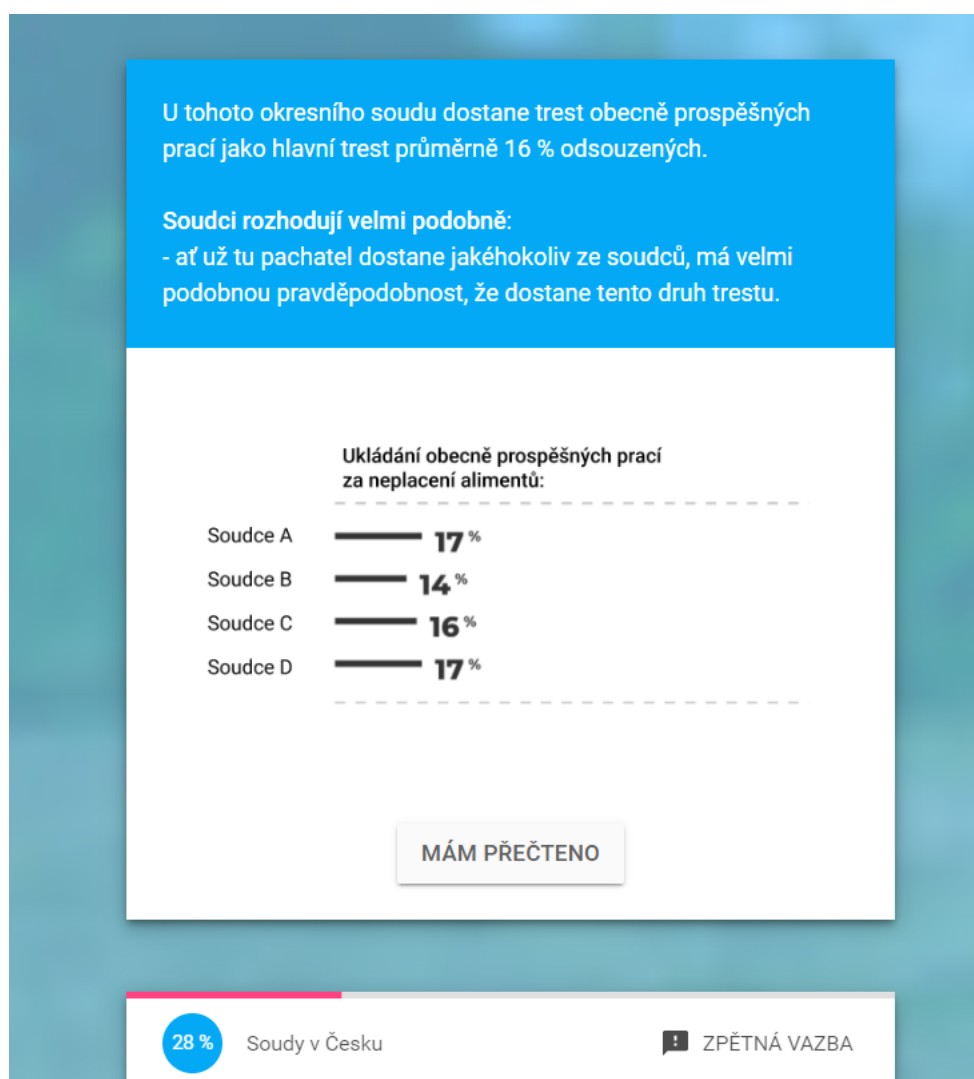
Slide 12. *The last question: Would you sign a petition that invites politicians to suggest particularizing sentencing principles? Such principles would assist judges in sentencing decision. (A preview of the petition was shown.)*

- *Want to read it*
- *Not interested at all*

Slide 13. *Great! That is the end of the questionnaire and very last thing (if slide 12: interested: and before we show you the petition promised), how credible, do you think, the information provided here is?*

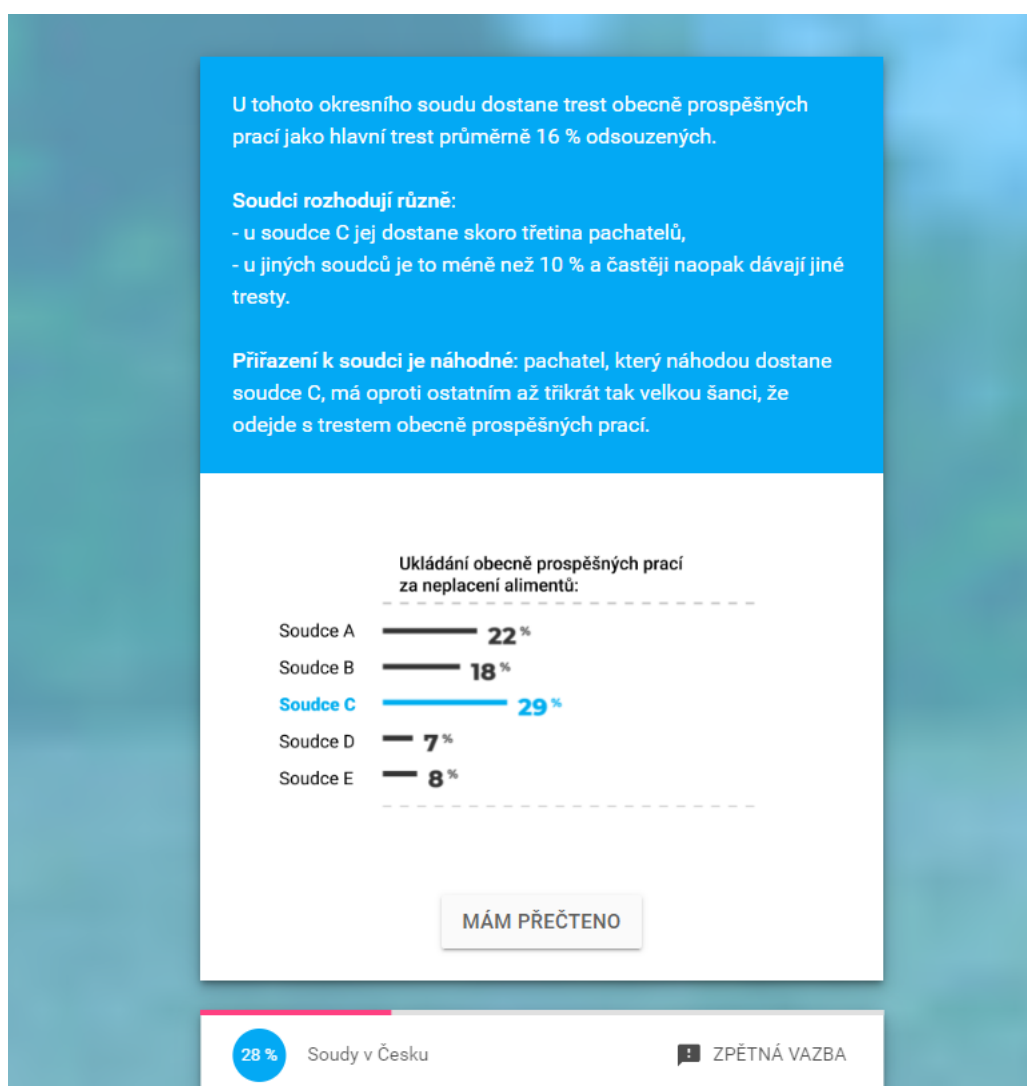
- *A great deal*
- *Quite a lot*
- *Not very much*
- *Not at all*

Figure 9: Control Slide



Notes: The control slide consists of a graph showing a frequency of community services used by four different judges. The graph is accompanied by a brief explanation of the graph.

Figure 10: Treatment Slide



Notes: The treatment slide consists of a graph showing a frequency of community services used by five different judges. The graph is accompanied by a brief explanation of the graph.

Petition

The aim of the following petition is to prompt political representation to particulates sentencing principles. The petition is addressed to members of the Committee on Constitutional and Legal Affairs of the Chamber of Deputies, Parliament of the Czech Republic and will be sent once there are at least 1000 signatures collected.

Should you be interested in signing this petition, leave us with your email address and we will send you the signature sheet.

To members of the Committee on Constitutional and Legal Affairs of the Chamber of Deputies, Parliament of the Czech Republic

Subject: Invitation to particulate sentencing principles

Sentencing decisions often crucially affect one's life. The difference between being sentenced to imprisonment and being given a suspended sentence has significant consequences for the offender, his family and friends.

A judge assigned to a case has an exclusive right to make the sentencing decision. Making such important decisions, the judge consider many circumstances related to the case and consequences of different types and length of punishments. Given the complexity of the decision, it is natural that the personality and experience of judge affect what sentence he will choose. It is right that criminal justice is built on an independence of judges and the principle of individualization of sentences.

It has become a widespread topic in the public discussion that more particularized sentencing principles that help judges in sentencing decisions may limit the role of personality of the judge and thus promote refinement of sentencing. In particular, it may lead to, among others, to establishing a non-binding instructions of how to proceed in sentencing decision that may not only help judges in the sentencing decisions, but also to a better understanding of what type of punishment and why was imposed by offenders and the general public.

We, the signatories of this petitions, would like to invite members of the Committee on Constitutional and Legal Affairs of the Chamber of Deputies, Parliament of the Czech Republic, to support our action.

Debriefing Letter

Last week you participated in a survey, in which we – on behalf of our client – informed you about shares of cases in which judges of one of the regional courts in the Czech Republic sentence offenders to community services for a crime of *failure to pay alimony*.

The information that was showed to you is truthful and describes a situation at one of the regional courts. However, the situation may not correspond to other regional courts. In the Czech republic, there are several dozens of regional courts that may differ in differences among judges in propensity to sentenced offenders to different types of punishments for different crimes.

Should you be interested in sentencing decision in the Czech Republic, there is a webpage *jaktrestame.cz* devoted to it.