# Pre-analysis plan

*Title:*

Is no news good news? Corruption enforcement and corruption perceptions in Ukraine

*Authors:*

Yuri Fedyk, Yuriy Gorodnichenko, Jonathan Lehne, Ilona Sologoub, and Giancarlo Spagnolo

*Date:*

February 19 2024

## Introduction

*Abstract:*

This study evaluates how individuals update their perceptions of corruption in response to new information about corruption enforcement. Theoretically, information about judicial investigations or convictions may simultaneously provide a negative signal about the extent of corruption and a positive signal about the level of enforcement. Our survey experiment is designed to measure these two effects and the net effect on corruption perceptions. It consists of an online survey of 7000 respondents in Ukraine, in which randomly selected sub-samples are exposed to information about corruption cases.

*Motivation*:

Corruption is – like most illegal transactions – hard to observe and quantify systematically. Given that the 'true' level of corruption is unobservable, corruption perceptions are widely used as a substitute in research and policy discussions. Perceptions have been shown to be biased measures of corruption (Olken 2009, Gutmann et al. 2020) but they matter in their own right. Domestic corruption perceptions may sway the outcome of elections while international corruption perceptions may, in the case of Ukraine, affect willingness to supply aid, foreign direct investment, reconstruction funding and the prospects of EU accession.

Over the past decade, Ukraine has implemented a range of reforms intended to combat corruption, including the creation of new, independent agencies tasked with anti-corruption enforcement. One consequence of their activities has been a steady stream of news about the discovery and investigation of corruption cases. The publication of news about prosecutions and convictions is an important component of law enforcement because it is a prerequisite for deterrence. News about corruption cases may, however, also affect corruption perceptions. Individuals who learn about a new corruption case may focus on the details of the crime and their perception of corruption in the country may worsen. Others may focus on the fact that it has been detected and is being prosecuted, causing their perceptions to improve. Our study aims to evaluate these responses empirically and in particular, to measure the net effect of news about enforcement on corruption perceptions.

Information about corruption cases can be delivered in different forms. Individuals may be most frequently exposed to news stories about a single corruption case. Such news reports may have a different effect from aggregate statistical information about the total number of prosecutions or convictions. For example, news stories about individual cases may encourage a focus on the crime while statistical information may be viewed through the lens of enforcement. Our study includes treatment arms that provide both kinds of information in order to evaluate whether the nature of the news affects the response.

*Research Questions*:

How does news about judicial corruption cases affect corruption perceptions?

Does information about individual cases have a different effect from aggregate statistical information?

## Survey

*Survey company*:

We have contracted the research agency Info Sapiens to conduct the survey. This company is based in Kyiv and has extensive experience conducting surveys in Ukraine.

*Survey format*:

The survey will be conducted online. It will be conducted in Ukrainian. It consists of 23 questions and we estimate that it takes around 20 minutes to complete.

*Structure of survey*:

1. Pre-amble and 6 questions about demographic characteristics
2. 3 questions that elicit priors about perceptions of corruption and enforcement
3. 3 questions on media consumption
4. Treatments
5. 3 questions that measure posteriors
6. 8 additional questions about corruption in Ukraine

*Pilot*:

A pilot survey with 49 respondents was carried out in January 2024.

*Sampling*:

The sample consists of an estimated 7000 respondents. These are drawn randomly from a large panel of over 50,000 potential respondents maintained by Info Sapiens. The composition of this larger panel is designed to be demographically representative of the Ukrainian population as of early 2022. However, we do not expect our sample to be representative of the entire Ukrainian population for several reasons: (i) there is selection into participation in an online survey (ii) a significant fraction of Ukraine's population is currently living abroad as refugees (iii) a significant

fraction of Ukraine's population is currently in combat or living in areas directly affected by combat.

*Randomisation*:

The sample will be split into 5 groups with an estimated 1400 respondents in each group. There is one pure control group and 4 treatment groups. The randomisation into treatment status will be carried out by Info Sapiens. Each of the respondents will be randomly assigned to one of five groups with the help of survey software.

*Follow-up survey*:

We will have one follow-up survey, where respondents are asked the same 3 questions about corruption and enforcement posteriors as they were in the original survey (section 5 in structure of survey above). We will also ask a question about volunteering and donating to Ukrainian institutions (the retrospective version of Q B16) in the original survey. The timing of this follow-up will be staggered, such that a randomly selected 1/3 of the respondents are interviewed after approximately 1 month, 1/3 are interviewed after 2 months, and 1/3 are interviewed after 3 months.

One of the main reasons for the follow-up survey, is to measure the persistence of the treatment effects. If there are no significant effects in our main specifications after the main survey, the follow-up survey will not be conducted.

## Empirical Strategy

*Primary outcome variables:*

1. Perception of corruption in Ukraine

Measured in question B7 of the survey. Respondents are asked how important the problem of corruption is for Ukraine on a scale from 1 (not at all important) to 10 (the most important problem).

2. Perception of corruption enforcement in Ukraine

Measured in question B9 of the survey. Respondents are asked how willing the government is to fight corruption in Ukraine on a scale from 1 (not at all willing) to 10 (determined to fight corruption).

In each case, respondents will have been asked similar questions with different phrasing, and an equivalent 10 point scale prior to treatment.

In the analysis we will report results for (i) variables that use the original 10-point scale and (ii) dummy variables that take the value of 0 for values below 6 and 1 otherwise.

*Secondary outcome variables:*

1. Corruption perceptions in individual sectors (Q B8)
2. Trend in corruption over past 3 years (Q B10)
3. Prior knowledge about anti-corruption agencies (Q B11)

4. Assessment of anti-corruption agencies' competence (Q B12)
5. Tolerance with regard to bribe-giving (Q B14)
6. Willingness to volunteer for or donate to Ukrainian institutions (Q B16)
7. Migration preferences (Q B17)

*Balance checks:*

We will conduct balance checks to test the randomisation using the responses to questions asked prior to treatment. We will check for balance based on the following 19 variables:

| Category | Number of variables and type |
| --- | --- |
| Female | Dummy variable |
| Age | Discrete variable |
| Married | Dummy variable |
| Number of children under 18 | Discrete variable |
| Main occupation | 8 dummy variables for each of the possible responses |
| Sector of occupation | 5 dummy variables for each of the possible responses |
| Prior about corruption perceptions | Discrete variable |
| Prior about enforcement perceptions | Discrete variable |

The balance tests will consist of t-tests that assess the difference in means between the control group and each of the treatment arms for each of the listed variables.

*Treatment effects:*

The specifications for the main results of the paper are as follows:

(1) $Corruption_i^{posterior} = a_0 + a_1 \times Corruption_i^{prior} + a_2 \times Enforce_i^{prior} + a_3 \times Corruption_i^{prior} \times I\{i \in Treat\ \#1\} + a_4 \times Enforce_i^{prior} \times I\{i \in Treat\ \#1\} + a_5 \times Corruption_i^{prior} \times I\{i \in Treat\ \#2\} + a_6 \times Enforce_i^{prior} \times I\{i \in Treat\ \#2\} + a_7 \times Corruption_i^{prior} \times I\{i \in Treat\ \#3\} + a_7 \times Enforce_i^{prior} \times I\{i \in Treat\ \#3\} + controls + error$

where $I\{i \in Treat\ \#X\}$ is an indicator variable equal to 1 if respondent $i$ is in treatment group X.

(2) $Enforce_i^{posterior} = b_0 + b_1 \times Corruption_i^{prior} + b_2 \times Enforce_i^{prior} + b_3 \times Corruption_i^{prior} \times I\{i \in Treat\ \#1\} + b_4 \times Enforce_i^{prior} \times I\{i \in Treat\ \#1\} + b_5 \times Corruption_i^{prior} \times I\{i \in Treat\ \#2\} + b_6 \times Enforce_i^{prior} \times I\{i \in Treat\ \#2\} + b_7 \times Corruption_i^{prior} \times I\{i \in Treat\ \#3\} + b_7 \times Enforce_i^{prior} \times I\{i \in Treat\ \#3\} + controls + error$

(3) $Corruption_i^{posterior} = a_0 + a_1 \times Corruption_i^{prior} + a_2 \times Corruption_i^{prior} \times I\{i \in Treat\ \#1\} + a_3 \times Corruption_i^{prior} \times I\{i \in Treat\ \#4\} + controls + error$

(4)  $Enforce_i^{posterior} = b_0 + b_1 \times Enforce_i^{prior} + b_2 \times Enforce_i^{prior} \times I\{i \in Treat\ \#1\} + b_3 \times Enforce_i^{prior} \times I\{i \in Treat\ \#4\} + controls + error$

Specifications (1) and (2) are designed to estimate the relationship between corruption perceptions, enforcement perceptions and observable signals (see theoretical framework below). Coefficients $(a_1, \dots, b_1, \dots)$ can be mapped into $a$ and $b$ in matrix $H$ in the theoretical framework.

Specification (3) and (4) are designed to compare the effect of (i) aggregate statistical information about corruption conviction to (ii) information about a specific case.

*Controls:*

We will report results excluding controls and including demographic controls (age, gender, number of children, and any variables that are unbalanced in the balance checks described above).

*Standard errors:*

We will report heteroskedasticity-robust standard errors. Randomisation into treatment happens at the individual level.

## Theoretical Framework

Consider the following setup modelling perceptions of corruption and law enforcement

$c$ = true level of corruption (unobservable)

$e$ = true level of enforcement (unobservable)

$m$ = number of corruption cases (from media or law enforcement)

$s$ = number of sentences/convictions on corruption cases

Suppose that the relationship between corruption and observable signals is given by

$m = a_c \times c + a_e \times e + \epsilon_m$

$s = b_c \times c + b_e \times e + \epsilon_s$

where $\epsilon_m \sim (0, \Sigma_m)$ and $\epsilon_s \sim (0, \Sigma_s)$ are independent noise and $a_e, a_c, b_e, b_c$ are some constants. Without loss of generality, we can assume $a_c > 0$ and $b_e > 0$. We can organize this in a matrix form

$$[m\ s\ ] = [a_c\ a_e\ b_c\ b_e\ ][c\ e\ ] + [\epsilon_m\ \epsilon_s\ ] = H[c\ e\ ] + [\epsilon_m\ \epsilon_s\ ]$$

We can define

$$y \equiv [m\ s\ ]$$

$$z \equiv [c\ e\ ]$$

5

$$\epsilon \equiv [\epsilon_m \ \epsilon_s]$$

$$H \equiv [a_c \ a_e \ b_c \ b_e]$$

to arrive at

$$y = Hz + \epsilon$$

$$\epsilon \sim ([0 \ 0], [\Sigma_m \ 0 \ 0 \ \Sigma_s]) = (0, \Sigma_{noise})$$

People have to infer corruption and enforcement from observables and they use Bayesian learning (Klaman filter) to do so.

Let prior be $z_0 \sim (\mu_z, \Sigma_0)$. That is, $z_0$ measures their prior beliefs about the degree of corruption (perception of corruption) and about the degree of law enforcement (perception of enforcement).

The posterior $z_1$ is given by

$$z_1 = z_0 + P \times (y - Hz_0)$$

where $P = \Sigma_0 \times H' \times (H \times \Sigma_0 \times H' + \Sigma_{noise})^{-1}$ is the gain of the Kalman filter. Note that some people will revise their beliefs up ($y > Hz_0$; i.e., signals are surprisingly high relative to their initial beliefs) and some people will revise their beliefs down ($y < Hz_0$; i.e., signals are surprisingly low relative to their initial beliefs).

Our information experiment effectively measures $P$ (the passthrough from signals to beliefs). The empirical implementation boils down to running the regressions described by (1) and (2) in the Treatment Effects section above.