1 Methodology

1.1 Data

To study corruption at the individual level, the AmericasBarometer (AB) survey from the Latin American Public Opinion Project (LAPOP) is used. This survey was administered in Ecuador in a face-to-face interview format from 2004 to 2019 at mostly two year intervals. It asks about several matters, including democracy, corruption, political processes, economic considerations, among others. Some of the data used in this paper comes from the copyrighted ®AmericasBarometer survey, financed by Universidad San Francisco de Quito. Most of the data comes from the open-access AmericasBarometer databases available in the LAPOP website. Table 1 presents descriptive statistics computed for all variables used in the study.

The empirical models estimated in this study will use the survey data from the 2014 and 2016 rounds in Ecuador, with $n_{2014} = 1489$ and $n_{2016} = 1545$. The survey is based on a multi-stage national probability design, with a stratification by region (Costa, Sierra, Amazonía). Each of these major strata were substratified by size of municipality and urban/rural areas (LAPOP, 2017b). The errors for each of these surveys, incorporating design effects, are $\pm 2.5\%$ and $\pm 1.9\%$, respectively (LAPOP, 2017a; LAPOP, 2017b). Both of the surveys are self-weighted, however, 95% confidence intervals for the descriptive statistics which are adjusted for design-effects are presented when relevant.

1.2 Empirical Models

The empirical analysis is concerned with the answers to the *EXC18* question in the AB interviews: "Do you think given the way things are, sometimes paying a bribe is justified?" (Moscoso, 2018, p.96), asked originally in Spanish. The question has been asked in all survey rounds in Ecuador and is the last one after a set of questions regarding corruption exposure and perception. The corruption tolerance variable (*ctol*) takes the value of 1 when the respondent answers "Yes", 0 when they answer "No" and for any other responses the observation is dropped from the models. All models

Table 1: Descriptive statistics for all variables used in the empirical models

Variable	Question code	2014		2016	
		Est.	SE	Est.	SE
Corruption tolerance	EXC18	13.59	1.39	27.18	1.21
Unemployment	OCUP4A	10.06	1.04	22.89	1.2
Confidence in the President	B21A	69.01	1.77	49.64	1.49
Approval of the President	M1	70.26	1.57	55.41	1.43
Economic situation (Worse)	IDIO2	22.93	1.26	51.76	1.45
No political wing	L1	21.49	2.11	8.67	0.74
Center	L1	42.58	1.92	45.7	1.49
Left	L1	22.23	1.25	22.46	1.24
Right	L1	13.7	1.16	23.17	1.15
Women	Q1	50.37	0.34	50.29	0.3
Age	Q2	39.41	0.17	38.64	0.22
Years of education	ED	10.67	0.15	11.43	0.14
Urban	UR	65.21	4.11	66.41	4.07
External political efficacy	EFF1	35.31	1.69	41.93	1.33
Internal political efficacy	EFF2	38.55	1.58	41.49	1.34
Participated in a protest	PROT3	6.82	0.89	4.67	0.55
Interest in politics	POL1	33.45	1.63	32.29	1.35
Perceives corruption	EXC7, EXC7NEW	70.29	1.74	83.49	0.97
Exposed to corruption	EXC 2,6,11,13,14,15,16	26.97	2.01	27.69	1.23

Note: Descriptive statistics table with estimates (Est.) and robust standard errors (SE), where age, years of education and the external and internal political efficacies are arithmetic means. All other variables are percentages, calculated for 2014 and 2016 as seen in Appendix A. Standard errors adjusted for design effects. Question codes come from the AB survey questionnaires. Data from the open-access AB databases.

have *ctol* as their dependent variable. Responses to other questions of the AB in these periods are used as regressors, and their encodings are explained in detail in Appendix A.

In order to identify the changes in public behavior which led to the increase in corruption tolerance, observations from both surveys are pooled and the following general model is estimated:

$$P(ctol = 1|X) = G(X\theta) = G\left[\beta_0 + \delta_0 y_{16} + R\beta + \delta_1(y_{16} \cdot x^*)\right] \tag{1}$$

where R is a vector of important explanatory variables for *ctol* and x^* is a key regressor whose change across time may have significantly influenced the rise of *ctol* between 2014 and 2016. This key regressor is interacted with a year dummy y_{16} which equals unity for 2016 observations. The complete regressors' vector X includes all variables in R and the interaction term. The vector θ includes the coefficients vector β as well as the intercepts β_0 and δ_0 and the δ_1 coefficient. G is the link function, which can be unity for a linear probability model, or be equal to the logit and probit functions.

The partial effect of the key regressor x^* on $P(ctol = 1|X)$ will be:

$$\frac{\partial P(ctol = 1|X)}{\partial x^*} = \frac{\partial G}{\partial \theta} \cdot \frac{\partial \theta}{\partial x^*} = G'(\theta) \cdot (\beta_{x^*} + \delta_1 y_{16}) \tag{2}$$

Therefore, the coefficient of interest in this study is $\hat{\delta}_1$, which would measure the ceteris paribus effect of a change in the key regressor x^* from 2014 to 2016 in the dependent variable *ctol*. If there has been a change in 2016 in the x^* which influences corruption tolerance, $\hat{\delta}_1$ should be statistically significant.

A $\hat{\delta}_1$ coefficient which is not statistically different from zero would mean that individuals with and without this key characteristic are equally likely to justify corruption across time. Additionally, if $\hat{\beta}_{x^*}$ and $\hat{\delta}_1$ have different signs but similar magnitudes, the "net" effect might approach zero.

To better understand these potential cases, the following single cross-section models are also estimated:

$$P(ctol = 1|R_y) = G(R_y \beta_y) = G[\beta_{y,0} + \beta_{y,x^*} + \beta_y X] \tag{3}$$

3

for either $y = 2014$ or $y = 2016$. The vector $\boldsymbol{R}_y$ incorporates important explanatory variables for period y, including the key variable x^* for the period in question. The magnitudes of the $\hat{\beta}_{y,x^*}$ should be similar between the two periods when $\hat{\delta}_1$ is not statistically different from zero in the pooled cross-sections model of Equation 1. Also, $\hat{\beta}_{y,x^*}$ for 2016 should not be statistically different from zero if $\hat{\beta}_{y,x^*}$ cancel each other out in the pooled cross-sections model.

Average partial effects tables are shown for all models estimated in this paper. The Results section includes only the logit estimations of each empirical model. LPM and probit estimations of the models are included in the appendices, along with the average partial effects for the probit models.

The individual level approach that all of these models use might be more empirically accurate than a cross-national approach which pools national averages across countries. As mentioned before, this approach is less likely to omit important variables (Bergh et al., 2017). Also, this approach might reflect general perceptions and incidence of corruption more accurately. This is because country-level indicators are based on opinions from experts whereas the AB proportions capture the opinion from all citizens (Morris, 2008). However, results found using this approach are likely less applicable to countries other than the one studied. Besides, since there is no tracking individuals across time using the AB survey panel-data methods cannot be implemented.

1.3 Incorporating design effects

All models use survey-weighting to adjust for the complex sample design effects, as suggested by Castorena (2021) for the use of AB survey data on research projects. In the Ecuadorian case, surveys from 2014 and 2016 are self-weighted, so the survey-weighting does not affect coefficient magnitudes or average partial effects. However, the design effects do change standard errors for all coefficients. Survey-weighted standard errors are presented in this paper for both model coefficients and average partial effects.