Identification of Electoral Model and Ballot Stuffing*

Anastasia Burkovskaya†

May 2018

This paper introduces a model of electoral choice that allows for derivation of joint distribution of turnout and voter share from unobservable joint distribution of costs of voting and preferences over candidates. Under a set of mild assumptions, we show non-parametric identification of joint distribution of costs of voting and preferences over candidates from observable data on single election/referendum. We also offer an extension of the model that allows for identification of ballot stuffing type of electoral fraud. In addition, we offer an empirical illustration of the model estimation using the 2011 Russia parliamentary election data.

Keywords: electoral preferences, ballot stuffing, non-parametric identification

1 Introduction

Elections and referenda are useful to make political decisions and to observe the preferences of the population. There are many instances in which we would like to know the preferences of the electorate. For example, a policy maker might be concerned about public opinion regarding gay marriage, or a government trying to build a future political agenda might seek insight on where society is heading. In the absence of

*I am deeply indebted to Rosa Matzkin for her guidance and support through the entire work on this paper. I appreciate discussions and comments from Kadir Atalay, Peter Exterkate and Stefano Fiorin. I would also like to thank audiences at USyd, ITAM, PUC-Chile, Universidad Torcuato di Tella, Universidad de los Andes and UCLA for helpful feedback. I am grateful to Vasily Korovkin for sharing the data on independent observers.

†School of Economics, University of Sydney: anastasia.burkovskaya@sydney.edu.au

strategic voting, electoral data delivers revealed preferences and, thus, is the most suitable data available to study the electorate's preferences. However, in countries where voting is not compulsory, many people forgo voting, and then we can only observe the preferences of the citizens that do participate in an election. Since selection bias is likely, it is impossible to observe the preferences of non-voters. However, campaign managers or government officials might still want to know what absentee voters are thinking. Two recent examples are the Colombia peace referendum and the 2016 US elections. In the Colombian referendum, a long sought-after peace deal was rejected at the polls, an unexpected outcome that left politicians without a clear electoral mandate. The Colombian government still had to conduct negotiations with the FARC, and had to try to do so without losing future popular support. In the 2016 US election, the unpredictable candidacy and victory of Donald Trump was a result of overlooked preferences of a large share of the population by both Republican and Democrat establishments. Presently, both parties are preparing for future elections. Candidates will need as much information as possible about the current electorate, including non-voters, because exceptional turnout of a specific type of population might lead to a victory. An example is the unusual spike in African-American turnout in the 2008 U.S. presidential elections due to the Obama's popularity.[1]

To ascertain preferences and predict changes in the voting behavior, we propose a structural model that allows for non-parametric identification of the preferences of the entire population from a single election data on turnout and voter share. The main advantages of our structural model are the ability to study a single election or referendum, as well as the lack of assumptions about preference formation and stability of electoral preferences over time. In addition, sometimes identification of electoral preferences is further complicated by the presence of electoral fraud, especially ballot stuffing – illegal addition of extra ballots. The 2011 Russia parliamentary elections (Enikolopov et al. (2013)) are an example of this type of fraud. We extend

[1] *2008 Surge in Black Voters Nearly Erased Racial Gap*, NY Times, July 20, 2009.

the proposed model to identify ballot stuffing in two different cases. The first assumes the known voting costs at a polling station, allowing for an identification of the exact amount of fraud. The second case assumes the existence of a "clean" randomized subsample (without electoral fraud) in one region, and allows for an identification of distribution of amount of fraud in a polling station. To our knowledge, this is the first paper that proposes a structural model that does not rely on assumptions about preference formation for evaluation of electoral preferences and fraud.

Political preferences have been studied by economists before. However, many papers either study turnout and voter share separately (Degan (2007), Kernell (2009), Merlo and de Paula (2017), Coate and Conlin (2004), McMurray (2013)), or build the model structure based on a spatial model[2] that assumes that each voter has a preferred policy and evaluates candidates based on the distances between proposed and preferred policies (Degan and Merlo (2011)). The main disadvantages of using the spatial model is that it requires construction of the ideology metric and data from several elections. This paper offers a different perspective on preference evaluation. We do not discuss the way of the formation of preferences and the variables the preferences depend on. The preferences in our model are exogenous and we avoid relying on the ideology index. We are interested in understanding various components (personal, local and regional) of voter preferences and voting costs in a single election. This is especially important in the cases of atypical elections or referenda.

In this paper, exogeneity of preferences accounts for different processes of preference formation. For example, if voters are ideological, as in the spatial model, then preference distribution is the distribution of relative distance from the candidates. If one adds additional assumptions about the relative ideology metric and uses the data from multiple elections, then it will be possible to identify the ideological distribution of the population as in Degan and Merlo (2011). Our model also includes the uncertain-voter model – a more confident voter derives higher utility from making

[2]The model was introduced in Downs (1957) and developed in Riker and Ordeshook (1968) and Hinich and Munger (1994).

the right choice, i.e., voting for a specific candidate. Personal preferences account for an asymmetry of information in this case. Thus, the identified preferences can be treated as a "reduced form" obtained from different models.

Another important contribution of the paper is evaluation of ballot stuffing. Existing literature explores three directions:

First, the analysis of different statistical irregularities in the data: Benford's law (Mebane (2008), Breunig and Goerres (2011)), unusual kurtosis of the distribution of electoral data (Klimek et al. (2012)), spikes in the distribution of votes (Kobak et al. (2016), Rozenas (2017)), and other inconsistencies. These methods are purely statistical and often test the presence of fraud, but do not evaluate its amount.

The second direction is the evaluation of fraud through both natural experiments (Cantu (2014)) and randomized assignment of independent observers (Enikolopov et al. (2013)). These papers explore how much fraud occurs in polling stations with observers, versus polling stations in the control group. Without the structural model researchers cannot infer the amount of fraud at a polling station level. Our structural model allows researchers to use data from randomized control trials for ballot stuffing estimation.

The third direction of this research is to fit a parametric model. This requires assumptions about parametric distribution, and what data is relevant to the estimation of parameters. For example, Levin et al. (2009) uses the data from the preceding electoral period coupled with an assumption on stability of electoral preferences over time. However, previous electoral data is not always available or reliable, and the electoral preferences are not always stable. Our model is more general, it is non-parametrically identified and does not require the aforementioned assumptions about preferences or historical data. In this paper, due to its simplicity, we provide a parametric empirical illustration to familiarize the reader with the model. However, our model allows for derivation of non-parametric estimators, which we plan to discuss in the future work.

Advantages of evaluating ballot stuffing with the proposed model include: the absence of parametric assumptions, and a flexible estimation procedure based on a "clean" subsample that uses readily available data from independent observers. This procedure can infer information about any polling station in the country. Finally, the estimation procedure based on known costs of voting does not require any "clean" data at all. To the best of our knowledge, this is a more general and applicable way to estimate ballot stuffing than other existing methods.

First, we propose a structural model of electoral choice. This paper supposes a voter needs to decide between two candidates. Similar to the probabilistic voting model (Lindbeck and Weibull (1987), Persson and Tabellini (2000)), we assume an individual's political preferences can be separated into three components: personal, local and regional. Additionally, we explore voter costs: A voter, who does not find candidates that different from one another, will not be willing to pay a high cost, like a long wait in line at a polling station.Thus, people who are close to indifferent between candidates will abstain from elections. To this author's knowledge this is the first time that the above approaches have been combined to describe electoral choice. In addition, we suppose that the local component and costs of voting are the same for all people in one polling station and regional characteristics affect all residents of that region in the same way. From these assumptions, we derive values of turnout and voter share in the population. This structure allows us to identify and estimate the preferences in the whole population under the following assumptions: (1) additivity of preference components, (2) independence of personal and regional components on other characteristics, and, (3) linearity of the regional component in regional characteristics. Note that these assumptions still allow for a more general structure than the models based on one or another way of preference formation.

Second, this model is particularly well-suited to describe and evaluate ballot stuffing.[3] In this kind of electoral fraud, the administrators of a polling station place addi-

[3]The model allows adding several other types of electoral fraud or electoral boycott.

tional ballots in the electoral urn. In addition, they adjust corresponding official lists of participants and records of turnout. We suppose that only one of the candidates has access to ballot stuffing[4]. This implies that the number of votes for another candidate is truthful in every polling station, allowing for the identification of distribution of personal preferences and the regional component. Note that the presence of ballot stuffing affects only the polling station variables – the local component and the costs of voting. If the costs are known, then the local component can be recovered and the exact amount of fraud is identified. If the costs are unknown, then the exact amount of fraud cannot be obtained. However, a "clean" subsample of the data may enhance the identification of joint distribution of costs and the local component. Additionally, if the amount of fraud is independent of costs given observables, it becomes possible to evaluate the distribution of the amount of fraud in each polling station.

Finally, we use the 2011 Russia parliamentary election data to illustrate how the model can be estimated and used for evaluation of the counterfactual voter shares as well as average ballot stuffing.

Section 2 introduces a model of electoral choice. Section 3 discusses the identification of the model. In Section 4, we provide identification of ballot stuffing and a counterfactual voter shares in the entire population. Section 5 consists of empirical illustration. All derivations of estimators and asymptotics are included in the Appendix.

2 Electoral model

There are two candidates, A and B, running for office. Each voter has preferences over candidates: Similarly to probabilistic voting in Persson and Tabellini (2000),

[4]This assumption applies to many cases in which ballot stuffing is accessible only to the incumbent party. The model can be extended to include situations in which both parties have ability to add extra ballots, but the researcher will need to know which polling station is controlled by which candidate for identification. However, we recognize that even with such extension we do not cover all possible situations.

voter i in a polling station j in region K chooses candidate A over B if

$$\sigma_A^{ijK} + \delta_A^{jK} + \mu_A^K > \sigma_B^{ijK} + \delta_B^{jK} + \mu_B^K,$$

where σ^{ijK} is a parameter of individual "pure" preferences towards the candidates, δ^{jK} is popularity of a candidate in the area of the polling station and is the same for one polling station j, μ^K is a regional effect in popularity of each candidate and it is a function of some observable characteristics of the region X_K, such as average income, level of education, share of old population, etc. Thus, $\mu^K = \tilde{h}(X_K)$. Note that the additive structure is not a strong assumption and it does not affect evaluation of counterfactuals, including electoral fraud, if non-parametric estimation is used.

In order to obtain the reduced form of the model, we define parameters of difference in preferences between candidates $\sigma^{ijK} = \sigma_B^{ijK} - \sigma_A^{ijK}$, $\delta^{jK} = \delta_B^{jK} - \delta_A^{jK}$ and $\mu^K = \mu_B^K - \mu_A^K = \tilde{h}_B(X_K) - \tilde{h}_A(X_K) \equiv h(X_K)$. Therefore, voter i in a polling station j in region K chooses candidate A over B if

$$\sigma^{ijK} + \delta^{jK} + h(X_K) < 0,$$

where σ^{ijK} is personal "pure" preference for the candidate B, δ^{jK} and μ_K are polling station j and regional effects on preferences, correspondingly. Moreover, $\mu^K = h(X_K)$, where $h(\cdot)$ is a continuous and monotone in all arguments function.

To include turnout in the model, we follow the empirical evidence that suggests voting costs affect electoral participation (Fujiwara et al. (2016), Leon (2017)). A voter chooses to participate in elections if the difference in her preferences from different candidates is higher than costs of participation:

$$|\sigma^{ijK} + \delta^{jK} + h(X_K)| \geq c^{jK}.$$

Participation costs c^{jK} are random and the same for all voters in the same polling station, but might be different across different polling stations. Costs might represent the length of line to vote, the weather, difficulty of obtaining a voter card, etc. Such representation of participation implies that if the preferences of a voter are close to indifference between the candidates, then she does not attend elections. And, in contrast, if a person has very strong preferences toward one or another candidate, then she comes to the polling station even when costs are high[5].

Generally, a polling station represents a small geographical area, in which the distance from a voter's home to the polling station is similar for all voters. This motivates the assumption about the fixed level of costs of voting at a polling station. We understand that several factors, even weather conditions, might feel different for individual participants. However, the electoral data is aggregated at a polling station level and it would be impossible to disentangle personal costs of voting from the preferences without personal level data or/and strong assumptions about the model structure and behavior.

Note that the above assumption applies well if people are more or less homogenous in terms of costs of voting in each polling station, while there can be variations across polling stations. For example, different age groups may prefer different suburbs. If the homogeneity does not hold, then the unique part of the voting costs for each person will be accounted in the personal component of the preferences. This does not cause any problems for electoral fraud estimation. However, one must be careful while evaluating counterfactuals related to change of costs of voting as the results might be under/overstated.

In addition, we assume that individual "pure" preferences σ^{ijK} are independent identically distributed variables with density $g(\cdot)$ and cumulative distribution $G(\cdot)$. Costs of voting c^{jK} and local preferences δ^{jK} are independent identically distributed variables with joint density $f_{\delta,c}(\cdot,\cdot)$.

[5] We recognize that this model does not account for "marginal voter" way of thinking, i.e., when voters believe that their vote does not matter and abstain from elections. The data shows that people vote and we do not attempt to contribute to the divisive question of why they do so.

Next we introduce "swing voters", σ_A^{jK} and σ_B^{jK}, in every polling station j of region K, who are indifferent between participating and not participating in elections:

$$\sigma_A^{jK} = -\delta^{jK} - \mu^K - c^{jK}$$
$$\sigma_B^{jK} = \sigma_A^{jK} + 2c^{jK}.$$

Notice that people with "pure" preferences $\sigma^{ijK} < \sigma_A^{jK}$ will vote for the candidate A, people with $\sigma^{ijK} > \sigma_B^{jK}$ will vote for the candidate B, and everybody in between the swing voters will abstain from elections. As a result, the number of people who vote for A in a polling station j in region K is $n_A^{jK} = \int_{-\infty}^{\sigma_A^{jK}} dG(x) = G(\sigma_A^{jK})$. The same number for candidate B is $n_B^{jK} = \int_{\sigma_B^{jK}}^{+\infty} dG(x) = 1 - G(\sigma_B^{jK})$. Thus, turnout in the polling station is $\tau^{jK} = 1 - G(\sigma_B^{jK}) + G(\sigma_A^{jK})$ and A's share of votes is $\pi_A^{jK} = \frac{n_A^{jK}}{n_A^{jK}+n_B^{jK}} = \frac{G(\sigma_A^{jK})}{1-G(\sigma_B^{jK})+G(\sigma_A^{jK})}$.

The only data available in any elections is voter share, π_A^{jK}, and turnout, τ^{jK}, across all polling stations j and all regions K. However, the following electoral variables can be easily recovered from the data:

$$G(\sigma_A^{jK}) = \pi_A^{jK}\tau^{jK}$$
$$G(\sigma_B^{jK}) = 1 - \tau^{jK} + G(\sigma_A^{jK}) = 1 - \tau^{jK} + \pi_A^{jK}\tau^{jK}.$$

In everything that follows, we denote the observable electoral variables $Y = G(\sigma_A^{jK})$ and $Z = G(\sigma_B^{jK})$, while vector $X \in \mathbb{R}^L$ is a vector of L regional characteristics.

3 Identification of the model

This section discusses how to identify unobservable $g(\cdot)$, $h(\cdot)$ and $f_{\delta,c}(\cdot,\cdot)$ from observable joint distribution of Y and Z conditional on X.

Assumption 1. *Personal preferences σ is independent on X, δ and c, its support in $\mathbb{R}$ is compact, and it has continuously differentiable density $g(\cdot)$, strictly increasing* *on the support cumulative distribution function $G(\cdot)$, and $E\sigma = 0$.*

Assumption 2. *Local preferences δ and costs of voting c have continuously differentiable joint density $f_{\delta,c}(\cdot,\cdot)$, cumulative distribution function $F_{\delta,c}(\cdot,\cdot)$, and they are independent on regional characteristics X.*

The joint distribution of swing voters across polling stations in region K with characteristics X is:

$$F_{\sigma_A,\sigma_B|X}(y,z) = P(\sigma_A < y, \sigma_B < z|X) = P(-\delta - c - h(X) < y, -\delta + c - h(X) < z) =$$
$$= F_{-\delta-c,-\delta+c}\left(h(X) + y, h(X) + z\right).$$

Now we obtain the joint distribution of Y and Z conditional on X, $F_{Y,Z|X}(\cdot,\cdot)$, as follows:

$$F_{Y,Z|X}(y,z) = P(G(\sigma_A) < y, G(\sigma_B) < z|X) = P(\sigma_A < G^{-1}(y), \sigma_B < G^{-1}(z)) =$$
$$= F_{-\delta-c,-\delta+c}\left(h(X) + G^{-1}(y), h(X) + G^{-1}(z)\right), \tag{1}$$

from which we derive

$$f_{Y,Z|X}(y,z) = \frac{f_{-\delta-c,-\delta+c}\left(h(X) + G^{-1}(y), h(X) + G^{-1}(z)\right)}{g(G^{-1}(y))g(G^{-1}(z))}$$

and

$$F_{Z|X}(z) = F_{-\delta+c}\left(h(X) + G^{-1}(z)\right).$$

Cumulative distribution function $F_{Y,Z|X}(\cdot,\cdot)$ of electoral variables conditional on the regional characteristics is observed. The right-hand side of the above equation is completely unknown and is to be identified. It is important to notice that the proposed model is not identified without additional assumptions. In order to see this, consider that there are true functions $F_{\delta,c}(\cdot)$, $G^{-1}(\cdot)$ and $h(\cdot)$. Suppose that

$F_{-\delta-c,-\delta+c}(t_1,t_2)=\frac{t_1}{t_2}$, then another functions $\tilde{h}(X)=2h(X)$ and $\tilde{G}^{-1}(y)=2G^{-1}(y)$ will generate the same data.

Theorem 1. *If Assumptions 1 and 2 hold, and there exists X_0 such that the value $h(X_0)=h$ is known, then $f_{\delta,c}(\cdot,\cdot)$, $g(\cdot)$ and function $h(X)$ are identified.*

The proof of the above theorem is provided in the Appendix. In this paper, we work with a stronger assumption on $h(X)$. We consider linear $h(X)=\beta' X$ and normalize coefficients β. The linear form is chosen to decrease the usually high dimensionality of regional characteristics X. We offer two different normalizations. One implies that β equals to the vector of average derivatives of the conditional cumulative distribution function of Z given X with respect to regional characteristics X. This normalization is more suitable for non-parametric estimation. Another normalization is imposed on β indirectly through function $g(\cdot)$ and should be used with parametric estimation.

Theorem 2. *Functions $f_{\delta,c}(\cdot,\cdot)$, $g(\cdot)$ and coefficients β are identified if Assumptions 1 and 2 hold, $h(X)=\beta' X$ and one of the following normalization conditions applies:*

1. $\beta_1=E_{X,Z}\frac{\partial F_{Z|X}(z)}{\partial X^1}$;
2. *the value of $g(G^{-1}(z))$ is known at some point z_0 inside the support.*

Proof. Notice that $g(\cdot)$ is a derivative of $G(\cdot)$, thus, $\frac{\partial G^{-1}(x)}{\partial x}=\frac{1}{g(G^{-1}(x))}$ and by taking partial derivatives of the equation (1), we obtain the following:

$$f_{Z|X}(z)=\frac{f_{-\delta+c}\left(h(X)+G^{-1}(z)\right)}{g(G^{-1}(z))}$$
$$\frac{\partial F_{Z|X}(z)}{\partial X^i}=f_{-\delta+c}\left(h(X)+G^{-1}(z)\right)h^i(X)=g(G^{-1}(z))f_{Z|X}(z)\beta_i,$$

where $h^i(X)=\beta_i$ denotes partial derivatives with respect to i-th element. The last equality is available for different values of X and z, so we integrate all the points over the joint density of X and z:

$$\int\int_0^1 \frac{\partial F_{Z|X}(z)}{\partial X^i} f(X,z)dzdX = \beta_i \int\int_0^1 g(G^{-1}(z)) f_{Z|X}(z) f(X,z)dzdX \tag{2}$$

$$\frac{\beta_i}{\beta_j} = \frac{\int\int_0^1 \frac{\partial F_{Z|X}(z)}{\partial X^i} f(X,z)dzdX}{\int\int_0^1 \frac{\partial F_{Z|X}(z)}{\partial X^j} f(X,z)dzdX} = \frac{E_{X,Z}\frac{\partial F_{Z|X}(z)}{\partial X^i}}{E_{X,Z}\frac{\partial F_{Z|X}(z)}{\partial X^j}}. \tag{3}$$

First, we will consider the normalization condition $\beta_1 = E_{X,Z}\frac{\partial F_{Z|X}(z)}{\partial X^1}$ and obtain identification of $h(\cdot)$: $\beta_i = E_{X,Z}\frac{\partial F_{Z|X}(z)}{\partial X^i}$.

Hence, β is identified. Now we apply the same logic and integrate over density $f(X)$ for given z, to obtain the following:

$$\int \frac{\partial F_{Z|X}(z)}{\partial X^i} f(X)dX = \beta_i g(G^{-1}(z)) \int f_{Z|X}(z) f(X)dX = \beta_i g(G^{-1}(z)) f_Z(z)$$

$$g(G^{-1}(z)) = \frac{\int \frac{\partial F_{Z|X}(z)}{\partial X^i} f(X)dX}{\beta_i f_Z(z)}.$$

However, note that the linearity of $h(\cdot)$ together with the proposed normalization of β, require that $\frac{\partial F_{Z|X}(z)}{\partial X^i}$ is a constant and equals to β_i. By taking this into account, we get

$$g(G^{-1}(z)) = \frac{1}{f_Z(z)}. \tag{4}$$

If instead of the condition on β, we assume that $g(G^{-1}(z))$ is known at some value z_0, then from (3) we obtain $\beta_i = aE_{X,Z}\frac{\partial F_{Z|X}(z)}{\partial X^i}$, where $a \neq 0$ is some constant. This implies that (4) will be $g(G^{-1}(z)) = \frac{1}{af_Z(z)}$, implying that $a = \frac{1}{f_Z(z_0)g(G^{-1}(z_0))}$. Thus, a, β and $g(G^{-1}(z))$ are identified. The rest of the proof works for both normalization conditions identically.

The identification of $\phi(x) = g(G^{-1}(x))$ provides us with the identification of $G(\cdot)$: Notice that $g(G^{-1}(x))$ is density at the point where cumulative distribution function $G(\cdot)$ takes value x. It implies that we know the structure of the density but we are missing the axis. However, expectation of personal preferences is assumed to be 0

(Assumption 1), which is sufficient for identification of $G(\cdot)$:

$$F_Z(z) = \int_0^z f_Z(u)du = \int_0^z \frac{1}{\phi(u)}du = \int_0^z \frac{1}{g(G^{-1}(u))}du = G^{-1}(z) + const.$$

Using the condition for expectation:

$$\int_0^1 G^{-1}(z)g(G^{-1}(z))dG^{-1}(z) = \int_0^1 G^{-1}(z)dz = 0$$

$$\int_0^1 \left[\int_0^z \frac{1}{\phi(u)}du - const\right] dz = 0.$$

Thus, unknown constant is

$$const = \int_0^1 \int_0^z \frac{1}{\phi(u)}dudz = \int_0^1 \int_x^1 \frac{1}{\phi(u)}dzdu = \int_0^1 \frac{1-u}{\phi(u)}du = \int_0^1 (1-u)f_Z(u)du.$$

Suppose a random variable U is distributed with density $f_U(u) = 2(1-u)$, if $u \in [0,1]$, and it is independent on X, then $const = 0.5E_U(f_Z(U))$. Thus, $G^{-1}(z) = F_Z(z) - 0.5E_U(f_Z(U))$. It follows that $G(\cdot)$ is identified from its inverse function.

Note that under linearity assumption on $h(X)$, we obtain that $f_{Y,Z|X}(y,z) = f_{Y,Z|\beta'X}(y,z)$. In addition, after obtaining functions $G^{-1}(\cdot)$ and $h(\cdot)$, it is possible to identify the joint density of swing voters:

$$f_{-\delta-c,-\delta+c}\left(\beta'X + G^{-1}(y), \beta'X + G^{-1}(z)\right) = f_{Y,Z|\beta'X}(y,z)g(G^{-1}(y))g(G^{-1}(z)).$$

Finally, by applying the transformation theorem, we derive the joint density of local preferences and costs of voting, δ and c:

$$f_{\delta,c}(x,y) = 2f_{-\delta-c,-\delta+c}(-x-y, -x+y).$$

□

4 Applications

4.1 Counterfactuals

In this section, we show how to identify the would be voter shares if the voting were mandatory and, hence, everyone showed up at the polling station. The model allows straightforward calculation of this counterfactual. By definition, at each polling station

$$\begin{cases} Y = G(-\delta - X\beta - c) \\ Z = G(-\delta - X\beta + c) \end{cases} \Rightarrow \begin{cases} -\delta - X\beta - c = G^{-1}(Y) \\ -\delta - X\beta + c = G^{-1}(Z) \end{cases}.$$

If c were to be 0, then we would observe $\tilde{Y} = \tilde{Z} = G(-\delta - X\beta)$. Hence,

$$\tilde{Y} = \tilde{Z} = G(0.5G^{-1}(Y) + 0.5G^{-1}(Z)).$$

4.2 Ballot Stuffing

This section discusses how to evaluate the amount of ballot stuffing using the developed model. Formally, ballot stuffing is defined as the illegal practice of one person submitting multiple ballots during a vote in which only one ballot per person is permitted. In non-democratic elections, ballot stuffing happens when the official staff that runs the voting illegally puts a number of pre-filled ballots into the ballot box. In addition, after the polling station closes, the staff illegally corrects the official turnout to make it consistent with the number of votes in the ballot box. Video evidence of ballot stuffing is publicly available for the 2017 Turkey referendum and the most of Russian elections since 2011.

In this paper, we assume that only candidate A has an opportunity to rig the election. In this case, the observables will be affected as follows:

Variable Y denotes the number of votes for the candidate A in a polling station.

If q is a number of additional ballots due to fraud, then we will observe variable $Y^f = Y + q$.

Variable Z will not be affected as $1 - Z$ represents the number of votes for the candidate B.

4.3 Identification

Theorem 3. *Function $g(\cdot)$ and coefficients β are identified if Assumptions 1 and 2 hold, $h(X) = \beta' X$ and one of the following normalization conditions is true:*

1. $\beta_1 = E_{X,Z} \frac{\partial F_{Z|X}(z)}{\partial X^1}$;

2. *the value of $g(G^{-1}(z))$ is known at some point z_0 inside the support.*

The proof follows the first part of the proof of Theorem 2.

The joint distribution of local preferences and costs $f_{\delta,c}(\cdot,\cdot)$ is not identified due to presence of fraud at the local level. Thus, we need some additional information. In this paper, we offer two potential solutions: (1) costs of voting are known in each polling station; or (2) there exists a clean (non-fraudulent) subsample of the data.

Theorem 4. *If cost of voting c is known in each polling station, then the realization of local preferences δ and the amount of fraud q are identified.*

Proof. The main argument of this theorem is very simple: Notice that by the definition, $Z = G(\sigma_B)$, where $\sigma_B = -\delta - h(X) + c$, hence, $\delta = c - h(X) - G^{-1}(Z)$.

By Theorem 3, $h(X)$ and $G^{-1}(\cdot)$ are identified. If c is known, then we find δ.

Note that $Y^f = Y + q$, where by definition $Y = G(\sigma_A)$ and $\sigma_A = -\delta - h(X) - c$. Thus,

$$q = Y^f - G(-\delta - h(X) - c) = Y^f - G(G^{-1}(Z) - 2c).$$

□

The known costs of voting might be a good assumption in situations where it is known which variables have impact on the attendance of the voters. An example would be the weather conditions – it is believed that heavy rainfalls affected the turnout in the Colombian peace referendum.[6] In alike cases, it is possible to relate costs of voting to the amount of rainfall in the area. Moreover, Fujiwara et al. (2016) show that rainfall on current and past election days affects voter turnout in the US presidential elections. Hence, costs of voting might be estimated and predicted.

Theorem 5. *The joint distribution of local preferences and costs $f_{\delta,c}(\cdot,\cdot)$ is identified if $f_{Y,Z|X=X^k}$ can be recovered from a "clean" subsample in a region with characteristics X^k. Moreover, if $q \perp\!\!\!\perp c|X,Z$, then the distribution of the amount of fraud q, $f_{q|X,Y^f,Z}(\cdot)$, is identified at every polling station.*

In some situations, a representable subsample of clean electoral data might be available. For example, independent observers might be randomized to some polling stations in one region. If their presence helps reduce fraud and as long as randomization is done carefully, then the data available from these polling stations will allow for estimation of $f_{Y,Z|X^k}(y,z)$. However, notice that in this situation it is not possible to identify the exact amount of fraud and only the distribution conditional on observables is available.

5 Empirical Illustration

Non-parametric identification allows for broad range of estimation procedures. Two main approaches toward estimation would be parametric and non-parametric. In this paper, we make parametric assumptions about distributions of the model variables, and estimate the model with the conditional MLE. This will demonstrate how to estimate the electoral model, counterfactuals and ballot stuffing.

[6] *Colombia just voted no on its plebiscite for peace. Here's why and what it means*, The Washington Post, October 3, 2016.

Below we develop an estimation procedure and apply it to evaluate the amount of ballot stuffing during the 2011 Russia parliamentary elections in several polling stations. The Russian political system consists of numerous parties. However, the country's regime is authoritarian, and the current political spectrum can be divided into two factions: United Russia, the ruling party and all other political parties. Hence, we combine all votes given for the various opposition parties to obtain the number of votes for the opposition. The official results report 49.3% as the voter share obtained by the United Russia with 60.1% turnout. Note that the below analysis is only for illustration purposes, and the obtained results should be taken with caution[7].

5.1 Data

The main data set is the results of the 2011 Russia parliamentary elections for each polling station. The data set is openly available on the website of the Central Electoral Commission of the Russian Federation[8]. In total we use 94,795 observations after exclusion of the voting on the foreign territory.

The choice of the unit for a region should be based on availability of data that is being used as controls. There are no requirements regarding the region size, and a region might consist of one polling station, if detailed information on regional characteristics is available. We use a Territorial Electoral Commission (TIK) as a region for the non-urban areas and we combine the data of several TIKs inside of the city to form a city-region. This procedure left us with 2,483 regions. Note that such choice of a region is not the best for Russia as TIKs do not always coincide with non-urban regional units used by the municipalities. Hence, for proper analysis one should form a region by including polling stations inside of the municipality rather than using the division suggested by the Electoral Commission.

[7] The proper estimation would require matching of over 90,000 data points with their municipalities and collection of regional characteristics for over 2,000 municipalities, for which data is often available in different formats. Such data collection is outside of the scope of this paper.

[8] http://www.izbirkom.ru/region/izbirkom

We take the total number of voters and the average size of a polling station in a region as regional characteristics X. We normalize them by dividing by the largest value in the sample, hence, all values of X are between 0 and 1[9]. Electoral variable Y is the number of votes for the United Russia divided by the total number of voters[10] in the polling station. To obtain variable Z, we divide the number of pro-opposition votes by the total number of voters and subtract it from 1.

We also use the data on clean polling stations from Enikolopov et al. (2013). The authors collect the data from 156 randomly assigned independent observers in Moscow, which revealed that even the presence of one independent observer reduced the average share of the United Russia from 47% to 36%. We use 75 polling stations that were reported as "no violations" by the independent observers to constitute our "clean" subsample.

5.2 Model

This section illustrates how the model can be estimated parametrically. Due to non-parametric identification, any additional assumption results in overidentification and, hence, a potentially large array of ways of estimation. Here we show only one of the approaches. We impose assumptions about the distribution of costs of voting, and personal and local components. Note that one could start with an assumption about the distribution of the final data set. In that case, the logic of the estimation would follow the non-parametric approach. The assumptions we impose below produce a linear model, which is the simplest model for the demonstration.

Assumption 3. *Suppose that* $\sigma \sim U[-0.5, 0.5]$.

[9] Note that the proper regional characteristics should include a large number of controls such as average income, share of old population, etc. This data can be obtained from the municipality websites.

[10] To obtain the total number of voters for each polling station we add up columns $n1$ (the number of registered voters) and $n13$ (the number of voters voted with the absentee certificates) and subtract columns $n12$ (the number of absentee certificates issued by the polling station) and $n15$ (the number of absentee certificates issued by the TIK) in the data set.

Assumption 3 implies that $G(x) = x + 0.5$ if $x \in [-0.5, 0.5]$. This assumption automatically takes care of the necessary normalization in Theorem 2. It also guarantees linearity of the electoral variables in their components.

Assumption 4. *Suppose that* $\begin{pmatrix} \delta \\ c \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_\delta \\ \mu_c \end{pmatrix}, \begin{pmatrix} \sigma_\delta^2 & \sigma_{\delta,c} \\ \sigma_{\delta,c} & \sigma_c^2 \end{pmatrix}\right)$.

Note that Assumptions 3 and 4 imply the following.

$$\begin{cases} Y = G(\sigma_A) = -\delta - c - X\beta + 0.5 \\ Z = G(\sigma_B) = -\delta + c - X\beta + 0.5 \end{cases} \Rightarrow \begin{pmatrix} Y \\ Z \end{pmatrix} |X \sim N(\mu_X, \Sigma_{YZ}),$$

where $\mu_X = \begin{pmatrix} -\mu_\delta - \mu_c - X\beta + 0.5 \\ -\mu_\delta + \mu_c - X\beta + 0.5 \end{pmatrix}$ and $\Sigma_{YZ} = \begin{pmatrix} \sigma_\delta^2 + 2\sigma_{\delta c} + \sigma_c^2 & \sigma_\delta^2 - \sigma_c^2 \\ \sigma_\delta^2 - \sigma_c^2 & \sigma_\delta^2 - 2\sigma_{\delta c} + \sigma_c^2 \end{pmatrix}$.
Hence, the data is conditionally normal with $L + 5$ parameters that we estimate by the conditional MLE. See the Appendix for the derivation of the MLE estimators and their asymptotics.

Table 1: Conditional MLE estimators

	μ_δ	μ_c	β_1	β_2	σ_δ^2	σ_c^2	$\sigma_{\delta c}$
parameter	-0.0699	0.1747	0.0278	1.1548	0.0252	0.0090	0.0106
standard error	0.0008	0.0004	0.0022	0.0105	0.0001	0.0001	0.0001

Table 1 presents the MLE estimates of the model parameters. Costs of voting have average $\mu_c = 0.1747$, which means that on average almost 35% of the voting population are absent due to these costs[11]. The local component has negative average, which implies that on average each locality preferences are biased toward the United Russia. Everything else equal, the United Russia receives the support of additional 14% of the total voting population due to the local preferences component. The total number of votes and the average size of the polling station positively affect the voter

[11]The official turnout is 60.1%.

share of the opposition, which indicates the greater support of the current regime in rural areas. Note also that the variance of costs of voting is significantly lower than the variance of the local component, however, the correlation between the two variables $\rho_{\delta c} = 0.7$ is strong. This suggests that the costs of voting are greater in the polling stations with the greater support of the opposition.

5.3 Counterfactual analysis

Evaluation of counterfactuals is straightforward. In this section, we consider the electoral preferences of the entire population or the voter shares that the candidates would receive if the voting were mandatory, i.e., $c = 0$. Generally, $Y_i^c = G(0.5G^{-1}(Y_i) + 0.5G^{-1}(Z_i))$.

In the Russian example, $G(\cdot)$ is a uniform cumulative distribution function, so the counterfactual value is simply $Y_i^c = 0.5Y_i + 0.5Z_i$ at each polling station. We recalculate this value for all polling stations in the data set and obtain that the voter share of the United Russia would be 50.03% instead of 49.31% in the 2011 parliamentary election if the voting were mandatory. Note that this result relies on uniformity of the personal component, which possibly does not hold in the Russian case.

5.4 Ballot Stuffing

First, note that under the proposed assumptions $Z|X \sim N(\mu_c - \mu_\delta - X\beta + 0.5, \sigma_Z^2)$. According to the identification results, β is identifiable, while μ_c and μ_δ cannot be separated as well as the variances. Hence, we estimate the coefficients β of the regional characteristics X using Z only. In this case, the conditional MLE is equivalent to the linear regression. The estimates are shown below in Table 2.

Notice that the impact of the total number of voters in the region on the electoral variables, β_1, is the same as in the MLE estimation of the entire sample. However, the effect of the average polling station size, β_2, is significantly larger. This has to

Table 2: Estimators of β

	const	β_1	β_2
coefficient	0.7550	0.0221	1.5986
standard error	0.0008	0.0021	0.0297

be because the ballot stuffing is correlated with the average polling station size. We believe this due to increased availability of the empty ballots in the larger polling stations. If the ballot stuffing occurs before the closure of the polls, then the smaller polling station is more likely to run out of ballots than a larger one.

Second, we fit the clean subsample into the normal distribution and obtain the following estimates of the parameters in Moscow.

Table 3: Distribution parameters in Moscow

	$\hat{\mu}_{Y\|\tilde{X}}$	$\hat{\mu}_{Z\|\tilde{X}}$	$\hat{\sigma}^2_{Y\|\tilde{X}}$	$\hat{\sigma}^2_{Z\|\tilde{X}}$	$\hat{\sigma}_{YZ\|\tilde{X}}$
parameter	0.1258	0.6270	0.0015	0.0023	-0.0001
standard error	0.0044	0.0056	0.0002	0.0004	0.0002

Third, we calculate estimators for the distribution parameters for a region with the characteristics $\bar{X}$ as follows.

$$\hat{\mu}_{Y|\bar{X}} = \hat{\mu}_{Y|\tilde{X}} + (\tilde{X} - \bar{X})\hat{\beta}$$
$$\hat{\mu}_{Z|\bar{X}} = \hat{\mu}_{Z|\tilde{X}} + (\tilde{X} - \bar{X})\hat{\beta}$$

We estimate these parameters for the city of Kostroma and obtain the values[12] reported in Table 4.

[12] Note that the conditional covariance matrix Σ_{YZ} is the same in each region, so we do not need to estimate it again.

Table 4: Distribution parameters in Kostroma

	$\hat{\mu}_{Y\|\tilde{X}}$	$\hat{\mu}_{Z\|\tilde{X}}$	$\hat{\sigma}^2_{Y\|\tilde{X}}$	$\hat{\sigma}^2_{Z\|\tilde{X}}$	$\hat{\sigma}_{YZ\|\tilde{X}}$
parameter	0.1535	0.6547	0.0015	0.0023	-0.0001
standard error	0.0044	0.0056	0.0002	0.0004	0.0002

Finally, we evaluate the average amount of ballot stuffing based on observables:

$$E(q|X, Z, Y^f) = Y^f - \mu_{Y|X,Z},$$

where

$$\mu_{Y|X,Z} = \mu_{Y|X} + \frac{\sigma_{YZ|X}}{\sigma^2_{Z|X}}(Z - \mu_{Z|X}) \text{ and } \sigma^2_{Y|X,Z} = \sigma^2_{Y|X} - \frac{\sigma^2_{YZ|X}}{\sigma^2_{Z|X}}.$$

See the Appendix for the asymptotic variance of this estimator.

We calculate the average ballot stuffing in two polling stations in Moscow: one "clean" (UIK 265) and one of unknown quality (UIK 856). In addition, we randomly pick two polling stations in Kostroma (UIK 213 and UIK 299) and evaluate the average fraud there. The results are shown in Table 5.

Table 5: Estimators of average ballot stuffing

	UIK 265	UIK 856	UIK 213	UIK 299
Y^f	0.1275	0.2934	0.1499	0.1964
average ballot stuffing	0.0011	0.1749	-0.0046	0.0404
standard error	0.0045	0.0121	0.0047	0.0058
reported voter share of United Russia	25.02%	53.97%	29.30%	33.68%
expected voter share of United Russia	24.85%	32.13%	29.93%	28.74%

UIK 265 in Moscow is part of the "clean" subsample and the fraud estimation confirms it. Similarly, we do not report any evidence of ballot stuffing in UIK 213 in

Kostroma. In contrast, UIK 856 in Moscow and UIK 299 in Kostroma show significant levels of ballot stuffing. In UIK 856, the administrative staff filled in ballots for 17.49% of the voter population of the polling station, which resulted in the increase of the voter share of the United Russia from 32.13% to 53.97% after taking the turnout into account. In UIK 299, the administrative personnel stuffed ballots for 4.04% of the voter population, which augmented the United Russia voter share from 28.74% to 33.68%.

6 Conclusion

We introduced a structural model of electoral preferences that accounts for turnout and voter share and does not rely on assumptions about preference formation. The advantage of the model is that it is well-suited for the evaluation of a single election and ballot stuffing. We showed how the model can be non-parametrically identified and parametrically estimated.

The future work might develop in three dimensions. First, it is possible to develop non-parametric estimators of the model. In this case, the estimation will rely on the weakest assumptions, which will deliver robustness. Second, the model can be modified to account for electoral boycott and other types of electoral fraud. Third, it might be possible to relax some of the independence assumptions.

References

[1] Breunig, C. and A. Goerres *"Searching for Electoral Irregularities in an Established Democracy: Applying Benfords Law Tests to Bundestag Elections in Unified Germany"*, Electoral Studies 30, 534-545, 2011.

[2] Burkovskaya, A. *"Non-parametric Estimation of Electoral Model and Ballot Stuffing,"* Working paper, 2018.

[3] Cantu, F. *"Identifying Irregularities in Mexican Local Elections"*, American Journal of Political Science 58 (4), 936-951, 2014.

[4] Coate, S. and M. Conlin *"A Group Rule-Utilitarian Approach to Voter Turnout: Theory and Evidence"*, American Economic Review 94 (5), 1476-1504, 2004.

[5] Degan, A. *"Candidate Valence: Evidence from Consecutive Presidential Elections"*, International Economic Review 48 (2), 457-482, 2007.

[6] Degan, A. and A. Merlo *"A Structural Model of Turnout and Voting in Multiple Elections"*, Journal of the European Economic Association 9 (2), 209-245, 2011.

[7] Downs, A. *"An Economic Theory of Political Action in a Democracy"*, Journal of Political Economy 65 (2), 135-150, 1957.

[8] Enikolopov, R., V. Korovkin, M. Petrova, K. Sonin and A. Zakharov *"Field Experiment Estimate of Electoral Fraud in Russian Parliamentary Elections"*, Proceedings of the National Academy of Sciences 110 (2), 448-452, 2013.

[9] Fujiwara, T., K. Meng and T. Vogl *"Habit Formation in Voting: Evidence from Rainy Elections"*, American Economic Journal: Applied Economics 8 (4), 160-188, 2016.

[10] Hinich, M.J. and M.C. Munger *"Ideology and the Theory of Public Choice"*, Ann Arbor: University of Michigan Press, 1996.

[11] Kernell, G. *"Giving Order to Districts: Estimating Voter Distributions with National Election Returns"*, Political Analysis 17 (3), 215-235, 2009.

[12] Klimek, P., Y. Yegorov, R. Hanel and S. Thurner *"Statistical Detection of Systematic Election Irregularities"*, Proceedings of the National Academy of Sciences 109 (40), 16469-16473, 2012.

[13] Kobak, D., S. Shpilkin, and M. S. Pshenichnikov *"Integer Percentages as Electoral Falsification Fingerprints,"* The Annals of Applied Statistics 10, 54-73, 2016.

[14] Leon, G. *"Turnout, political preferences and information: Experimental evidence from Peru,"* Journal of Development Economics 127, 56-71, 2017.

[15] Levin, I., G.A. Cohn, P.C. Ordeshook and R.M. Alvarez *"Detecting Voter Fraud in an Electronic Voting Context: An Analysis of the Unlimited Reelection Vote in Venezuela"*, Proceedings of the 2009 conference on Electronic voting technology/workshop on trustworthy elections, USENIX Association, 2009.

[16] Lindbeck, A. and J.W. Weibull *"Balanced-Budget Redistribution as the Outcome of Political Competition"*, Public Choice 52, 273-297, 1987.

[17] Mebane, W.R., Jr. *"Election Forensics: The Second-Digit Benfords Law Test and Recent American Presidential Elections"*, Election Fraud: Detecting and Deterring Electoral Manipulation. Washington, D.C.: Brookings Institution Press, 162-81, 2008.

[18] Merlo, A., and A. de Paula *"Identification and Estimation of Preference Distributions When Voters are Ideological"*, Review of Economic Studies 84(3), 1238-1263, 2017.

[19] McMurray, J. *"Aggregating Information by Voting: The Wisdom of the Experts versus the Wisdom of the Masses"*, Review of Economic Studies 80, 277-312, 2013.

[20] Persson, T. and G. Tabellini *"Political Economics: Explaining Economic Policy"*, MIT Press, 2000.

[21] Riker, W. and P. Ordeshook *"A Theory of the Calculus of Voting"*, The American Political Science Review 62 (1), 25-42, 1968.

[22] Rozenas, A. *"Detecting Election Fraud from Irregularities in Vote-Share Distributions,"* Political Analysis 25 (1), 41-56, 2017.

A Appendix

A.1 Proof of Theorem 1

Notice that $g(\cdot)$ is a derivative of $G(\cdot)$, thus, $\frac{\partial G^{-1}(x)}{\partial x} = \frac{1}{g(G^{-1}(x))}$ and by taking partial derivatives of the equation (1), we obtain the following:

$$f_{Z|X}(z) = \frac{f_{-\delta+c}\,(h(X) + G^{-1}(z))}{g(G^{-1}(z))}$$
$$\frac{\partial F_{Z|X}(z)}{\partial X^i} = f_{-\delta+c}\left(h(X) + G^{-1}(z)\right) h^i(X) = g(G^{-1}(z)) f_{Z|X}(z) h^i(X)$$

where $h^i(\cdot)$ denotes partial derivatives with respect to i-th element.

Take the last expression for pairs (X, z_1) and (X, z_2) and obtain the following ratio.

$$\frac{\frac{\partial F_{Z|X}(z_1)}{\partial X^i}}{\frac{\partial F_{Z|X}(z_2)}{\partial X^i}} = \frac{g(G^{-1}(z_1))}{g(G^{-1}(z_2))} \frac{f_{Z|X}(z_1)}{f_{Z|X}(z_2)}.$$

Functions $\frac{\partial F_{Z|X}(z)}{\partial X^i}$ and $f_{Z|X}(z)$ are observable, hence, we identify the ratio $\frac{g(G^{-1}(z_1))}{g(G^{-1}(z_2))}$. Now pick a point z_0 and denote $g(G^{-1}(z_0)) = t$, then

$$g(G^{-1}(z)) = t \frac{g(G^{-1}(z))}{g(G^{-1}(z_0))}.$$

In addition, $g(\cdot)$ is a density, thus,

$$\int_0^1 g(G^{-1}(z))dz = t \int_0^1 \frac{g(G^{-1}(z))}{g(G^{-1}(z_0))} dz = 1,$$

implying that t is identified together with $g(G^{-1}(z))$.

Identification of $g(\cdot)$ and $G^{-1}(\cdot)$ under $E\sigma = 0$ is shown in detail in *Theorem* 2.

Now it is possible to identify the joint density of swing voters: Fix X at the value

X_0, then

$$f_{-\delta-c,-\delta+c}\left(h+G^{-1}(y),h+G^{-1}(z)\right)=f_{Y,Z|X_0}(y,z)g(G^{-1}(y))g(G^{-1}(z)).$$

Finally, by applying the transformation theorem, we derive the joint density of local preferences and costs of voting, δ and c:

$$f_{\delta,c}(x,y)=2f_{-\delta-c,-\delta+c}(-x-y,-x+y).$$

A.2 Proof of Theorem 5

Note that

$$f_{-\delta-c,-\delta+c}\left(h(X^k)+G^{-1}(y),h(X^k)+G^{-1}(z)\right)=f_{Y,Z|X=X^k}(y,z)g(G^{-1}(y))g(G^{-1}(z)).$$

By Theorem 3, $h(X)$ and $g(\cdot)$ are identified. It also implies that $G^{-1}(\cdot)$ and $g(G^{-1}(\cdot))$ can be recovered (see proof of Theorem 2). Thus, the only part that prevents us from identifying $f_{-\delta-c,-\delta+c}(\cdot,\cdot)$ is $f_{Y,Z|X=X^k}(y,z)$, and the existence of the "clean" data solves this problem. Finally, $f_{\delta,c}(\cdot,\cdot)$ can be easily obtained from

$$f_{\delta,c}(x,y)=f_{-\delta-c,-\delta+c}(-x-y,-x+y).$$

Also note that $f_{Y,Z|X}(\cdot,\cdot)$ is identified for any region:

$$f_{Y,Z|X}(y,z)=\frac{f_{-\delta-c,-\delta+c}\left(h(X)+G^{-1}(y),h(X)+G^{-1}(z)\right)}{g(G^{-1}(y))g(G^{-1}(z))}.$$

Notice that the condition $q\perp\!\!\!\perp c|X,Z$ is equivalent to $Y\perp\!\!\!\perp q|X,Z$. In addition, we observe $\bar{Y} = Y + q$ at every polling station and we know $f_{Y|X,Z}(\cdot)$. Hence,

$$F_{q|X,\bar{Y},Z}(t) = P(q \leq t | Y + q = \bar{Y}, X, Z) = P(\bar{Y} - Y \leq t | q = \bar{Y} - Y, X, Z) = P(Y \geq \bar{Y} - t | X, Z)$$
$$= 1 - F_{Y|X,Z}(\bar{Y} - t) = 1 - F_{Y|X,Z}(\bar{Y} - t).$$

Thus, cdf $F_{q|X,\bar{Y},Z}(\cdot)$ is identified.

A.3 MLE derivations

In this section, we derive MLE estimators of the structural model for the case of jointly normal electoral variables. Hence, $Y, Z|X \sim N(\mu, \Sigma)$, where $\mu = \begin{bmatrix} \mu_Y \\ \mu_Z \end{bmatrix}$ and $\Sigma = \begin{bmatrix} \sigma_Y^2 & \sigma_{YZ} \\ \sigma_{YZ} & \sigma_Z^2 \end{bmatrix}$. The model structure implies that

$$\mu_Y = -\mu_\delta - \mu_c - X\beta + 0.5$$
$$\mu_Z = -\mu_\delta + \mu_c - X\beta + 0.5$$
$$\sigma_Y^2 = \sigma_\delta^2 + 2\sigma_{\delta,c} + \sigma_c^2$$
$$\sigma_Z^2 = \sigma_\delta^2 - 2\sigma_{\delta,c} + \sigma_c^2$$
$$\sigma_{YZ} = \sigma_\delta^2 - \sigma_c^2.$$

Our purpose is to estimate $\theta = (\mu_\delta, \mu_c, \beta', \sigma_\delta^2, \sigma_c^2, \sigma_{\delta,c})'$.

First, we write down the log-likelihood of the data:

$$\log \mathcal{L} = -n \log(2\pi) - 0.5n \log |\Sigma_{YZ}| - 0.5|\Sigma_{YZ}|^{-1} \sum_{i=1}^{N} \left[\sigma_Z^2 (Y_i + \mu_\delta + \mu_c + X_i\beta - 0.5)^2 - \right.$$
$$\left. -2\sigma_{YZ}(Y_i + \mu_\delta + \mu_c + X_i\beta - 0.5)(Z_i + \mu_\delta - \mu_c + X_i\beta - 0.5) + 2\sigma_Y^2 (Z_i + \mu_\delta - \mu_c + X_i\beta - 0.5)^2 \right]$$

By taking partial derivatives, we get the following system of equations.

$$(\sigma_c^2 - \sigma_{\delta,c})\sum_{i=1}^{N} Y_i + (\sigma_c^2 + \sigma_{\delta,c})\sum_{i=1}^{N} Z_i + 2\sigma_c^2 \sum_{i=1}^{N} X_i\beta + 2n((\mu_\delta - 0.5)\sigma_c^2 - \mu_c\sigma_{\delta,c}) = 0$$

$$(\sigma_\delta^2 - \sigma_{\delta,c})\sum_{i=1}^{N} Y_i - (\sigma_\delta^2 + \sigma_{\delta,c})\sum_{i=1}^{N} Z_i + 2n(\mu_c\sigma_\delta^2 - (\mu_\delta - 0.5)\sigma_{\delta,c}) = 0$$

$$(\sigma_c^2 - \sigma_{\delta,c})\sum_{i=1}^{N} X_i^j Y_i + (\sigma_c^2 + \sigma_{\delta,c})\sum_{i=1}^{N} X_i^j Z_i + 2\sigma_c^2 \sum_{i=1}^{N} X_i^j X_i\beta +$$
$$+2((\mu_\delta - 0.5)\sigma_c^2 - \mu_c\sigma_{\delta,c})\sum_{i=1}^{N} X_i^j = 0 \text{ for } j = \overline{1,k}.$$

$$\sum_{i=1}^{N}[\sigma_{\delta,c}(Y_i - Z_i + 2\mu_c) - \sigma_c^2(Y_i + Z_i + 2\mu_\delta + 2X_i\beta - 1)]^2 = 4n|\Sigma_{\delta,c}|\sigma_c^2$$
$$\sum_{i=1}^{N}[\sigma_\delta^2(Y_i - Z_i + 2\mu_c) - \sigma_{\delta,c}(Y_i + Z_i + 2\mu_\delta + 2X_i\beta - 1)]^2 = 4n|\Sigma_{\delta,c}|\sigma_\delta^2$$

$$\sum_{i=1}^{N}[\sigma_\delta^2\sigma_{\delta,c}(Y_i - Z_i + 2\mu_c)^2 - (\sigma_{\delta,c}^2 + \sigma_c^2\sigma_\delta^2)(Y_i - Z_i + 2\mu_c)(Y_i + Z_i + 2\mu_\delta + 2X_i\beta - 1) +$$
$$+\sigma_c^2\sigma_{\delta,c}(Y_i + Z_i + 2\mu_\delta + 2X_i\beta - 1)^2] = 4n|\Sigma_{\delta,c}|\sigma_{\delta,c}$$

Denote $a_i = Y_i - Z_i + 2\mu_c$ and $b_i = Y_i + Z_i + 2\mu_\delta + 2X_i\beta - 1$, then the above system can be rewritten as follows.

$$\sigma_c^2 \sum_{i=1}^{N} b_i - \sigma_{\delta,c}\sum_{i=1}^{N} a_i = 0 \tag{5}$$

$$\sigma_\delta^2 \sum_{i=1}^{N} a_i - \sigma_{\delta,c}\sum_{i=1}^{N} b_i = 0 \tag{6}$$

$$\sigma_c^2 \sum_{i=1}^{N} X_i^j b_i - \sigma_{\delta,c} \sum_{i=1}^{N} X_i^j a_i = 0 \text{ for } j = \overline{1,k} \tag{7}$$

$$\frac{\sigma_{\delta,c}^2}{\sigma_c^2} \sum_{i=1}^{N} a_i^2 - 2\sigma_{\delta,c} \sum_{i=1}^{N} a_i b_i + \sigma_c^2 \sum_{i=1}^{N} b_i = 4n|\Sigma_{\delta,c}| \tag{8}$$

$$\sigma_\delta^2 \sum_{i=1}^{N} a_i^2 - 2\sigma_{\delta,c} \sum_{i=1}^{N} a_i b_i + \frac{\sigma_{\delta,c}^2}{\sigma_\delta^2} \sum_{i=1}^{N} b_i = 4n|\Sigma_{\delta,c}| \tag{9}$$

$$\sigma_\delta^2 \sum_{i=1}^{N} a_i^2 - \left(\sigma_{\delta,c} + \frac{\sigma_\delta^2 \sigma_c^2}{\sigma_{\delta,c}}\right) \sum_{i=1}^{N} a_i b_i + \sigma_c^2 \sum_{i=1}^{N} b_i = 4n|\Sigma_{\delta,c}| \tag{10}$$

First, note that Equation (5) and Equation (6) give us $\sum_{i=1}^{N} a_i = \sum_{i=1}^{N} b_i = 0$ as long as $|\Sigma_{\delta,c}| \neq 0$. Hence, we obtain that $\hat{\mu}_c = 0.5(\bar{Z} - \bar{Y})$ and $\hat{\mu}_\delta = 0.5(1 - \bar{Y} - \bar{Z} - 2\bar{X}\beta)$. By plugging these values back into a_i and b_i, we derive

$$a_i = (Y_i - \bar{Y}) - (Z_i - \bar{Z})$$
$$b_i = (Y_i - \bar{Y}) + (Z_i - \bar{Z}) + 2(X_i - \bar{X})\beta.$$

Note that a_i depends only on the observable data, while β is the only parameter vector in b_i. Then by subtracting Equation (8) from Equation (9), we get $\sum_{i=1}^{N} b_i^2 = \frac{\sigma_\delta^2}{\sigma_c^2} \sum_{i=1}^{N} a_i^2$ and plug it into Equation (8) and Equation (10), obtaining $\sum_{i=1}^{N} a_i b_i = \frac{\sigma_{\delta,c}}{\sigma_c^2} \sum_{i=1}^{N} a_i^2$. Now by plugging it back into Equation (8) or Equation (10), we derive

$\sum_{i=1}^N a_i^2 = 4n\sigma_c^2$, $\sum_{i=1}^N a_i b_i = 4n\sigma_{\delta,c}$, and $\sum_{i=1}^N b_i^2 = 4n\sigma_\delta^2$. Hence, we have

$$\hat{\sigma}_c^2 = \frac{1}{4n}\sum_{i=1}^N ((Y_i - \bar{Y}) - (Z_i - \bar{Z}))^2$$
$$\hat{\sigma}_\delta^2 = \frac{1}{4n}\sum_{i=1}^N ((Y_i - \bar{Y}) + (Z_i - \bar{Z}) + 2(X_i - \bar{X})\hat{\beta})^2.$$

Now we use k equations eq. (7) together with $\sum_{i=1}^N a_i b_i = 4n\sigma_{\delta,c}$ to get the system of $k+1$ linear equations with $k+1$ unknowns, which are β and $\sigma_{\delta,c}$, then $\begin{bmatrix}\hat{\beta}\\ \hat{\sigma}_{\delta,c}\end{bmatrix} = A^{-1}b$, where A is a $(k+1)\times(k+1)$ matrix with the following elements:

$$A_{t,m} = 2\hat{\sigma}_c^2 \sum_{i=1}^N X_i^t (X_i^m - \bar{X}^m) \text{ if } t, m = \overline{1,k}$$
$$A_{k+1,m} = 2\sum_{i=1}^N ((Y_i - \bar{Y}) - (Z_i - \bar{Z}))(X_i^m - \bar{X}^m) \text{ if } m = \overline{1,k}$$
$$A_{m,k+1} = -\sum_{i=1}^N X_i^t ((Y_i - \bar{Y}) - (Z_i - \bar{Z})) \text{ if } t = \overline{1,k}$$
$$A_{k+1,k+1} = -4N$$

and b is a vector with $k+1$ elements:

$$b_t = -\hat{\sigma}_c^2 \sum_{i=1}^N X_i^j ((Y_i - \bar{Y}) - (Z_i - \bar{Z})) \text{ if } t = \overline{1,k}$$
$$b_{k+1} = -\sum_{i=1}^N ((Y_i - \bar{Y})^2 - (Z_i - \bar{Z})^2).$$

Now we are left to derive the asymptotic variance of the MLE estimators. Define the score function as $s(Y,Z|X,\theta) = \frac{\partial \log f_{Y,Z|X}(Y,Z)}{\partial \theta}$, which is $(k+5)$-vector as follows.

$$\begin{bmatrix} -\frac{1}{2|\Sigma_{\delta,c}|}\left((\sigma_c^2 - \sigma_{\delta,c})Y + (\sigma_c^2 + \sigma_{\delta,c})Z + 2(\sigma_c^2 X\beta + (\mu_\delta - 0.5)\sigma_c^2 - \mu_c\sigma_{\delta,c})\right) \\ -\frac{1}{2|\Sigma_{\delta,c}|}\left((\sigma_\delta^2 - \sigma_{\delta,c})Y - (\sigma_\delta^2 + \sigma_{\delta,c})Z - 2(\sigma_{\delta,c} X\beta + (\mu_\delta - 0.5)\sigma_{\delta,c} - \mu_c\sigma_\delta^2)\right) \\ -\frac{X^1}{2|\Sigma_{\delta,c}|}\left((\sigma_c^2 - \sigma_{\delta,c})Y + (\sigma_c^2 + \sigma_{\delta,c})Z + 2(\sigma_c^2 X\beta + (\mu_\delta - 0.5)\sigma_c^2 - \mu_c\sigma_{\delta,c})\right) \\ \dots \\ -\frac{X^k}{2|\Sigma_{\delta,c}|}\left((\sigma_c^2 - \sigma_{\delta,c})Y + (\sigma_c^2 + \sigma_{\delta,c})Z + 2(\sigma_c^2 X\beta + (\mu_\delta - 0.5)\sigma_c^2 - \mu_c\sigma_{\delta,c})\right) \\ -\frac{\sigma_c^2}{2|\Sigma_{\delta,c}|} + \frac{1}{8|\Sigma_{\delta,c}|^2}\left(\sigma_{\delta,c}(Y - Z + 2\mu_c) - \sigma_c^2(Y + Z + 2\mu_\delta + 2X\beta - 1)\right)^2 \\ -\frac{\sigma_\delta^2}{2|\Sigma_{\delta,c}|} + \frac{1}{8|\Sigma_{\delta,c}|^2}\left(\sigma_\delta^2(Y - Z + 2\mu_c) - \sigma_{\delta,c}(Y + Z + 2\mu_\delta + 2X\beta - 1)\right)^2 \\ \frac{\sigma_{\delta,c}}{|\Sigma_{\delta,c}|} - \frac{1}{4|\Sigma_{\delta,c}|^2}\left(\sigma_\delta^2\sigma_{\delta,c}(Y - Z + 2\mu_c)^2 - (\sigma_{\delta,c}^2 + \sigma_\delta^2\sigma_c^2)(Y - Z + 2\mu_c)(Y + Z + 2\mu_\delta + 2X\beta - 1) + \sigma_c^2\sigma_{\delta,c}(Y + Z + 2\mu_\delta + 2X\beta - 1)^2\right) \end{bmatrix}$$

Then the asymptotic variance of the MLE estimator can be calculated as the inverse of the information, i.e., $V_\theta = (E[s(Y,Z|X)s(Y,Z|X)'])^{-1}$. We use the analog estimator of the variance in our calculations.

A.4 Derivation of asymptotics of the clean subsample estimators

We assume that $\begin{pmatrix} Y \\ Z \end{pmatrix} |X = X^k \sim N(\mu, \Sigma)$, where $\mu = \begin{pmatrix} \mu_Y \\ \mu_Z \end{pmatrix}$ and $\Sigma = \begin{bmatrix} \sigma_Y^2 & \sigma_{YZ} \\ \sigma_{YZ} & \sigma_Z^2 \end{bmatrix}$. In addition, the density is estimated parametrically:

$$f_{Y,Z|X=X^k}(y,z) = \frac{1}{2\pi|\Sigma|^{0.5}} \exp\left\{-\frac{1}{2}\left(\begin{pmatrix} y \\ z \end{pmatrix} - \mu\right)' \Sigma^{-1} \left(\begin{pmatrix} y \\ z \end{pmatrix} - \mu\right)\right\}.$$

Five parameters are estimated from the clean data $\theta = (\mu_Y, \mu_Z, \sigma_Y, \sigma_Z, \sigma_{YZ})'$ as follows.

$$\hat{\mu}_Y = \frac{1}{N_k}\sum_{i=1}^{N_k} Y_i = \bar{Y} \text{ and } \hat{\mu}_Z = \frac{1}{N_k}\sum_{i=1}^{N_k} Z_i = \bar{Z}$$

$$\hat{\sigma}_Y^2 = \frac{1}{N_k - 1}\sum_{i=1}^{N_k}(Y_i - \bar{Y})^2 \text{ and } \hat{\sigma}_Z^2 = \frac{1}{N_k - 1}\sum_{i=1}^{N_k}(Z_i - \bar{Z})^2$$
$$\hat{\sigma}_{YZ} = \frac{1}{N_k - 1}\sum_{i=1}^{N_k}(Y_i - \bar{Y})(Z_i - \bar{Z}).$$

We want to derive V_{f^K} and show that $\sqrt{n}(\hat{f}_{Y,Z|X=X^k}(y,z) - f_{Y,Z|X=X^k}(y,z)) \xrightarrow{d} N(0, V_{f^K})$.

First, we trivially use multivariate CLT and obtain that $\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} N(0, V_\theta)$.

To derive V_θ, note the following: $\hat{\mu}_Y \sim N\left(\mu_Y, \frac{\sigma_Y^2}{N_k}\right)$ and $\hat{\mu}_Z \sim N\left(\mu_Z, \frac{\sigma_Z^2}{N_k}\right)$, and

$$\frac{(N_k-1)\hat{\sigma}_Y^2}{\sigma_Y^2} \sim \chi^2(N_k - 1) \Rightarrow V\left(\frac{(N_k-1)\hat{\sigma}_Y^2}{\sigma_Y^2}\right) = 2(N_k - 1) \Rightarrow V(\hat{\sigma}_Y^2) = \frac{2\sigma_Y^4}{N_k - 1}.$$

By analogy, $V(\hat{\sigma}_Z^2) = \frac{2\sigma_Z^4}{N_k - 1}$. In addition, due to independence $\text{cov}(\hat{\mu}_Y, \hat{\sigma}_Y^2) = \text{cov}(\hat{\mu}_Z, \hat{\sigma}_Z^2) = 0$. And $\text{cov}(\hat{\mu}_Y, \hat{\mu}_Z) = \frac{\sigma_{YZ}}{N_k}$.

Due to normality $\hat{\mu}_Y \perp\!\!\!\perp (Y_i - \bar{Y})$, $\hat{\mu}_Y \perp\!\!\!\perp (Z_i - \bar{Z})$, $\hat{\mu}_Z \perp\!\!\!\perp (Y_i - \bar{Y})$ and $\hat{\mu}_Z \perp\!\!\!\perp (Z_i - \bar{Z})$. Hence, $\text{cov}(\hat{\mu}_Y, \hat{\sigma}_Z^2) = \text{cov}(\hat{\mu}_Z, \hat{\sigma}_Y^2) = \text{cov}(\hat{\mu}_Y, \hat{\sigma}_{YZ}) = \text{cov}(\hat{\mu}_Z, \hat{\sigma}_{YZ}) = 0$.

Thus, we are left to obtain $V(\hat{\sigma}_{YZ})$, $\text{cov}(\hat{\sigma}_Y^2, \hat{\sigma}_Z^2)$, $\text{cov}(\hat{\sigma}_Y^2, \hat{\sigma}_{YZ})$ and $\text{cov}(\hat{\sigma}_Z^2, \hat{\sigma}_{YZ})$. To calculate those, we will need the following elements.

$$E(Y_i Z_i) = \sigma_{YZ} + \mu_Y \mu_Z$$
$$E(Y_i^2 Z_i) = (\sigma_Y^2 + \mu_Y^2)\mu_Z + 2\mu_Y \sigma_{YZ}$$
$$E(Y_i Z_i^2) = (\sigma_Z^2 + \mu_Z^2)\mu_Y + 2\mu_Z \sigma_{YZ}$$
$$E(Y_i^2 Z_i^2) = (\sigma_Y^2 + \mu_Y^2)\mu_Z^2 + (\sigma_Y^2 + \mu_Y^2)\sigma_Z^2 + 4\mu_Y\mu_Z\sigma_{YZ} + 2\sigma_{YZ}^2$$

Hence,

$$V(Y_i Z_i) = (\sigma_Y^2 + \mu_Y^2)\mu_Z^2 + (\sigma_Y^2 + \mu_Y^2)\sigma_Z^2 + 4\mu_Y\mu_Z\sigma_{YZ} + 2\sigma_{YZ}^2 - (\sigma_{YZ} + \mu_Y\mu_Z)^2 =$$
$$= \sigma_Y^2\mu_Z^2 + \sigma_Y^2\sigma_Z^2 + \mu_Y^2\sigma_Z^2 + 2\mu_Y\mu_Z\sigma_{YZ} + \sigma_{YZ}^2.$$

By analogy, one can obtain

$$V(\bar{Y}\bar{Z}) = \frac{1}{N_k}\sigma_Y^2\mu_Z^2 + \frac{1}{N_k^2}\sigma_Y^2\sigma_Z^2 + \frac{1}{N_k}\mu_Y^2\sigma_Z^2 + \frac{2}{N_k}\mu_Y\mu_Z\sigma_{YZ} + \frac{1}{N_k^2}\sigma_{YZ}^2.$$

Also

$$\operatorname{cov}(Y_iZ_i, Y_iZ_j) = E(Y_i^2Z_i)\mu_Z - E(Y_iZ_i)\mu_Y\mu_Z = ((\sigma_Y^2 + \mu_Y^2)\mu_Z + 2\mu_Y\sigma_{YZ})\mu_Z -$$
$$-(\sigma_{YZ} + \mu_Y\mu_Z)\mu_Y\mu_Z = \sigma_Y^2\mu_Z^2 + \mu_Y\mu_Z\sigma_{YZ}$$

and

$$\operatorname{cov}(Y_iZ_i, \bar{Y}\bar{Z}) = \frac{1}{N_k^2}\operatorname{cov}\left(Y_iZ_i, \sum_i\sum_j Y_iZ_j\right) = \frac{1}{N_k^2}\operatorname{cov}\left(Y_iZ_i, Y_iZ_i + \sum_{j\neq i} Y_iZ_j + \sum_{j\neq i} Y_jZ_i\right)$$
$$= \frac{1}{N_k^2}\left(V(Y_iZ_i) + (N_k-1)\operatorname{cov}(Y_iZ_i, Y_iZ_j) + (N_k-1)\operatorname{cov}(Y_iZ_i, Y_jZ_i)\right)$$
$$= \frac{1}{N_k^2}(\sigma_{YZ}^2 + \sigma_Y^2\sigma_Z^2 + N_k\sigma_Y^2\mu_Z^2 + N_k\sigma_Z^2\mu_Y^2 + 2N_k\mu_Y\mu_Z\sigma_{YZ}).$$

Now we can derive $V(\hat{\sigma}_{YZ})$:

$$V(\hat{\sigma}_{YZ}) = \frac{1}{(N_k-1)^2}(N_kV(Y_iZ_i) + N_k^2V(\bar{Y}\bar{Z}) - 2N_k^2\operatorname{cov}(Y_iZ_i, \bar{Y}\bar{Z}))$$
$$= \frac{1}{(N_k-1)^2}(N_k(\sigma_Y^2\mu_Z^2 + \sigma_Y^2\sigma_Z^2 + \mu_Y^2\sigma_Z^2 + 2\mu_Y\mu_Z\sigma_{YZ} + \sigma_{YZ}^2) + N_k\sigma_Y^2\mu_Z^2 +$$
$$+\sigma_Y^2\sigma_Z^2 + N_k\mu_Y^2\sigma_Z^2 + 2N_k\mu_Y\mu_Z\sigma_{YZ} + \sigma_{YZ}^2 - 2(\sigma_{YZ}^2 + \sigma_Y^2\sigma_Z^2 + N_k\sigma_Y^2\mu_Z^2 +$$
$$+N_k\sigma_Z^2\mu_Y^2 + 2N_k\mu_Y\mu_Z\sigma_{YZ})) = \frac{1}{N_k-1}(\sigma_Y^2\sigma_Z^2 + \sigma_{YZ}^2).$$

In order to find covariances of the estimators of variances we will need the following.

$$\operatorname{cov}(Y_i^2, Z_i^2) = 4\mu_Y\mu_Z\sigma_{YZ} + 2\sigma_{YZ}^2$$

$$\operatorname{cov}(Z_i^2, (\bar{Y})^2) = \frac{1}{N_k^2}\operatorname{cov}\left(Z_i^2, \sum_i Y_i^2 + \sum_j \sum_{k\neq j} Y_j Y_k\right) = \frac{1}{N_k^2}\operatorname{cov}\left(Z_i^2, Y_i^2 + 2Y_iY_j\right)$$

$$= \frac{1}{N_k^2}(4N_k\mu_Y\mu_Z\sigma_{YZ} + 2\sigma_{YZ}^2)$$

By analogy, $\operatorname{cov}(Y_i^2, (\bar{Z})^2) = \frac{1}{N_k^2}(4N_k\mu_Y\mu_Z\sigma_{YZ}+2\sigma_{YZ}^2)$ and $\operatorname{cov}((\bar{Y})^2, (\bar{Z})^2) = \frac{1}{N_k^2}(4N_k\mu_Y\mu_Z\sigma_{YZ}+2\sigma_{YZ}^2)$. Hence,

$$\operatorname{cov}(\hat{\sigma}_Y^2, \hat{\sigma}_Z^2) = \frac{1}{(N_k-1)^2}\left(N_k\operatorname{cov}(Y_i^2, Z_i^2) - N_k^2\operatorname{cov}(Z_i^2, (\bar{Y})^2) - N_k^2\operatorname{cov}(Y_i^2, (\bar{Z})^2) + N_k^2\operatorname{cov}((\bar{Y})^2, (\bar{Z})^2)\right)$$

$$= \frac{1}{(N_k-1)^2}(N_k(4\mu_Y\mu_Z\sigma_{YZ} + 2\sigma_{YZ}^2) - 4N_k\mu_Y\mu_Z\sigma_{YZ} - 2\sigma_{YZ}^2 - 4N_k\mu_Y\mu_Z\sigma_{YZ} - 2\sigma_{YZ}^2 +$$

$$+4N_k\mu_Y\mu_Z\sigma_{YZ} + 2\sigma_{YZ}^2) = \frac{2}{N_k-1}\sigma_{YZ}^2.$$

Finally, for the last covariance, we need to calculate the following values.

$$E(Y_i^3 Z_i) = (\sigma_Y^2 + \mu_Y^2)\mu_Y\mu_Z + 2\mu_Y\mu_Z\sigma_Y^2 + 3(\sigma_Y^2 + \mu_Y^2)\sigma_{YZ}$$

$$\operatorname{cov}(Y_i^2, Y_iZ_i) = (\sigma_Y^2 + \mu_Y^2)\mu_Y\mu_Z + 2\mu_Y\mu_Z\sigma_Y^2 + 3(\sigma_Y^2 + \mu_Y^2)\sigma_{YZ} - (\sigma_Y^2 + \mu_Y^2)(\sigma_{YZ} + \mu_Y\mu_Z)$$

$$= 2\mu_Y\mu_Z\sigma_Y^2 + 2(\sigma_Y^2 + \mu_Y^2)\sigma_{YZ}$$

$$\operatorname{cov}(Y_iZ_i, (\bar{Y})^2) = \frac{1}{N_k^2}\operatorname{cov}(Y_iZ_i, Y_i^2 + 2(N_k-1)Y_iY_j) = \frac{1}{N_k^2}(2\mu_Y\mu_Z\sigma_Y^2 + 2(\sigma_Y^2 + \mu_Y^2)\sigma_{YZ}$$

$$+2(N_k-1)\mu_Y(\sigma_Y^2\mu_Z + \mu_Y\sigma_{YZ})) = \frac{1}{N_k^2}(2N_k\mu_Y\mu_Z\sigma_Y^2 + 2N_k\mu_Y^2\sigma_{YZ} + 2\sigma_Y^2\sigma_{YZ})$$

$$\mathrm{cov}(Y_i^2, \bar{Y}\bar{Z}) = \frac{1}{N_k^2}\mathrm{cov}(Y_i^2, Y_iZ_i + (N_k-1)Y_iZ_j + (N_k-1)Y_jZ_i)$$

$$= \frac{1}{N_k^2}(2\mu_Y\mu_Z\sigma_Y^2 + 2(\sigma_Y^2+\mu_Y^2)\sigma_{YZ} + (N_k-1)\mu_Z((\mu_Y^3+3\mu_Y\sigma_Y^2) - \mu_Y(\mu_Y^2+\sigma_Y^2))$$

$$+(N_k-1)((\sigma_Y^2+\mu_Y^2)\mu_Y\mu_Z + 2\mu_Y^2\sigma_{YZ} - \mu_Y\mu_Z(\sigma_Y^2+\mu_Y^2)))$$

$$= \frac{2}{N_k^2}(N_k\mu_Y\mu_Z\sigma_Y^2 + \sigma_Y^2\sigma_{YZ} + N_k\mu_Y^2\sigma_{YZ})$$

$$\mathrm{cov}((\bar{Y})^2, \bar{Y}\bar{Z}) = \frac{2}{N_k^2}(N_k\mu_Y\mu_Z\sigma_Y^2 + \sigma_Y^2\sigma_{YZ} + N_k\mu_Y^2\sigma_{YZ})$$

Thus,

$$\mathrm{cov}(\hat{\sigma}_Y^2, \hat{\sigma}_{YZ}) = \frac{1}{(N_k-1)^2}(N_k\mathrm{cov}(Y_i^2, Y_iZ_i) - N_k^2\mathrm{cov}(Y_iZ_i, (\bar{Y})^2) - N_k^2\mathrm{cov}(Y_i^2, \bar{Y}\bar{Z}) + N_k^2\mathrm{cov}((\bar{Y})^2, \bar{Y}\bar{Z}))$$

$$= \frac{1}{(N_k-1)^2}(2N_k\mu_Y\mu_Z\sigma_Y^2 + 2N_k(\sigma_Y^2+\mu_Y^2)\sigma_{YZ} - 2N_k\mu_Y\mu_Z\sigma_Y^2 - 2N_k\mu_Y^2\sigma_{YZ} - 2\sigma_Y^2\sigma_{YZ}$$

$$-2N_k\mu_Y\mu_Z\sigma_Y^2 - 2\sigma_Y^2\sigma_{YZ} - 2N_k\mu_Y^2\sigma_{YZ} + 2N_k\mu_Y\mu_Z\sigma_Y^2 + 2\sigma_Y^2\sigma_{YZ} + 2N_k\mu_Y^2\sigma_{YZ})$$

$$= \frac{2}{N_k-1}\sigma_Y^2\sigma_{YZ}.$$

By analogy, $\mathrm{cov}(\hat{\sigma}_Z^2, \hat{\sigma}_{YZ}) = \frac{2}{N_k-1}\sigma_Z^2\sigma_{YZ}$. Finally, we obtain V_θ:

$$V_\theta = \begin{bmatrix} \sigma_Y^2 & \sigma_{YZ} & 0 & 0 & 0 \\ \sigma_{YZ} & \sigma_Z^2 & 0 & 0 & 0 \\ 0 & 0 & 2\sigma_Y^4 & 2\sigma_{YZ}^2 & 2\sigma_{YZ}\sigma_Y^2 \\ 0 & 0 & 2\sigma_{YZ}^2 & 2\sigma_Z^4 & 2\sigma_{YZ}\sigma_Z^2 \\ 0 & 0 & 2\sigma_{YZ}\sigma_Y^2 & 2\sigma_{YZ}\sigma_Z^2 & \sigma_{YZ}^2 + \sigma_Y^2\sigma_Z^2 \end{bmatrix}.$$

Hence, $\sqrt{N_k}(\hat{\theta}-\theta) \xrightarrow{d} N(0, V_\theta)$.

A.5 Estimation of parameters of other than the clean subsample region

In the clean subsample, the distribution parameters suggest that

$$\begin{cases} \mu_{Y|\tilde{X}} = -\mu_\delta - \mu_c - \tilde{X}\beta + 0.5 \\ \mu_{Z|\tilde{X}} = -\mu_\delta + \mu_c - \tilde{X}\beta + 0.5 \end{cases}.$$

On the other hand, distribution means of the electoral variables with the regional characteristics $\bar{X}$ can be calculated similarly, i.e.,

$$\begin{cases} \mu_{Y|\bar{X}} = -\mu_\delta - \mu_c - \bar{X}\beta + 0.5 \\ \mu_{Z|\bar{X}} = -\mu_\delta + \mu_c - \bar{X}\beta + 0.5 \end{cases}, \text{ hence, } \begin{cases} \mu_{Y|\bar{X}} = \mu_{Y|\tilde{X}} + (\tilde{X} - \bar{X})\beta \\ \mu_{Z|\bar{X}} = \mu_{Z|\tilde{X}} + (\tilde{X} - \bar{X})\beta \end{cases}.$$

We use analog estimators in this case. The distribution variance estimator is the same as in the clean subsample due to equality of the conditional variances in the normal distribution. Thus, we are left to derive the asymptotic covariance of all the estimators together. Note that the clean subsample parameters are estimated from the clean subsample, which is a very small part of the entire sample (75 points out of 94,795), hence, the estimators variance will be dominated by the estimators derived from the clean subsample, and the covariance between estimators can be ignored. Thus, the asymptotic variance is

$$V_{\theta|\bar{X}} = \begin{bmatrix} \sigma_Y^2 + (\tilde{X} - \bar{X})V_\beta(\tilde{X} - \bar{X})' & \sigma_{YZ} + (\tilde{X} - \bar{X})V_\beta(\tilde{X} - \bar{X})' & 0 & 0 & 0 \\ \sigma_{YZ} + (\tilde{X} - \bar{X})V_\beta(\tilde{X} - \bar{X})' & \sigma_Z^2 + (\tilde{X} - \bar{X})V_\beta(\tilde{X} - \bar{X})' & 0 & 0 & 0 \\ 0 & 0 & 2\sigma_Y^4 & 2\sigma_{YZ}^2 & 2\sigma_{YZ}\sigma_Y^2 \\ 0 & 0 & 2\sigma_{YZ}^2 & 2\sigma_Z^4 & 2\sigma_{YZ}\sigma_Z^2 \\ 0 & 0 & 2\sigma_{YZ}\sigma_Y^2 & 2\sigma_{YZ}\sigma_Z^2 & \sigma_{YZ}^2 + \sigma_Y^2\sigma_Z^2 \end{bmatrix}.$$

Hence, $\sqrt{N_k}(\hat{\theta}_{\bar{X}} - \theta_{\bar{X}}) \xrightarrow{d} N(0, V_{\theta|\bar{X}})$.

Lemma 1. *Suppose that*

$$\sqrt{N}\left(\begin{bmatrix}\hat{\mu}_Y\\ \hat{\mu}_Z\\ \hat{\sigma}_Y^2\\ \hat{\sigma}_Z^2\\ \hat{\sigma}_{YZ}\end{bmatrix}-\begin{bmatrix}\mu_Y\\ \mu_Z\\ \sigma_Y^2\\ \sigma_Z^2\\ \sigma_{YZ}\end{bmatrix}\right)\xrightarrow{d} N(0,V_\theta),$$

$\mu_{Y|Z}=\mu_Y+\frac{\sigma_{YZ}}{\sigma_Z^2}(Z-\mu_Z)$ *and* $\sigma_{Y|Z}^2=\sigma_Y^2-\frac{\sigma_{YZ}^2}{\sigma_Z^2}$, *then*

$$\sqrt{N}\left(\begin{bmatrix}\hat{\mu}_{Y|Z}\\ \hat{\sigma}_{Y|Z}^2\end{bmatrix}-\begin{bmatrix}\mu_{Y|Z}\\ \sigma_{Y|Z}^2\end{bmatrix}\right)\xrightarrow{d} N(0,W),$$

where $W=\Gamma V_\theta\Gamma'$ *and*

$$\Gamma=\begin{bmatrix}1 & -\frac{\sigma_{YZ}}{\sigma_Z^2} & 0 & -\frac{\sigma_{YZ}(Z-\mu_Z)}{\sigma_Z^4} & \frac{Z-\mu_Z}{\sigma_Z^2}\\ 0 & 0 & 1 & -\frac{\sigma_{YZ}^2}{\sigma_Z^4} & -\frac{2\sigma_{YZ}}{\sigma_Z^2}\end{bmatrix}.$$

Proof. The proof is straightforward from the Δ-method. $\square$