EXPLAINING POLITICAL SURPRISES (AKA MAKING METHODOLOGY FUN): DETERMINANTS OF VOTING IN UKRAINIAN PRESIDENTIAL ELECTIONS

Florin N Fesnic
Center for the Study of Democracy, "Babeş-Bolyai" University Cluj, Romania
fesnic@democracycenter.ro

The majority of social science majors dislike statistics and research methods. Things can, and should be, different; (social) statistics and research methods can be interesting and fun. To do just that, I currently work on an applied methodology book project. Here I present a modified version of a draft chapter for the book, in which I answer the question why, in Ukraine, there was a strong negative correlation between the regional shares of the vote for Leonid Kuchma in two consecutive presidential elections. The modelling of political competition as two-dimensional (one socioeconomic, one ethno-linguistic) reveals the dominance of the latter dimension in Ukrainian politics. In addition to statistical and methodological lessons, the paper offers substantive lessons, relevant for the important role (positive or negative) that institutional design, policy-making and elite behavior can play in an ethnically divided emerging democracy.

INTRODUCTION

Regional differences in voting patterns in national elections are ubiquitous; we can find them in both established and emerging democracies. Sometimes, the roots of these differences are primarily socioeconomic (e.g., the English regions of the UK), while in other cases they are mostly cultural/ethnic (e.g., Canada or Belgium) and, finally, sometimes both economic and non-economic determinants play an important role (e.g., Romania; arguably, the non-English parts of the UK as well). While some of these examples may be debatable, the general point is that, across the group of countries where we can observe significant regional differences in electoral outcomes, we can think of the structural determinants of voting along a continuum, ranging from (almost) purely socioeconomic determinants at one end to ethno-linguistic determinants at the other.

What follows is an analysis of presidential elections held in Ukraine from 1994 to 2010 – more specifically, an analysis of regional (oblast)-level vote as a function of the two aforementioned factors, socioeconomic development and ethno-linguistic diversity. The results of the analysis reveal useful lessons for students who employ statistics to study social phenomena, from both statistical and research design perspectives (for instance, showing the importance of control variables and longitudinal analyses), and from a substantive perspective (illustrating the importance of in-depth knowledge of the cases under scrutiny). Last, but certainly not least, the process of revisiting Ukraine's electoral history in the last two decades offers general lessons for political elites and institutional designers.

METHODS AND DATA

The dependent variable in the analyses that follow are aggregate (county, or oblast)-level results (percentage of the total vote) of the "centrist"[1]/Russophile presidential candidates in the second round of the 1994, 1999, 2004[2] and 2010 presidential elections (Leonid Kuchma in the first two cases, Viktor Yanukovych in the last two). For two of the three independent variables (percentage urban population and percentage Russian-speaking population), I use data from the 2001 official census. A third independent variable is "West": a dummy variable for the seven oblasts that, up until the end of the First World War, were part of the Austro-Hungarian Empire[3].

[1]Throughout this paper, my use of "left", "right" and "center" is restricted to denoting solely the positions of candidates on socioeconomic policies.
[2]For 2004, when there was a second (repeat) runoff, the data refers to this second runoff election.
[3]The seven oblasts are Chernivtsi, Ivano-Frankivsk, Lviv, Rivne, Ternopil, Volyn and Zakarpattia. For a more detailed theoretical justification of this division, see, for instance, Roper and Fesnic 2003.

In: J. Engel (Ed.), *Promoting understanding of statistics about society.*
Proceedings of the Roundtable Conference of the International Association of Statistics Education (IASE),
July 2016, Berlin, Germany. iase-web.org/Conference_Proceedings.php

To analyse the impact of "development" (i.e., urbanization) and ethnicity on the vote I use linear regression, starting with bivariate regression[4] (the impact of regional development on support for "centrist" candidates, Table 1), then adding ethno-linguistic composition (as well as region) as control variables (Table 3). To offer a more nuanced view of the relationship between all these variables, I also add a table with bivariate correlations between all the variables used in the analysis (Table 2).

THE IMPACT OF DEVELOPMENT AND ETHNICITY IN UKRAINIAN ELECTIONS

The argument that "development" (one's well-being) has a significant impact on voting behaviour has a long tradition in political science (e.g., Lipset 1960). Moreover, scholars of development argue that urbanization is one of its most significant driving forces, as well as a measure of it (Handelman 2006). Therefore, it is reasonable to test whether urbanization affects voting, particularly in countries where we see a wide variation in both variables. Table 1 presents the results of this analysis for Ukraine, to which I added the results of two recent presidential elections in Romania, another post-Communist country (and also neighbour of Ukraine), as well as the US, a very different country from Ukraine (politically, economically, and socially).

Table 1. "Development" (percent urban population) and regional voting in three countries

	Ukraine (n=27)				*Romania (n=41)*	*US (n=50)*
	Kuchma		Yanukovych			
	1994	1999	2004	2010	Iohannis 2014	Obama 2008
urbanization	1.06	−0.39	1.30	1.11	0.43	0.21
standard error	0.24	0.19	0.25	0.24	0.15	0.06
significance	0.000	0.06	0.000	0.000	0.006	0.001
Adjusted R2	0.42	0.11	0.50	0.41	0.16	0.21

As we can see, urbanization has a significant impact on the vote in all three cases, but it is by far the largest in Ukraine. This impact is larger by a magnitude of 4 to 5 times than in the other two countries when we take into account both the coefficient(s) for the independent variable (b's) and the distribution of the variable in each country (i.e., if we multiply the b with the standard error). Nonetheless, if the results from 2004 and 2010 are similar, which is what we would normally expect when the same candidate runs in two consecutive elections (and they resemble the results from 1994), the change in the results from 1994 and 1999 is rather puzzling. If the correlation between Yanukovych's share of votes in the two elections is 0.99, the correlation between Kuchma's shares is -.80!

If anything, the massive shift in the pattern of regional support for Kuchma from 1994 to 1999 becomes even more puzzling if we try to analyse the two elections in conventional, left-right terms. In both elections (runoffs), his opponents (the incumbent president Leonid Kravchuck in 1994, and then the leader of the Communist Party, Petr Simonenko) were located to his left:

[4] As one of the reviewers of the initial draft pointed out, this approach may be looked upon as unconventional (why start with bivariate regression? And why development?) My justification for doing so (may or may not be enough for a reason) is that this is less of a first (or second) draft of an academic paper, and more of a draft of a textbook ("Levitt's *Freakonomics* (2006) meets Landman's *Issues and Methods* (2003)"), aiming to make statistics and methodology user-friendly and fun. In this respect, the paper resembles more a detective novel, where the mystery is solved only at the end, than the typical academic paper, where the reader is informed about the main results from the very beginning. My goal was to present the initial results (i.e., the strong impact of development in the initial bivariate model, impact that vanishes completely after controlling for ethnolinguistic structure), as a puzzle that needs to be solved using statistical methods, on the one hand, and substantive knowledge about the case, on the other.

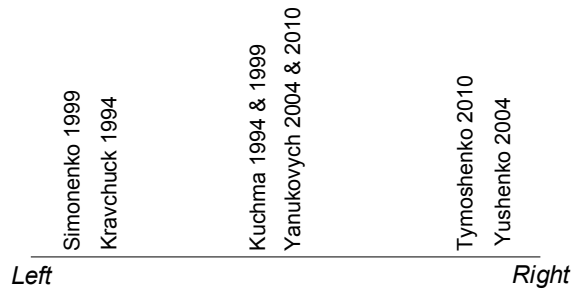

Figure 1. The relative left-right positions of presidential candidates in Ukraine, 1994-2010

However, if we also take into account the fact that Kravchuck was a Ukrainian nationalist, while Simonenko was the leader of the Communist Party, the most Russophile of all major parties in the country, things start to make sense. What explains the shift in the pattern of support for Kuchma from 1994 to 1999 is the *relative* change of his position along the second, Russophile versus nationalist position. In 1994, his opponent was a nationalist, thus he became the natural choice of Russophiles; the situation of 1999 was a mirror image of 1994:

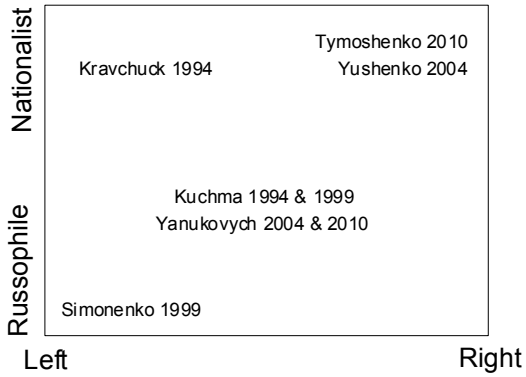

Figure 2. The mapping of presidential candidates along two dimensions

The results in Table 2 offer additional insights into the regional differences in Ukraine. Russian-speaking population is concentrated in the largest cities, such as Luhansk and Donetsk, which are located in Eastern Ukraine (Figure 3). That being the case, it is quite possible that language (plus region), rather than urbanization, truly explain the results in Table 1. Indeed, such an explanation is consistent with both Figure 2 and the results from Table 2 (see also Figures 3, 4 and 5):

Table 2. Correlation matrix of the variables employed in the analysis

	Kuchma 1994 (%)	Kuchma 1999 (%)	Yanukovych 2004 (%)	Yanukovych 2010 (%)	Urbanization (%)	Russian speakers (%)
Kuchma vote in 1999	−0.80**					
Yanukovych vote in 2004 (2nd runoff)	0.84**	−0.50**				
Yanukovych vote in 2010	0.86**	−0.55**	0.99**			
Urban population (1997)	0.67**	−0.37	0.72**	0.70**		
Russian-speaking (2001)	0.81**	−0.40*	0.94**	0.90**	0.73**	
Western Ukraine (seven oblasts)	−0.79**	0.91**	−0.54**	−0.58**	−0.57**	−0.45*

*Correlation is significant at the 0.05 level (2-tailed)
**Correlation is significant at the 0.01 level (2-tailed)

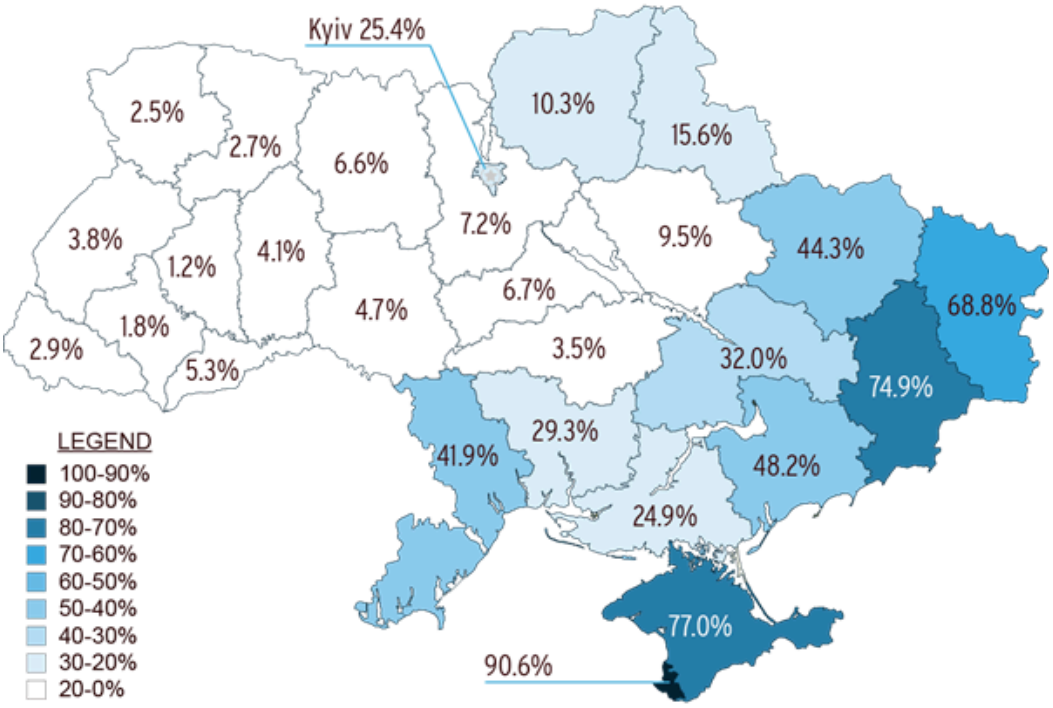

Figure 3. The share of Russian-speaking population across Ukrainian regions[5]

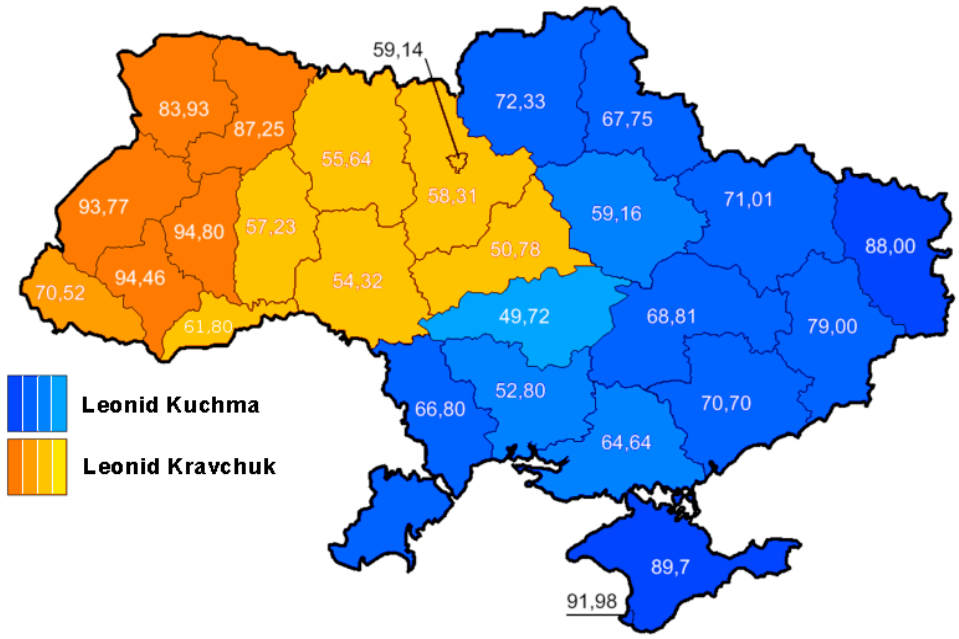

Figure 4. Regional differences in the results of the 1994 presidential election in Ukraine[6]

[5] Source: Young 2015 (based on 2001 Census data).
[6] Source: Babych 2009.

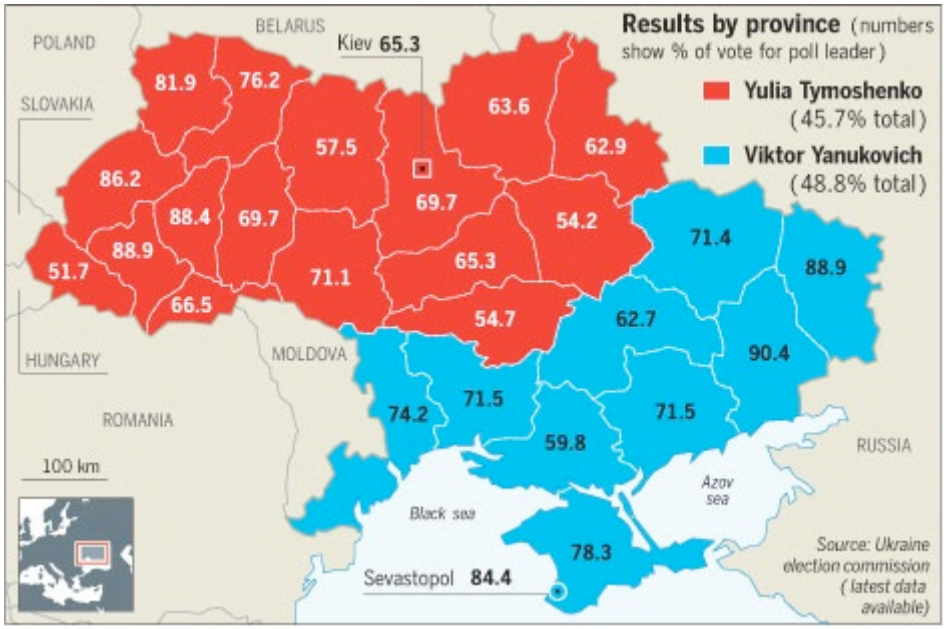

Figure 5. Regional differences in the results of the 2010 presidential election in Ukraine[7]

The analysis ends with Table 3 below, which offers a summary of the impact of development (urbanization), language, and region in the second round of the four presidential elections conducted in Ukraine from 1994 to 2010. The table presents both the unstandardized coefficients (b's) and standardized coefficients (β's) of linear regression. Taking into account the very special circumstances in which the 2014 presidential elections were conducted[8], I decided against including them in the analysis.

Table 3. Urbanization, language, region and voting in Ukrainian presidential elections

	Kuchma 1994	*Kuchma 1999*	*Yanukovych 2004*	*Yanukovych 2010*
Urbanization (b)	−0.22	0.37**	0.01	−0.08
Urbanization (β)	−0.138	0.361	0.004	−0.046
Russian-speaking (b)	0.64***	−0.13	0.97***	0.82***
Russian-speaking (β)	0.645	−0.201	0.866	0.832
West (b)	−34.9***	39.9***	−10.2	−13.9*
West (β)	−0.579	1.027	−0.150	−0.232
Adjusted R2	0.87	0.87	0.88	0.84

*p = 0.05
**p = 0.01
***p = 0.001

The results in Table 3 show that, with the notable exception of 1999, urbanization did not have, in fact, much of an impact. While region had a large impact in the 1990's, its impact has become less prominent in recent elections. As suspected, not only that language has the largest impact overall; in recent years, it has become almost the *sole* determinant of election results. For 2004, the bivariate regression using language (Russian) as the only predictor of the vote for Yanukovych has an adjusted R2 of 0.87, while in 2010, the adjusted R2 of the similar regression is "only" 0.84. While such results may be interesting, even exciting, for a neutral researcher, they do not bode well for the future of a country – which is, in fact, exactly what has happened.

[7] Source: Wagstyl & Olearchyk 2010.
[8] In essence, I think conducting an analysis of the 2014 election using the same framework as for the elections that preceded it would have been problematic because voting in some of the most important regions was either impossible (Crimea) or very difficult (Eastern Ukraine).

CONCLUSION

The longitudinal analysis of structural determinants of voting behaviour in Ukrainian presidential elections provide students who use statistics for studying social phenomena some useful lessons about the role of control variables, of the importance of not putting too much weight on any single cross-sectional analysis, and the substantive lesson of knowing well your cases. But it is equally useful in terms of lessons for political elites and institutional designers. It is unclear, for instance, to what extent majoritarian, winner-take-all institutions such as a strong, directly elected president had a role in the events that are unfolding right now in Ukraine – but they clearly did not help. Statistics and social research cannot, in and by themselves, make or break a democracy, but they can be helpful in highlighting problems. After that, it is the job of decision-makers to pay attention – or not. If they chose the latter, they do so at their own peril, but much more importantly, of their countries as well.

My participation at the IASE 2016 Roundtable Conference was supported by a joint grant from the World Bank Trust Fund for Statistical Capacity Building, ISI, and IASE.

REFERENCES

Babych, V. (2009). Ukraine presidential elections 1994, second round (Вибори 1994.png). <https://commons.wikimedia.org/wiki/File:%D0%92%D0%B8%D0%B1%D0%BE%D1%80%D0%B8_1994.png>. Accessed June 5, 2016.

Handelman, H. (2006). *The Challenge of Third World Development*, 4th ed. Prentice Hall.

Landman, T. (2003). *Issues and Methods in Comparative Politics: An Introduction*. London: Routledge.

Levitt, S.D. & Dubner, S.J. (2006). *Freakonomics*. London: Penguin.

Lipset, S.M. (1960). *Political Man: The Social Bases of Politics*. Garden City, New York: Anchor Books.

Roper, S. & Fesnic, F. (2003). Historical Legacies and Their Impact on Post-Communist Voting Behavior. *Europe-Asia Studies* 55, 1, 119-131.

Wagstyl, S. & Olearchyk, R. (2010). Yanukovich poised for victory in Ukraine. Financial Times (February 8). <http://www.ft.com/cms/s/0/d5a53eb2-14df-11df-8f1d-00144feab49a.html#axzz4AjCS5nsf>. Accessed June 5, 2016.

Young, T. (2015. 10 maps that explain Ukraine's struggle for independence. *Brookings Now* blog (May 21). <http://www.brookings.edu/blogs/brookings-now/posts/2015/05/21-ukraine-maps>. Accessed June 5, 2016.

DATASET (Ukraine):
<https://db.tt/iX0Q2QhO>
<https://dl.dropboxusercontent.com/u/19532654/ukraine_data_2010.sav>