

Spatial spillover effects associated with the Republicans' growing support

Tommaso Rondani

1 Introduction

The 2024 US presidential elections saw an outcome which had been considered unlikely by most analysts. In particular, incumbent Vice President Kamala Harris achieved the unenviable result of not being able to outperform in any county the former Democrat candidate Joe Biden. Therefore, it would be of interest to understand some of the factors that may affect an election at the county level and, most importantly, their spillover effect onto neighbouring counties.

As a matter of fact, [Süm24] identifies a large number of contextual factors based on aggregate data. One of the most obvious factors is the ethnic composition of the area, with the proportion of immigrants and its variation over time. [Jan+19] shows that the relationship between the proportion of non-Western immigrants and the electoral results of radical right parties is linear at macro levels. A rapid change in the ethnic proportion is also likely to increase the support for the radical right, as described in [Kau17].

Moreover, a large quantity of studies cited in [Süm24] report that the level of economic development of the area under study – such as median income, unemployment rate, and level of education – does not show to have a clear effect on the performance of radical right-wing parties during elections.

The analysis performed by [Süm24] revolves around understanding the magnitude of the halo effect, which is also called indirect effect in spatial

modelling. In other words, [Süm24] attempts to quantify the effect originated by the characteristic of neighbouring areas. In this scenario, at the neighbourhood level, an increase in the presence of immigrants and an increase in the ethnic diversity are both positively associated to the support for radical right-wing parties, as is remarked in [MG20]. It is still unclear whether this relation is still valid at a macro level.

Furthermore, [Süm24] underlines the importance of social psychological effects in the growing support for radical right-wing parties. As a matter of fact, especially at the micro level, individuals' preferences are influenced by the neighbours' expected reaction, so the presence of social pressure and social resonance could be of interest. Nevertheless, as it is likely to have a lower effect at the county level, the understanding of this phenomenon is outside the scope of this analysis.

The aim of this paper is the replication of the analysis done by [Süm24] in the context of the 2024 presidential election in the US at county level. Therefore, the percentage of votes obtained by the Republican party in each county is modelled through the median income – expressed in thousands of dollars –, the percentage of low-educated people – measured as the percentage of people who have not finished high school –, the unemployment rate, the percentage of immigrants, and the variation in the percentage of immigrants in four years. For the purpose of this paper, due to its electoral programme, considering the Republican party as a radical right-wing party appears to be a fair equivalence, as also stated by [MG20]. The educated guess of the author is that the analysis is expected not to perform well due to the intrinsic differences between the population of Sweden and the US.

The outline of this paper is as follows. First, the data used for the analysis is thoroughly described, citing its sources. Then, following [Elh10], the most suitable spatial regression model is fit to the data. Eventually, the results obtained are discussed with respect to the research question.

Variable	Min	1st Quartile	Median	Mean	3rd Quartile	Max
Republican_votes	6.47	57.95	70.05	66.74	79.03	95.90
Median_income	25.43	54.89	63.50	65.94	73.29	178.71
Non_employed	0.000	3.300	4.500	4.716	5.800	31.100
Low_educated	0.00	7.84	10.52	11.62	14.63	55.10
Immigrants_23	0.000	0.100	0.200	0.316	0.400	13.900
Delta	-8.800	-0.100	0.000	-0.02167	0.100	9.600

Table 1: Exploratory preliminary analysis

2 Description of data

The data used in this analysis is collected from three different sources. First, the electoral results of the Republican Party – in terms of percentage of total votes – is obtained via [Pre24] through Wikipedia, as it was more feasible to scrape. However, since its results are available only at the state level, Alaska is not considered as subdivided into counties, but as a whole.

Data related to the aforementioned independent variables is publicly available at [Bur23a], which contains the results of the American Community Survey of 2023. The 5-year version has been used, since it includes the most reliable data for all counties in the analysed year.

Eventually, the geographical data is collected from [Bur23b]. In particular, the geometry of the counties is contained in a shapefile where the projection used is NAD83 (or EPSG:4269), which utilises the North American Datum of 1983. For the reason explained before, the counties in Alaska are joined together into a *multipolygon*.

The main descriptive statistics for the variables in the dataset used in this analysis are reported in Table 1. It is worth recalling that *median income* is measured in thousands of dollars, whereas the other covariates are percentages. *Delta* represents the variation of the percentage of immigrants between 2019 and 2023.

The large amount of datasets collected are – at the beginning – cleaned,

as [Bur23a] provides them in a scheme which is odd and difficult to process. Then, the necessary data for the analysis is selected and, eventually, the cleaned datasets are joined together with the dataframe with the geometry of the counties.

3 Data analysis

In order to display the presence of spatial autocorrelation, a Moran's I test has been performed. Due to the fact that the test statistic shows a value close to 0.60, it is clear that there is a spatial pattern in the distribution of votes of the Republican party. In other words, similar electoral results may be found in contiguous counties. Comparable results may be inferred from the Moran's I test applied to the residuals of a linear regression model.

The algorithm devised by [Elh10] is followed in order to retrieve the optimal model. First, a linear model based on OLS estimation has been fit to the data. Then, a Lagrange multiplier test has been conducted in order to verify the presence of spatial dependence.

As both null hypotheses on both the spatial error and the spatial lag are rejected, a Spatial Durbin Model has been used and a Likelihood-ratio test has been performed to assess whether the SDM may be reduced to a Spatial Autoregressive Model or to a Spatial Error Model. However, since the null hypotheses on the restriction of the parameters of the SDM are rejected, it has been found that the SDM is the model that is most suitable to the data. A simple comparison of the AICs of the considered models would have been sufficient to reach the same conclusions. Nevertheless, the model is not able to capture the spatial autocorrelation, since there is a significant residual autocorrelation, as was expected beforehand.

Table 2 clearly shows that all direct regressors are significant, whereas only two of the lag regressors are significant. In particular, the two lag regressors related to immigration are affected by the fact that it was unfeasible to remove Western immigrants.

Variable	Estimate	Std. Error	z value	Pr(> z)
Median_income	0.085728	0.016634	5.1539	2.551e-07
Non_employed	-1.353005	0.089312	-15.1492	< 2.2e-16
Low_educated	0.530308	0.047678	11.1226	< 2.2e-16
Immigrants_23	-4.073476	0.357811	-11.3844	< 2.2e-16
Delta	3.055774	0.391649	7.8023	5.995e-15
lag.Median_income	-0.237099	0.020257	-11.7043	< 2.2e-16
lag.Non_employed	-0.253884	0.150414	-1.6879	0.091430
lag.Low_educated	-0.188608	0.069842	-2.7005	0.006924
lag.Immigrants_23	-0.716226	0.717920	-0.9976	0.318454
lag.Delta	-0.934152	0.824705	-1.1327	0.257336

Table 2: SDM coefficients and statistical significance

Variable	Direct	Indirect	Total
Median_income	0.0498	-0.5442	-0.4944
Non_employed	-1.5943	-3.6543	-5.2485
Low_educated	0.5666	0.5495	1.1161
Immigrants_23	-4.7901	-10.8544	-15.6445
Delta	3.2957	3.6341	6.9298

Table 3: Impact measures of the SDM

The magnitude of the direct, indirect, and total effect for the SDM is reported in Table 3. On the one hand, as expected, there is a positive direct effect for the median income, the percentage of low-educated individuals, and for the variation of immigrants. Nevertheless, it is surprising to discover that the unemployment rate and the percentage of immigrants appear to have a negative direct effect. On the other hand, the magnitude of the indirect effect is similar to the direct one for the percentage of low-educated individuals and the variation of immigrants, whereas direct and indirect effect are of opposite sign for *median income*. Moreover, unemployment rate and percentage of immigrants have, once again, a negative indirect effect. Some of these results appear counterintuitive and will be explained in detail later.

A Local Moran’s I test has been conducted as well, but the result has been deemed non-informative, since all of the influential counties are of the *low-low* type. The small number of influential counties is caused by the fact that the p-value adjustment is obtained via the Bonferroni method, which is heavily affected by the number of counties considered.

4 Discussions and conclusions

As mentioned in [MG20], the fact that immigration negatively affects the electoral results of the Republican party may be explained noticing that immigrants are less likely to vote for radical right-wing parties, regardless of their ethnicity. Moreover, it is worth stating that immigrants tend to settle in an area where the inhabitants share the same values as them. Therefore, it is even less surprising to find a strong negative indirect effect. Nevertheless, the assumption of a linear relationship between immigration and support for the Republican party may be a restriction and more analyses are required. For instance, a quadratic relationship might be able to capture a larger amount of spatial correlation.

It is also important to note that the indirect effect linked to immigration has a larger magnitude than the direct effect. In other words, the votes

obtained by the Republican party in a given county are negatively affected more by a higher percentage of immigrants in a neighbouring county rather than the percentage in the county itself.

Furthermore, the variation of the percentage of immigrants carries a similar magnitude for both direct and indirect effect. This result is unsurprising and may originate from the *group threat* theory mentioned both in [MG20] and [Süm24]. However, [Kau17] states that such threat is likely to fade over time.

Another unsurprising result is offered by the similar effects – in magnitude and direction – provided by the percentage of low-educated people. Rather than at a macro level, this fact is more likely to influence voting behaviour at a micro level. As a matter of fact, a valid explanation of this phenomenon may be the aforementioned social pressure and social resonance. It would also be of interest to verify the presence of correlation of this covariate with unemployment rate and median income.

One of counterintuitive results which was mentioned before is the negative effect – both direct and indirect – related to the unemployment rate. One would expect that a county with a high unemployment rate would be poorer and more prone to favour the Republican party. Nevertheless, college students and young people, who are likely to be left-leaning, tend to be a large component of the unemployed people.

Median income is the only covariate in which the direct effect and the indirect effect have opposite signs. Given their magnitudes, one may easily induce that having a wealthier neighbouring county negatively affects the support for the Republican party. However, due to the large inequality present in the US, it might have been more interesting to select – as is done by [MG20] – the percentage of people whose income is lower than 25 and greater than 100 thousands dollars.

The indirect – or halo – effects reported in Table 3 are either comparable or significantly greater in magnitude with respect to the direct effect. In particular, the halo effect induced by the percentage of immigrants reaches

the highest magnitude and the highest fraction of the total effect. Therefore, among the considered variables, immigration is the most influential especially when it occurs in neighbouring counties. It is also worth noting that there exists only one covariate for which the effects have opposite signs, but, since its direct effect is close to zero, it does not represent a counterexample.

4.1 Limitations

It is important to note that the fact that some factors have not been considered in the previous analysis may have had a non-negligible influence on the outcome. In particular, there is a difference in the ethnic homogeneity of the population in the US, as – according to [Bur23a] – it is a mixture of white, black, hispanic, asian, and bi-racial people. Therefore, a dishomogeneity index, perhaps based on the Gini index, may be added to the analysis. This covariate would be expected to improve the performance of the model, since [Süm24] notes that ethnic homogeneity in a county increases support for radical right-wing parties. For instance, [MG20] utilises the percentage of non-Hispanic white people and the percentage of non-English speakers as covariates.

Moreover, due to the lack of data, it would have been interesting to add the number of refugees, which is likely to have made an impact on the electoral result, in the covariate related to immigration. Similar reasonings may be used for data concerning criminality, especially micro-criminality. Also the fact that a discrete amount of immigrants prefer moving to an area where a similar culture and similar political opinions may be found is likely to have had an effect on the election.

Another issue related to immigration is that considering only non-Western immigrants may be more theoretically coherent, as they are more likely to come from countries with a lower GDP and with largely different culture and traditions. However, the assumption of an equivalence between non-Western and non-white immigrants would only represent a rough – and likely incorrect – approximation and the non-clarity of the data has provided a sufficient

reason not to pursue it.

A further limitation of this analysis is related to voter turnout, due to the fact that this specific aspect has not been considered, as it is not reliable given the peculiar system of registering to be able to cast the vote.

In conclusion, as expected, the analysis performed in this paper does not capture entirely the spatial autocorrelation of the data. As a matter of fact, the idea behind the replication of the study presented in [Süm24] was to manifest the fact that it is largely unlikely to find a common general model which can be used in different countries. Taking into account the intrinsic characteristics of a population and its cultural peculiarities is mandatory to perform a reliable analysis.

Code

The code is available at <https://github.com/tommyrond/Spatial-spill-over-effect-in-the-US>.

References

- [Bur23a] United States Census Bureau. *American Community Survey data*. 2023. URL: <https://data.census.gov/>.
- [Bur23b] United States Census Bureau. *Cartographic Boundary Files*. 2023. URL: <https://www.census.gov/geographies/mapping-files/time-series/geo/cartographic-boundary.html>.
- [Elh10] J. Paul Elhorst. ‘Applied Spatial Econometrics: Raising the Bar’. In: *Spatial Economic Analysis* (2010).
- [Jan+19] Heleen J. Janssen, Maarten van Ham, Tom Kleinepier and Jaap Nieuwenhuis. ‘A Micro-Scale Approach to Ethnic Minority Concentration in the Residential Environment and Voting for the Radical Right in The Netherlands’. In: *European Sociological Review* (2019).

- [Kau17] Eric Kaufmann. ‘Levels or changes?: Ethnic context, immigration and the UK Independence Party vote’. In: *Electoral Studies* (2017).
- [MG20] Jennifer A. Miller and Tony H. Grubestic. ‘A Spatial Exploration of the Halo Effect in the 2016 U.S. Presidential Election’. In: *Annals of the American Association of Geographers* (2020).
- [Pre24] Associated Press. *2024 Presidential Election Results*. 2024. URL: <https://apnews.com/projects/election-results-2024/?office=P>.
- [Süm24] Dávid Sümeghy. ‘Spatial spillover effects associated with the Sweden Democrats’ growing support’. In: *Scandinavian Political Studies* (2024).