Levi, Eugenio; Patriarca, Fabrizio

Working Paper

An exploratory study of populism: the municipality-level predictors of electoral outcomes in Italy

GLO Discussion Paper, No. 430

Provided in Cooperation with:
Global Labor Organization (GLO)

Suggested Citation: Levi, Eugenio; Patriarca, Fabrizio (2019) : An exploratory study of populism: the municipality-level predictors of electoral outcomes in Italy, GLO Discussion Paper, No. 430, Global Labor Organization (GLO), Essen

This Version is available at:
https://hdl.handle.net/10419/208373

An exploratory study of populism: the municipality-level predictors of electoral outcomes in Italy

Eugenio Levi* and Fabrizio Patriarca**

Abstract

We present an exploratory machine learning analysis of populist votes at municipality level in the 2018 Italian general elections, in which populist parties gained almost 50% of the votes. Starting from a comprehensive set of local characteristics, we use an algorithm based on BIC to obtain a reduced set of predictors for each of the two populist parties (Five-Star Movement and Lega) and the two traditional ones (Democratic Party and Forza Italia). Differences and similarities between the sets of predictors further provide evidence on 1) heterogeneity in populisms, 2) if this heterogeneity is related to the traditional left/right divide. The Five-Star Movement is stronger in larger and unsafer municipalities, where people are younger, more unemployed and work more in services. On the contrary, Lega thrives in smaller and safer municipalities, where people are less educated and employed more in manufacturing and commerce. These differences do not correspond to differences between the Democratic Party and Forza Italia, providing evidence that heterogeneity in populisms does not correspond to a left/right divide. As robustness tests, we use an alternative machine learning technique (lasso) and apply our predictions to France as to confront them with candidates' actual votes in 2017 presidential elections.

Keywords: Voting, Populism, Economic insecurity, Political Economy

JEL Codes: D72, F52, G01, J15, O33, Z13

* Free University of Bozen-Bolzano, Faculty of Economics and Management, eugenio.levi@unibz.it

** University of Modena and Reggio Emilia, Department of Economics "Marco Biagi"

Introduction

There is a group of rising stars on the world's political stage who do not object to being called populist. Among them, the US President Donald Trump, the Italian Interior Minister Matteo Salvini, the President of the French National Rally Marine Le Pen, Hungary's Prime Minister Viktor Orban. What they have in common is that they all use an anti-elite rhetoric and exploit the decline of traditional players in the political arena[1]. However, the more we move towards an empirical observation, the more populism appears as a multi-faceted phenomenon that is very hard to define. We can find both left-oriented and right-oriented populists, socially conservative and socially liberal, pro-taxes and against taxes, pro-environment and environmentally-neutral populists, and so on (see Kaltwasser et al., 2017, for a full account on different populist parties). Therefore, understanding the underlying factors of the so-called rise of populism in mature democracies is still an open issue and a difficult one.

Many papers in political science and political economy have addressed the question by testing different theories on populism. These theories encompass a clash between 'winners' and 'losers' of globalization, resentment over inequalities, a 'cultural backlash', expressive or protest vote, decay in the quality of politicians. Our approach is different. Because the phenomenon is so multi-faceted, we think an exploratory analysis can shed some light into populism and its consequences on the political landscape. Therefore, we look at the 'where' of populism and we go 'from the particular to the general' to test the predictive power of theories on populism: we use data-driven methods over a comprehensive set of local characteristics to find the municipality-level predictors of populism. Moreover, the decline of traditional parties is not often studied alongside the emergence of populist parties. However, these two phenomena are not just two sides of the same coin. What is needed is to understand whether populist parties are simply capturing an electorate on the run from corresponding 'establishment' parties on the traditional political spectrum – a sort of 'revolving doors' effect -, or whether they are building up an electorate of their own by changing the political landscape altogether – a sort of 're-grouping' effect.

[1] Populism historically emerged as a political movement that idealized rural life in Russia at the end of the 19th century. See Mudde (2004) or Encyclopaedia Britannica for a definition of populism. It is interesting to note that the same definition of populism varies across time and is quite different among these sources. For example, the first populist parties were very ambiguous in terms of their opinion on democracy, while today they mainly claim that they want to defend democracy against financial elites, immigration, etc. What all these definitions agree on is that populists share a rhetoric that champions the people, conceived as an incorruptible unity, against the elites.

Our case study is the 2018 general election in Italy that was held on the 4th of March of that year. In these elections, not only the populist parties boosted their votes, going from 25,55% to 32% for Five-Stars Movement (M5S) and from 4,08% to 17% for Lega, but they were then able to form a coalition government together. Italy is a key member of the European Union – one of the biggest economies and a founding member. It is also the main country whose future is at stake in the European political debate, and is indeed considered historically as a forerunner of European political turmoil, as it was for the rise of Fascism in the 1920s. In Italy there are two distinct and more relevant parties which can be classified as populist, the Five-Stars Movement (left-oriented) and Lega (right-oriented), and two main 'establishment' parties, the Democratic Party (left-oriented) and Forza Italia (right-oriented). This permits us to compare the drivers of populism along two axes: heterogeneity in populisms and populism vs. establishment.

We use a dataset of 36 predictors at municipality level over a total of 7,908 observations. Municipalities range from very small towns, with only a few hundred inhabitants, to very large cities, like Rome, Milan and Naples that have a population that ranges between 1 and 3 million. These predictors cover a wide range of socio-economic characteristics and can be grouped in 4 categories: socio-demographic measures that capture the social environment; variables capturing the presence of immigrants of different ethnicities[2]; measures capturing the economic structure of the municipality; and variables describing quality of life and welfare based on measures on population density and public services.

The methodology follows a recent paper by Sascha Becker et al. (2017) on Brexit. We use a machine-learning method based on BIC to select which predictors best explain the share of votes for populist and traditional parties in each category of characteristics, then we run a final round over these selected variables to identify the best overall models. In this way, we identify the key features for the electoral base of each of the four main Italian parties in 2018 and their dynamics relative to the previous election. It is to note that the beta coefficients of our final regressions cannot be interpreted in relation to internal validity, but they are useful for prediction as they have better out-of-sample model fit than usual OLS models. We exploit this to test the extent to which our models are able to catch more general ongoing political processes across countries. More specifically, we project the electoral base of all the 4 parties in France at province level and analyse the model's potential to explain the results of the

[2] Google Trends data shows very neatly that immigration was a hot topic compared to other topics in the debate that preceded the elections. In fact, an immigrant was killed in Macerata on the 3rd of February by an extreme-right activist, giving rise to a huge debate on the topic of immigration.

French 2017 Presidential election. We also carry a supplementary analysis at the regional level, as municipality characteristics highly overlap with regional ones.

From our results, the populists' electoral bases have different dimensions from the ones traditionally marking the divide between left- and right- wing. Such new dimensions are not common to M5S and Lega, and neither a new divide nor similarities emerge, providing evidence for a 're-grouping' effect. Besides, two different and clear patterns emerge. Lega is mainly characterized along economic dimensions, including those concerning inequalities and globalization, having an advantage in industrial and commercial areas, in richer municipalities and with an electorate of lower education level. The distribution of M5S's electoral base reflects much more the effect of a generational conflict (over 65 vs. the rest), with many features indicating a detachment from the institutional setting – lower quality of public services, higher rates of crime and unemployment. Along these unique and parallel dimensions, they both increase their support, while the traditional divide vanishes since the crisis of the two traditional parties Democratic Party and Forza Italia concerns exactly the core of their electoral base. All variables related to immigration, surprisingly, are not relevant enough to make it to the last model for any of the parties we have considered. The results on France strongly confirm the robustness of the model and suggest further interpretation and new directions for research.

We want to stress that we do not provide a causal interpretation to our results and that we do not directly investigate political attitudes of the voters. With our analysis, we cannot provide definitive evidence in support of any theory on populism. In particular, we deal with the demand side of populism only, as we do not provide any specific argument as to why populists insist on specific narrative cleavages or on the political economy of their policy choices. Even taking these limitations into account, our paper still derives interesting propositions to be tested in future research.

This paper fits into an emerging literature on populism in political economy. In particular, Rodrik (2018) provides an interesting theoretical framework and a valid literature review over the rise of populism. He claims that globalization explains most of the votes for populist parties while salient political cleavages – income or culture – determine the political orientation of the chosen populist party – left-oriented or right-oriented, respectively. There are some empirical papers that support these theories. Between them, Autor et al. (2016) stress that penetration of imports from China is a decisive factor to explain the increasing polarization in US politics. The outcome of the Brexit referendum and votes for populist and far-right parties across Europe were explained across the same lines by Colantone and Stanig

(2018a, 2018b, 2019), by Becker et al. (2017) and by Alabrese et al. (2019). A few papers explore more in general the correlation between economic insecurity and populist votes. Using 2002-2014 European Social Survey data, Guiso et al. (2017) find that an index based on the principal component of recent unemployment, income distress, and having low-skills while working in the low-tech manufacturing industry underlies populism across Europe. Bloise et al. (2019) also point to economic factors and inequalities. Algan et al. (2017) find the same evidence in connection with the variations in unemployment, which in their data seems to go in tandem with trust (or lack thereof) towards national and European institutions. Another source of economic insecurity is the recent rise in automation in production processes (Anelli et al., 2019).

Some political scientists claim that culture matters more than economic outcomes in explaining the rise of populism. Norris and Inglehart (2019) put forward a 'cultural backlash' theory, according to which votes for populist parties come from old and white people reacting to the affirmation of progressive cultural values within society. In the political sciences other interesting explanations are put forward. For example, Cramer's (2016) field study showed how in rural Wisconsin support for Donald Trump was generated by resentment over policies that favoured ethnic minorities in big cities. In this sense, resentment derives from a feeling of detachment from traditional political messages: people in low population density areas just feel abandoned. This is consistent with a theoretical model and some of the evidence provided by Di Tella and Rotemberg (2018) that shows how people living in rural areas and low educated people in urban areas may choose an incompetent leader over a competent one because of betrayal aversion. Using survey data, several other papers in political science find that a cultural reaction to progressive values, economic resentment and political disenchantment all seem to matter in explaining the emergence of populism across different countries (Mudde, 2007, Oesch, 2008, Arzheimer, 2009, Akkermann et al., 2014, Spruyt et al., 2016, Inglehart and Norris, 2016, Margalit, 2019).

The paper is organized as follows. Section I presents the methodology, Section II gives a detailed overview of the data, specifying what it represents in political terms, Section III gives the results and Section IV discusses and concludes.

I. METHODOLOGY

As our aim is to analyse the municipality-level drivers of votes for M5S, Lega, Democratic Party and Forza Italia in a comprehensive way and not to provide causal explanations, we focus on the cross-sectional variations only. This is because it includes time-invariant variables along time-variant ones and so long-term/structural drivers to the rise of populism. On the contrary, focusing on differences would have highlighted the short-run determinants[3]. We have a wide range of variables – 36 – divided in 4 groups and we want to select the variables within each group that are more relevant in order to run a final model in which only the most relevant drivers are included.

A. Best selection procedure

We exploit machine-learning methods to select which municipality-level drivers best predict votes for the 4 main Italian parties at the 2018 general elections and the differences in votes with the previous 2013 general elections. Our methodology closely follows that by Becker et al. (2017)[4]. As best selection procedure, we use an algorithm that runs OLS regressions on all possible combinations of predictors k in each group, computes the BIC (Bayesian Information Criterion) and selects the predictors k^* whose associated OLS has the lowest possible BIC[5]. Finally, considering the selected predictors for each group, we run a final race among the k^* predictors that give us a best overall model for each of the considered parties characterized by k^{**} predictors. The information criterion here plays the key role of making the selection between predictors:

$$BIC = -2ln(L) + k\ ln(n)$$

[3] For example, when considering the economic structure of a municipality, which is very stable over time, being able to use as data occupational structure by industry becomes relevant to our purpose. On a more practical note, given that census data collection in Italy only happens once every 10 years, the amount of data that is available on a cross-sectional level is much higher than over repeated time periods.

[4] Of course, we acknowledge that, as the field of machine-learning methods is rapidly expanding, there are many feasible selection procedures (see Mullainathan and Spiess, 2017, for an overview). There are older techniques, such as stepwise regression, and newer ones, such as Lasso/Ridge regression, random forests, etc. We chose BIC because it can be computed analytically without resorting to any approximation and it gives the smaller amount of 'ad hoc' parameters' choices to the researcher. Anyway, we also provide results using Lasso regressions, as we will explain later in more detail.

[5] We prefer BIC over AIC (Akaike Information Criterion) because it places more weight on the penalty term, therefore providing sparser models.

In which *n* is the number of observations, *k* the number of parameters and *L* is the maximized value of the likelihood function of the model. This criterion provides a way to identify, across a finite set of models and provided that *k* is sufficiently low to make it computationally feasible, the optimal combination of variables over a trade-off between efficiency of the model, represented by *L*, and its parsimony, represented by *k*.

B. Lasso regression

An alternative best selection procedure we will use is Lasso regression (Tibshirani, 1996)[6]. The estimator of a Lasso minimizes the residual sum of squares subject to a constraint on the sum of the absolute value of coefficients estimates. It corresponds at introducing a penalty term λ in the unconstrained minimization problem:

$$min_{\beta_0,\ \boldsymbol{\beta}} \left\{ \frac{1}{N} \sum_{i=1}^{N} (y_i - \beta_0 - \boldsymbol{x}_i \boldsymbol{\beta})^2 + \lambda \sum_{j=1}^{k} |\beta_j| \right\}$$

The penalty term drives to zero some of the *k* coefficients. To determine λ, we decided to resort to cross-validation: the dataset is split into a number of different samples – in our case, 100 – and λ is chosen as to maximize out-of-sample prediction.

The difference among the BIC and the Lasso models can be exploited by deriving the predictors that are common to the two models and obtaining an even smaller subset of relevant drivers of populism. We use this subset also to see if our models capture an underlying process on populism across countries. In particular, we record the values of the coefficients associated to the k^{**} predictors in Italy and predict the votes for the 'corresponding' parties in France. It is like having predicted votes for M5S, Lega, Democratic Party and Forza Italia in other countries, as if they were part of the electoral competition there, except that we assume that these parties have their actual names, so for the French 2017 presidential election Jean-Luc Melanchon, Marine Le Pen, Emmanuel Macron

[6] The good thing about Lasso is that it is computationally not so demanding, so that we can run it over our full set of predictors and obtain results in a few minutes. This is because the minimization problem has a convex form, so coordinate descent algorithms (Fu, 1998) can easily solve it.

and Francois Fillon, respectively. Then, we compute correlations between these predictions and the actual votes and we find the goodness of fit of these models using F-statistics.

II. DATA

As we are using model selection methods to provide our results, the choice of the data is not irrelevant. So, we will define the predictors, present their sources, and briefly discuss the theoretical hypotheses that we associate to them also given the parties' political stance on the matter (see Table 1 for an overview of the variables and their source).

<Table 1>

Our units of observation are municipalities. At the time of the 2018 elections, there were 7,954 municipalities. We excluded those in the Valle d'Aosta region, because of a different electoral system,[7] therefore we have a final total of 7,908 observations.
We will start by discussing our main outcome variables, i.e. votes for M5S, Lega, Forza Italia and Democratic Party; then we will move on to analysing in more detail the municipality-level drivers we are considering. We decided to group these drivers in 4 categories:

a) Socio-demographics characteristics;
b) Presence of immigrants;
c) Economic characteristics;
d) Measures related to welfare and to quality of life.

We add to these predictors, as fixed variables, the 19 regional dummies.

[7] Given the small size of the region, Valle d'Aosta adopts a first-past-the-post system, while for the rest of Italy the electoral law is a mixture between first-past-the-post and proportional systems. See https://www.bloomberg.com/news/articles/2018-03-02/italy-is-rolling-out-anew-electoral-system-here-s-how-it-works for a description of the 2017 electoral law that was in use during the 2018 elections. Overall, the period between the choice of the candidates and the electoral date was so short (29th of January to 4th of March) and the constituencies so large that the proportional part was overtly salient in the election.

A. Votes

M5S, Lega, Democratic Party and Forza Italia are by far the biggest parties in Italy. Taken together, they collected more than 80% of the votes at the 2018 general election (Figure 1). We provide here some arguments to justify how we position these parties on the scales of populism and left vs. right.

M5S was founded in 2007. Its rallying cry, especially in the first few years of its life, was "Vaffanculo" – Italian for "Fuck off" – meaning that the movement aimed at getting rid of the traditional political elites[8]. It is very hard to assess whether M5S is left- or right-oriented, as they do not take a clear stance on many issues, like immigration, European integration, neo-fascism, etc (Corbetta and Vignati, 2013, Diamanti, 2014). However, given that they advocate the need for more participatory democracy, stronger social and environmental policies, such as a minimum income scheme, most commentators, including the Chapel Hill Survey, position them on the left side of the political spectrum (Inglehart and Norris, 2016). In 2013 they first took part in a general election receiving 25,5% of the votes, then increased them to 32% in 2018, thus becoming the main Italian party.

Lega is the oldest existing party (in terms of direct continuity), as it was founded in 1989. From 2013, under the leadership of Matteo Salvini it became a nationalistic, anti-tax and anti-immigrant far-right party under the slogan "Italian First" and proposing a flat tax reform. In 2016 Lega founded a new online magazine explicitly called "The Populist". After an abrupt decline in votes from 2008 to 2013 (from 8,30% to 4%), Lega quite unexpectedly became the biggest party in the centre-right coalition in 2018 with 17% of the votes.

Democratic Party (PD) is a traditional left-wing party, founded in 2007 from a merge between the two main centre-left parties. At the 2018 general elections, it was the incumbent party. It is part of the European Socialist Party and has always been considered as the Italian equivalent of a social-democratic party. From its foundation onwards, it has always been losing votes in general elections, from 37,5% in 2008 to 25,4% in 2013 to 18% in 2018.

< Figure 1 >

Forza Italia was founded in 1993 by Silvio Berlusconi, who has always been the leader of the party under its different names (Forza Italia, Popolo delle Libertà and then Forza Italia again).

[8] After 2012 researches on corruption on Google in Italy, compared to all other searches, almost doubled.

Even though some commentators consider it as a prototype of a modern populist party (among them, Diamanti and Lazar, 2018), it is a traditional right-wing party in many ways. It is part of the European Popular Party, it is against taxes and the welfare state, quite liberal but with a penchant for conservative social policies on issues that are sensitive for the Catholic electorate. Its decline in the past 10 years, coinciding with the personal troubles of Berlusconi himself, has been relentless, as it moved from 46,8% in 2008 to 29,2% in 2013 to 14% in 2018.

B. Socio-demographic characteristics

Demographic characteristics. If cultural theories on populism are correct, older and married people should vote more for populist parties (Inglehart and Norris, 2016). Traditional parties have supported liberal social policies over the years, therefore older people and bearers of more traditional values may feel alienated by these parties' stance on policy measures that go in the direction of building a more tolerant society, in matters that concern, for example, gender equality, LGBT rights, or immigration[9]. Another issue that affects old and young people differently is that of pensions, which was a central topic of the 2018 electoral campaign, as both Lega and M5S promised to reduce the retirement age by rolling back the 2011 pensions' reform.

In order to capture these potential features of populism, we consider two variables for age: the percentage of population under 35 years of age and the percentage of population over 65 years of age. We consider as proxies for the relevance of traditional values the percentage of population that is married, the percentage of divorced people over the married ones, and the number of buildings per capita belonging to the Catholic Church that are used for living.

Social characteristics. Political scientists very much stress that populist parties are mostly voted by socially vulnerable people. This is because they may feel more frustrated over their own life and so more inclined to blame the elite and/or an outside enemy for their own condition (Oesch, 2008, Lazar and Diamanti, 2018). Furthermore, education is traditionally considered as a fundamental predictor of the demand for populism. Less educated people usually coincide with low skilled workers, are therefore more exposed to competition by immigrants and automation, and work in sectors that are more exposed to foreign trade

[9] For example, during the last parliamentary term Democratic Party approved with much controversy a law on civil partnerships, chiefly designed to grant rights to gay and lesbian couples, which may have caused a backlash from the electorate closer to the Catholic Church.

(Belanger and Pinard, 1991, Colantone and Stanig, 2018a, Anelli et al., 2019). They may also be less rational in their political choices, more traditional in their values, and less informed over political issues (Weakliem, 2002). Therefore, municipalities experiencing a higher degree of social problems and less education should vote more for Lega and M5S. We consider four predictors for social characteristics: rent share, percentage of crimes, the share of population with secondary education and job commuting.

C. Presence of immigrants

In the 2018 electoral campaign immigration figured as a prominent topic. The debate was sparked by an assassination attempt of a group of immigrants by a deranged far-right activist, an act of revenge for a supposed sexual crime. Lega in particular has focused on the negative consequences of immigration. In concordance, many results in the economic literature find that an increase in the number of immigrants causes an increase in votes for far-right parties in most European countries (Dustman et al., 2018, Halla et al., 2017, Levi et al., 2020), and this holds for Italy too (Barone et al., 2016, Caselli et al., 2019). However, the literature that specifically focus on populism provides mixed results on this, suggesting that migration may be part of a populist narrative more than a determinant of votes at local level (Colantone and Stanig, 2018a).

We both consider the total percentage of immigrants over the total population and percentage of different ethnic immigrant groups: African, from Eastern European countries, from Muslim countries and Chinese. There is indeed evidence that cultural distance increases hostility towards immigrants, so that non-European immigration boost anti-immigrant votes more than European or Anglo-Saxon immigration (Harmon, 2017, Brunner and Khun, 2018).

D. Economic characteristics

While traditionally the divide between political left and right has mainly concerned economic issues, there is increasing evidence showing that economic insecurity and inequalities are at least partly responsible for the emergence of populist parties on both sides of what would typically be the left-right cleavage (Kriesi et al., 2006, Guiso et al., 2017). An appeal to the anonymous and disillusioned worker, who feels threatened by globalization, by technological change and by immigrants, is often part of the rhetoric of populist parties and should point both to sectors that are more exposed to these threats, like manufacturing, basic services and

high-import penetration industries (Colantone and Stanig, 2018), and to vulnerable people either in the job market or because of their income level. On a different note, theories on cultural backlash or detachment seems to suggest that people who are employed in agriculture vote more for populist parties (Inglehart and Norris, 2016, Cramer, 2016). Furthermore, M5S's main proposal on economic issues was a basic income scheme for unemployed people.

We consider features related to economic sectors and to the job market, inequalities and import penetration. For economic sectors, we select the percentage of people working in 4 sectors at municipality level: manufacturing, commerce, basic services and advanced services. The omitted category is agriculture. On features of the job market, we chose the size of the companies and the duration of contracts, unemployment and youth unemployment, and the ratio between employees and self-employed workers. As features related to income, we consider both the average income, within-municipality inequalities (as described by the Gini index) and poverty, which we define as the percentage of people in a municipality earning less than 10,000 Euros per year[10]. Last but not least**,** we also consider the direct role of globalization by using net exports from the rest of the world and from European countries[11].

E. Welfare and quality of life

Even though socio-demographic and socio-economic characteristics are usually considered the best predictors for the emergence of populism, we consider additional predictors related to welfare and to the quality of life. This choice has two reasons. First, we want to have the broadest set of municipality-level features and let the data speak for itself in deciding which ones are populism's drivers. Second, evidence for Rome and Milan and for other countries (the UK, Germany, the USA, etc.) suggests that big cities behave differently from small towns and this could be accounted for by differences in the quality of life or by aggregation effects (Johnston et al., 2007).

We try to capture a big city/small town divide with a set of three predictors: municipality population and two measures of population density. Measures of population density in our

[10] Predictors on income are correlated among themselves in a specific way. From our data municipalities with higher average income are less likely to have a higher percentage of poor people (correlation is -0.88) but there is almost no relation with within-municipality inequalities (correlation is 0.07). However, there seems to be a correlation between within-municipality inequalities and the percentage of poor people (correlation is 0.30).

[11] We try to disentangle a possible predictive effect of this variable from net exports with other European countries. This is because both M5S and Lega, while very ambiguous on the matter, were more inclined to be against European institutions and did not completely rule out the possibility of exiting the Eurozone. Note that net exports to all over the world and to Europe at province level are highly correlated (0.94), especially as the European ones in our data account for 58% of the total net exports.

dataset are the percentage of people living in detached houses over the total population and the percentage of people living in buildings with more than 9 flats. An additional predictor we consider that captures urban features is the index of seismic risk with three levels of risk.
For welfare, we use as predictor the sum of all measures related to unemployment that substitute for work income[12]. Then, the average number of beds in hospital per capita and life expectancy constitute our measures of the quality of health services. Number of kilometres of railroads per square kilometre at province level are our measure of infrastructure quality. We only consider double track railway lines, as usually single lines are considered a sign of bad infrastructure quality[13].

III. RESULTS

In this section, we first analyse our results by outcomes for each party (Table 2 and 3). After having completed our main analysis, we turn to Lasso regression, to predicting French votes and to regional macro-drivers of populism.

< Table 2 >

< Table 3 >

A. Results by party

Socio-demographic and socio-economic variables seem to mostly characterize votes for the 4 key parties in Italian politics. For M5S, in the final model we have 4 sociodemographic, 3 economic and 2 quality of life predictors, for Lega 4 socio-demographic, 10 economic and 1 quality of life predictors, for Democratic Party 5 socio-demographic, 5 economic and 2 quality of life predictors and for Forza Italia 6 socio-demographic, 3 economic and 3 quality of life predictors. The sectorial composition of employment and the population density are the only ones that are relevant for all the parties. Strangely enough, no variable on the presence of immigrants is a good predictor, not even by looking at the different ethnicities. In the first levels of the algorithm immigration from China survived with a negative sign on Lega, which is the opposite evidence than what is usually found in the literature.

[12] *CIGO* and *CIGS* in 2017, *Indennità di Mobilit*à and *NASPI* in 2016 and *ASPI* in 2015

[13] Even though railway lines were mainly built at the beginning of the 20th century, the quality of public trains has improved over the last 20 years so we expect that where there are more railroads people will vote less for populist parties.

First, let us consider Lega and M5S vis-à-vis theoretical hypotheses on populism (see Table 2). Support for Lega comes from less educated and less densely populated municipalities. It gets more votes in municipalities where occupation in commerce and manufacturing is higher but occupation in advanced services is lower, where import penetration is higher if from the rest of the world but lower if from European countries and the size of the firms is smaller. Also, characteristics related to income are good predictors of Lega: municipalities with lower income and higher inequalities are more prone to vote for the nationalistic party. These results are all consistent with theories on populism, as the relevance of lower education points to cultural backlash, the economic sectors, income and import penetration to economic insecurity/globalization, lower population density to political detachment. However, the negative sign of other predictors such as crimes, job commuting, poverty rate and unemployment point in another direction, i.e. Lega performs better in safer municipalities where the job market is more efficient. So, ultimately, evidence for the predictive power of economic insecurity for Lega is mixed.

M5S thrives in municipalities where the percentage of married people, crimes per capita, the share of unemployed and of people working in basic services are higher. This is clearly consistent with theories on economic insecurity and, partly, with the idea that culturally more traditional people are more in favour of M5S. The relevance of health services, as more beds in hospitals per capita correspond to less votes for M5S, is a hint that political detachment may play a role too. However, support is higher among young people and lower among old people, which strongly contradicts theories on cultural backlash or on resentment/detachment. Also the higher percentage of votes in less densely populated municipalities speaks against these theories. Then, predictors related to income, to other economic sectors and to import penetration are all irrelevant for M5S. Interestingly and not consistently with theories on populist votes as protest votes, turnout has more or less the same predictors as votes for M5S (positive on young people, married couples and crime, negative on the elderly).

Two main results come from our analysis in terms of populism vs. traditional parties and of different types of populisms.

First, the sets of variables characterizing M5S and Lega are very different from each other. There do not seem to be common predictors for different populisms. In particular, for Lega the demographic variables and public services are not selected, as opposed to M5S. On the other side, some socio-economic variables as income, inequality, poverty, globalization and

education are selected only for Lega. Besides, for the dimensions selected for both parties the evidence points in opposite directions: coefficients have different signs on unemployment and crime rates, different economic sectorial characterization and type of housing. Based on the 'where' of populism, different populisms emerge. The first one, represented by Lega, prevails in smaller, richer, less-educated, industrial and commercial, and safer municipalities. The second one, represented by M5S, thrives in bigger, younger, jobless and more dangerous municipalities that have more people employed in services.

Second, this difference between populisms does not seem to fit well with a traditional left/right divide as expressed by predictors for Democratic Party and Forza Italy. Lega is no similar and no symmetrical to Forza Italia. The common features of the centre-right coalition are limited to the rate of crime, education and job commuting, all with a negative coefficient. Moreover, variables on import penetration mark a distinction between Lega and all traditional parties.

Equally, M5S is no similar and no symmetrical to Democratic Party. They have opposite features on age, on marriage status, on unemployment and on the quality of health services. However, if we look at crimes per capita, at economic sectors – with a slip for Democratic Party on more advanced services -, and at population density – with a slip for Democratic Party on buildings with more then 9 flats, we find some similarities. M5S has also similarities with Forza Italia on demographic variables. In general, a negative effect of old age and a positive one of unemployment mark the distinction of M5S with both traditional parties.

The only clear difference between populism on one side and 'establishment' parties on the other side rests in how votes are evolving from the 2013 general elections (see Table 3). In fact, predictors for variations in votes for M5S and Lega coincide with those for the levels and the sign of the coefficients are the same. The only differences are education and seismic risk with a negative sign and railroads with a positive sign for M5S and the share of married people with a positive sign for Lega. Overall, it seems that populist parties are boosting their votes in their strongholds. On the contrary, Democratic Party and Forza Italia are losing more votes where the municipality characteristics better predict votes for them, as most of the drivers on levels come here with an opposite sign. This holds for the percentage of old people, advanced services and health services per capita for Democratic Party, for the percentage of married and less educated people, of job commuters, commerce and the extension of the railways for Forza Italia.

B. Lasso regression

In Table 3, we present a comparison between the selected predictors with the best selection procedure and those with a Lasso regression. Reassuringly, most predictors are the same, as the Lasso procedure mainly selects a subgroup of the best selection models' predictors. For M5S, only commerce and beds in hospitals per capita drop out. For Lega, the screen through Lasso is more informative: it leaves in crimes per capita and education as socio-demographic predictors, manufacturing, commerce and advanced services as economic sectors, the Gini index as income predictor, and the percentage of detached houses as indicator of the quality of life. When we discuss the results in the next section and derive some implications for studies on populism, we are mainly going to take into consideration these shared predictors.

< Table 4>

C. Predictions on France

We consider the subset of shared predictors of the 4 parties in Italy – the outcomes of both the best selection procedures and of the Lasso regressions – and predict votes for France at the level of "department", so for 97 data points[14]. There is a good correspondence between the fitted and the actual values. The correlations in Table 4 are positive and significant for most of the parties: they are 0.36 for Lega and Marine Le Pen (1% level), 0.22 for M5S and Jean Luc Melanchon (5% level), 0.55 for Democratic Party and Emmanuel Macron (1% level). Only between Forza Italia and Francois Fillon we don't find any correspondence, as the correlation is only 0.06. Fitted values based on Lega are also significantly negatively correlated with all the other main presidential candidates, Melanchon included, while fitted values based on M5S are not correlated with any other presidential candidate, Le Pen included.

Interestingly, predictions based on Democratic Party significantly positively correlate with Francois Fillon too (0.36 at 1% level), while the actual values of Macron and Fillon are not

[14] We exclude the overseas Department from the list. Data for the 2017 presidential election was collected from the Minister of Internal Affairs, the rest of the predictors were taken from INSEE except from crimes per capita and age, which come from 2016 Eurostat data. In place of railway lines, we use the kilometres of motorways over a defined area, such as calculated by the Ministry of the Environment in 2010.

significantly correlated with each other. The F-tests based on the predicted values all support the existence of this correlation at the same level of significance, confirming that our models have good descriptive power also when it comes to France, supporting our results both on populism vs. establishment and on different populist orientations.

< Table 5 >

D. Regional effects analysis

Regions alone explain most of the variability in votes across municipalities. They do it mainly because of an overlap of municipalities and regional characteristics. This is born out by comparisons in R^2s of regression models with regional dummies only, the best overall models and regressions models without regional dummies. For M5S they are 0.64, 0.70 and 0.50 respectively, for Lega 0.76, 0.84 and 0.70, for Democratic Party 0.29, 0.40 and 0.23, for Forza Italia 0.36, 0.42 and 0.26.

Figure 2 displays the coefficient of the regional dummies. The baseline region is the region where the average regional votes more closely correspond to average votes at national level (Marche for Lega, Lazio for M5S and Forza Italia, Friuli-Venezia Giulia for Democratic Party). Then regions are divided in 6 categories based on the percentage distance in steps of 5-10 percent from the baseline region. The evidence on the regional dummies is suggestive of macro-drivers of votes for populist parties. It points to the North-South divide being important in explaining votes for Lega and M5S even after taking into consideration characteristics at municipality level. It is quite striking that, after considering municipality-level drivers, Lega receives between 10 and 20 percent more in Northern regions such as Trentino-Alto Adige, Lombardia and Veneto and M5S in Southern regions such as Sicily and Sardinia. On the contrary, Forza Italia and Democratic Party are more homogeneous in their distribution of votes at national level. Forza Italia is slightly stronger in Southern regions and Democratic Party still has a peculiarly high share of votes in traditionally left-wing Regions in the Centre of Italy (Tuscany, Emilia-Romagna and Umbria).

We do not have enough regions in Italy to provide unequivocal explanations for these regional effects[15]. Theoretically, there may be a role for local politics, as the Italian regions

[15] We did some correlations between votes and some regional characteristics. We find that explanations in terms of income inequalities and social capital are both potentially true, as there is a high correlation between

have increasing power in some policy domains (health, public transportation, etc.); or for the persistency in votes at a regional level; or for the long-term political history of the North-South divide (for an explanation in terms of social capital, see Putnam, 1994, for an explanation in terms of wages and employment, see Boeri et al., 2017).

< Figure 2 >

V. DISCUSSION AND CONCLUSIONS

We have performed a data driven analysis of Italian general elections to obtain a reduced set of predictors of the four main parties' electoral outcomes. Although such machine learning techniques can't substitute for purely descriptive or fine causal analysis, differences and similarities between the sets of predictors provide informative evidence on the features of the emergence of populism.

We find two main qualifications to the emergence of populism: heterogeneity in populisms and the absence of a "revolving door" effect. Let us start by discussing heterogeneity in populisms. Take Lega. It advocates right-wing economic policies, as this populism is against taxes and in favour of tax amnesties, and it is anti-immigrant and nationalistic. Looking at the relevant predictors, resentment against the fiscal role of the State, a globalization backlash and a cultural backlash seem to be potentially good explanations for votes for Lega. M5S instead asks for direct democracy, proposes basic income as a solution to economic issues and is pro-environment: by our predictors, social vulnerability and detachment may play a bigger role in its success. This distinction still holds when we apply the predictors to France and assess their predictive power on votes for Marine Le Pen and Jean Luc Melanchon. Ultimately, our evidence suggests that different forms of populism compete in the electoral arena without being directly alternative, as it is in the traditional right-left dimension. Similarly to firms competing in monopolistic competition, they do not compete in the (same) market, they compete within (different) markets. This is not to say that these populisms may not have underlying common features. Most of them probably are on the supply side: they both have some authoritarian elements in their organization and, in their narratives, they

M5S and, in opposite direction, Lega with the Gini index, income, turnout at elections and the share of undeclared work. Persistency in votes seems also be the case for all parties except M5S.

mistrust the State, traditional democratic and civic organizations. They present themselves as "new" parties in a period of crisis in fight against the "old" elites.

When comparing their predictors to Forza Italia and Democratic Party, similarities and differences emerge alike. There seems to be no straight "revolving doors" effect. However, it also emerges that traditional parties are losing votes in their strongholds while populist parties are gaining in their "new" constituencies. Combining these qualifications is suggestive of political re-grouping along different lines than the pure left/right traditional divide.

Last but not least, even if we cannot give a causal interpretation to each specific predictor, our paper still derives some additional interesting propositions to be tested in future research. For example, our results on age suggest to verify how much cultural backlash is associated only to right-wing populism, while a different divide on age – young people voting for populists - could derive from an increasing sense of detachment from traditional parties. This may explain the recent rise of left-wing Podemos in Spain or of France Insoumise in France. Another example is that economic insecurity may boost votes for different parties depending on its attributes: even with a well-functioning job market, economic threats due to automation and globalization may favour right-wing parties, while, even in the absence of a direct threat from globalization, demand for protection in the aftermath of an economic crisis may boost votes for left-wing populist parties instead. Not only can we suggest that these effects may create a demand for populism, but we can also hint that while the former favours a type of populism that builds on the public's aversion to the economic policies that followed the 2008 economic crisis and proposes tax cuts, the latter favours a type of populism that highlights the need to improve the quality of democracy and is fuelled by the corruption of the traditional actors on the public sphere.

BIBLIOGRAPHY

Akkerman, A., Mudde, C., & Zaslove, A. (2014). How populist are the people? Measuring populist attitudes in voters. *Comparative political studies*, *47*(9), 1324-1353.

Alabrese, E., Becker, S. O., Fetzer, T., & Novy, D. (2019). Who voted for Brexit? Individual and regional data combined. *European Journal of Political Economy*, *56*, 132-150.

Algan, Y., Guriev, S., Papaioannou, E., & Passari, E. (2017). The European trust crisis and the rise of populism. *Brookings Papers on Economic Activity*, *2017*(2), 309-400.

Anelli, M., Colantone, I., & Stanig, P. (2019). We were the robots: Automation and voting behavior in Western Europe. *BAFFI CAREFIN Centre Research Paper*, (2019-115).

Arzheimer, K. (2009). Contextual factors and the extreme right vote in Western Europe, 1980–2002. *American Journal of Political Science*, *53*(2), 259-275.

Arzheimer, K. (2016). Electoral Sociology–who Votes for the Extreme Right and Why–and When?. In *The Populist Radical Right: A Reader*, *277*.

Autor, D., Dorn, D., Hanson, G., & Majlesi, K. (2016). *Importing political polarization? The electoral consequences of rising trade exposure* (No. w22637). Cambridge, MA: National Bureau of Economic Research.

Barone, G., D'Ignazio, A., de Blasio, G., & Naticchioni, P. (2016). Mr. Rossi, Mr. Hu and politics. The role of immigration in shaping natives' voting behavior. *Journal of Public Economics*, *136*, 1-13.

Becker, S. O., Fetzer, T., & Novy, D. (2017). Who voted for Brexit? A comprehensive district-level analysis. *Economic Policy*, *32*(92), 601-650.

Belanger, S., & Pinard, M. (1991). Ethnic movements and the competition model: some missing links. *American Sociological Review*, 446-457.

Bloise F., Chironi D., & Pianta M. (2019). Inequality and elections in Italian regions. *MPRA Paper*, n. 96416.

Brunner, B., & Kuhn, A. (2018). Immigration, Cultural Distance and Natives' Attitudes Towards Immigrants: Evidence from Swiss Voting Results. *Kyklos*, *71*(1), 28-58.

Caselli, M., Fracasso, A., & Traverso, S. (2019). Globalization and electoral outcomes. Evidence from Italy. *Economics & Politics*, 1– 36 forthcoming doi:10.1111/ecpo.12147

Colantone, I., & Stanig, P. (2018a). Global competition and Brexit. *American political science review*, *112*(2), 201-218.

Colantone, I., & Stanig, P. (2018b). The trade origins of economic nationalism: Import competition and voting behavior in Western Europe. *American Journal of Political Science*, *62*(4), 936-953.

Colantone, I., & Stanig, P. (2019). The Surge of Economic Nationalism in Western Europe. *Journal of Economic Perspectives*, *33*(4), 128-51.

Corbetta, P., & Vignati, R. (2013). Left or right? The complex nature and uncertain future of the 5 Stars movement. *Italian Politics and Society*, *72*(73), 53-62.

Cramer, K.J. (2016). *The Politics of Resentment: Rural Consciousness in Wisconsin and the Rise of Scott Walker*. University of Chicago Press.

Diamanti, I. (2014). The 5 Star Movement: a political laboratory. *Contemporary Italian Politics*, *6*(1), 4-15.

Diamanti, I., & Lazar, M. (2018). *Popolocrazia. La metamorfosi delle nostre democrazie*. Laterza Editore.

Dustmann, C., Vasiljeva, K., & Damm, A. P. (2018). Refugee migration and electoral outcomes. *The Review of Economic Studies*.

Fu, W. J. (1998). Penalized regressions: the bridge versus the lasso. *Journal of computational and graphical statistics*, *7*(3), 397-416.

Guiso, L., Herrera, H., Morelli, M., & Sonno, T. (2017). Populism: Demand and Supply. *CEPR Discussion Paper*, No. DP11871.

Halla, M., Wagner, A. F., & Zweimüller, J. (2017). Immigration and voting for the far right. *Journal of the European Economic Association*, *15*(6), 1341-1385.

Harmon, N. A. (2018). Immigration, ethnic diversity, and political outcomes: Evidence from Denmark. *The Scandinavian Journal of Economics*, *120*(4), 1043-1074.

Inglehart, R. F., & Norris, P. (2016). Trump, Brexit, and the rise of populism: Economic have-nots and cultural backlash. *HKS Faculty Research Working Paper Series*, RWP16-026.

Johnston, R., Jones, K., Propper, C., & Burgess, S. (2007). Region, local context, and voting at the 1997 general election in England. *American Journal of Political Science*, *51*(3), 640654.

Kaltwasser, C. R., Taggart, P. A., Espejo, P. O., & Ostiguy, P. (Eds.). (2017). *The Oxford handbook of populism*. Oxford University Press.

Kriesi, H., Grande, E., Lachat, R., Dolezal, M., Bornschier, S., & Frey, T. (2006). Globalization and the transformation of the national political space: Six European countries compared. *European Journal of Political Research*, *45*(6), 921-956.

Levi, E., Mariani, R. D., & Patriarca, F. (2020). Hate at first sight? Dynamic aspects of the electoral impact of migration: the case of Ukip. *Journal of Population Economics*, *33*(1), 1-32.

Margalit, Y. (2019). Economic Insecurity and the Causes of Populism, Reconsidered. *Journal of Economic Perspectives*, *33*(4), 152-70.

Mudde, C. (2004). The populist zeitgeist. *Government and opposition*, *39*(4), 541-563.

Mudde, C. (2007). *Populist radical right parties in Europe* (Vol. 22, No. 8). Cambridge: Cambridge University Press.

Mullainathan, S., & Spiess, J. (2017). Machine learning: an applied econometric approach. *Journal of Economic Perspectives*, *31*(2), 87-106.

Norris, P., & Inglehart, R. (2019). *Cultural backlash: Trump, Brexit, and authoritarian populism*. Cambridge University Press.

Oesch, D. (2008). Explaining workers' support for right-wing populist parties in Western Europe: Evidence from Austria, Belgium, France, Norway, and Switzerland. *International Political Science Review*, *29*(3), 349-373.

Rodrik, D. (2018). Populism and the Economics of Globalization. *Journal of International Business Policy*, 1-22.

Spruyt, B., Keppens, G., & Van Droogenbroeck, F. (2016). Who supports populism and what attracts people to it?. *Political Research Quarterly*, *69*(2), 335-346.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, *58*(1), 267-288.

Weakliem, D. L. (2002). The effects of education on political opinions: An international study. *International Journal of Public Opinion Research*, *14*(2), 141-157.

Table 1: Municipality characteristics

Characteristics	**Variables (and their source)**
Socio-demographic	Share of residents under 35 and over 65 (source: ISTAT 2018), share of married and divorced persons married (ISTAT 2018), the number of ecclesiastical cohabitations per capita (Census 2011), crimes per capita by province (Ministry of the Interior 2016), the rate of secondary schooling (Census 2011), share of residents who rented an apartment (Census 2011) and who is commuter for work reasons (Census 2011)
Presence of immigrants	Share of immigrants (ISTAT 2018), share of African immigrants (ISTAT 2018), share of Chinese immigrants (ISTAT 2018), share of Eastern European immigrants (ISTAT 2018) share of immigrants from Muslim-majority countries (ISTAT 2018)
Socio-economic	share of residents working in manufacturing, commerce, basic services and advanced services (Census 2011), share of small businesses, that is with less than 15 employees, by province (INPS 2016), unemployment rate (Census 2011), youth unemployment rate (Census 2011), relationship between self-employed workers and employees (Ministry of Finance 2016), share of workers with unlimited duration contracts by province (INPS 2016), income average, Gini's inequality index, poverty rate calculated on those who have an annual income of less than € 10000 (2016 Ministry of Finance), net exports from Europe and the rest of the world by province (Coeweb ISTAT 2016)
Welfare and the quality of life	Population (ISTAT 2018), share of residents in single-family dwellings and in buildings with more than 9 apartments (Census 2011), risk index seismic (Civil Protection), life expectancy by province (ISTAT 2016), income support per capita by province (INPS 2015-2016), beds in hospital per capita by province (ISAT 2016), square kilometres of double tracks on total surface area by province (Ministry of Infrastructure 2005)

Figure 1: Geographical distribution of votes for M5S, Lega, Democratic Party and Forza Italia

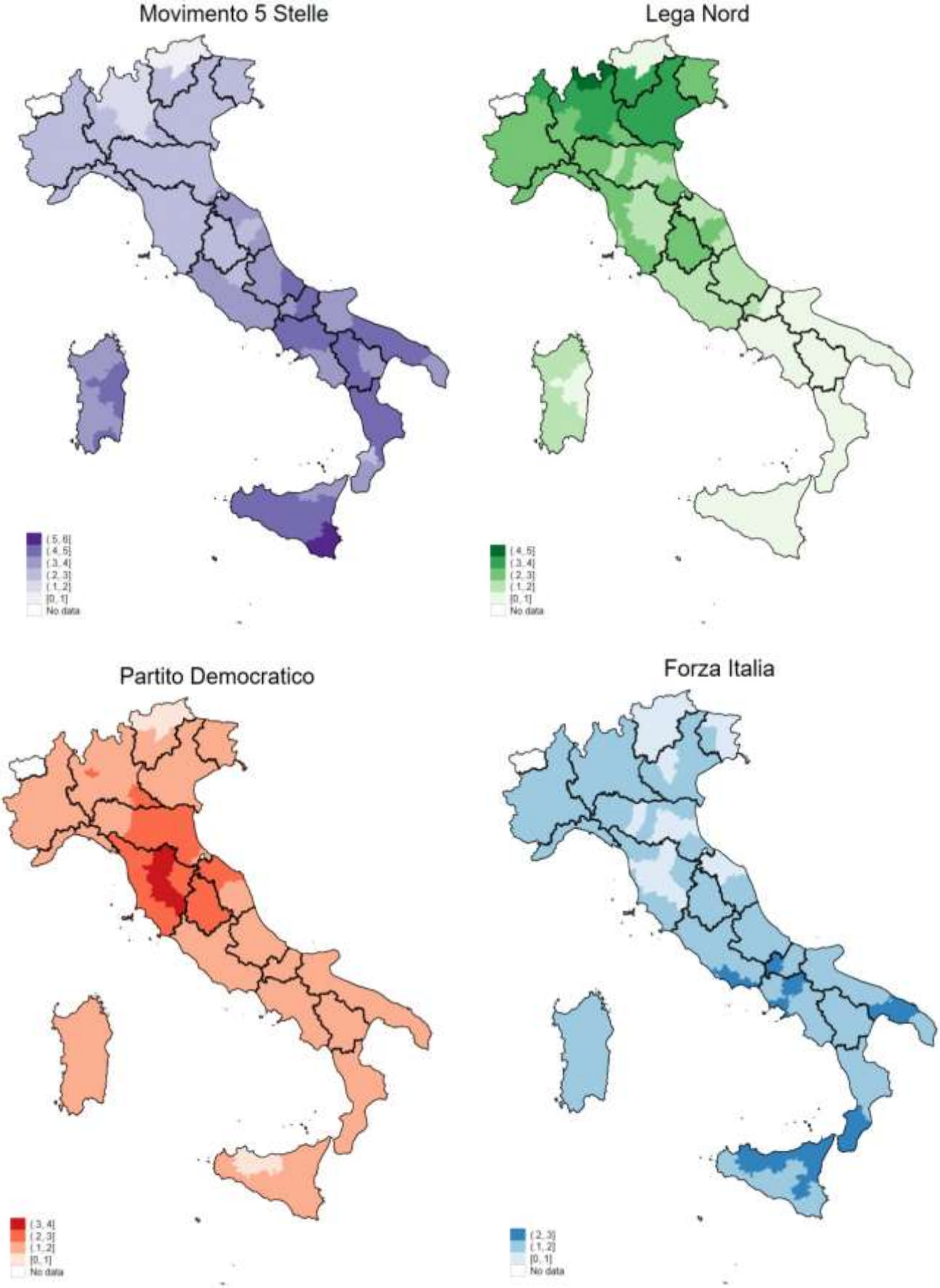

Figure 2: Geographical maps of regional residuals for M5S, Lega, Democratic Party and Forza Italia

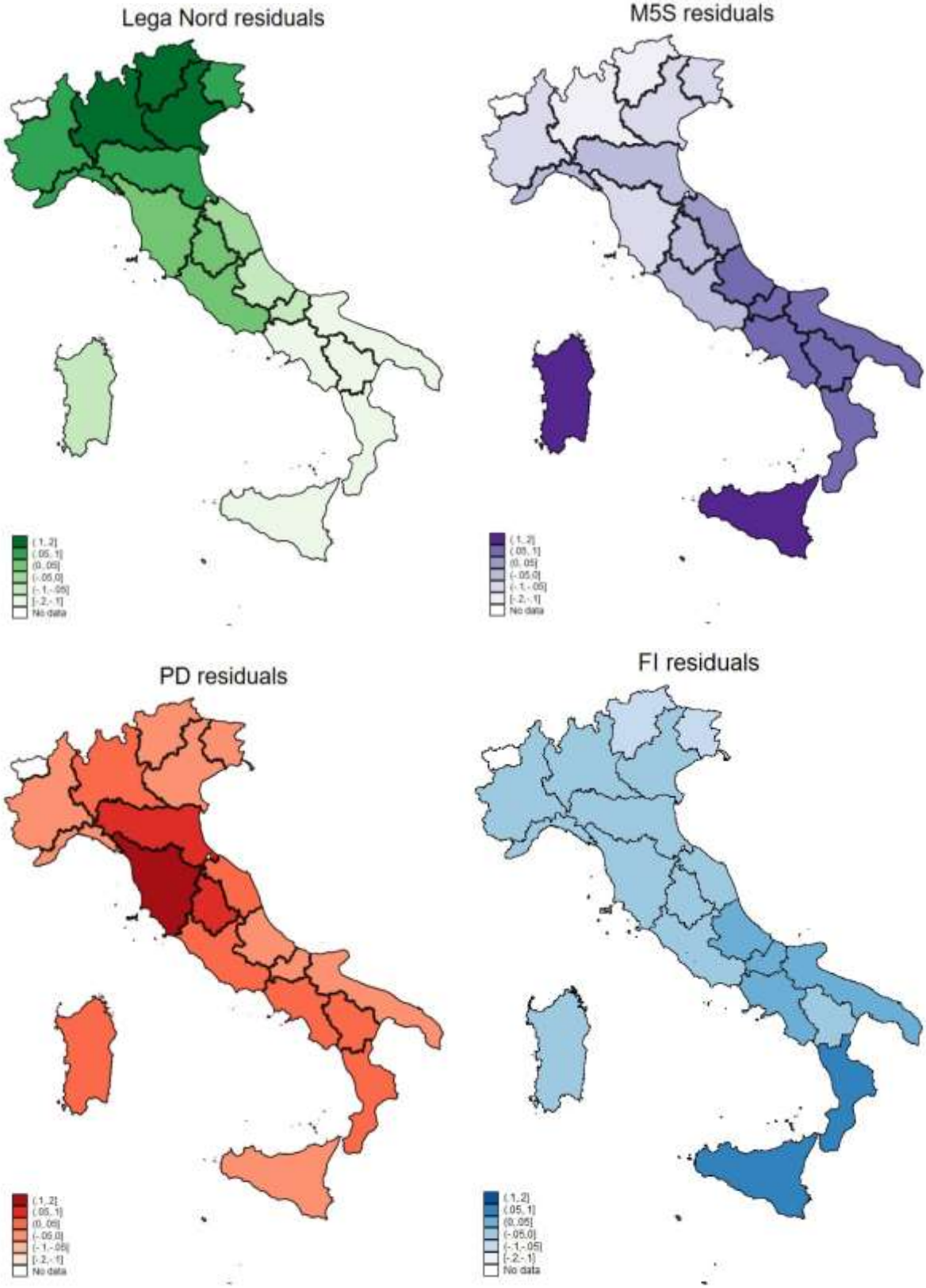

Table 2: Best predictors for 2018 Italian general elections

	M5S	Lega Nord	Democratic Party	Forza Italia	Turnout
Age: 18-35	0.167***	0.0529		0.222***	0.325***
	(0.0436)	(0.0399)		(0.0293)	(0.0343)
Age: 65+	-0.191***		0.218***		-0.260***
	(0.0211)		(0.0149)		(0.0153)
Share of flat-tenants			-0.445***	0.424***	-0.660***
			(0.0622)	(0.0624)	(0.0553)
Crimes per capita	0.676***	-1.057***	1.000***	-0.790***	0.417***
	(0.0749)	(0.0761)	(0.0827)	(0.0691)	(0.0811)
Married	0.300***		-0.141***	0.0429***	0.332***
	(0.0331)		(0.0188)	(0.0123)	(0.0235)
High School Degree		-0.0982***	0.0488***	-0.0992***	0.0855***
		(0.0185)	(0.0126)	(8.68e-03)	(8.15e-03)
Job commuters		-0.0866***		-0.0994***	
		(0.0265)		(0.0122)	
Net exports to Europe		-1.39e-05***	1.76e-05***		1.43e-05***
		(2.19e-06)	(2.01e-06)		(1.73e-06)
Net exports		7.75e-06***	-5.60e-06***	-3.37e-06***	-5.56e-06***
		(1.36e-06)	(1.24e-06)	(5.61e-07)	(1.12e-06)
Manufacturing		0.378***			
		(0.0265)			
Commerce	-0.247***	0.252***	-0.378***	0.346***	
	(0.0319)	(0.0499)	(0.0277)	(0.0300)	
Basic services	0.898***				
	(0.0830)				
Advanced services		-0.262***	0.507***		
		(0.0927)	(0.0672)		
Small firms		0.0244***			
		(0.00315)			
Permanent contracts				-0.100***	0.0726***
				(0.0113)	(0.0108)
Unemployment	0.625***	-0.230***	-0.369***		
	(0.0630)	(0.0370)	(0.0392)		
Income		-6.64e-06***			
		(7.11e-07)			
Poverty share		-0.142***			
		(0.0274)			
Gini index		0.278***			
		(0.0324)			
Population					-6.19e-08***
					(1.71e-08)
Seismic Risk					0.000568***
					(0.000211)
Housing: houses	-0.0319***			0.0157***	-0.0112***
	(0.00461)			(0.00327)	(0.00353)

Housing: more than 9 flats		-0.0278***	0.0483***		
		(0.00858)	(0.00712)		
Beds in hospitals	-0.000377***		0.000471***	0.000335***	
	(0.000108)		(8.98e-05)	(8.20e-05)	
Railroads				25.57***	
				(2.555)	
Constant	0.221***	0.206***	0.0806***	0.121***	0.673***
	(0.00961)	(0.0159)	(0.00803)	(0.00654)	(0.00847)
Observations	7,908	7,759	7,908	7,908	7,908
R-squared	0.705	0.844	0.403	0.421	0.628

Robust standard errors in parentheses. Regional dummies are included (but not reported). Source: authors' calculations.

*** p<0.01, ** p<0.05, * p<0.1

Table 3: Best predictors for 2018-2013 variation in votes in Italy general elections

	M5S 2018 - 2013	Lega 2018 - 2013	Democratic Party 2018 - 2013	Forza Italia 2018 - 2013	Turnout 2018 - 2013
Age: 65+			-0.0961***	0.136***	0.0515***
			(0.0183)	(0.0209)	(0.0171)
Crimes per capita		-0.504***			-0.370***
		(0.0526)			(0.0456)
Married		0.0886***		-0.205***	-0.0827***
		(0.0209)		(0.0273)	(0.0235)
High School Degree	-0.0652***	-0.126***		0.0866***	-0.0741***
	(0.0112)	(8.89e-03)		(0.0157)	(6.55e-03)
Job commuters				0.0776***	-0.0452***
				(0.0205)	(0.00996)
Net exports with Europe				3.93e-06***	
				(9.73e-07)	
Net exports		-6.99e-06***			3.51e-06***
		(5.81e-07)			(4.39e-07)
Manufacturing			0.0851***		
			(0.0243)		
Commerce	-0.286***	0.143***	0.209***	-0.207***	
	(0.0464)	(0.0340)	(0.0334)	(0.0408)	
Basic services			-0.316***		
			(0.0722)		
Small firms			-0.0224***		
			(0.00444)		
Permanent contracts			0.132***		
			(0.0150)		
Unemployment	0.277***			-0.235***	
	(0.0512)			(0.0480)	
Youth Unemployment					-2.26e-06
					(1.88e-06)
Income				5.05e-06***	
				(7.21e-07)	
Poverty share				0.188***	
				(0.0333)	
Gini index			0.238***	-0.315***	
			(0.0161)	(0.0358)	
Seismic risk	-0.000219***				
	(1.93e-05)				
Housing: houses	-0.0230***				
	(0.00450)				
Housing: more than 9 flats		-0.0629***			
		(0.00620)			
Beds in hospitals			-0.000272***	0.000454***	
			(8.67e-05)	(7.96e-05)	
Railroads	23.85***		-12.64***	-12.58***	

	(1.995)		(2.333)	(2.233)	
Constant	0.0953***	0.195***	-0.123***	-0.152***	0.0441***
	(0.00739)	(0.0109)	(0.0178)	(0.0195)	(0.0160)
Observations	7,281	7,281	7,267	7,267	7,183
R-squared	0.736	0.693	0.231	0.211	0.191

Robust standard errors in parentheses. Regional dummies are included (but not reported).
Source: authors' calculations

*** p<0.01, ** p<0.05, * p<0.1

Table 4: Best predictors vs. Lasso regression or 2018 Italy general elections

	M5S	Lega	Democratic Party	Forza Italia	Turnout
Predictors by the best selection model	Age: 18-35, Age: 65+, Crimes per capita, Married, Commerce, Basic services, Unemployment, Housing: houses, Beds in hospital	Age: 18-35, Crimes per capita, High School Degree, Job Commuters, Net exports with Europe, Net exports, Manufacturing, Commerce, Advanced services, Small firms, Unemployment, Income, Poverty share, Gini index, Housing: more than 9 flats	Age: 65+, Share of flat-tenants, Crimes per capita, Married, High School Degree, Net exports with Europe, Net exports, Commerce, Advanced services, Unemployment, Housing: more than 9 flats, Beds in hospital	Age: 18-35, Rent share, Crimes per capita, Married, High School Degree, Job Commuters, Net exports, Commerce, Unlimited duration contracts, Housing: houses, Beds in hospital, Railroads	Age: 18-35, Age: 65+, Share of flat-tenants, Crimes per capita, Married, High School Degree, Net exports with Europe, Permanent contracts, Population, Sismic Risk, Housing: villas
Predictors by Lasso regression	Age: 18-35, Age: 65+, Crimes per capita, Married, Basic services, Unemployment, Poverty share, Gini Index, Housing: houses	Crimes per capita, High School Degree, Manufacturing, Commerce, Basic services, Advanced services, Gini index, Housing: more than 9 flats, Railroads	Age: 65+, Crimes per capita, Married, Divorced, High School Degree, Job Commuters, Commerce, Advanced Services, Unemployment, Gini Index, Housing: more than 9 flats, Life Expectancy	Age: 65+, High School Degree, Commerce, Poverty share, Gini index, Life Expectancy, Railroads	Age: 18-35, Age: 65+, Married, Divorced, High School Degree, Job Commuters, Manufacturing
Shared predictors	Age: 18-35 (+), Age: 65+ (-), Crimes per capita (+), Married (+), Basic services (+), Unemployment (+), Housing: houses (-)	Crimes per capita (-), High School Degree (-), Manufacturing (+), Commerce (+), Advanced services (-), Gini index (+), Housing: more than 9 flats (-)	Age: 65+ (+), Crimes per capita (+), Married (-), Commerce (-), Advanced services (+), Unemployment (-), Housing: more than 9 flats (+)	High School Degree (-), Commerce (+), Railroads (+)	Age: 18-35 (+), Age: 65+ (-), Married (+), High School Degree (-)

Signs in parentheses are the directions of the coefficients. Source: authors' calculations

Table 5: Correlations between actual and fitted values in 2017 France presidential elections

	LEGA	M5S	DEMOCRATIC PARTY	FORZA ITALIA
MARINE LE PEN	0.363***	-0.0211	-0.545***	-0.307**
JEAN-LUC MELANCHON	-0.247*	0.216*	0.118	0.429***
EMMANUEL MACRON	-0.361***	0.0486	0.552***	0.160
FRANÇOIS FILLON	-0.215*	0.0537	0.363***	0.0630

Source: authors' calculations