



The Blank and the Null: An examination of non-conventional voting choices

RODRIGO MARTINS

CeBER and Faculty of Economics of the University of Coimbra

CeBER Working Papers

No. 6

2017

The Blank and the Null: An examination of non-conventional voting choices

Rodrigo Martins[♦]

Faculty of Economics, University of Coimbra, Portugal
and Centre for Business and Economics Research (CeBER), Portugal

Abstract

This paper analyses the determinants of null and blank voting at the 2011 Portuguese legislative elections. An extensive datasets at the parish level and a fractional regression model estimator are used to estimate both voting alternatives. The results found point to some common explanatory patterns as well as to important differences between the two choices. Evidence also indicates that the performance of the local economy, especially unemployment, is important but only for the explanation of blank variations and in more urban areas, where more sophisticated voters reside. Furthermore, results point to the presence of a relevant degree of persistence in both choices and indicate that past electoral choices influence both voting choices in a way that seems to suggest the existence of protest motives.

Keywords: invalid voting; blank ballots; null ballots; Portugal; elections

JEL classification: D72, H7.

[♦] Rodrigo Martins: Faculty of Economics, University of Coimbra, Av. Dias da Silva 165, 3004-512 Coimbra, Portugal. Tel.: +351 239 790578; E-mail: rodrigom@fe.uc.pt.

1. Introduction

The number of annulled and blank votes in stable democracies has traditionally been very small thus usually regarded as a residual phenomenon without consequences to the democratic system. However invalid voting (blanks plus nulls) has more than doubled since the 1970s and since the year 2000 the average percentage of invalid votes in national elections across the world has been estimated at 4,3%, which is not an irrelevant figure¹. This means that, on election day there are a significant number of citizens who actually pay the costs of going to the ballots and choose to intentionally cast a blank or a null vote, consequently depriving themselves from the opportunity to affect the outcome. Although vast, the literature on voting behavior presents an extremely limited number of papers dealing with the answer to this intriguing behavioral puzzle. The majority of scholar explanations for invalid ballots revolve around three dimensions. There is the idea that unfortunate socioeconomic features generates incompetent voters; Alternatively we find the “protest vote” motive, where political disaffection mobilizes voters to cast invalid ballots; and some studies suggest a third explanation related to a set institutional characteristics that can increase the complexity of the electoral systems. While the studies on invalid voting have yielded some interesting findings about one or several of these dimensions no consensus has yet been reached. Also, the number of studies on the subject is quite limited and clearly the reality of individual countries and their invalid voting profiles are under researched.

This article uses an extensive dataset to investigate the determinants of blank and null voting at the 2011 Portuguese legislative elections, and contributes to the voting literature in five major aspects. First, it provides an empirical analysis for a European country that registers relatively low levels of invalid voting. As far as we know, no recent paper is found addressing this subject on the context of a European country and major elections. Additionally, the particular time period examined makes a very interesting laboratory to study the phenomenon of invalid voting since it is characterized by significant social, economic and political unrest.

Second, this paper raises the theoretical question of what determines blank/null voting through the perspective of the famous voting equation that serves as anchor to most of the

¹ Source: International IDEA’s Voter Turnout Database.

literature that studies voting behavior and tries to contextualize the main contributions found in the literature within this rational model.

Third, the article also provides a careful analyzes of the theoretical foundations regarding the effect of economic conditions on non-traditional voting options, as well as a thorough empirical examination. The role of the economy is clearly under explored in invalid voting studies and this investigation attempts to contribute to a further understanding of the economic motives.

Fourth, both null and blank equations are estimate which allows the investigation of differences and similarities between the two. Comparative studies systematically use invalid votes (blanks plus null), by the most part because a significant portion of the countries do not report blank and nulls separately, and most other papers examine only one or the other type of choice². The results found here point to some common patterns as well as important differences between these two alternatives.

Finally, this paper explores two new explanatory possibilities for null and blank voting: the existence of habit-forming patterns; and whether past electoral choices can influence the likelihood of casting an invalid vote. The results found are quite interesting and help improve our understanding of why some people cast void or spoiled ballots.

The remainder of this article is organized as follows. Section 2 presents a brief contextual review of the relevant literature and discusses the role of the economy. Section 3 describes the data and presents the econometric model. The main results and some robustness checks are presented and discussed in section 4 and Section 5 concludes.

2. Theoretical approaches

When trying to analyze voting patterns the starting question is why people vote in the first place. In the particular case of this study the focus is on why some of those that vote do it on non-traditional alternatives. Although this topic has been scarcely studied, several alternative explanations have been provided in the literature over the years, however without clear formal linkages to known theoretical models of voting behavior. In this section, we start by touring these arguments in a different, more contextual perspective, framing these

² On exception is Driscoll and Nelson (2014).

“atypical” voting options within the instrumental approach that is the most widespread theoretical model and that actually tries to be a theory of everything when it comes to voting choices. After, we discuss the particular role of the economy in this scenario.

2.1. *Non-traditional voting choices and the instrumental model*

Downs (1957) model of instrumental voting posits that rational individuals weight the benefits of voting (B) times the probability that their vote will be decisive (P) against the respective overall costs (C), thus deriving measurable utility from the act of voting (R). This relationship can be portrayed through a famous equation:

$$R = P \times B - C. \tag{1}$$

The equation simply states that if $P \times B > C$ then the individual will vote and the B part explains in whom this vote will occur³; however if the costs exceed the weighted benefits the individual will not vote. In the eventuality of costs and benefits matching then it would be indifferent for the individual to vote or not. Downs concludes that since the likelihood of being decisive (P) in a general election is infinitesimal the trouble of going to the ballots systematically exceeds the expected payoff, meaning that for most people the optimal solution is to abstain. Several solutions were proposed to solve this well-known paradox of voting mainly focusing on non-instrumental reasons to vote (D), that added a new element to the initial voting equation:

$$R = P \times B - C + D. \tag{2}$$

The D part is assumed to include those positive satisfactions derived from the act of voting like the possibility to avoid the collapse of democracy (Downs 1957), the idea that people feel content by fulfilling their civic duty (Riker and Ordeshook 1968) and other subjective rewards that have been put forward over the years (See Geys 2006 for an encompassing survey on the rational theories of voter turnout).

³ According to Downs, in a two-party system, B represents the absolute difference between the utility the individual is expected to obtain in the future from the two possible scenarios: The victory of one or the other party.

Now, where do the blank and null voting options and theories fit in all this? Although null voting has no explicit place in the equation, the model has formal room for blank voting provided that there is a positive utility derived from the act of voting and indifference between the available alternatives ($R > 0$ and $B = 0$). The required zero utility for B may be more frequent than one would expect, since the process of generating it (and also P) is complex requiring significant computational capabilities and needing a substantial amount of information. As such some people probably do not engage in their determination, and so the part of instrumental benefits is essentially zero. This means that, for example, low literacy, inexperience and alienation from politics may help explain high levels of blank votes⁴ and the scope of these variables can also be extended to clarify the incidence of annulled votes through the common sense explanation that people can simply make mistakes or get confused when filing their voting bulletin and that some are more prone to it than others. Some papers like Power and Roberts, 1995; Power and Garand, 2007 explore and find support for this idea that social economic characteristics lead to incompetent voters, and thus to more invalid votes.

Some authors claim that invalid votes, are fundamentally a function of the institutional design of democracies. They draw attention to the formal and legal aspects typically examined by turnout studies⁵, and focus on the impact they have on voting efficiency as for instance, bicameralism, mandatory voting, and electoral disproportionality, among others. Some evidence on this is found for Brazil by Power and Roberts (1995) revealing important institutional factors determining invalid voting and by McAllister and Makkai (1993) when comparing invalid votes for the Australian senate and house elections⁶. Overall, in the context of the instrumental model, the notion that some electoral features make voting more or less simple actually means electoral schemes that, by the most part, make benefits (B and P) more or less easier to compute.

Within the instrumental model political fragmentation should also increase the probability of casting a spoiled vote since the presence of more parties/candidates at one election should make it harder to compute B and harder for people to feel they are electing

⁴ Driscoll and Nelson (2014) report that blank voting is more common among those with lower levels of formal education.

⁵ See, for instance, Blais (2006) for a survey on the turnout determinants.

⁶ Also for Australia, Hill and Young (2007) report that the complexity of legislative elections is a relevant source for invalid votes. See also Kouba and Lysek (2016).

the actual incumbent, due to the possibility of coalition governments⁷. Ugglá (2008) finds evidence that more fragmented contests increase the percentage of invalid votes. Additionally, He points out that the more uncertain an electoral outcome seems to be the less invalid votes occur. Kouba and Lysek (2016) also find strong evidence that electoral competitiveness affects invalid voting variations. This aspect clearly belongs to the P part of the voting equation as the probability of being decisive rests critically on two factors: the number of voters; and how close the electoral race is. However, since the instrumental model requires $B = 0$ for admitting the possibility of blank/null votes, P should actually be irrelevant, as benefits are zero no matter the value of P . To make sense of this contradiction between the model and actual empirical results one solution is to admit that, at least for some individuals, their voting equation does not multiply the benefits of voting by the probability of being decisive; rather it adds them, as suggested by Blais (2000). In this perspective people that are in some way aware of the increasingly value of their vote as the electoral race gets closer may have their probability of casting a null/blank vote decreased. For Uglá (2006) competent and interested voters seem to be those that use invalid voting as protest in low competition elections. As such blank and null votes, rather than being a question of incompetence or electoral design, can actually be viewed as a form of protest, fostered by political disaffection and/or bad (or worsening) economic conditions. This hypothesis falls within the expressive part (D) of the instrumental equation. For the pure rational voter, even if the D part includes strong protest motives blank and null vote are still trapped to the $B = 0$ scenario. This means that the expected utility provided by each candidate is required to be the same; however, for the individual both candidates can be equally good or equally bad, and this makes the protest sentiments expressed in D probably a function of what determines B . So for some voters the difference between the overall quality of the candidates may be irrelevant, rather they prefer to focus on the sum of these utilities as it gives them an overall sense of future leadership value and helps to determine D . If some voters strip from the instrumental equation the non-expressive benefits then B is irrelevant in order to make blank/null choices rational. Some degree of party preference may occur, however it is possible

⁷ Blais and Dobrzynska (1998) find evidence that when the effective number of parties is higher people tend to participate less. Nevertheless if we think outside the formal model the resulting effect may be the opposite as more available alternatives can increase the chances for individuals to find a compatible party (Blais and Carty 1990).

that the strong dominance of protest feelings can drive the individual to non-traditional options. But to make room for the null voting option we need to rethink the expressive motives within D . We must admit, for instance, positive satisfactions with democracy and positive dissatisfactions with democracy. Both feelings are positive in this context since they both increase the value of the vote, however, the first increases the predisposition to more conventional voting options, while the later increases the probability of non-traditional outcomes occurring. Either way the choice between protesting through blank or null voting probably relates strongly to the intensity of the disaffection some voters feel and to the capability to compute the instrumental benefits of voting in parties/candidates, things very difficult to measure and thus to test or include in empirical models. Driscoll and Nelson (2014) found evidence to support the idea that many blank and null votes are casted intentionally and driven by political concerns and that null voting was more common among politically sophisticated individuals.

If studies repeatedly find evidence that competent voters are involved in non-traditional voting options then the hypothesis that there may be a process of habit forming should at least be considered as possible and tested. Matsusaka and Palda (1999) advocate that past voting behavior is an important predictor for current electoral choices. Also, some authors like Plutzer (2002) and Gerber et al. (2003) suggest that the question of whether to vote or not can be analyzed as a process of habit formation. This paper explores the possibility that blank/null votes may also exhibit some degree of persistence, a topic that, as far as we known, has not yet been empirically tested or theoretically proposed.

2.2. The role of economic conditions

Systematically economic variables have been considered as describing the social context in the sense that they capture features like social marginality and exclusion. As such raising unemployment is expected to be positively correlated with casting null/blank ballots. However, if we go through the studies of party vote, government vote and turnout the economy is consistently considered an explanatory dimension by itself (see for instance, Paldam 2004). If parties are elected to deliver well-being and economic conditions are the key aspect of prosperity then the economic dimension, and especially unemployment, should be much more than another group of variables describing the social context for any kind of voting

choice. So far only three studies have used unemployment in the analysis of non-conventional voting options. Ugglá (2008) and Kouba and Iysek (2016) found it had no apparent significant effect on the number of invalid bulletins, while Galatas (2008) reports that, at odds with the social context hypothesis, an increase in unemployment seems to decrease blank voting.

In the turnout literature, we find a debate with two contending theories that defend opposite effects regarding the impact of economic conditions on turnout but both present compelling arguments that can be extended to predict the effects on blank and null variations. The first argues that people under economic strain are more prone to vote, in what Schlozman and Verba (1979) called a “mobilization” effect. In this view, people are encouraged to be more active politically (vote, protest, lobby...) because they blame the government for their economic hardships and wish to manifest their discontent regarding the government’s policies, therefore economic duress should foster turnout. If people are more committed to the electoral process then they are more prone to vote within the traditional options, thus an increase in unemployment or general adverse economic conditions should reduce invalid votes. The second theory, more related to the social context hypothesis, argues the opposite, considering that especially the poor and the unemployed exhibit a strong reduction in their ability to participate in politics. It assumes that voters respond to adverse economic conditions by withdrawing from the political process. The reason is that people facing economic hardships face serious problems, therefore they tend to focus their efforts and attention on solving them, thus paying less attention to external matters like politics (see, for instance, Rosenstone 1982). For those that face economic problems but still want to vote the likelihood of casting an invalid vote should increase. The lack of information and attention on electoral issues makes it harder for people to perceive and even compute the instrumental benefits related to conventional voting options. Furthermore, voting blank or null can be seen by these people as a form of protesting against the adverse economic conditions provided by the government and by the democratic system in general. Martins and Veiga (2013) find evidence that both mobilization and withdrawal effects seem to affect turnout in Portugal and that these effects depend heavily on the magnitude of the economic hardships.

One thing that is common to these arguments is that they focus on those voters directly affected by these economic hardships. However, one could argue that, for instance, the impact of unemployment in the percentage of blanks is much more than the sum of the electoral reaction of the unemployed. Everybody reacts to higher levels of unemployment:

some lose their jobs while others become worried it will happen to them or to their family or to their friends. Basically most voters know that increases in unemployment are bad news and if they know that then many people should react electorally to shifts in unemployment, not just the unemployed. Unemployment works differently from other social context variable like education or age and voters seem to view unemployment as a social issue, as well as an economic issue (see Blount 2002). So the importance and impact of the economy relies critically not just on how many people are unemployed and how they react but also on how much each and every voter knows and understand about economic issues. If we return to the rational voting equation one can find alternative arguments to the effect of the economy, although we also end up recognizing the legitimacy of both positive and negative effects expectations. In the instrumental equation, the economy is as probably the most important determinant of B . Perceiving economic benefits, however, is computationally challenging, being economics the complex science that it is, and time demanding in terms of information. Knowing the real economic benefits of voting increases the instrumental value of the vote thus it should reduce the likelihood of blank/null vote occurring. But this effect is conditioned, because the difficulties of economic voting make it only assessable to sophisticated voters. So provided that voters have some degree of sophistication, it is expected that economic hardships increases the attention paid to the economy, therefore decreasing blank/null votes. However, non-sophisticated voters still know something about the economy even if it is only their personal economy, which helps create an overall felling of satisfaction or dissatisfaction. As such, the place for these economic considerations is probably the D part of the equation and economic duress should foster dissatisfaction and consequently blank and null voting. This can be viewed as a sort of economic protest voting. Overall, the effect of economic conditions under this theoretical framework seems to depend heavily on the dominant profile of the voters. The 2011 Portuguese elections are a particularly interesting moment to analyze the effect of the economy on invalid votes. With a “bailout” over the Portuguese debt, raising unemployment, important spending cuts and tax hikes this was a moment in the country’s history where economic issues took central stage on the media and on people’s lives and the reality of economic hardships or the prospect of it was in everyone’s mind. So, in this context it is reasonable to assume that even less sophisticated voters can become sufficiently proficient in economic subjects to incorporate them in their voting decisions, thus increasing the expectations of finding a negative correlation with null and blank voting.

3. Data and Model specification

In order to examine the determinants of spoiled and blank voting a dataset was constructed for the 2011 legislative elections covering all Portuguese parishes⁸ (*freguesias*). Currently there are 4260 parishes in Portugal with population ranging from 31 inhabitants to more than 66000. The choice to analyze this particular election was made because 2011 was a census year, thus offering a significant amount of reliable information at the parish level that for other years does not exist. Also the particular political, economic and social scenario involving this election makes it a very interesting case study. In March 2011 The socialist Prime-Minister Jose Socrates tried to introduce a mixed of spending cuts and tax hikes (Stability and Growth Pact) that had been demanded by the EU to offer a bailout over Portugal's debt levels. The proposed measures were flunked by the Portuguese parliament and following the vote in parliament, PM José Socrates stepped down triggering new elections. Before these scheduled elections, Socrates finally did make a request to the EU and the IMF for a bailout at the beginning of April 2011. The June election that followed was then the third “post-bailout” general election in Europe and the economic scenario in Portugal at the time was impending bankruptcy, record levels of unemployment and public debt, and the country's sovereign bond yield also hitting record highs. José Socrates did run for office again leading the Socialist party to a severe defeat in an election that that gave the right-wing parties a clear mandate to rule the country for the following years. The 2011 legislative elections registered historical low levels of turnout (below 60%) and a 2,66% record high percentage of blank votes.

Portugal is a typical multiparty system where the Assembly of the Republic is the unicameral parliament, currently composed of 230 deputies elected for a four year period by direct suffrage. Parties present closed and blocked lists of candidates in each district and there is a proportional transformation of votes into seats using the Hondt method. Over the years two parties have dominated Portuguese politics, namely the Socialist Party (PS) a center-left party and the Social Democratic Party (PSD) a center-right party. One or the other have always been in office whether in a single party government or leading coalitions.

⁸ Each district is partitioned into municipalities which in turn are subdivided into a variable number of parishes (*freguesias*). These are the lowest administrative units in Portugal.

To analyze this 2011 scenario, electoral data was obtained from the electoral administration of the Secretariat general for internal affairs (*Secretaria Geral da Administração Interna*) and the remaining data used in the paper was collected from the Portuguese National Institute of Statistics (*Instituto Nacional de Estatística – INE*). Descriptive statistics for all variables are reported in the annex.

The empirical model used has an autoregressive component and is of the following form:

$$y_{E,i} = \alpha + \rho y_{E-1,i} + \beta Elect_{Ei} + \gamma Econ_{Ei} + \delta Social_{Ei} + \theta ddist_i + \varepsilon_i \quad (3)$$

where $y_{E,i}$ is the proportion of interest, namely blanks/nulls as a share of the actual voters in parish $i = 1, \dots, 4260$ at election year $E = 2011, 2009$. *Elect*, *Econ*, and *Social* represent, respectively, the vectors of the electoral, economic and social-demographic variables. The coefficient on the lag of the dependent variable (ρ) measures its persistence and prevents the model from being purely static; ε_i is the error term and *ddist* is a vector of district dummies capturing general spatial differences at the district level and accounting for the fact that votes are transformed into seats at this geographical level.

For the baseline model the vector of electoral variables includes a standard measure of competitiveness defined as the difference in vote percentages between the largest party and its closest challenger (*Winning margin*); the Laakso and Taegepera (1979) index for the effective number of political parties (*number of parties*); and the incumbent vote shares at the previous election (*Gov. vote*). Positive variations of the first two are expected to foster blank/null voting, although admissibly fragmentation can assume the opposite sign. As far as we know past government voting has never been used to explain non-conventional voting choices. The idea here is that disappointment with the government's performance can be stronger for those that actually voted for him in the previous election⁹. If invalid voting is strongly related to expressive motives as many scholars argue, then former government supporters that cannot find a credible political alternative may have an increased likelihood to express their dissatisfaction through invalid voting. People that feel betrayed by the government may extend the mistrust that follows to the entire political system, as such the

⁹ The "negative voting" theory (see Lau 1982) reinforces this argument as it posits that the motivation to politically punish governments is greater than the motivation to politically reward them.

variable will capture this specific protest motive if parishes where the government had more support tend to exhibit a higher number of invalid votes.

The set of economic covariates includes the unemployment rate and, as a *proxy* for local income, the municipal per capita purchasing power¹⁰. The expected effect of the economy can go either way as discussed previously.

The vector of variables depicting the social context includes the proportion of people over 65 years old (*Pop. over 65*), the percentage of people with a university degree (*Higher education*) and the percentage of those that cannot read or write (*Illiteracy*). If we assume the incompetence hypothesis then a higher percentage of university graduates and a lower illiteracy rate should reduce the number of invalid votes. However if we assume that non-conventional voters are sophisticated protesters then one ought to expect that, at least a higher education, should increase the likelihood of casting a spoiled or unmarked ballot. The population density (*Pop. density*) is also introduced in the model as a measure of general urbanization and it is expected to be inversely correlated with invalid votes.

The percentages of blank and null votes are traditionally computed as the ratio between the number of blank\ null votes and the total number of votes casted, therefore, by definition, both measures are bounded between 0 and 1 and this statistical propriety cannot be overlooked econometrically. This means that the effect of any particular explanatory variable (x_i) on a dependent variable y , $0 \leq y \leq 1$, cannot be constant all over its range. Consequently standard linear regression estimators like OLS are not adequate and the predicted values they generate are not guaranteed to lie in the required unit interval. One common alternative to handle the bounded nature of y is to apply the logit transformation which yields the following model that can be estimated as a linear function

$$\log\left(\frac{y}{1-y}\right) = X\beta + \varepsilon \quad (4)$$

where β is the $k \times 1$ vector of variable coefficients that we aim to estimate, X is a $1 \times k$ vector of independent variables and ε the stochastic error term (see Maddala 1983 for a discussion on the subject). However, neither zeros nor ones can be included in this estimation

¹⁰ This is an index (national average equals to 1) constructed by the Portuguese National Institute of Statistics that, based on a set of indicators obtained from a factor analysis model, characterizes the Portuguese municipalities from a purchasing power perspective.

strategy, as the log-odds ratio is not well defined for those values. Since over 4% (6%) of our sample of blank (null) percentages are zeros, this method does not seem to fit the profile of our data as it would create a truncation problem or the need for *ad hoc* adjustments to the data.

Since there is no reason to assume that the zeros in our sample are created via a different process than all the other proportions, the nonlinear fractional regression model (FRM) developed by Papke and Wooldridge (1996) seems to be the most adequate method to estimate equation 3. This method deals with fractional dependent variables defined in the unit interval including the respective borders and assumes that $E(Y|X) = G(X\beta)$, where $G(\cdot)$ is a known linear function satisfying the unit interval. The authors show that the most efficient way to estimate the parameters β is to assume a Bernoulli distribution of Y conditional on X and to maximize the quasi-likelihood function

$$ll(\beta) = y \log[G(X\beta)] + (1 - y) \log[1 - G(X\beta)]. \quad (5)$$

Furthermore, they indicate that the resulting quasi-maximum likelihood estimator for β will always be consistent, regardless of the true distribution of Y conditional on X ¹¹.

4. Empirical results

The findings of this study are reported and carefully discussed in this section. At first it is considered a set of explanatory variables and two alternative estimators; then we dig deeper into the potential differences between urban and non-urban areas. Additionally, some robustness checks are provided at the end of this section using alternative explanatory variables.

4.1. Initial estimates

Table 1 presents the results from the estimation of the baseline model. Because the data used is aggregate and parishes vary significantly in their size, estimations are weighted

¹¹ See Ramalho, Ramalho and Murteira (2011) for a survey on the methods for dealing with fractional regression models.

by the size of the electorate¹². Additionally, VIF (*Variance inflation factors*) were calculated for the model and, as expected due to the significant number of observation in the sample, no evidence of severe multicollinearity was found. The results from the fractional regression model (FRM) are reported, but since it is reasonable to assume that the error terms can be correlated across the blank and null equations, and since some papers do to not distinguish between the two, the SUR (*seemingly unrelated regression*) estimates are also reported. Despite the choice goes to the FRM - for the reasons explained above – it is considered a good practice, at this stage, to report the results from other “competing” estimators. As to the overall estimates found with FRM¹³ and SUR the results are much similar and consistent across both estimation methods.

[Insert Table 1 around here]

Looking at the results provided by the FRM, we observe that electoral, economic and socio-demographic variables seem to be simultaneously relevant to explain blank voting. As for the spoiled votes, the electoral dimension takes centre stage along with some social characteristics, but no economic effects are found.

The electoral variables essentially seem to work in the same way for both voting choices and exhibit the expected signs: more parties and reduced competitiveness increase blank and null voting percentages. Also government vote in the previous election is strongly significant and positively related to both voting options implying that parishes more supportive of the government in the previous election exhibit higher proportions of blanks and nulls. This suggests the presence of protest motives when casting an invalid vote. Probably, former incumbent supporters unhappy with the government’s performance see invalid voting as a form of expressing their dissatisfaction. The effect found is more relevant for blank votes as the estimated positive effect found for a one percentage point variation on government support is 2,1 percent when compared to the estimated 0,43 percent effect found for null votes.

Although the results for the electoral variables seem to suggest that both voting options respond to the same stimuli and therefore one could sum the blank and null votes and estimate a single equation on invalid votes, when we turn to the other explanatory

¹² This also helps reduce the interpretation problem known as the “ecological fallacy”.

¹³ The coefficients reported for the FRM are the average marginal effects.

dimensions we find quite a different story. Economic conditions have no explanatory power when it comes to spoiled votes but are strongly relevant in explaining the percentage of blanks. Unmarked ballots seem to be less in wealthier regions and a one percentage point increase in unemployment is on average expected to decrease blank vote in 2,8 percentage points, meaning that Individuals are more likely to vote for the existing parties in areas with high unemployment, rather than cast blank ballots. This result is in line with the findings of Galatas (2008) for the provincial elections in Ontario. One explanation is that in the context of adverse economic conditions people are encouraged to be more active politically (mobilization effect) and thus more prone to vote for existing parties. This can also suggest that, probably sophisticated economic voters are less likely to cast a blank ballot. Complementary, as suggested before, the specific economic context of the 2011 elections, namely the widespread adverse economic conditions witnessed, may have contributed to an increase in the number of individuals concerned with the economy thus “sophisticating” more voters.

Regarding the social variables, as expected an increase in the percentage of those more experienced voters (*pop. over 65*) reduces the likelihood of casting a blank or null vote. However, the effect of education works differently for both voting options. On the one hand, higher education and lower illiteracy foster blank vote supporting the idea that unmarked votes are related to protest motives and sophisticated voters; on the other hand, parishes with higher percentages of university graduates seem to exhibit lower proportions of null votes which suggests that more educated voters that engage in invalid voting seem to prefer the blank option to the null option.

Across all estimations presented in this paper the lagged dependent variable is found to have a strong statistical significance. This presents evidence that some degree of persistence exists in both blank and null votes that is important to account for when estimating equations for invalid voting options. The results suggest that on average around 58% (37%) of a parish’s percentage of blank (null) votes is transported from one election to the other. Obviously, we cannot conclude that 58% (37%) of current blank (null) voters are repeating the electoral choice they made in the previous election, however, the results are a strong indicator that there may exist for a relevant percentage of individuals the habit of casting a blank/null vote.

4.2. Robustness check

Urbanization variables are included regularly as explanatory variables in equations that study invalid vote variations (see, for instance Knack and Kropf 2003; Power and Roberts 1995). This paper takes a different approach using the distinction between urban, developing (averagely urban) and rural parishes to divide the sample and estimate separate equations for the three types of parishes. The reasons why the sample is split are: (1) most socio-demographic and economic explanatory variables are substantially different between rural and urban areas thus their impact and importance on the explanation of the dependent variables may differ; (2) Furthermore, typically urban areas register Higher literacy rates, better wages, more exposure to all sort of cultural and political events, higher contact with politically relevant information, higher population density that helps intensify social networks and easier access to political rallies, protest marches and electoral information. Therefore, one could safely assume that the typical urban voter has a more sophisticated profile than the average rural voter.

Data on the distinction between urban, developing and rural parishes are made available by Portuguese National Institute of Statistics (INE) and the results for the sample splits are reported in table 2 for both the blank and null voting equations.

[Insert Table 2 around here]

Overall the results are consistent with those presented in the previous table, though some interesting patterns emerge from the urban–rural divide. As to blank voting, electoral variables remain relevant for all types of parishes, however the negative impact of unemployment seems to go from important in urban areas to less important in developing parishes to irrelevant in rural areas. If the typical rural voter is less sophisticated than the urban voter then this results suggests that adverse economic conditions only seem to reduce the likelihood of casting a blank vote for those more sophisticated voters, reinforcing the theoretical arguments behind the negative effect hypothesis. The explanatory importance of the education variables prevails across estimations however, only in urban areas does a higher education impacts negatively the null vote. Overall, the results for the education variables seem to reinforce the idea that people with a university degree intending to vote but not in any of the available parties, clearly prefer to vote blank rather than null.

If we look at the null vote determinants' in non-urban areas, besides the influence of electoral variables and some degree of persistence nothing more seems to help us understand this voting choice. Furthermore, the overall importance of the explanatory variables in all six equations and the pseudo-R2 reported suggest that we are more successful in explaining blank variations than null ones and that we seem to know more about both phenomenon's in urban areas than in non-urban areas.

Some further robustness inspections were made to the baseline empirical model to confirm the reported results and to explore some other potential effects. These results are presented in table 3.

[Insert Table 3 around here]

In Columns 1 and 4 the share of votes for the largest opposition party in the previous election (*Social Democratic party*, PSD) was added to the blank and null regressions to check if the effect of government vote in the previous election could be party related and not government related as assumed. Since the votes for the opposition party proved to be statistically insignificant and the impact of the votes for the Socialist incumbent remained virtually the same as in table 1, the effects found seem to originate from the fact that the Socialist party was in office at the time of the 2011 elections. In columns 2 and 4 the variables capturing the social context were replaced by other *proxies*, namely the percentage of active people working in the primary and tertiary sectors¹⁴ and the average age in each parish. The sign and significance found for these variables confirmed the results presented in table 1. Additionally, the crime rate by municipality and the percentage of women over the total population were also included as regressors providing some interesting results. Raising criminality reduces the blank vote suggesting that social marginality and exclusion fosters party vote, therefore it seems that those people directly affected by a more perilous and "decadent" environment prefer to resort to a more active democratic participation. Strikingly gender seems to be relevant as a one percentage point increase in the number of women is expected to depress null vote in 2,3 percentage points. As the null vote is a more aggressive

¹⁴ People working in the tertiary sector are associated with a higher education while those working on the primary sector relate more to a lower formal education. These variables refer to the classification of economic activities in three sectors: the primary, the secondary and the tertiary. The tertiary sector is characterized by the production of services and the primary the production of raw materials.

way of protesting, even requiring the individual to spoil the voting bulletin, maybe it is a less appealing voting choice for women.

Finally, in columns 3 and 6 some alternative variables to characterize the economic environment are used, namely the average housing rents and the employment rate. The results found corroborate the previous evidence since they highlight the importance of the labor market for blank voting and the irrelevance of the economic environment for the explanation of null vote variations.

5. Conclusion

This paper investigates the determinants of non-conventional voting choices at the 2011 Portuguese legislative elections. Taken as a whole, the empirical model is more successful in predicting levels of blank votes than null votes and it seems to capture more about both phenomenon's in urban areas than in non-urban areas. Since the choice of explanatory variables in empirical works is based on theories that assume some sort of rational to the act of casting a blank or a null vote, probably in areas where we find more sophisticated voters (urban areas) the covariates are more successful in capturing the variations of the dependent variable.

While the empirical investigation seems to support the idea that shifts in electoral variables such as less competitiveness, more political fragmentation and higher percentages of government supporters in the previous election, all increase blank and null percentages, it also suggests that the similarities between both voting options almost end there and that the differences found are important and advise against thinking that both options are "twins". For the blank vote model we found support for hypotheses relating to each cluster of independent variables, whereas for spoiled votes that was not the case. One noteworthy situation of this happening was the impact of economic conditions that was strongly relevant in explaining the percentage of unmarked ballots but irrelevant when it came to spoiled votes. Contrary to all studies that investigate non-traditional voting choices, this paper takes considerable theoretical and empirical attention to the role of economic conditions. The results highlight the importance of variables related to conditions in the labor market suggesting that a raise in unemployment strongly depresses blank voting percentages. It seems that economic hardships are encouraging some people to be more active politically

and more committed to the electoral process thus more prone to vote within the traditional options, rather than leading them to protest against these same adverse economic conditions by voting blank. A further analysis of the phenomenon points to the economy being a relevant predictor of blank voting variations in urban areas but not in rural areas. This result raises the suspicion that economically sophisticated voters tend to prefer traditional voting options when facing local adverse economic conditions, and that for those less sophisticated voters the economy has no real importance neither in the choice to vote blank nor in the decision to vote null.

No clear evidence was found supporting the idea that invalid voting is actually a question of voters' ignorance or confusion. Although the complexity of the electoral process, measured by the number of relevant parties, appears to foster invalid voting the results for the education variables seem to suggest the opposite indicating that areas with a higher percentage of university graduates have more unmarked ballots and less spoiled ballots and that increasing levels of illiteracy reduces the percentage of blanks.

Aside from the analysis of the Portuguese reality and the particular focus on the economy, there are two other unique aspects of this study. The first is the investigation to the existence of habit-forming patterns, and results strongly suggest that for some Portuguese voters casting a blank or a null ballot may be a recurrent practice. The second is the investigation of whether past electoral choices can influence the likelihood of casting an invalid vote. The evidence found seems to corroborate this connection, as parishes more supportive of the government in the previous election were found to exhibit higher percentages of blanks and nulls. One interpretation of this result is that former government supporters that feel betrayed by the government tend to become, at least in the short term, highly suspicious of the entire political system, thus protesting against it through non-conventional voting choices. A final interesting feature we came across was a gender bias: results with aggregate data seem to indicate that men are more prone to spoil their bulletin than women.

Bibliography

Blais, A. 2006. What Affects Voter Turnout?. *Annual Review of Political Science*, 9, 111-25.

- Blais, A. 2000. *To Vote or Not to Vote: The merits and Limits of Rational Choice Theory*. Pittsburgh: University of Pittsburgh Press.
- Blais, A. and A. Dobrzynska. 1998. Turnout in Electoral Democracies. *European Journal of Political Research*, 33, 239-61.
- Blais, A. and K. Carty. 1990. Does Proportional Representation Foster Voter Turnout?. *European Journal of Political Research*, 18, 167-81.
- Blount, S. 2002. Unemployment and Economic Voting. *Electoral Studies*, 21, 91-100.
- Downs, A. 1957. *An Economic Theory of Democracy*: Harper Collins Publishers.
- Driscoll, A. and M. Nelson. 2014. Ignorance or Opposition? Blank and Spoiled Votes in Low-Information, Highly Politicized Environments. *Political Research Quarterly*, Vol. 67(3) 547– 561.
- Galatas, S. 2008. "None of the above?" Casting blank ballots in Ontario provincial elections. *Politics & Policy*, 36:3, 448-73
- Gerber, A., D. Green, and R. Shachar. 2003. Voting may be Habit-Forming: Evidence from a randomized field experiment. *American Journal of Political Science*, 47:3, 540-50.
- Geys, B. 2006. "Rational" Theories of Voter Turnout: A Review. *Political Studies*, 4, 16-35.
- Hill, L. and S. Young. 2007. Protest or error? Informal voting and compulsory voting. *Australian Journal of Political Science*, 42:3, 515-21.
- Knack, S. and M. Kropf. 2003. Voided ballots in the 1996 Presidential Election: A county level analysis. *Journal of Politics*, 65(3), 881-97.
- Kouba, K. and J. Lysek. 2016. Institutional determinants of invalid voting in post-communist Europe and Latin America. *Electoral Studies*, 41, 92-104.
- Laakso, M. and R. Taagepera. 1979. "Effective number of parties: A measure with application to West Europe." *Comparative Political Studies*, 12, 3-27.
- Lau, R. 1985. "Two explanations for negative effects in political behavior." *American Journal of Political Science*, 29:1, 119-38.
- Maddala, G. S. 1983. *Limited dependent and qualitative variables in Econometrics*, Cambridge University Press, Cambridge.
- Martins, R. and F. J. Veiga. 2013. Economic performance and turnout at national and local elections, 57, 429–48.
- McAllister, I. and T. Makkai. 1993. Institutions, society or protest? Explaining invalid votes in Australian elections. *Electoral Studies*, 12:1, 23-40.

- Matsusaka, J. and F. Palda. 1999. "Voter Turnout: How much can we explain?" *Public Choice*, 98, 431-46.
- Paldam, M. 2004. Are Vote and Popularity Functions Economically Correct?, in C.K. Rowley & F. Schneider (Eds.), *The Encyclopedia of Public Choice* (pp. 49-59). Netherlands: Kluwer Academic.
- Papke, L.E. and J.M. Wooldridge. 1996. Econometric methods for fractional response variables with an application to 401(k) plan participation rates. *Journal of Applied Econometrics* 11(6): 619–632.
- Plutzer, E. 2002. "Becoming a Habitual Voter: Inertia, Resources, and Growth in Young Adulthood." *American Political Science Review*, 96:1, 41-56.
- Power, T.J. and J.T. Roberts. 1995. Compulsory voting, invalid ballots, and abstention in Brazil. *Political Research Quarterly*, 48:4, 795-826.
- Power, T.J. and J.C. Garand. 2007. Determinants of invalid voting in Latin America. *Electoral Studies*, 26, 432-44.
- Ramalho E., Ramalho, J.S. and J.M.R. Murteira. 2011. Alternative estimating and testing empirical strategies for fractional regression models, *Journal of Economic Surveys* Vol. 25, No. 1, pp. 19–68.
- Riker, W., P. Ordeshook. 1968. A theory of the calculus of voting. *The American Political Science Review*, Vol. 62, No. 1. (Mar), 25-42.
- Rosenstone, S. 1982. Economic Adversity and Voter Turnout. *American Journal of Political Science*, 26:1, 25-46.
- Schlozman, K. and S. Verba. 1979. *Injury to Insult: Unemployment, Class, and Political Response*. Cambridge: Harvard University Press.
- Ugla, F. 2008. Incompetence, alienation, or calculation? Explaining levels of invalid ballots and extra-parliamentary votes. *Comparative Political Studies*, 41:8, 1141-64.

LIST OF TABLES

Table 1. Determinants of Blank and Null voting: 2011 Portuguese Legislative Elections

	BLANK		NULL	
	SUR (1)	FRM (2)	SUR (3)	FRM (4)
ELECTORAL				
Winning margin	0.0149*** (7.353)	0.0167*** (4.521)	0.0051*** (3.548)	0.0053** (2.465)
Number of parties	0.0038*** (10.526)	0.0042*** (7.055)	0.0011*** (4.462)	0.0012*** (2.788)
Gov vote (prev. election)	0.0200*** (11.588)	0.0211*** (7.569)	0.0043*** (3.490)	0.0046*** (3.285)
ECONOMIC				
Unemployment	-0.0260*** (-8.191)	-0.0285*** (-5.564)	-0.0027 (-1.177)	-0.0031 (-1.215)
Purchasing power	-0.0014*** (-3.302)	-0.0014** (-2.181)	-0.0000 (-0.040)	0.0001 (0.217)
SOCIAL				
Pop. over 65	-0.0061*** (-3.401)	-0.0078*** (-3.421)	-0.0020 (-1.551)	-0.0032** (-2.038)
Higher education	0.0092*** (4.236)	0.0085*** (2.594)	-0.0109*** (-6.931)	-0.0139*** (-6.150)
illiteracy	-0.0113*** (-3.329)	-0.0172*** (-3.751)	0.0004 (0.184)	-0.0009 (-0.322)
Pop. density	-0.0002*** (-3.897)	-0.0002** (-2.427)	0.0000 (0.884)	0.0000 (0.956)
PERSISTENCE				
Blank/null prev. election	0.6235*** (34.169)	0.5840*** (19.451)	0.3657*** (25.208)	0.3005*** (17.340)
Observations	4,258	4,258	4,258	4,258
R2/Pseudo-R2	0.508	0.288	0.332	0.116

Notes: Clustered standard errors by municipality are in parentheses; significance levels at which the null hypothesis is rejected: ***, 1%; **, 5%, and *, 10%. All estimations include a constant term and a complete set of district dummies. For FRM estimations the reported coefficients are the average marginal effects.

Table 2. Determinants of Blank and Null voting: from urban to rural areas

VARIABLES	BLANK			NULL		
	Urban (1)	Developing (2)	Rural (3)	Urban (4)	Developing (5)	Rural (6)
ELECTORAL						
Winning margin	0.0128*** (2.818)	0.0169*** (3.260)	0.0134** (2.454)	0.0011 (0.453)	0.0117*** (3.571)	0.0094*** (2.696)
Number of parties	0.0038*** (5.416)	0.0034*** (3.389)	0.0040*** (3.529)	0.0010** (2.227)	0.0022*** (3.255)	0.0010 (1.417)
Gov vote (prev. election)	0.0201*** (4.216)	0.0232*** (5.896)	0.0162*** (5.194)	0.0017 (0.882)	0.0092*** (3.802)	0.0068*** (3.080)
ECONOMIC						
Unemployment	-0.0363*** (-4.796)	-0.0196** (-2.142)	-0.0046 (-0.707)	-0.0025 (-0.764)	-0.0035 (-0.528)	0.0001 (0.029)
Purchasing power	-0.0009 (-1.260)	0.0038* (1.910)	-0.0011 (-0.423)	0.0007* (1.933)	-0.0015 (-0.706)	-0.0007 (-0.428)
SOCIAL						
Pop. over 65	-0.0083** (-2.508)	0.0000 (0.002)	0.0042 (0.954)	-0.0062*** (-2.895)	0.0032 (0.801)	0.0005 (0.161)
Higher education	0.0090** (2.263)	0.0226** (2.173)	0.0919*** (7.033)	-0.0115*** (-4.611)	-0.0106 (-1.492)	-0.0182 (-1.309)
illiteracy	-0.0020 (-0.238)	-0.0403*** (-4.528)	-0.0113* (-1.856)	-0.0038 (-0.874)	-0.0032 (-0.494)	-0.0049 (-1.080)
Pop. density	-0.0001 (-1.200)	-0.0066* (-1.870)	-0.0003 (-0.029)	0.0000 (0.701)	-0.0027 (-0.960)	-0.0018 (-0.216)
PERSISTENCE						
Blank/null prev. election	0.6875*** (14.988)	0.5383*** (15.432)	0.4642*** (11.665)	0.4846*** (14.299)	0.2848*** (9.141)	0.2641*** (12.605)
Observations	1050	1024	2184	1050	1024	2184
Pseudo-R2	0.456	0.427	0.279	0.247	0.114	0.111

Notes: Clustered standard errors by municipality are in parentheses; significance levels at which the null hypothesis is rejected: ***, 1%; **, 5%, and *, 10%. All estimations include a constant term and a complete set of district dummies. The reported coefficients are the average marginal effects.

Table 3. Determinants of Blank and Null voting: alternative predictors

	BLANK			NULL		
	(1)	(2)	(3)	(4)	(5)	(6)
ELECTORAL						
Opposition (prev. election)	-0.0031 (-1.046)			-0.0004 (-0.194)		
SOCIAL						
Criminality		-0.0053*** (-3.696)			0.0000 (0.044)	
Women		-0.0070 (-0.690)			-0.0206*** (-3.482)	
Age		-0.0001** (-2.206)			0.0000 (0.988)	
Tertiary sector		0.0076*** (3.940)			0.0002 (0.121)	
Primary sector		-0.0102*** (-2.867)			-0.0034* (-1.857)	
ECONOMY						
Employment			0.0286*** (5.493)			0.0032 (1.234)
Housing rents			0.0041 (1.471)			-0.0007 (-0.422)
Winning margin	0.0163*** (4.365)	0.0148*** (4.115)	0.0151*** (4.306)	0.0053** (2.401)	0.0062*** (3.002)	0.0054*** (2.593)
Number of parties	0.0039*** (5.922)	0.0040*** (6.383)	0.0040*** (7.012)	0.0011** (2.491)	0.0011*** (2.870)	0.0012*** (2.900)
Gov. vote (prev. election)	0.0189*** (5.554)	0.0191*** (7.550)	0.0207*** (7.511)	0.0043** (2.186)	0.0067*** (4.289)	0.0046*** (3.264)
Unemployment	-0.0292*** (-5.645)	-0.0349*** (-7.293)		-0.0032 (-1.201)	0.0028 (0.935)	
Purchasing power	-0.0014** (-2.162)	0.0007 (1.090)		0.0001 (0.218)	-0.0016*** (-3.599)	
Pop. density	-0.0002** (-2.457)	-0.0001** (-2.122)	-0.0002*** (-2.660)	0.0000 (0.946)	-0.0000 (-0.665)	0.0000 (0.982)
Pop. over 65	-0.0073*** (-2.998)		-0.0080*** (-2.842)	-0.0032* (-1.954)		-0.0033** (-2.019)
Higher education	0.0088*** (2.648)		0.0038 (1.514)	-0.0139*** (-6.243)		-0.0135*** (-7.366)
illiteracy	-0.0169*** (-3.676)		-0.0172*** (-3.745)	-0.0008 (-0.309)		-0.0009 (-0.328)
Blank/null prev. election	0.5831*** (19.339)	0.5747*** (19.176)	0.5836*** (19.271)	0.3003*** (17.316)	0.3239*** (17.331)	0.3006*** (17.309)
Observations	4258	4253	4257	4258	4253	4257
Pseudo R2	0.288	0.295	0.287	0.116	0.108	0.116

Notes: Clustered standard errors by municipality are in parentheses; significance levels at which the null hypothesis is rejected: ***, 1%; **, 5%, and *, 10%. All estimations include a constant term and a complete set of district dummies. The reported coefficients are the average marginal effects.

ANNEX

Table A.1. Descriptive statistics

Variable	Obs	Mean	Std.,Dev.	Min	Max
Blank	4258	0,0246	0,0134	0	0,1135
Blank (prev. election)	4260	0,0152	0,0097	0	0,1071
Null	4258	0,0145	0,0096	0	0,1026
Null (prev. election)	4260	0,0150	0,0097	0	0,1099
Winning margin	4260	0,1665	0,1115	0	0,4834
Number of parties	4260	3,3986	0,6496	2,0366	5,6473
Government (prev. election)	4260	0,3681	0,1051	0	0,8353
Opposition (prev. election)	4260	0,3394	0,1230	0	0,9219
Unemployment	4260	0,1243	0,0523	0	1
Employment	4259	0,8760	0,0508	0,4333	1
Housing rents	4260	0,1622	0,0970	0	0,7800
Purchasing Power	4260	0,787	0,2406	0,4983	2,1688
Higher education	4260	0,2619	0,1134	0,0394	0,7500
Pop. over 65	4260	0,0657	0,0521	0	0,4513
Illiteracy	4260	0,2350	0,0658	0,0702	0,6531
Pop. density	4260	0,4976	1,6736	0,0009	29,4954
Criminality	4260	0,3230	0,1079	0,0960	1,0490
Women	4260	0,5198	0,0218	0,3497	0,6634
Age	4260	45,7261	6,3890	28,390	69,340
Tertiary sector	4256	0,5741	0,1502	0,0714	0,9619
Primary sector	4258	0,1063	0,1281	0	1

Notes: Units of measure: Housing Rents: thousands of euros; Age: years; Pop..density: thousands inhabitants; Criminality: hundreds of thousands inhabitants; Number of parties: parties; All other variables are percentages between zero and one.

Variability: Income, Criminality, Winning margin and number of parties vary across municipalities; all other variables vary across parishes.



CEBER WORKING PAPERS

(Available on-line at www.uc.pt/go/ceber)

-
- 2017-06 *The Blank and the Null: An examination of non-conventional voting choices – Rodrigo Martins*
- 2017-05 *Where is the information on USD/Bitcoins hourly price movements?- Helder Sebastião, António Portugal Duarte & Gabriel Guerreiro*
- 2017-04 *The response of non-price competitiveness and productivity due to changes in passed income gaps. Evidence from the OECD countries - Pedro André Cerqueira, Micaela Antunes & Elias Soukiazis*
- 2017-03 *Dutch Disease in Central and Eastern European Countries - João Sousa Andrade & António Portugal Duarte*
- 2017-02 *On the gains of using high frequency data and higher moments in Portfolio Selection- Rui Pedro Brito, Hélder Sebastião & Pedro Godinho*
- 2017-01 *Growth adjustments through non-price competitiveness and productivity. A cumulative causation approach- Elias Soukiazis, Micaela Antunes & Pedro André Cerqueira*

A série CeBER Working Papers foi iniciada em 2017.