Measuring Electoral Persistence in Multiparty Systems: A Multivariate Fractional Regression Approach

Rodrigo Caldeira de Almeida Martins
Centre for Business and Economics CeBER and Faculty of Economics, University of Coimbra

José Maria Ruas Murteira
Centre for Business and Economics CeBER and Faculty of Economics, University of Coimbra

CeBER Working Papers
No. 5 / 2026

CeBER is funded by the Foundation for Science and Technology, I.P.

Measuring Electoral Persistence in Multiparty Systems: A Multivariate Fractional Regression Approach

Rodrigo Martins [(*)]

Faculty of Economics, University of Coimbra, and CEBER

José Murteira

Faculty of Economics, University of Coimbra, and CEBER

Abstract

Since most of the world's democracies feature multiparty systems, the analysis of aggregate electoral data involving three or more parties is of particular interest to political scientists. To handle this type of situation, the present paper adopts a quasi-maximum likelihood approach in the estimation of regression models of multiparty vote shares, which accommodates any number of parties and different functional forms, while relying only on correct specification of conditional means of shares. Rather than introducing a new estimator, this framework is used to quantify and assess the extent of covariate-independent vote persistence in multiparty elections, evincing how the well-known Dogit specification can be used to operationalize and estimate this persistence within a compositional setting. In particular, the distinction between the share of votes that systematically responds to political, economic, and sociodemographic stimuli and the share that remains stable across local contexts can be formulated as an estimation and testing problem using aggregate data. The strategy set forth in the paper is illustrated with cross-sectional data on aggregate vote shares from a Portuguese legislative election.

[(*)] Rodrigo Martins (Corresponding author), Faculty of Economics, University of Coimbra, Av. Dias da Silva, 165, 3004-512 Coimbra, Portugal. E-mail: rodrigom@fe.uc.pt. This work was financed by National Funds of the Portuguese Foundation for Science and Technology (FCT) within the projects UIDB/05037/2020.

1 Introduction

It is a known fact that the vast majority of the world's democracies feature elections involving more than two relevant parties. Consequently, in order to properly analyse aggregate electoral results in most countries, an adequate statistical tool is required that can accommodate the multivariate nature of such data. In addition, the fractional, unit-sum, inherent definition of electoral party results poses relevant issues that challenge conventional models and estimators, such as linear univariate and seemingly unrelated regressions (SUR) models, and least squares estimation. Furthermore, the usage of statistical frameworks that approximate the idiosyncrasies specific to voting behaviour is of special interest to political scientists. One important characteristic, in particular, is that a significant portion of individuals tends to vote always in the same way, being captive to one specific electoral alternative. Paldam (2004) refers that, in a typical election, 80% of the people belong to the permanent part of voting, i.e. they actually do not respond electorally to political, economic and sociodemographic stimuli. To explain these patterns of committed voting there is a vast longstanding literature focusing on the prevalence of electoral choices manly through partisan identifiers but also by age cohorts and group loyalties (see, for instance, Campbell *et al.*, 1960, and Green *et al.*, 2002). As it stands, the traditional econometric strategies used to estimate vote or turnout functions with aggregate data are essentially targeting the swing voters, thus supposing that the remaining part of the electorate is captured by the constant term, when the function used is in levels or, when using first-differenced data, that these committed voters are simply expelled.

The contribution of the present text to the extant literature is twofold. Firstly, the paper focuses on the measurement of covariate-independent vote persistence in multiparty settings, using a flexible econometric framework based on a quasi-maximum likelihood (QML) method to estimate models of aggregate vote shares in a multiparty environment and to assess how the importance of this persistent component varies across political alternatives. This method proves easy to use and only requires correct specification of the responses' conditional means, given covariates. For this reason, it is more robust than approaches that rely on a fully correct likelihood specification, such as maximum likelihood (ML) estimation. As is well known, a fully parametric approach (*e.g.*, Dirichlet-based ML—see Kotz *et al*., 2000, Ch. 49) is not, in general, robust to misspecification of the conditional likelihood. One useful exception occurs when the conditional means of the responses are correctly specified and the adopted likelihood is a member of the linear exponential family (LEF)—the case with the QML method here proposed. Actually, QML has been around for quite some time, introduced decades ago in the Statistics and Econometrics literature (Gouriéroux *et al*., 1984; McCullagh *et al*., 1989, and the references therein). The use of this estimator with univariate fractional data, based on the Bernoulli p.f., was proposed by Papke et al. (1996; 2008); more recently, the QML estimator using the multivariate Bernoulli (MB) p.f. was also advocated as a suitable approach for multivariate share regressions (Mullahy, 2015; Murteira *et al*., 2016).

Secondly, the paper presents an application to multiparty vote shares of one generalization of the multinomial Logit, the Dogit model (Gaudry *et al*., 1978), that constitutes a particularly interesting alternative to the former, as it embodies the distinction of the proportion of voters actually responding electorally to those political, economic and sociodemographic stimuli that are included in the model as covariates—the proportion of swing voters—from those shares of the total vote that are received by the competing parties, regardless of those stimuli. The estimation strategy operationalizes "electoral stickiness" within a coherent compositional framework that most papers on the subject tend to ignore. Throughout the paper, the covariate-independent component of the Dogit specification is interpreted as a reduced-form measure of persistent or sticky electoral support across parishes. Although such persistence may reflect partisan attachment or long-term political alignment, it may also capture other stable contextual factors not explicitly included among the covariates.

The next section details the proposed regression models for multiparty aggregate electoral data, as well as the QML approach to model estimation and testing. In the ensuing text, this general strategy is illustrated with cross-sectional data from the Portuguese 2011 general elections, covering all 4260 Portuguese parishes (*freguesias*). In this study, estimation results strongly suggest that a significant proportion of the electorate is "captive" to one of the political alternatives and that, at odds with the traditional responsibility hypothesis, left governments, and left parties, in general, seem to benefit electorally from increases in unemployment. Following comments on these empirical results, the paper is concluded in Section 4, which also contains suggestions for future research, hinted at by the present work.

2 Regression Models and Estimation Methods

Framework

Suppose that the number of contending parties in a given election is equal to $M \geq 2$, and let $\boldsymbol{y} \equiv (y_1, \ldots, y_M)'$ denote the vector of vote shares received by these parties. By definition, the components of $\boldsymbol{y}$ are elements of the unit $(M-1)$-simplex

$$\mathcal{S}^{M-1} \equiv \left\{ \boldsymbol{y} \in \boldsymbol{\mathcal{R}}^M : \sum_{m=1}^{M} y_m = 1, y_m \geq 0, \forall m \right\}.$$

Researchers' main interest often lies in the estimation of the conditional means of the elements of $\boldsymbol{y}$, given a set of explanatory variables, $\mathrm{E}(\boldsymbol{y}|\boldsymbol{X})$. Frequently, $\mathrm{E}(\boldsymbol{y}|\boldsymbol{X})$ is specified as a vector of functions of covariates' indices, that is, $\mathrm{E}(\boldsymbol{y}|\boldsymbol{X}) \equiv \boldsymbol{G}(\boldsymbol{X\beta};\boldsymbol{\eta})$, where $\boldsymbol{G}(\cdot)$ denotes a column vector of M functions, and $\boldsymbol{\beta}$ and $\boldsymbol{\eta}$ represent vectors of parameters ($\boldsymbol{\eta}$ denotes parameters that are not coefficients of the covariates). Formally,

$$\boldsymbol{G}(\boldsymbol{X\beta};\boldsymbol{\eta}) \equiv [G_1(\boldsymbol{X\beta};\boldsymbol{\eta}), \ldots, G_M(\boldsymbol{X\beta};\boldsymbol{\eta})]',$$

with each G_m denoting the conditional mean of the m-th share, that is, $G_m(\boldsymbol{X\beta};\boldsymbol{\eta}) \equiv \mathrm{E}(y_m|\boldsymbol{X})$. Given the compositional nature of the elements of $\boldsymbol{y}$, the components of $\boldsymbol{G}$ are also defined in the unit simplex, that is, $G_m(\boldsymbol{X\beta};\boldsymbol{\eta}) \geq 0$, $\forall m$, and $\sum_{m=1}^{M} G_m(\boldsymbol{X\beta};\boldsymbol{\eta}) = 1$. In this expression, $\boldsymbol{X} \equiv (\boldsymbol{x}_1', \ldots, \boldsymbol{x}_M')'$ denotes an $(M \times K)$-matrix of explanatory variables with m-th row $\boldsymbol{x}_m' \equiv (x_{m1}, \ldots, x_{mK})$ $(1 \times K)$, $m = 1, \ldots, M$, and $\boldsymbol{\beta}$ is a column K-vector of population parameters, so that $\boldsymbol{X\beta} = (\boldsymbol{x}_1'\boldsymbol{\beta}, \ldots, \boldsymbol{x}_M'\boldsymbol{\beta})'$ is a column vector of M indices of covariates. To simplify the notation, $\mathrm{E}(\boldsymbol{y}|\boldsymbol{X})$ or its elements will often be referred to without explicit mention of their arguments: for instance, $\boldsymbol{G} \equiv \boldsymbol{G}(\boldsymbol{X\beta};\boldsymbol{\eta})$ and $G_m \equiv G_m(\boldsymbol{X\beta};\boldsymbol{\eta})$ (possibly with a superscript – *e.g.*, G_m^L).

Finally, for estimation purposes, a sample of N observations on all variables is supposed available to the researcher. Accordingly, y_{im}, $\boldsymbol{y_i}$, $\boldsymbol{x_{im}}$ and $\boldsymbol{X_i}$, $i = 1, \dots, N$, denote the i-th observation on, respectively, y_m, $\boldsymbol{y}$, $\boldsymbol{x_m}$ and $\boldsymbol{X}$. For simplicity, the i-th replica of $G_m(\boldsymbol{X_i\beta};\boldsymbol{\eta}) \equiv \mathrm{E}(y_{im}|\boldsymbol{X_i})$ will often be represented by G_{im}, with the corresponding estimate denoted as $\hat{G}_{im} \equiv G_m\left(\boldsymbol{X_i}\hat{\boldsymbol{\beta}};\hat{\boldsymbol{\eta}}\right)$.

Conditional Mean Models

The models used for binary responses in the univariate case are also employed to describe the conditional mean of fractional responses. Accordingly, the specifications adopted to model the probability of an individual choosing between M mutually exclusive modalities may also be employed to describe $\mathrm{E}(\boldsymbol{y}|\boldsymbol{X})$ in the multivariate case, since they satisfy the bounded, unit-sum nature of the shares' means.

Multinomial Logit Model

In a multivariate context, special attention has been devoted to the well-known multinomial Logit specification, which can be generally expressed as

$$G_m^L = \exp(\boldsymbol{x'_m\beta}) \Big/ \sum_{l=1}^{M} \exp(\boldsymbol{x'_l\beta}), \quad m = 1, \dots, M. \tag{1}$$

This formulation is general enough to encompass alternative- (that is, party-) invariant covariates and alternative-specific parameters, if interactions of alternative-specific indicators with alternative-invariant explanatory variables are included as covariates.

The models used in the present application (Section 3) include only party-invariant covariates (varying only with the statistical units), with party-specific coefficients: this particular case will henceforth be denoted as MNL. The MNL model can be formulated as

$$G_m^L \equiv \mathrm{E}(y_m|\boldsymbol{x}) = \exp(\boldsymbol{x}'\boldsymbol{\beta_m})\Big/\Big[1+\sum\nolimits_{l=1}^{M-1}\exp(\boldsymbol{x}'\boldsymbol{\beta_l})\Big], \quad m=1,\dots,M-1, \tag{2}$$

with $\boldsymbol{x}' \equiv (1, x_1, x_2, \dots, x_K)$ the vector of explanatory variables (allowing for intercepts) and each of the $M-1$ parameter vectors defined as $\boldsymbol{\beta_m} \equiv (\beta_{m0}, \beta_{m1}, \dots, \beta_{mK})'$, $\quad m = 1, \dots, M-1$. Then, for the M-th alternative one has

$$G_M^L = 1 - \sum\nolimits_{l=1}^{M-1} G_l^L = 1\Big/\Big[1+\sum\nolimits_{l=1}^{M-1}\exp(\boldsymbol{x}'\boldsymbol{\beta_l})\Big].$$

The multinomial Logit specification implies a uniform pattern in the way mean shares react to changes in one particular vote mean share. From (1), the cross elasticity of vote mean share of party m with respect to that of party l, $m \neq l$, can be expressed as

$$\varepsilon_{ml} \equiv (\partial G_m^L/\partial G_l^L)(G_l^L/G_m^L) = -G_l^L,$$

which means that every expected share, other than l, suffers an identical proportional change that only depends on the initial mean share for party l.[1]

Unlike in linear models, the marginal effects of covariates' changes on the means of fractional responses are not constant but depend on covariates' values. Under MNL, the relation between coefficients' and marginal effects' signs is somehow complex; for a unit increase in a continuous covariate x_k, expression (2) yields

$$\frac{\partial G_m^L}{\partial x_k} = \begin{cases} G_m^L\left[\left(\beta_{mk} - \sum\nolimits_{l=1}^{M-1} G_l^L \beta_{lk}\right)\right], & m = 1,\dots,M-1 \\ -\sum\nolimits_{l=1}^{M-1} \dfrac{\partial G_l^L}{\partial x_k}, & m = M \end{cases},$$

[1] This result is parallel to the well-known 'independence of irrelevant alternatives' (IIA) property of MNL, as a discrete choice model.

where the second line reflects the compositional nature of the responses.[2] Naturally, for discrete covariates differences of the conditional means should replace derivatives.

To estimate causal effects, one can, following the usual practice, either evaluate the above formulae with parameters' estimates at 'representative' values of the covariates (*e.g.*, sample averages) or use average partial effects (APE), computed as averages of marginal effects across all sample individuals. For a generic covariate z,

$$\widehat{\text{APE}} = N^{-1} \sum\nolimits_{i=1}^{N} \frac{\partial \hat{G}_{im}^{L}}{\partial z_i},$$

with $\partial \hat{G}_{im}^{L}$ denoting evaluation of partial derivatives at parameters' estimates, $\hat{\boldsymbol{\beta}}$.

Dogit Model

As previously mentioned, there is reason to suspect that a significant fraction of aggregate voting outcomes is not influenced by the political, economic and sociodemographic stimuli that are included in the model as covariates. In this case, the MNL approach is not fully appropriate. Rather, one specification that seems particularly adapt to deal with this possibility is provided by the Dogit model, a generalization of the multinomial Logit devised by Gaudry and Dagenais (1979) as a discrete choice model, introduced in the area of transportation research. In general, the Dogit can be expressed as

$$G_m^D = \left[\exp(\boldsymbol{x}_{\boldsymbol{m}}'\boldsymbol{\beta}) + \eta_m \sum\nolimits_{l=1}^{M} \exp(\boldsymbol{x}_{\boldsymbol{l}}'\boldsymbol{\beta})\right] \Big/ \left[\left(1 + \sum\nolimits_{l=1}^{M} \eta_l\right) \sum\nolimits_{l=1}^{M} \exp(\boldsymbol{x}_{\boldsymbol{l}}'\boldsymbol{\beta})\right] =$$

---

(2) Note that, given the compositional nature of the responses and their means, partial effects add up to zero across alternatives:

$$\sum\nolimits_{m=1}^{M} G_m = 1 \Rightarrow \sum\nolimits_{l=1}^{M} \frac{\partial G_m^L}{\partial x_{lk}} = 0.$$

$$\eta_m/\left(1+\sum\nolimits_{l=1}^{M}\eta_l\right)+\left\{1-\sum\nolimits_{l=1}^{M}\left[\eta_l/\left(1+\sum\nolimits_{l=1}^{M}\eta_l\right)\right]\right\}G_m^L = \tag{3}$$

$$(\eta_m+G_m^L)/\left(1+\sum\nolimits_{l=1}^{M}\eta_l\right), \tag{4}$$

$$\eta_m \geq 0, \qquad m=1,\dots,M,$$

where G_m^L denotes the multinomial Logit conditional mean of y_m. Evidently, the latter results as a special case of the Dogit when $\boldsymbol{\eta}=\mathbf{0} \Leftrightarrow \eta_m=0, \forall m$.

As suggested by expression (3), the Dogit affords an interpretation of each party's mean share as the sum of two terms: the first, expressed by $\eta_m/(1+\sum_{l=1}^{M}\eta_l)$, represents that parcel of the m-th party's mean aggregate share independent of covariates; the second part evinces the remainder of the mean shares, $1-\sum_{l=1}^{M}[\eta_l/(1+\sum_{l=1}^{M}\eta_l)]$, that depends on relevant stimuli (represented by appropriate covariates) and is dispersed among parties according to an MNL specification, G_m^L. Following Ben-Akiva (1977), the Dogit can also be viewed as a model that allows for each party to receive a 'captive' vote share, as measured by the first term in the mean share specification. This latter share can be the result of unobserved socio-demographic covariates such as, *e.g.*, ethnic attachment or the result of institutional designs such as low district magnitudes that deter voters from switching to uncompetitive alternatives.

As the empirical application in Section 3 utilizes the MNL specification, the Dogit model considered in the study also includes party-invariant covariates and party-specific associated coefficients. Write this special case as

$$G_m^D=\left[\exp(\boldsymbol{x}'\boldsymbol{\beta_m})+\eta_m\sum\nolimits_{l=1}^{M}\exp(\boldsymbol{x}'\boldsymbol{\beta_l})\right]\Big/\left[\left(1+\sum\nolimits_{l=1}^{M}\eta_l\right)\sum\nolimits_{l=1}^{M}\exp(\boldsymbol{x}'\boldsymbol{\beta_l})\right]. \tag{5}$$

The Dogit differs from the Logit with respect to yet another issue, alluded to above. While the latter implies a uniform pattern of substitution among alternatives, the Dogit model yields a more complex pattern, depending on the level of each alternative. The cross elasticity of alternative m with respect to alternative l can be expressed as

$$\varepsilon_{ml} = -(G_l^D / G_m^D) \left[\left(1 + \sum_{l=1}^{M} \eta_l \right) G_m^D - \eta_m \right],$$

involving not only the mean share of alternative l (as in the MNL) but also that of alternative m.

Marginal effects of covariates' changes on the Dogit conditional means can be generally expressed as

$$\frac{\partial G_m^D}{\partial z} = \frac{\partial G_m^L}{\partial z} \Big/ \left(1 + \sum_{l=1}^{M} \eta_l \right). \tag{6}$$

In accordance with the Dogit additive separability of mean shares into constant and covariates' dependent terms, marginal effects only influence the latter part of the conditional mean. Expectably, this can lead to significant differences in marginal effects, when compared to those obtained under the MNL approach.

Conditional Means Models—General Remark

The multinomial Logit model can be viewed as a special case of several other specifications, usually designed to overcome the former's limitations, namely the IIA restriction and corresponding uniform pattern of substitution. Among these extensions, the Nested Logit and the Random Parameters Logit are among the most often cited in the econometric literature—namely as models of discrete choice probabilities (see, *e.g.*, Train, 2009, and references therein). These and other generalizations of MNL, as well as other approaches not nesting the latter (*e.g*. multinomial Probit) can be also applied to fractional data. The present paper is limited to the multinomial Logit and Dogit specifications for the sake of brevity and, more importantly, because these models offer an illustrative example of how to address one specific issue of particular interest in the area of Political Analysis. Naturally, different models can be used if deemed appropriate to address other issues of interest to political researchers.

Estimation and Inference

As previously mentioned, the parameters of the model for $\mathrm{E}(\boldsymbol{y}|\boldsymbol{X})$ can be estimated by full information ML, with some conditional distribution of $\boldsymbol{y}$, given $\boldsymbol{X}$, $f_{\boldsymbol{Y}|\boldsymbol{X}}$. Nonetheless, there are more robust estimators available, not requiring full specification of the response's conditional law. One such approach is the QML method, a maximum likelihood-type estimator that only requires correct formal specification of $\boldsymbol{G}$, provided the adopted likelihood is a LEF member—see Gouriéroux, *et al.*, 1984. In practice, although $f_{\boldsymbol{Y}|\boldsymbol{X}}$ may be unknown (the usual situation faced by researchers), with some reasonable assumption on $\boldsymbol{G}$, the parameters of the latter can be consistently estimated through maximization of a LEF likelihood.

In the present context, a natural choice for the likelihood is provided by the MB p.f.—see Johnson, *et al.* (1997, Ch. 36)—appropriate when each individual selects one among M possible alternatives. Let a_m, $m = 1, \dots, M$, denote a binary indicator equal to 1 if alternative m is chosen ($a_m = 0$ otherwise). Let $\pi_m \equiv \Pr(a_m = 1) = E(a_m)$; then, the MB p.f. can be expressed as $f_{\boldsymbol{A}}(\boldsymbol{a}) = \prod_{m=1}^{M} \pi_m^{a_m}$, with $\boldsymbol{a} \equiv (a_1, \dots, a_M)'$. In a regression framework, the parameters π_m are usually replaced by conditional probabilities of success, given covariates. With observations on fractional variables, $(y_{i1}, \dots, y_{iM})'$, and covariates, $\boldsymbol{X_i}$, $i = 1, \dots, N$, substituting G_{im} for π_m, and y_{im} for a_m, the i-th term of the quasi-likelihood can be written as

$L_i = \prod_{m=1}^{M} G_{im}^{y_{im}}$. This yields the individual contribution to the (quasi-) log-likelihood

$$LL_i \equiv \log L_i = \sum_{m=1}^{M} y_{im} \log G_{im} = \sum_{m=1}^{M-1} y_{im} \log(G_{im}/G_{iM}) + \log G_{iM}, \qquad (7)$$

where $G_{iM} = 1 - \sum_{m=1}^{M-1} G_{im}$ and the last expression stresses the LEF form of the log-likelihood. Notice that this cannot be the 'true' likelihood, as the y_m are continuous fractional random variables, whereas the initial a_m are binary variables. Naturally, G_{im} is to be replaced by the adopted conditional mean specification: for instance, G_{im}^{L} or G_{im}^{D} under, respectively, multinomial Logit or Dogit.

Denote the parameters of $\boldsymbol{G}$ in general as $\boldsymbol{\theta} \equiv (\boldsymbol{\beta}', \boldsymbol{\eta}')'$. With independent observations on $\boldsymbol{y}$ and $\boldsymbol{X}$, the QML estimator, $\widehat{\boldsymbol{\theta}} \equiv \left(\widehat{\boldsymbol{\beta}}', \widehat{\boldsymbol{\eta}}'\right)'$, obtained through maximization of $\sum_{i=1}^{N} LL_i$, is consistent and asymptotically normal, provided that $\boldsymbol{G}$ is correctly specified (Gouriéroux, *et al.*, 1984).(3) The covariance matrix of $\widehat{\boldsymbol{\theta}}$ is not the covariance matrix of the full information ML estimator because, in general, QML does not rest on the 'true' likelihood. Consequently, the misspecification-robust 'sandwich' form of the covariance matrix estimator should be used.(4)

As a Dogit-nested model, the MNL specification can be tested against the former by considering the parametric null hypothesis $H_0\colon \boldsymbol{\eta} = \boldsymbol{0}$. Usually, to this effect, one would consider any of the three classical test procedures—Wald, likelihood ratio or score (LM) tests. In the present case, however, only the LM test retains its

(3) Expression (6) suggests one cautionary note, regarding estimation of the Dogit's marginal effects. At first glance, one might conclude from (6) that these effects can be directly obtained from the corresponding MNL's $\boldsymbol{\beta}$ estimates, once $\boldsymbol{\eta}$ is estimated. This conclusion is erroneous due to the fact that the Dogit likelihood does not factor into functionally separable terms, one involving only $\boldsymbol{\beta}$ and one with $\boldsymbol{\eta}$ alone. Therefore, one cannot estimate the latter without a valid estimator of the former.

(4) The expressions for the covariance matrix estimator under particular models for $\boldsymbol{G}$ (*e.g.*, MNL or Dogit) are somewhat intricate in the multivariate case. In view of the fact that, to the authors' knowledge, all econometrics packages feature instructions that produce the corresponding estimates, those expressions are omitted here (the interested reader is directed to one of several Econometrics textbooks available, *e.g.*, Cameron, *et al.*, 2005, Ch. 15.12).

usual asymptotic properties and can, therefore, be used in the usual manner. This is because the value of the parameter vector under the null (multinomial Logit)—$\boldsymbol{\eta} = \mathbf{0}$—is located on the border of the parameter space of the alternative model (Dogit) (Andrews, 2001); recall expression (4) and the non-negativity of the Dogit parameters, $\eta_m, m = 1, \dots, M$.

Actually, the test of the multinomial Logit against the Dogit can be easily implemented as a test of omitted variables within the Logit conditional means. Therefore, to implement the test, one can simply estimate a multinomial Logit model twice, including those additional covariates in the second-step estimation and subsequently assessing their statistical relevance.(5) As shown by Tse (1987), the Dogit can be approximated by an augmented multinomial Logit model, with additional covariates easily obtained from the first-step estimation of the Logit conditional means. The approximation to the Dogit can be written, in general, as

$$G_{im}^{A} \equiv \exp\left[\boldsymbol{x}'_{\boldsymbol{im}}\boldsymbol{\beta} + \zeta_m\left(1/\hat{G}_{im}^{L}\right)\right] / \sum_{l=1}^{M} \exp\left[\boldsymbol{x}'_{\boldsymbol{il}}\boldsymbol{\beta} + \zeta_l\left(1/\hat{G}_{il}^{L}\right)\right] \approx G_{im}^{D}, \tag{8}$$

where ζ_m, $m = 1, \dots, M$, designate parameters and $\hat{G}_{im}^{L}$ denotes the first-step estimate of G_m^L evaluated at the i-th observation in the sample, $i = 1, \dots, N$. Consequently, with $\boldsymbol{\zeta} \equiv (\zeta_1, \dots, \zeta_M)'$, the test of the multinomial Logit against the Dogit is a test of H_0: $\boldsymbol{\zeta} = \mathbf{0}$, that can be carried out as, *e.g.*, a conventional Wald test of the joint significance of the parameters ζ_m, $m = 1, \dots, M$, in this augmented Logit specification. Under H_0 the Wald statistic is asymptotically distributed as a chi-squared random variable with M degrees of freedom. In any event, given that the estimation method is QML, the robust Wald test should be used.

(5) Actually, this strategy can be used to test the multinomial Logit against any alternative model, nesting the former or not—see Murteira (2016) for a general treatment of the subject.

In the present context, rejection of the null hypothesis means that, at least for one contending party, the corresponding vote expected share is, to some degree, rigid and not simply a mean response to those stimuli that are included in the model. Accordingly, rejection of the Logit specification suggests that it is reasonable to proceed with estimation of the Dogit model. Meanwhile, if deemed appropriate, restricted versions of the latter can obviously be entertained. For instance, if, for a given party, electoral results in all circles are reasonably assumed to depend entirely on the model's covariates, one can take the corresponding parameter in $\boldsymbol{\eta}$ (and $\boldsymbol{\zeta}$) to be zero, *a priori* (with the number of restrictions in the null hypothesis reduced by one).

The following empirical analysis exploits cross-sectional variation in parish-level electoral outcomes and covariates that are either measured prior to the election (e.g. unemployment rates) or correspond to slow-moving socio-demographic characteristics (e.g. age structure, education, sectoral employment, population density). As such, concerns of simultaneity or reverse causality do not arise in this context, since electoral outcomes cannot plausibly affect past economic conditions or the demographic composition of parishes. The estimation strategy therefore does not suffer from standard endogeneity issues arising from feedback effects between explanatory variables and vote shares. At the same time, as with all analyses based on aggregate data, the estimated effects should be interpreted as conditional associations at the parish level, and not as direct statements about individual voting behaviour. In particular, the Dogit model's covariate-independent components capture all sources of systematic vote share variation not explained by the included covariates, which may reflect persistent partisan attachment, historical political alignments, or other unobserved contextual factors at play at the local level.

3 Empirical Application: The 2011 Portuguese Legislative Elections.

Political Context

Portugal's democracy features a typical multiparty political system, with a unicameral parliament, *Assembleia da República* (Assembly of the Republic), currently composed of 230 deputies elected for a four-year term by direct suffrage. Parties present closed and blocked lists of candidates in each district, with votes proportionally converted into seats through the Hondt method.

Over the last four decades, two parties have dominated Portuguese politics, namely the Socialist Party (PS), a centre-left party, and the Social Democratic Party (PSD), a centre-right party. One or the other have always been in office, either in a single party government or leading coalitions. Since the 1990's the number of political parties represented in Parliament has been fairly stable and without excessive fragmentation, ranging from four to six parties.

The June 2011 general elections were triggered by the resignation of the socialist Prime-Minister José Sócrates in March of that year, after a failed attempt to bring Parliament to introduce a mix of spending cuts and tax hikes (known as "Stability and Growth Pact") demanded by the European Union (EU) as a prerequisite to offer a bailout over Portugal's debt levels. Eventually, Portugal did issue an official request to the EU and the International Monetary Fund (IMF) for a bailout, and the June elections were then the third "post-bailout" general elections in Europe. The economic scenario in Portugal at the time was one of impending bankruptcy, with record levels of unemployment and public debt, and the sovereign bond yield hitting record highs. Prime Minister Sócrates did run for office again, leading

PS to a severe defeat in an election that afforded right-wing parties a clear mandate to rule the country for the following years.

The overall national results produced by these legislative elections are summarized in Table 1. Although there were seventeen parties running for these elections, only five of them eventually got seats in parliament, representing 91.5% of the national vote. All other political forces, with a joint share of 4.28%, can basically be viewed as a competitive fringe. The five major parties, that ran in every electoral circle, are PS, PSD – both already mentioned – the Democratic and Social Centre-People's Party (CDS-PP), a pre-electoral coalition formed by the Portuguese Communist Party (PCP) and the Ecologist Party (PEV), and Left Block (BE). Following electoral results, the right-wing parties, PSD and CDS-PP, formed a coalition that ruled the country until mid-2016.

Table 1
National Results at the 2011 Portuguese Legislative Elections

Party	Number of Deputies	Number of Votes	Share
BE (Left Bloc)	8	288 923	5.17%
PCP/PEV (Communist Party/Ecologist Party)	16	441 147	7.90%
PS (Socialist party)	74	1 566 347	28.05%
PSD (Social Democratic Party)	108	2 159 181	38.66%
CDS-PP (Democratic and Social Centre/People's Party)	24	653 888	11.71%
Other Parties	0	247 499	4.28%

Source: *Comissão Nacional de Eleições – CNE* (National Elections Commission).

Variables and Data

In view of the number of major political forces running for the 2011 elections, six voting options are considered in the choice set available to the electorate: to vote for one of the five parties that achieved parliamentary representation (BE, PCP/PEV, PS, PSD and CDS-PP) or to vote for one of the remaining political forces (termed "Other Parties", denoted as OP). This choice set yields $M = 6$ fractional dependent variables with unit-sum, thus defined on the unit 5-simplex, $\mathcal{S}^5$. Neither blank nor null votes were accounted for in this study, so the voting system considered here is strictly related to party vote.

Portugal is organized, for administrative purposes, into districts that are, in turn, partitioned into municipalities. These are subdivided into a variable number of parishes (*freguesias*), the lowest administrative units in Portugal. To illustrate the application of the methods described in the previous section, the compositional dataset used in the study comprises observations on aggregate vote shares for each of the six political alternatives described above, obtained in the 4260 parishes in Portugal. Thus, in short, the models and estimator employed in the study take $M = 6$ (number of alternative dependent variables) and $N = 4259$ (sample size).(6) Table A1 (in the Appendix) reports summary statistics on the response variables used in the present study. Note that, as displayed in the table, the data set includes null values for all response variables; as previously mentioned—see footnote (3)—the presence of zeros (or ones) in the sample does not pose any computational difficulty because the QML estimator handles all possible values within the simplex in exactly the same way.

The analysis of this specific election comes with the particular advantage of 2011 being a census year, thus offering a significant amount of reliable information at the parish level that otherwise would hardly be available. Hence, the data set

(6) The number of Portuguese parishes totals 4260; due to one missing observation, the available sample comprises $N = 4259$ observations.

includes parish-level observations on several explanatory variables which were chosen in order to capture the effects of the economic and social environment on vote percentages of the various political forces. These variables include the unemployment rate ($unempr$), the proportion of residents over 65 years old ($perc65$), the fraction of illiterate persons ($illit$), the proportion of active people working in the Tertiary Sector ($tertiary$) and, as a general measure of urbanization, the population density ($popdens$), measured in thousand people per square kilometre. Table A1 displays descriptive statistics for all the covariates used in the study.[7]

Regarding this choice of regressors, the lagged vote shares are not included since they provide a direct measure of electoral inertia rather than a structural decomposition of vote shares. While such variables capture persistence through correlation with past outcomes, they do not distinguish between the components of electoral support that react to observed covariates and those that remain stable through local contexts. By comparison, the Dogit specification allows vote shares to be decomposed into a covariate-responsive component and a covariate-independent component, which can be interpreted as a measure of persistent electoral support. Adding lagged vote shares would consequently absorb part of the persistence that the model is designed to quantify, fogging the distinction between predictive persistence and the underlying structure of electoral support.

(7) Electoral data were obtained from the electoral administration of *Secretaria Geral da Administração Interna* (General Secretariat for Internal Affairs) and data on covariates were collected from the census operation of *INE – Instituto Nacional de Estatística* (Portuguese National Institute of Statistics).

Regression Models

The present study starts with the specification of an MNL regression model for the six dependent variables defined previously, with conditional means given the co-variates described above. All explanatory variables are individual- (that is, parish-) specific and do not vary with the different parties. Thus, from (2) the MNL conditional mean for the i-th statistical unit (parish) is given by

$$E(y_{im}|\boldsymbol{x_i}) = G_{im}^L = \begin{cases} \exp(\boldsymbol{x_i'\beta_m})\Big/\Big[1+\sum_{l=1}^{5}\exp(\boldsymbol{x_i'\beta_l})\Big], & m=1,\dots,5, \\ 1\Big/\Big[1+\sum_{l=1}^{5}\exp(\boldsymbol{x_i'\beta_l})\Big], & m=6, \end{cases} \tag{9}$$

with

$$\boldsymbol{x_i'} \equiv (1, unempr_i, perc65_i, illit_i, tertiary_i, popdens_i), \quad i = 1, \dots, 4260,$$

and five conformable vectors of coefficients, $\boldsymbol{\beta_m} \equiv (\beta_{m0}, \beta_{m1}, \dots, \beta_{m5})'$, $m = 1, \dots, 5$. These $(30 = 5 \times 6)$ coefficients are estimated through QML, using an MB quasi-log-likelihood with individual contribution obtained from (7), replacing G_{im} with the above MNL expression, G_{im}^L.

Upon estimation of model (9), the MNL is tested against the Dogit by estimating (8) and assessing the joint statistical significance of the additional covariates in the approximate model. As reported below, the null hypothesis (MNL) is clearly rejected so the Dogit model—expression (4)—is subsequently estimated by QML. The following subsection reports and interprets the estimated results for both models.

Empirical Results

Table A2 in the Appendix displays the QML estimates and misspecification-robust standard errors for all the parameters of the MNL and Dogit models. As mentioned above, estimation of the Dogit was prompted by a clear rejection of the null hypothesis $H_0: \zeta_1 = \cdots = \zeta_6 = 0$, through a Wald test, using the robust covariance matrix of the QML estimates of the coefficients in the approximate model, G_{im}^A—check expression (8). The resulting value of the Wald test statistic is approximately 68.04, with near zero p-value (with reference to the chi-squared law with six degrees of freedom, χ_6^2).

The coefficients' estimates and standard errors presented in Table A2 are not very informative, *per se*, given the nonlinearity of the estimated models. Nonetheless, it seems worth noting, at the outset, that adopting the Dogit induces, as expected, differences in many coefficients' estimates, when compared to the MNL model's estimates (in sign, magnitude and/or significance)—note that if the true conditional means are of the Dogit-type, only the estimates of this model's parameters are consistent (not those from MNL). Naturally, these differences are reflected in the estimated APEs, computed below under each model. Given the clear rejection of the MNL in favour of the Dogit, the latter's estimates should reasonably be deemed more reliable.

In what regards the $\boldsymbol{\eta}$ parameters, these are all individually significant (in addition to being jointly significant, as attested by the test of MNL against the Dogit). Note that each of these estimates does not provide a direct estimate of the covariates-independent part of the vote mean share for the corresponding party. Rather, each quota is to be computed as

$$\hat{\eta}_m \Big/ \left(1 + \sum_{l=1}^{6} \hat{\eta}_l\right), \qquad m = 1, \dots, 6.$$

Accordingly, the Dogit decomposition should not be interpreted as directly identifying individual-level partisan loyalty, but rather as distinguishing between vote-share variation associated with observed covariates and a persistent residual component. Table 2 displays the resulting estimates, together with corresponding standard errors (obtained through the delta method).

Table 2
Estimated Captive Shares

Party	BE	PCP/PEV	PS	PSD	CDS-PP	OP
Estimated Captive Share (percentage points)	1.27	1.24	25.90	34.81	9.51	6.18
Standard Error	.0016	.0048	.0052	.0087	.0027	.0032
Captive Share / Total Vote (%)	24.6	15.7	92.3	90.0	81.2	-

All estimated mean quotas are clearly significant—$\mathcal{N}(0,1)$ p-values are all close to zero—which suggests that, for every political force under study, a significant part of its parish vote mean share does not depend on the model's explanatory variables. The corresponding six estimated mean shares total about 78.91%, therefore, on average, only 21.09% ($= 100\% - 78.91\%$) of the vote in each parish is estimated to be distributed by the six political forces in response to the included covariates (according to the six specified MNL conditional means)—a sizeable rebuttal of the MNL null model. Clearly, the two major Portuguese parties (PS and PSD) take most of the estimated shares, while results for the far-left parties (PCP/PEV and BE) suggest a relatively low support base. The term 'captive vote share' used here should be interpreted as a reduced-form measure of covariate-independent electoral persistence rather than as a direct structural estimate of individual partisan attachment. Naturally, the magnitude of these estimated persistent shares depends on the covariates included in the specification and may in part reflect stable regional, historical, or socio-political factors not explicitly modelled. To better assess the relative importance of the covariate-independent component across political forces, in the last line of Table 2 it is reported the estimated captive shares as a percentage of each party's total vote thus assessing the relative importance of the captivity estimate. In this sense, the Dogit decomposition does more than measure the general magnitude of covariate-independent voting, highlighting also potential differences in the *structure* of electoral support across parties. Results reveal sizable heterogeneity across parties with a distinctive pattern. While the estimated persistent component is representative of much of the vote for the two larger parties (around 92% for PS and 90% for PSD), for the smaller parties these percentages drop significantly, particularly on the left, probably consistent with more available alternatives within that ideological space. These findings imply that the distinction between "permanent" and responsive vote components is not uniform across the party system, reflecting differences in how parties are embedded within broader ideological configurations. One plausible interpretation is that electoral variation tends to happen inside relatively well-defined ideological blocks. This is coherent with theories of electoral competition (Downs, 1957) and empirical work on party systems organized along stable ideological blocks (Adams et al., 2005).

Table 3 displays the estimated APEs (percentage points—p.p.) under the MNL and Dogit models.(8) Although the MNL's estimates are probably not consistent, it seems useful to confront both sets of results for each pair of explanatory and response variable.

(8) For each explanatory variable, the sum of APEs across the six political forces may not be strictly zero due to rounding errors.

Table 3
Estimated Average Partial Effects under Dogit and MNL Models

Covariate	Party	Dogit	MNL
unempr	**BE**	.0292	.0231
	PCP/PEV	.1392	.1421
	PS	.2469	.3308
	PSD	-.2538	-.3625
	CDS-PP	-.1531	-.1153
	OP	-.0085	-.0182
perc65	**BE**	-.0425	-.0557
	PCP/PEV	-.0074	-.0418
	PS	-.1211	-.0441
	PSD	.2244	.2018
	CDS-PP	-.0340	-.0375
	OP	-.0194	-.0227
illit	**BE**	-.0529	-.0507
	PCP/PEV	.1333	.0938
	PS	-.0982	-.0487
	PSD	.0661	.0269
	CDS-PP	-.0875	-.0363
	OP	.0390	.0150
tertiary	**BE**	.0454	.0526
	PCP/PEV	.1137	.1177
	PS	-.0527	-.0439
	PSD	-.1659	-.1821
	CDS-PP	.0143	.0104
	OP	.0452	.0453
popdens	**BE**	.2519	.0193
	PCP/PEV	.6554	.0748
	PS	.3300	.2171
	PSD	-1.5786	-.4755
	CDS-PP	.1855	.1755
	OP	.1558	-.0113

Estimated APEs in percentage points. All estimates significant at 1%.

Overall, the contents of the table are expected to provide an approximate answer to the following main concern: "how does a unit change in a covariate affect that segment of a party's mean share that is influenced by the set of observed explanatory variables?" Under MNL, the figures in the table obviously estimate point changes within the total party vote (100%); on the other hand, under the Dogit model, the reported APEs refer to a much lesser segment of this total, estimated at approximately 21%. In any event, again one should bear in mind that the MNL's estimates are very likely unreliable, in view of this model's strong rejection; therefore, the figures in the last column of Table 3 should be viewed with skepticism.(9) Consequently, the general overview of the reported results focuses on the Dogit estimates. It is noted, at the outset, that all estimates in Table 3 are statistically significant at very low levels (standard errors are omitted from the table, in view of this uniformity).

Upon inspection, several interesting results seem noteworthy. First, and foremost, under the Dogit a unit increase (one p.p.) in unemployment is estimated to increase the parish vote mean share of the government party (PS) by 0.2469 p.p. This finding is fundamentally at odds with the notion that voters punish (or reward) governments if the economy goes badly (well). This so-called "responsibility hypothesis" has firmly established itself in the literature as the principal reason for economic voting.(10) Nevertheless, the present results stray in a different direction, suggesting the idea that voters seem to react to higher unemployment by voting in those parties they believe are more concerned or prepared to combat the problem—traditionally left-wing parties, as suggested by Hibbs (1977). Rattinger (1981, 1991) calls this the "clientele hypothesis" and finds evidence that surges in unemployment do not affect left-wing governments; they may even improve their electoral results. Further inspection of the results for the unemployment covariate

(9) For instance, under Dogit an increase of one percentage point (p.p.) in *unempr* is estimated to cause an increase of .0292 p.p. in the parish vote mean share of BE—whereas, under the MNL, a (doubtful) .0231 p.p. increase estimate is obtained. As another example, under Dogit a unit increase in *popdens* (1000 more people per square kilometre) is estimated to cause a decrease of 1.5786 p.p. in the parish vote mean share of PSD, as opposed to an estimated decrease of .4755 p.p. under MNL

(10) Incidentally, the studies of Veiga, *et al.* (2004a, 2004b, 2010) suggest that this reward/punishment mechanism is present in the Portuguese reality).

reveals yet another pattern backing this idea: left-wing parties (PS, PCP/PEV and BE) seem to benefit electorally from raises in unemployment, while the opposite occurs for the other parties. Whether these findings are a consequence of the particular socio-economic environment of the 2011 Portuguese elections, or do they relate to the particular econometric methodology employed is unclear and deserves further investigation.

A second interesting result found on Table 3 relates to the overall effect of the variable $perc65$, with a one p.p. increase of this covariate estimated to improve PSD's parish vote mean share by 0.2244 p.p., while penalizing all other parties. Since PSD is the most important opposition party, and strongest alternative to the government, this finding seems to indicate that parishes with a higher percentage of older voters tend to punish the current government by voting for the party that is in better electoral conditions to succeed, not considering, for the most part, other political options. In some sense, this ties to the idea of a dichotomous electoral rationality where individuals only choose to vote for or against the incumbent, an assumption very characteristic of the literature on government voting. Interestingly, the inverse occurs when we take a closer look at the effect of $popdens$ (population density, a general proxy for urbanization). In this case PSD is the only party expected to be electorally penalized by positive variations of this covariate.

Results also seem to suggest that parishes with more percentage of people working in the tertiary sector show a preference for parties not typically accustomed to being in office. PS and PSD seem to get less support in areas with higher percentages of people working in the tertiary sector, while all other political forces benefit from it. The notion that people working in this sector are often associated with higher levels of education and are, therefore, probably more sophisticated in their decision process and less “myopic” in their retrospective evaluations, may help explain this result.

4 Concluding Remarks

The present paper proposes the application of the general QML method to the estimation of regression models for aggregate electoral data. The suggested approach is both simple and rich enough to encompass several different models, rendering it useful to address the regression analysis of multivariate fractional data, the typical case with aggregate vote shares.

Among other choices, use of the multinomial logit model provides a simple and appropriate framework, not only for the estimation of marginal effects of covariates on vote shares but, also, for the statistical assessment of maintained hypotheses of interest in the Political Science methodology. The empirical study here included illustrates this potential by nesting the multinomial logit within the Dogit model, thereby allowing for the distinction between the so-called swing vote, mean-dependent on various kinds of explanatory variables, and that fraction of vote shares which is due to some sort of (otherwise unexplained) vote stickiness across parishes. The proposed general method is illustrated with electoral results obtained in the 2011 Portuguese legislative election. Naturally, the substantive electoral patterns identified in the empirical application may partly reflect the exceptional political and economic circumstances surrounding the 2011 Portuguese legislative election. Results show that, jointly, the captive estimates amount on average to roughly 80% of total vote shares, thus being broadly in line with numbers previously reported in the literature for the proportion of stable, or non-swing voters.

Though the two concepts are not directly comparable, given the aggregate nature of the data used in the present analysis, this similarity suggests that the proposed framework captures a meaningful dimension of electoral persistence.

Also, the proposed framework not only quantifies aggregate vote persistence, but it is also a useful tool for comparing the relative importance of stable and responsive electoral support across political alternatives. Each party's captive estimate when expressed relative to its own total vote shares reveals two blocks. Major parties exhibit high levels of persistence (around 90%), while in smaller parties the percentages are much lower. This reflects differences in how parties are positioned within broader ideological configurations.

Naturally, different issues of political interest can be addressed through this general approach. Among others, a random parameters' approach, with marginalization of unobserved heterogeneity, or the consideration of some sort of hierarchy in party grouping (*e.g.*, left- versus right-wing parties) and consequent aggregate vote correlation, both lead to well-known generalizations of the multinomial logit, testable through the present methodology. Accordingly, these—or other—issues provide a natural ground for subsequent research

Appendix

The present Appendix displays results that support the empirical application of Section 3, not directly needed in the main text exposition. These results include descriptive statistics for all the variables used in the study (Table A1) and parameters' estimates and associated standard errors, obtained under the MNL and Dogit models (Table A2).

Table A1
Summary Statistics of the Variables used in the Empirical Application

Variable	Average	St. Deviation	Minimum	Maximum
$\boldsymbol{y_1}$ (BE)	3.48	2.28	.00	33.33
$\boldsymbol{y_2}$ (PCP/PEV)	5.51	7.28	.00	71.43
$\boldsymbol{y_3}$ (PS)	28.31	9.98	.00	70.09
$\boldsymbol{y_4}$ (PSD)	44.29	13.15	.00	94.83
$\boldsymbol{y_5}$ (CDS-PP)	10.63	4.79	.00	62.28
unempr	12.41	5.06	.00	56.67
perc65	26.18	11.32	3.94	75.00
illit	23.50	6.58	7.02	65.31
tertiary	57.37	15.10	.00	96.19
popdens	49.77	167.38	.09	2949.54

Measurement units
$y_1 - y_5$, $unempr$, $perc65$, $illit$, $tertiary$: percentage points;
$popdens$: thousand people per square kilometre.

Table A2
Estimates of the Parameters of the Dogit and MNL Models

Response	Covariate	Parameter	Dogit			MNL		
			Est	SE		Est	SE	
y_1		β_{10}	1.801	.673	***	-.798	.074	***
(BE)	***unempr***	β_{11}	2.145	.763	***	.922	.223	***
	perc65	β_{12}	-1.017	.710		-1.328	.131	***
	illit	β_{13}	-5.072	1.333	***	-1.646	.260	***
	tertiary	β_{14}	-.687	.580		.949	.070	***
	popdens	β_{15}	.033	.015	**	.007	.003	**
y_2		β_{20}	.141	.698		-1,893	.130	***
(PCP/PEV)	***unempr***	β_{21}	3.930	.941	***	2.838	.485	***
	perc65	β_{22}	1.073	.754		-.482	.201	**
	illit	β_{23}	.742	1.297		1.517	.387	**
	tertiary	β_{24}	-.205	.630		1.573	.139	***
	popdens	β_{25}	.054	.017	***	.015	.005	***
y_3		β_{30}	4.467	1.104	***	1.591	.060	***
(PS)	***unempr***	β_{31}	12.2297	2.557	***	1.401	.170	***
	perc65	β_{32}	-4.420	2.301	*	.138	.106	
	illit	β_{33}	-7.129	2.575	***	-.362	.213	*
	tertiary	β_{34}	-5.502	.965	***	-.742	.060	***
	popdens	β_{35}	.048	.051		.009	.003	***
y_4		β_{40}	4.489	.781	***	2.228	.058	***
(PSD)	***unempr***	β_{41}	-2.952	1.340	**	-.600	.190	***
	perc65	β_{42}	4.557	1.105	***	.756	.102	***
	illit	β_{43}	-1.430	1.538		-.130	.213	
	tertiary	β_{44}	-5.451	.855	***	-1.007	.060	***
	popdens	β_{45}	-.337	.096	***	-.009	.004	**
y_5		β_{50}	4.728	.933	***	.831	.068	***
(CDS-PP)	***unempr***	β_{51}	-13.465	2.978	***	-.854	.218	***
	perc65	β_{52}	-2.104	1.171	*	-.061	.116	
	illit	β_{53}	-10.680	3.134	***	-.532	.250	**
	tertiary	β_{54}	-1.438	.843	*	-.487	.068	***
	popdens	β_{55}	.082	.022	***	.018	.005	***
y_1		η_1	.060	.011	***			
y_2		η_2	.059	.026	**			
y_3		η_3	1.230	.100	***			
y_4		η_4	1.650	.153	***			
y_5		η_5	.451	.037	***			
y_6		η_6	.293	.032	***			

y_6: Other Parties (OP). */**/*** : significant at the 10%/5%/1% level.

References

Adams, J., S. Merrill III, and B. Grofman. 2005. *A Unified Theory of Party Competition: A Cross-National Analysis Integrating Spatial and Behavioral Factors*. Cambridge: Cambridge University Press.

Aitchison, J. 1982. "The Statistical Analysis of Compositional Data (with discussion)". *Journal of the Royal Statistical Society, Series B (Statistical Methodology)* 44:2: 139–77.

Andrews, D. W. K. 2001. "Testing when a Parameter is on the Boundary of the Maintained Hypothesis". *Econometrica* 69: 683–784.

Campbell, A., P. Converse, W. Miller, D. Stokes. 1960. *The American Voter*. New York: Wiley.

Downs, A. 1957. *An Economic Theory of Democracy*. New York: Harper and Row.

Gaudry, M., M. Dagenais. 1979. "The dogit model". *Transportation Research* 13: B: 105–11.

Green, D. P., B. Palmquist, E. Schickler. 2002. *Partisan Hearts and Minds: Political Parties and the Social Identities of Voters*. New Haven: Yale University Press.

Gouriéroux, C., A. Monfort, A. Trognon. 1984. "Pseudo Maximum Likelihood Methods: Theory". *Econometrica* 52: 681–700.

Hibbs, D. A. Jr. 1977. "Political Parties and Macroeconomic Policy". *American Political Science Review*, 71:4 (December): 1467–87.

Honaker, J., J. Katz, G. King. 2002. "A Fast, Easy, and Efficient Estimator for Multiparty Electoral Data". *Political Analysis* 10:1 (January): 84-100.

Katz, J., G. King. 1999. "A Statistical Model for Multiparty Electoral Data". *American Political Science Review* 93:1 (January): 15–32.

Kotz, S., N. Balakrishnan, N. Johnson. 2000. *Continuous Multivariate Distributions*. Vol. 1. New York: Wiley.

Luce, R. D. 1959. *Individual Choice Behavior: A Theoretical Analysis*. New York: Wiley.

McCullagh, P., J. Nelder. 1989. *Generalized Linear Models*. 2nd ed. London: Chapman and Hall.

Mullahy, J. "Multivariate Fractional Regression Estimation of Econometric Share Models". *Journal of Econometric Methods* 4:1: 71-100.

Murteira, J. 2016. "Goodness of Link Tests for Multivariate Regression Models". *Communications in Statistics – Theory and Methods* 45:24 (February): 7367-75.

Murteira, J., J. Ramalho. 2016. "Regression Analysis of Multivariate Fractional Data". *Econometric Reviews* 35:4: 515-52.

Paldam, M. 2004. "Are Vote and Popularity Functions Economically Correct?". In *The Encyclopedia of Public Choice*, eds. C.K. Rowley, F. Schneider. Dordrecht: Kluwer Academic Publishers, 49-59.

Papke, L. E., 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 (November): 619-32.

Papke, L. E., J. M. Wooldridge. 2008. "Panel Data Methods for Fractional Response Variables with an Application to Test Pass Rates". *Journal of Econometrics* 145: 121–133.

Rattinger, H. 1981. "Unemployment and the 1976 Election in Germany: Some Findings at the Aggregate and the Individual Level of Analysis". In *Contemporary Political Economy*, eds. D.A. Jr. Hibbs, H. Fassbender. Amsterdam: North-Holland Publishing Company, 121-36.

Rattinger, H. 1991. "Unemployment and Elections in West Germany". In *Economics and Politics: The Calculus of Support*, eds. H. Norpoth, M. Lewis-Beck, J. Lafay. Ann Arbor: Michigan University Press, 49-63.

Tomz, M., J. Tucker, J. Wittenberg. 2002. “An Easy and Accurate Regression Model for Multiparty Electoral Data”, *Political Analysis* 10:1(January): 66-83.

Veiga, F., L. Veiga. 2004a. "The Determinants of Vote Intentions in Portugal". *Public Choice* 118: 341-64.

Veiga, F., L. Veiga. 2004b. "Popularity Functions, Partisan Effects, and Support in Parliament". *Economics and Politics* 16:1(January): 101-15.

Veiga, F., L. Veiga. 2010. "The Impact of Local and National Conditions on Legislative Elections Results”. *Applied Economics* 42(13): 1727-34.