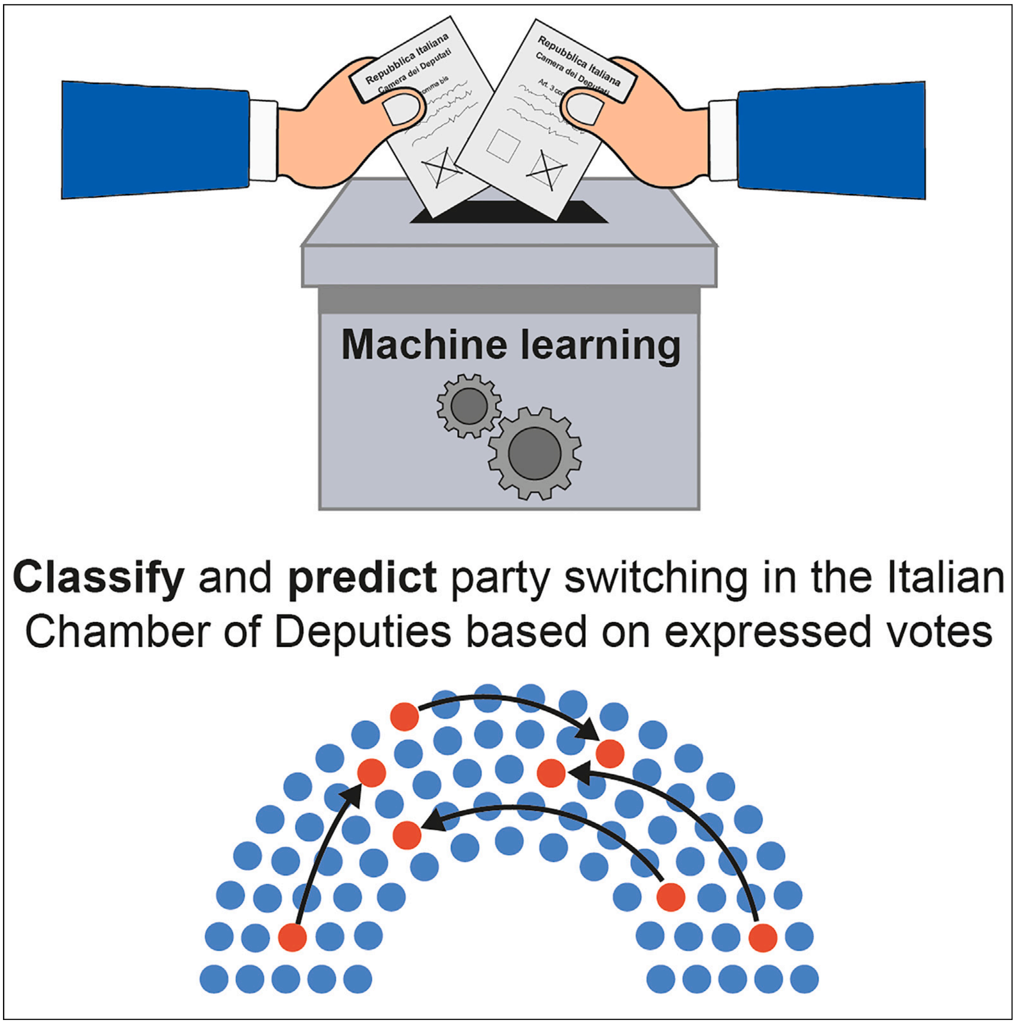
iScience

CellPress
OPEN ACCESS

Article

Predicting party switching through machine learning and open data



Nicolò Meneghetti, Fabio Pacini, Francesca Biondi, Marina Cracchiolo, Emanuele Rossi, Alberto Mazzoni, Silvestro Micera

silvestro.micera@epfl.ch

Highlights

Party switching is common in Italian Parliament, involving almost 1/3 of Deputies

A machine learning classifier accurately predicts party switchers from voting data

Party switchers tend to engage more frequently in secret ballots

Declining agreement with their party's majority predicts future switchers

Meneghetti et al., iScience *26*, 107098
July 21, 2023
https://doi.org/10.1016/j.isci.2023.107098

iScience

CellPress OPEN ACCESS

Article

Predicting party switching through machine learning and open data

Nicolò Meneghetti,[1,2] Fabio Pacini,[3,4] Francesca Biondi,[3] Marina Cracchiolo,[1,2] Emanuele Rossi,[3,5] Alberto Mazzoni,[1,2,7] and Silvestro Micera[1,2,5,6,7,8,*]

SUMMARY

Parliament dynamics might seem erratic at times. Predicting future voting patterns could support policy design based on the simulation of voting scenarios. The availability of open data on legislative activities and machine learning tools might enable such prediction. In our paper, we provide evidence for this statement by developing an algorithm able to predict party switching in the Italian Parliament with over 70% accuracy up to two months in advance. The analysis was based on voting data from the XVII (2013–2018) and XVIII (2018–2022) Italian legislature. We found party switchers exhibited higher participation in secret ballots and showed a progressive decrease in coherence with their party's majority votes up to two months before the actual switch. These results show how machine learning combined with political open data can support predicting and understanding political dynamics.

INTRODUCTION

Party switching refers to changes in party affiliation among elected legislators. This behavior is present in several democracies such as Denmark, the United States, Australia, Brazil, Ecuador, the European Parliament, France, Japan, Poland, and Italy.[1–3] Party switching is increasingly frequent in Italy,[1,4] and the Italian XVII legislature (from March 5th 2013 to March 22nd 2018) has indeed set the record for the number of members of parliaments (MPs) who switched their political affiliation: in Italy's lower house, the Chamber of Deputies, almost one out of three legislators changed their parliamentary group affiliation while serving their office.

Scholars have reported the detrimental effect of political party switching on parties' stability and, as a direct consequence, on democratic sustainability.[5] Floor crossings can indeed dramatically change the power balance within the Parliament. This can in turn result in the distortion of electoral results or weakened opposition parties, leading to excessive power concentration.[6] Efficiently predicting the likelihood of party switching by MPs would not only allow us to potentially know in advance if a given legislative proposal will receive expected support but it would also provide foresight on the overall stability of the political system.

Party switching was proposed to constitute a potential threat to public confidence in the transparency of the political system,[7] as it makes it difficult for the voters to navigate with ease among party labels in the search for ideal candidates. Given these premises, it is easy to appreciate how party switching can constitute a risky endeavor from the point of view of politicians. Voters do not appreciate it, as they were reported to penalize legislators who switch their parties.[2,8,9] As such, it is both legitimate and intriguing to inquire about the underlying human reasons of individual legislators driving the phenomenon of party switching.

Previous studies have explored various determinants of party switching, such as office-seeking motivations[7,10–12]—which proposes that MPs switch parties to gain more office perks within the parliament or the government, and policy-seeking motivations[13,14]—the assumption that MPs might change political party if this might push forward their political agenda. Other works investigated the features characterizing party switchers focusing instead on the political reasons leading to the switch,[3,4,7,10,14–19] often modeled within a multidimensional space of political opinions. In particular, a relevant previous study about Italian party switching[1] was based on a parametric political science approach, heavily reliant on political

[1]The Biorobotics Institute, Scuola Superiore Sant'Anna, 56025 Pisa, Italy

[2]Department of Excellence for Robotics and AI, Scuola Superiore Sant'Anna, 56025 Pisa, Italy

[3]Dirpolis Institute, Scuola Superiore Sant'Anna, 56025 Pisa, Italy

[4]Department of Linguistics, Literature, History, Philosophy and Law (DISTU), Tuscia University, 01100 Viterbo, Italy

[5]Health Interdisciplinary Center, Scuola Superiore Sant'Anna, 56025 Pisa, Italy

[6]Bertarelli Foundation Chair in Translational Neural Engineering, Center for Neuroprosthetics and Institute of Bioengineering, Ecole Polytechnique Federale de Lausanne, Lausanne, Switzerland

[7]These authors contributed equally

[8]Lead contact

*Correspondence: silvestro.micera@epfl.ch

https://doi.org/10.1016/j.isci.2023.107098

assumptions and considerations.[20–23] These heuristic models of political opinions are however prone to misinterpretations.[24]

In this paper, we investigated the mechanisms underlying party switching with a novel approach, which was developed to efficiently predict it with the largest possible advance, rather than focusing on its determinants. We extracted a dataset that includes the voting records and the personal information of each member of the Italian Lower House, the Chamber of Deputies (for details see STAR Methods). Then, combining information theory and classification algorithms, we tried answering three different questions: (i) can the voting features (e.g., the fraction of abstentions, of votes in favor or cast in secret ballots, and so on. See STAR Methods for the full list) discriminate between party switchers and single-party MPs? (ii) do all these voting features share the same contribution in the classification of these two groups of legislators? (iii) is it possible to predict whether a legislator will leave the party based on the past voting record? In other words, our goal was to use machine learning tools to analyze the vast amount of parliamentary data and to quantitatively determine whether party switching is a spontaneous and unpredictable event, or if it can be predicted based on specific past features of the deputies. Additionally, we aimed at identifying the most effective way to achieve such a prediction.

The work will start with a study of the prototypical features of party switchers in the XVII legislature of the Italian Chamber of Deputies. Then, we will move to the identification and evaluation of an algorithm predicting future party switches, which, differently from an *a posteriori* classification, could play a relevant role in political practices. In addition, we will test the generalizability of our approach also to the XVIII legislature. Finally, we will discuss how the properties of the designed algorithms can shed light on the reasons underpinning party switching.

RESULTS

Voting features differ between party switcher and single-party members of parliament

To investigate the features characterizing the MPs who changed their political party (party switchers, PS), we compared voting and non-voting features (see STAR Methods) of PS with those of single-party MPs. Specifically, we focused on those MPs who held a seat in the Chamber of Deputies, the lower house of the Italian Parliament. Therefore, when we refer to "MPs" throughout this work, we are exclusively referring to Deputies. We defined as party switchers those Deputies involved in at least one change of political group, but the analysis focused only on the first party switch. Further following changes of political affiliation were not considered. We extracted from the official open data platform of the Italian Chamber of Deputies the personal information and the votes each member has expressed during the XVII legislature (in office from March the 5th 2013 to March the 22nd 2018. See STAR Methods for details). During this term, and generally in contemporary Italy politics,[4] party switching was a frequent phenomenon: during the XVII legislature, indeed, 31% of MPs changed at least once their political group affiliation during their mandate (Figure 1A).

For each voting session in the Chamber of Deputies, we extracted seven features for each Deputy (see STAR Methods for details): (i) whether the MP was present or absent, (ii) whether the MP was present but abstained from voting, (iii) whether the vote was cast in a secret ballot, (iv, v) whether the MP voted in opposition or in favor, (vi) whether the MP voted in agreement with the final decision of the majority of the Chamber, and (vii) whether the MP voted in agreement with the majority of his/her/their party. The probability density functions of these features were then compared between PS and single-party MPs (Figure 1B). Median values were significantly different only for the following features: fraction of votes in agreement with own party's majority, of votes in favor of approved legislations, of abstentions from voting, and participation in secret ballots (Wilcoxon rank-sum test, alpha level set to 0.01). As some distributions appeared bimodal, we additionally performed Kolgomorov-Smirnov test on such voting pattern distributions. We found all seven voting feature distributions to be significantly different (Kolgomorov-Smirnov, alpha level set to 0.01) between PS and single-party MPs. Note that the distribution of votes in favor of approved legislations is bimodal in single-party MPs and unimodal in PS.

We additionally extracted for each MP non-voting features, i.e., independent from the votes expressed in the Chamber: birth region, election region, education level, gender, number of served terms, and birth year. Statistically significant differences between PS and single-party MPs were found in the election and birth region, the number of served terms, the gender, and the birth year (Figure 1C, Wilcoxon rank-sum,

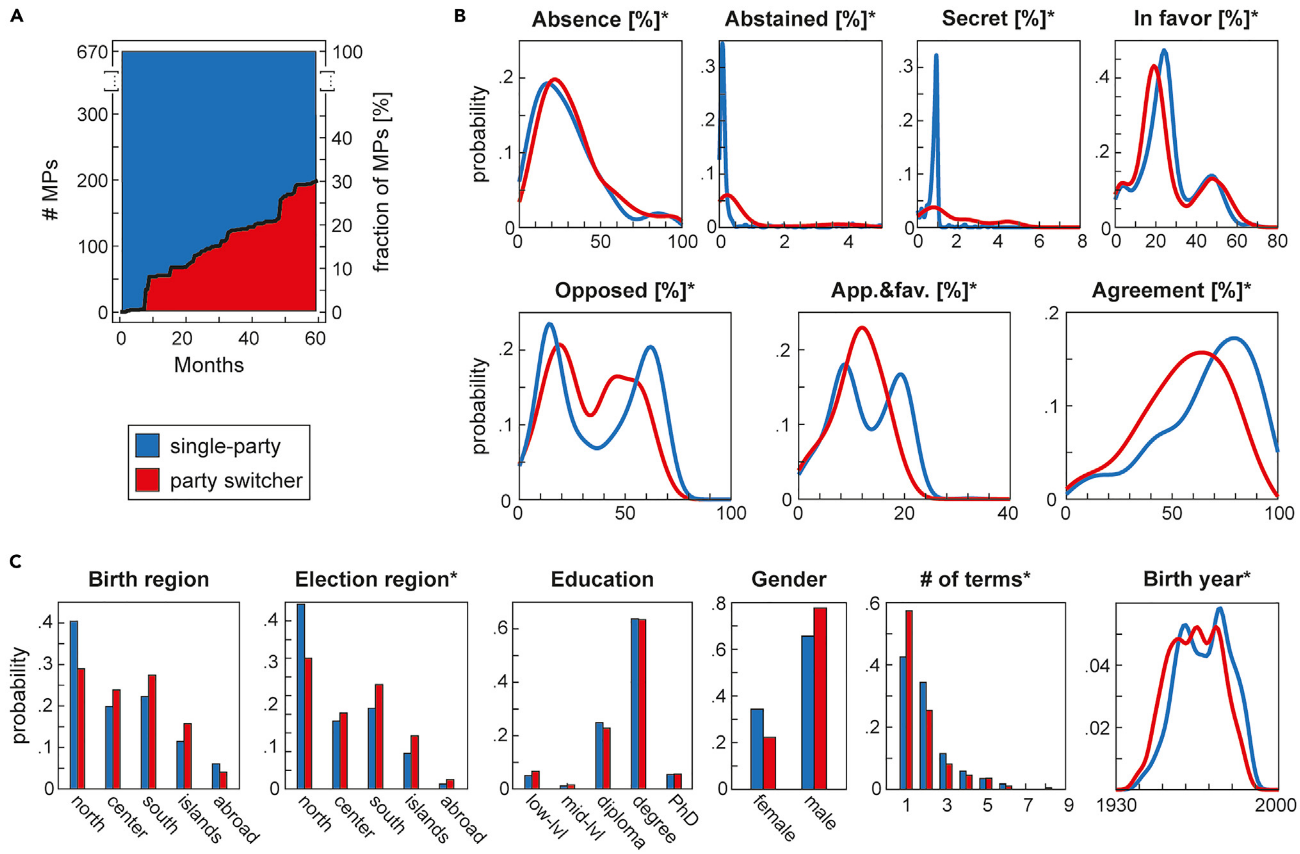


Figure 1. Voting and non-voting features of the parliamentarians of the Italian Chamber of Deputies in the XVII legislature

(A) Party switchers cumulative distribution across the XVII legislature, in the text referred to as *PS(t)*. The colored areas represent the fraction of single-party (blue) and party switcher (red) Deputies.

(B) Distribution (at the end of the XVII legislature) of voting features in single-party (blue) and party switcher (red) MPs. Asterisks indicate a significant difference in Wilcoxon rank-sum test (RS): absence RS: $p > 0.05$; agreement RS: $p < 0.001$; abstained RS: $p < 0.001$; Approved and in favor RS: $p = 0.008$; opposed RS: $p = 0.1$; in favor RS: $p = 0.02$; secret RS: $p < 0.001$. All distributions differed in terms of the Kolgomorov-Smirnov test (alpha level set to 0.01).

(C) Probability distribution of non-voting features in single-party (blue) and party-switcher (red) MPs. Asterisks indicate a significant difference in the Wilcoxon rank-sum test (RS) for the birth year and the chi-square for the other features (χ^2): birth region χ^2: $p = 0.03$; election region χ^2: $p = 0.01$; education χ^2: $p = 0.89$; gender χ^2: $p = 0.002$; number of terms χ^2: $p < 0.001$; birth year RS: $p < 0.001$. Election region, number of served terms, and birth year statistically differed in terms of the Kolgomorov-Smirnov test (alpha level set to 0.01).

and χ^2, alpha level set to 0.01). Kolgomorov-Smirnov found instead significant differences only among the election region, the number of served terms, and the age.

Of note, we found the results remained consistent if we considered only those MPs who changed their political affiliation in bulk (i.e., more than 10 in the same week, usually near to a government change or the formation of a novel party) or those that did it individually (see Figure S1). For this reason, we will merge these two families of party switchers for the remainder of this work.

A machine learning algorithm discriminates party switchers from single-party MPs

The presence of significant differences in voting and non-voting features between the two groups suggested the possibility to build a machine learning algorithm to discriminate between PS and single-party MPs (Figure 2A). To this aim, we implemented a random forest algorithm (see STAR Methods) to discriminate, at each instant within the legislature, the MPs who changed their party prior to the considered time point (PS(t), see Figure 1A) from those who did not (non-PS(t)). It is important to note that the non-PS(t) group includes both single-party MPs and those who will switch their political affiliation at some point following t. The input to the random forest algorithm to discriminate between PS(t) and non-PS(t) consisted of the 6 non-voting features, and the 8 voting features computed from the beginning of legislature until

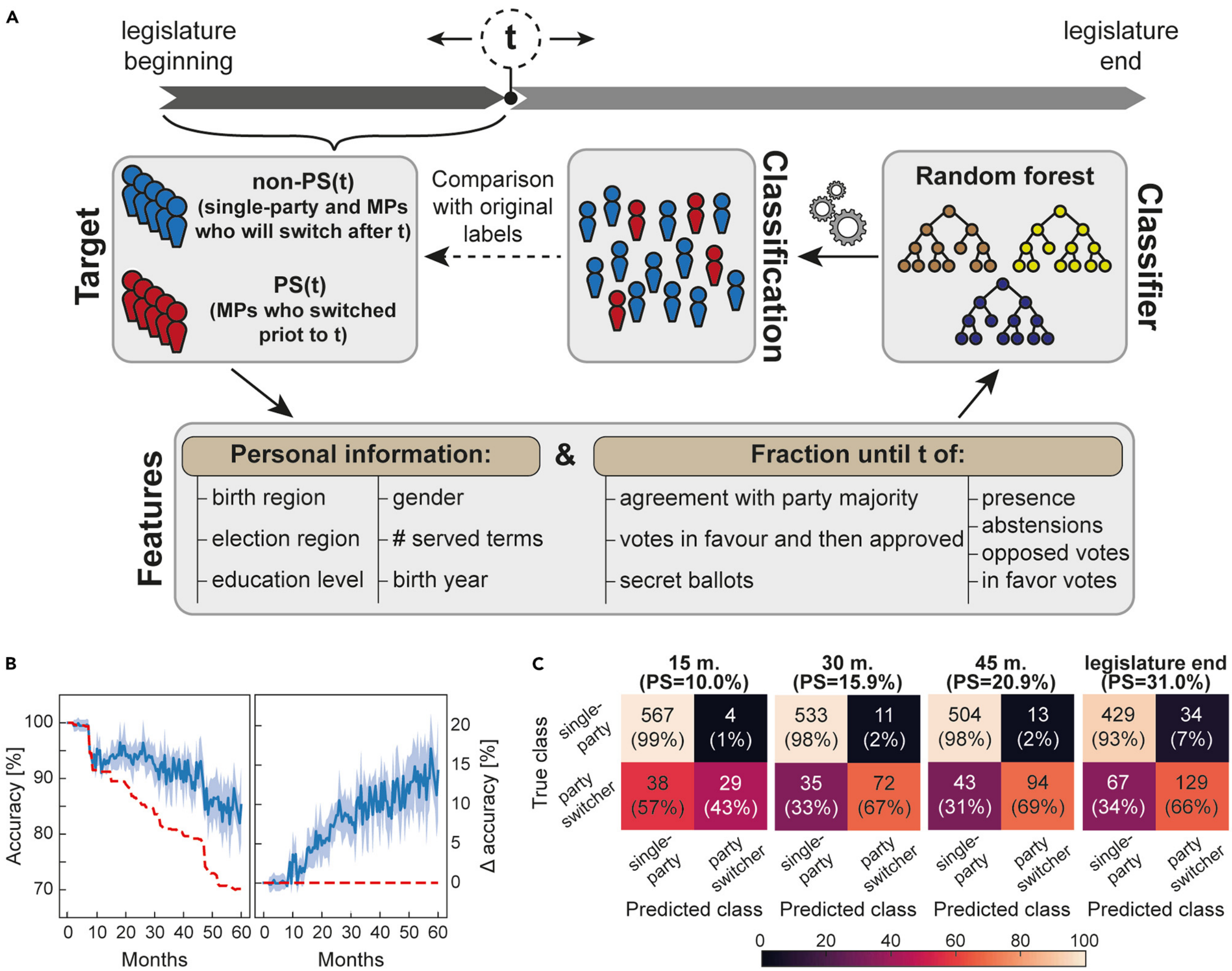


Figure 2. Classification of party switchers based on the past votes

(A) Schematic illustration of the classification algorithm. For each time *t* across the XVII Italian legislature, we collected the MPs who changed their political affiliation before *t*; the remaining ones were treated as single-party MPs. Between the beginning of the XVII legislature and time *t*, we gathered 6 non-voting and 7 voting features (See STAR Methods for details).

(B) (left) Classification accuracy throughout the legislature of PS(t) and non-PS(t). The dashed red line indicates chance level. Shaded regions indicate standard deviation of 5-fold cross-validation. (right) Classification accuracy is as in the left panel but expressed as an increase with respect to the chance level.

(C) Confusion matrices of the classification of single-party and party switcher Deputies across different months (m.) from the beginning of the XVII legislature.

time point t (Figure 2A). Please note that in those cases in which the considered time t extends beyond the party switch day for a given MP, only the votes until the party switch are considered.

At the beginning of the legislature, the two groups were very unbalanced (i.e., there are very few MPs who already changed their party), and the temporal window in which computing the average fraction of voting patterns (i.e., the average fraction of absent, opposed, absent votes, etc.) is too short to ensure robust discrimination. The classification accuracy became significantly higher than the chance level only for $t > 12$ months (t-test, $p < 0.05$ with Bonferroni correction), improving as the size of PS(t) increased along with the XVII legislature (Figure 2B. Linear regression on the Δ classification accuracy: slope = 0.27%/month, $p < 0.001$ two-sided t-test). The specificity of the party switchers reached almost 2 out of 3 correctly classified MPs during the last 6 months of the legislature (65.8 ± 1.7%, mean ± std, Figure 2C. "std" stands for standard deviation). The sensitivity of the single-party parliamentarians was instead higher, with an average value of 96.2 ± 5.7% throughout the whole legislature, mean ± std (Figure 2C). Accordingly,

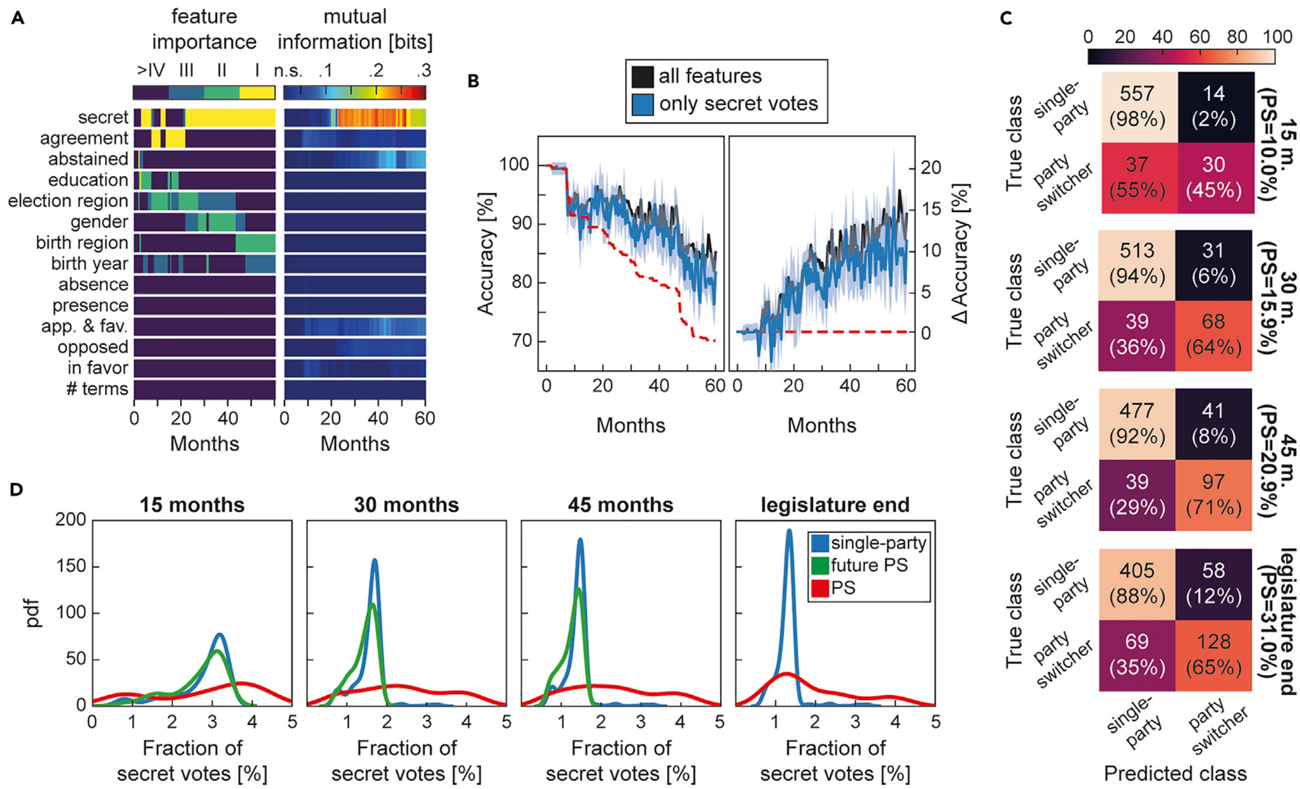


Figure 3. The fraction of secret voting is sufficient to discriminate between party switchers and single-party MPs

(A) (left) Specific feature relevance ranking in the classification of single-party and party switcher MPs (I first, II second, III third, and IV fourth). (right) Mutual information of the features. "n.s." stands for non-significant mutual information.

(B) (left) Classification accuracy throughout the legislature of single-party and party switcher MPs when employing all the features (black line) or only the fraction of secret votes (blue line). (right) Classification accuracy as in left but expressed as an increase with respect to chance level. Shaded regions indicate standard deviation of 5-fold cross-validation.

(C) Confusion matrices of the classification of single-party MPs and party switchers based solely on the fraction of secret votes across different months (m.) from the beginning of the XVII legislature.

(D) Distribution of the fraction of secret votes in single party MPs (blue line), party switchers (red line), and MPs who will change their party in future moments of the legislature (green line).

the F1-score remained above about 90% throughout the whole legislature (reaching its minimum of 89.25 $\pm$ 2.6%, mean $\pm$ std, at the end of the legislature, see Figure S2A).

Party switchers engage more frequently in secret voting

We investigated the relative relevance of every feature in the classification of party switchers. The maximum relevance minimum redundancy (*mrmr*) algorithm[25] (see STAR Methods) identified the fraction of votes in agreement with party majority as the most relevant feature for the classification task in the first 20 months, and the fraction of secret votes as the most relevant feature for the rest of the XVII legislature (Figure 3A left). Mutual information analysis (see STAR Methods) revealed, however, the level of agreement with the party majority to carry very limited encoding power (0.04 $\pm$ 0.01 bits, mean $\pm$ std) for the discrimination of the two MPs groups throughout the whole legislature. The fraction of secret votes was instead the feature carrying the largest information content after 20 months (0.21 $\pm$ 0.03 bits, mean $\pm$ std, for $t > 20$ months, Figure 3A right). Additionally, mutual information was also very low for non-voting features (0.004 $\pm$ 0.003 bits, mean $\pm$ std, Figure 3A right).

We then tested whether the information contained in the fraction of secret votes was significant enough to distinguish the two groups of Deputies by itself. We found that a random forest classifier trained solely on the fraction of secret votes performed similarly to the same classification algorithm but trained on all of the 14 features (Figure 3B, difference of accuracy [1.4 $\pm$ 3.2]%, mean $\pm$ std). The specificity of the PS was above

2 out of 3 correctly classified MPs during the last 6 months of the legislature ([67.5 $\pm$ 0.9]%, mean $\pm$ std), similar to the classifier trained with all the features (Figure 3C). The sensitivity, instead, slightly decreased as compared to the complete classifier: the rate of correctly classified single-party MPs had an average value of [91.7 $\pm$ 4.2]% (mean $\pm$ std) throughout the whole legislature (and dropped to [87.8 $\pm$ 0.8]%, in the last six months, mean $\pm$ std, Figure 3C). A similar trend was also observed when considering the F1-score (reaching at the end of the legislature a value of 86.2 $\pm$ 1.9%, mean $\pm$ std, see Figure S2B).

We found this classification performance to be explained by the fact that for $t > 20$ months, PS(t) participated progressively more frequently in secret ballots, as compared to single-party MPs (Figure 3D). Note that non-PS(t) includes both single-party MPs (i.e., those MPs who stuck with their political affiliation for the whole legislature) and PS($t^*>t$) (i.e., the "future switchers"): considering, however, these two groups separately did not affect the results (Figure 3D). Further information concerning the political role of secret ballots in Italian parliamentary politics can be found in Methods S1.

Voting features predict upcoming party switching

So far, we determined which features identified party switcher at a given time. We next investigated whether, and to what extent, the extracted features retained predictive power in the discrimination between single-party MPs and future party switchers. In other words, we explored the possibility of predicting switches occurring after the time *t* based on voting and non-voting features defined over time windows preceding *t*.

To test our features against this prediction task, we adopted the procedure illustrated in Figure 4A (see STAR Methods for details). Briefly, we paired each party switcher with a randomly selected (without replacement) single-party Deputy. For both components of each pair, we considered first the votes over the same 60 days prior to the switch of the PS component of the pair. The 60-day voting and non-voting features of every pair were then randomly assigned to the training or the test set with a probability of 0.5. We performed then a classification on this dataset of PS and single-party MPs with a random forest. Note that crucially we are now predicting the occurrence of a party switch only based on the MPs' non-voting features and the votes immediately preceding the switch itself (by no more than 60 days). We achieved a highly significant accuracy of [73.3 $\pm$ 3.7]% (Figure 4B left end and Figure 4C top left, t-test $p < 0.001$). This means that based on a 60-day observation window, our algorithm leads to a robust estimate of whether an MP is just about to change party. Additionally, we also found that setting the considered time window in number of voting sessions (as opposed to days) led to similar results in terms of prediction accuracy (see Figures S3A and S3B).

We furtherly accessed the relative contribution of voting and non-voting features in the prediction accuracy. Prediction accuracy of the random forest solely trained on voting patterns reached [71.1 $\pm$ 3.3]%, while it went down to [58.2 $\pm$ 6.5]% when it was exclusively dependent on non-voting features.

We assessed the possible contribution of votes occurring earlier than 60 days before the switch by repeating the prediction task while sliding the considered time window toward the past (i.e., moving away from the party switch). As expected, the accuracy of the classifier trained only on non-voting features did not display any significant temporal structure (Figure 4B, augmented Dickey-Fuller test, $p < 0.001$), as these features were time-invariant. The accuracy of the classification algorithm trained on both voting and non-voting features or exclusively on voting features was stable up to the [-4, -2] months window, indicating the robustness of the results. The prediction accuracy decreased only as the considered window was moved further away from the party switch (Figure 4B). Specifically, as we moved the considered time window away from the party switch by 6 months, the accuracy of the algorithm trained on the expressed votes approached the accuracy of the algorithm trained on the non-voting features (mean accuracy of non-voting features across the whole legislature: [57.6 $\pm$ 6.7]%; mean accuracy of voting features after moving the considered window of 6 months before party switching: [62.6 $\pm$ 4.5]%, mean $\pm$ std). Interestingly, the classifier adopting either voting, non-voting, or both features never reached the random guess level of 50%, hence indicating a residual difference between single-party MPs and party switchers that did not depend on the timing of party leave.

The fraction of votes in agreement with the party majority predicts party switch

As done for the classification algorithm, we then investigated the relative importance of every feature for the prediction task. The *mrmr*) algorithm found that the fraction of votes in agreement with the party majority was the most relevant feature for the prediction. This feature's predictive power, however, decreased as the 60-day window was moved to over 4 months before the party switch (Figure 5A left). When the window was

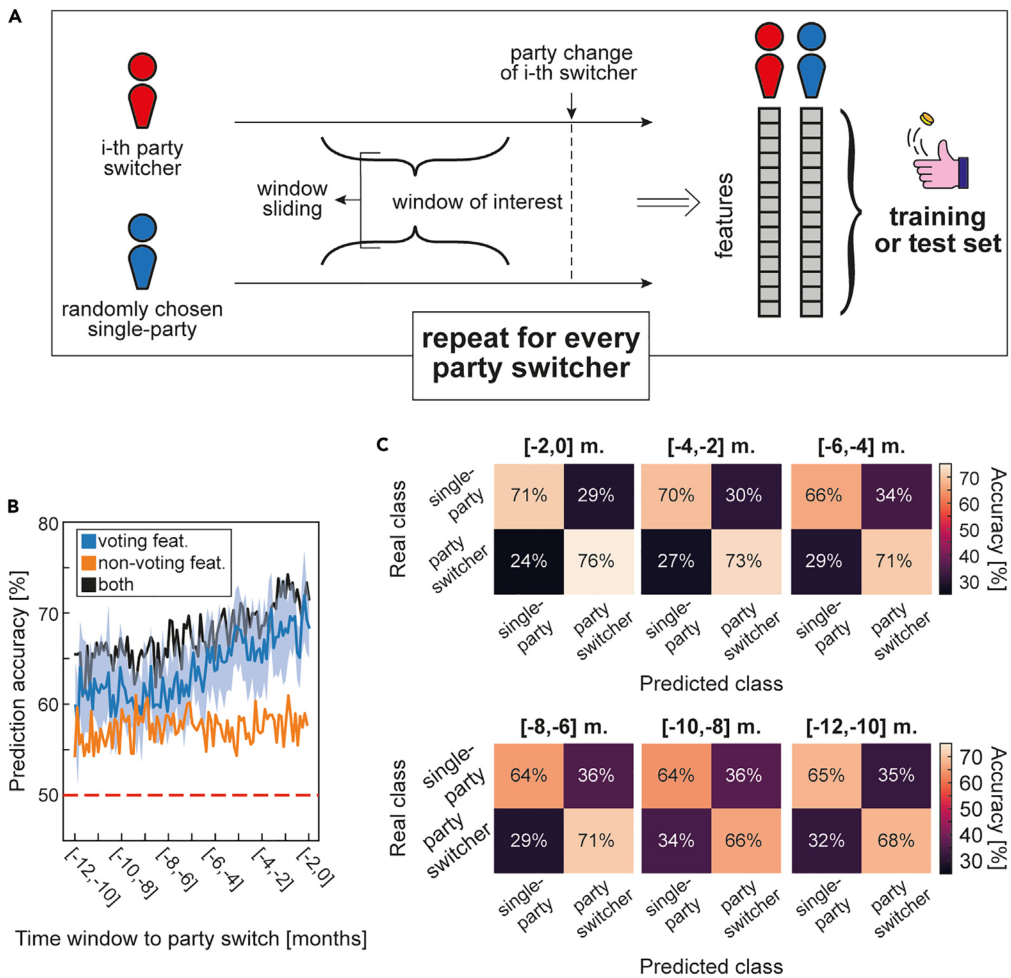


Figure 4. Prediction of party switching in the Italian Chamber of Deputies

(A) Schematic illustration of the pipeline used for predicting party switching. Briefly, we randomly (without replacement) coupled a single-party MP with a PS. We considered their votes expressed within 60 days prior to the party change of the party switcher. The resulting couple was then randomly assigned to the training or test set with a probability of 0.5. This process was repeated while sliding the 60-day window toward the past (see STAR Methods for full details).

(B) Prediction accuracy of single-party and party switcher MPs extracting voting information from 60-day windows set at different times before the party switch. The algorithm was trained on only voting (blue line), non-voting (orange line), or both (black line) set of features. The dashed red line indicates the random guess level. Mean accuracy levels until within a month to the date of party change were: (i) voting: [70.1 ± 3.5]%; non-voting: [58.6 ± 7.2]%; both: [72.2 ± 3.9]%. Mean ± std. Maximum accuracy levels were: (i) voting: [71.6 ± 4.6]%; non-voting: [60.1 ± 7.6]%; both: [73.7 ± 4.6]%. Max ± std. Shaded regions indicate standard deviation of 5-fold cross-validation.

(C) Confusion matrices of the prediction task described in B.

moved further back in time, the *mrmr* algorithm relied instead on non-voting features, such as the birth year and the number of served terms (associated with lower decoding performance). Mutual information analysis (Figure 5A right) revealed the non-voting features to carry very limited information about the prediction task. The fraction of votes in agreement with the party majority within 4 months prior to party switching carried, instead, the larger information content for the prediction task, in agreement with the *mrmr* algorithm.

We consequently tested whether the information contained in the fraction of votes in agreement with the party majority was sufficient to predict future switches. We found that the temporal evolution of the median of the votes in agreement of PS began to drop below the single-party MPs approximately 6 months prior to party change (Figure 5B).

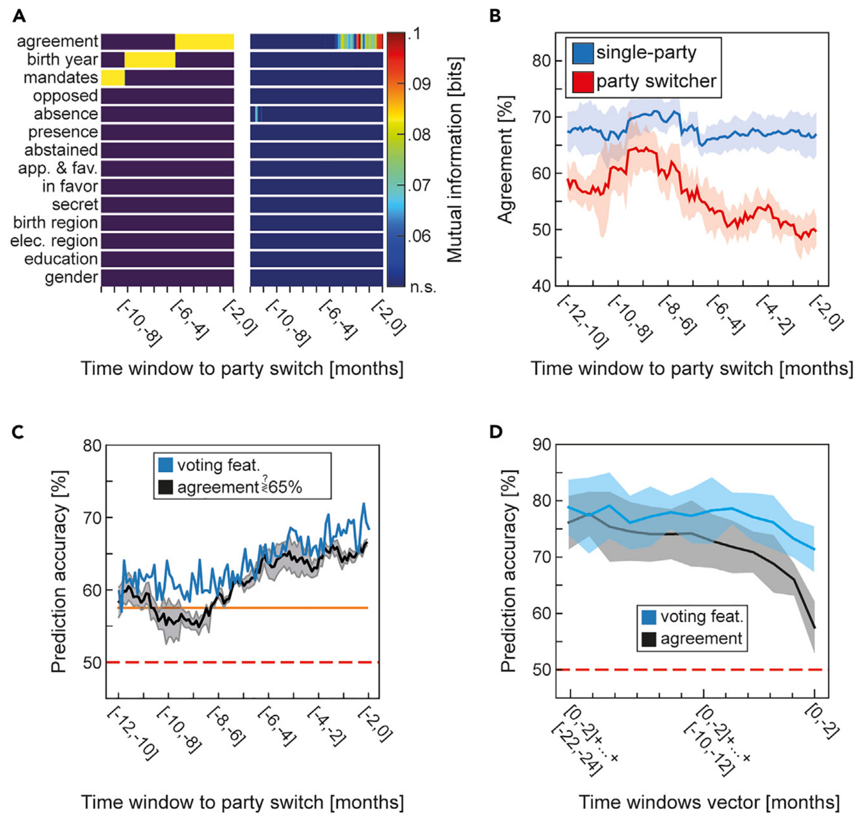


Figure 5. Prediction of party switching employing the fraction of votes in agreement with the party majority

(A) (left) Most important feature (in yellow) as computed by the *mrmr* (maximum relevance minimum redundancy) algorithm in the prediction of single-party and party switcher MPs. (right) Mutual information of the features. "n.s." stands for non-significant mutual information.

(B) Median of the fraction of votes in agreement with the party majority of single-party (blue line) and party switcher MPs (red line) (computed in a 60-day window) as a function of the translation backward in time to the party switch. The shaded region indicates confidence interval.

(C) Prediction accuracy of single-party and party switcher using voting information extracted from a 60-day window. The prediction was computed as a function of the translation of said window, starting from the party switch day and proceeding backward in time. We adopted two algorithms: (i) random forest trained on every voting feature (blue line); (ii) a simple binary yes-no question: "is the fraction of votes in agreement with party majority greater than 65%? If yes: single-party MP; else: party switcher MP" (black line). Mean accuracy levels within one month to the date of party change based only on the simple rule was [64.1 $\pm$ 0.6]%.

(D) Prediction based on concatenated values of features observed over different and non-overlapping intervals of time. The rightmost point indicates the prediction accuracy of PS and single-party MPs considering all the voting features (blue trace) or just the level of agreement (black trace) in a 60-day window preceding the party switch (i.e., [-2,0]). Proceeding toward the leftmost point, each x axis tick represents the prediction accuracy when consistently concatenating the features extracted from non-overlapping 60-day windows moving away from the party switch event. The procedure continues until the leftmost point in which the investigation starts from 24 months before the party switching day. The dashed red line indicates the chance level of 50%. The shaded regions indicate standard deviation.

A simple classification rule ("Is the fraction of votes in agreement on a 60-day window lower than 65%? Then the MP is a party switcher") performed similarly to the random forest algorithm trained on the whole set of voting features, reaching a prediction accuracy of [65.6 $\pm$ 0.3]% (Figure 5C, difference of accuracy [4.4 $\pm$ 1.7]%, mean $\pm$ std).

As the fraction of votes in agreement brought information even two months ahead of the switch, we wondered if considering a longer history in voting could improve our prediction algorithm. We developed then a second algorithm based on the whole previous history up to a given starting time, with features again integrated over two months bins. We found that, when considering all features, the maximum prediction accuracy was [79.8 $\pm$ 5.7]% (against [73.7 $\pm$ 4.5]% considering only the two months preceding the switch) and was achieved considering a history of 6 months, after which no further significant improvement was achieved (Figure 5D, black line). Interestingly, very similar values of [77.7 $\pm$ 5.6]% were achieved by exclusively considering the temporal evolution of the agreement level between the Deputy and his/her/their party majority, although this required considering 12 months of votes (Figure 5D, blue line).

Note that it is important to take into account previous history as a vector with the evolution in time of the features, as simply integrating over longer time windows does not improve the prediction accuracy (Figure S3).

Application of party switchers classification to the XVIII legislature

We examined the generalizability of the results obtained from the XVII legislature. We found the following one, the XVIII legislature, to be suitable for this test as, interestingly, party switching played a quantitatively similar role: 32% of the deputies changed at least once their parliamentary group affiliation during their office in the XVIII legislature (Figure S4A).

Following the same procedure outlined for the previous legislature (see STAR Methods), we extracted the voting and non-voting features described above also for the XVIII legislature of the Italian Chamber of Deputies (in office from March 23rd 2018 to October 12th 2022 – see Figures S4B and S4C for the distribution of the features at the end of the legislature).

Following the pipeline outlined in Figure 2A, we fed these features into a random forest to discriminate party switcher (PS(t)) from single-party (non-PS(t)) MPs. Similar to the previous legislative term, we found classification accuracy to increase along the XVIII legislature, becoming significantly higher than the chance level for $t > 16.5$ months (Figures 6A and 6B t-test, $p < 0.05$ with Bonferroni correction). Classification specificity of party switcher MPs reached 66.6 $\pm$ 1.5% during the last 6 months of the legislature, while single-party sensitivity averaged 98.8 $\pm$ 1.7% in the same period (Figure 6C).

We accessed the importance played by each feature in this classification. The *mrmr* algorithm in the first half of the XVIII legislature identified as the most relevant features the fraction of secret votes, the level of agreement with the party's majority, and the fraction of abstentions from voting. Mutual information revealed, however, these features to retain a limited amount of encoding power in this time frame (Figure 6C. For $t < 25$ months: agreement MI: 0.007 $\pm$ 0.008 bits; secret MI: 0.005 $\pm$ 0.012 bits; abstentions MI: 0.008 $\pm$ 0.010 bits). In the second half of the XVIII legislature, instead, both *mrmr* and mutual information identified the fraction of votes cast in secret ballots as the foremost significantly important feature (for $t > 25$ months secret votes MI: 0.21 $\pm$ 0.03 bits). Accordingly, training the classification algorithm solely on secret voting yielded comparable results to the one instructed on the whole set of voting and non-voting features, both in terms of classification accuracy (difference of accuracy was 0.2 $\pm$ 1.8% throughout the whole XVIII legislature, Figure 6D), specificity (67.3 $\pm$ 3.4% during the last 6 months of the legislature, Figure 6E), and sensitivity (97.6 $\pm$ 1.9% during the last 6 months of the legislature, Figure 6E).

Similar to what was observed for the XVII, the importance retained by the fraction of secret voting in the XVIII legislature could be explained by the preference of party switchers to cast their votes in secret ballots as compared to their single-party colleagues (Figure 6F).

Application of party switching prediction to the XVIII legislature

Following the same procedure outlined in Figure 4A, we applied the prediction workflow to the XVIII legislature. In the XVII legislature, we showed how prediction accuracy decreased as the considered time window in which information was extracted was moved further away from the party switch. In the XVIII legislature, instead, we found the prediction accuracy to retain little or no temporal structure in relation to the party switch (Figure 7A, augmented Dickey-Fuller test, $p < 0.05$). Accordingly, both the *mrmr* algorithm and mutual information identified the MPs' birth regions as the most relevant feature for the prediction (Figure 7B), regardless of how distant the 60-day window was set from the party switch. The selection as the foremost relevant feature of a non-voting (hence time-invariant) personal characteristic explains the aforementioned lack of temporal patterns observed in the prediction accuracy.

We found the importance covered by this feature to be due to the asymmetric distribution of the birth region of the two MPs groups across the Italian territory: party switchers were indeed more frequently born in

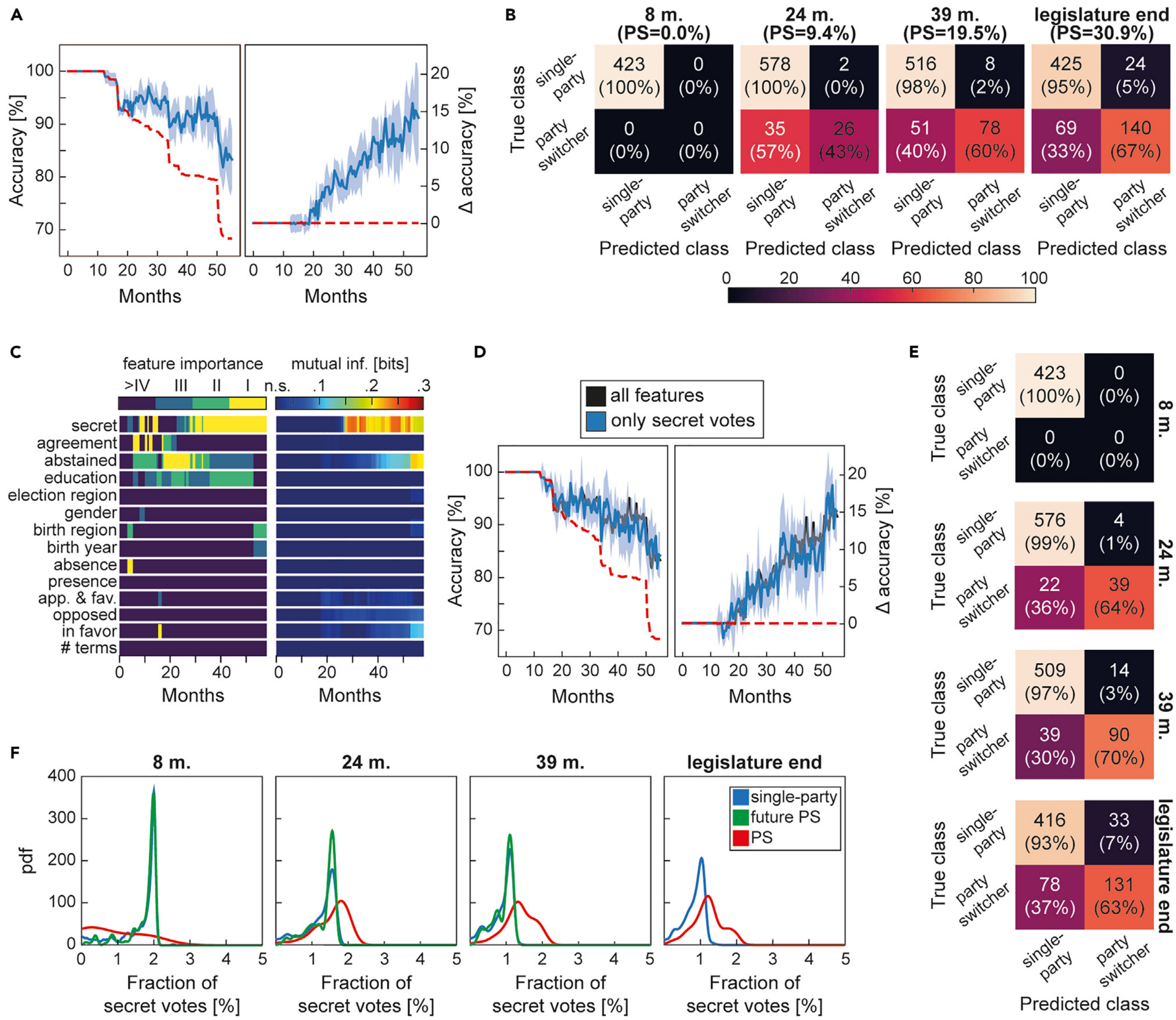


Figure 6. Classification of party switchers based on the past votes in the XVIII legislature of the Italian Chamber of Deputies

(A) (left) Classification accuracy throughout the XVIII legislature of PS(t) and non-PS(t). The dashed red line indicates the chance level. Shaded regions indicate standard deviation of 5-fold cross-validation. (right) Classification accuracy as in left but expressed as an increase with respect to chance level.

(B) Confusion matrices of the classification of party switcher and single-party MPs at different months (m.) from the beginning of the XVIII legislature.

(C) (left) Specific feature relevance ranking in the classification of single-party and party switcher Deputies (I: first, II: second, III: third, and IV: fourth). (right) Mutual information of the features in bits. "n.s." stands for non-significant mutual information (see STAR Methods).

(D) (left) Classification accuracy throughout the legislature of single-party and party switcher MPs when training the random forest classifier on all the features (black line) or only the fraction of secret votes (blue line). (right) Classification accuracy as in left but expressed as increase with respect to chance level. Shaded regions indicate standard deviation of 5-fold cross-validation.

(E) Confusion matrices of the classification of single-party and party switcher Deputies based exclusively on the fraction of secret votes across different months (m.) from the beginning of the XVIII legislature.

(F) Probability density function of the fraction of secret votes in single-party MPs (blue line), party switchers (PS, red line), and MPs who will change their party in future moments of the legislature (future PS, green line).

southern regions and on islands (Figure 7C). A such peculiar imbalance was explained by considering the political events that characterized the XVIII legislature. The party of relative majority "M5S" (Figure 7D) underwent a radical transformation during the XVIII legislature, resulting in a disproportionally high number of defections: the political party distribution of party switchers was accordingly biased toward "M5S" (Figure 7E). Incidentally, the birth region of M5S MPs was also disproportionate toward southern regions and

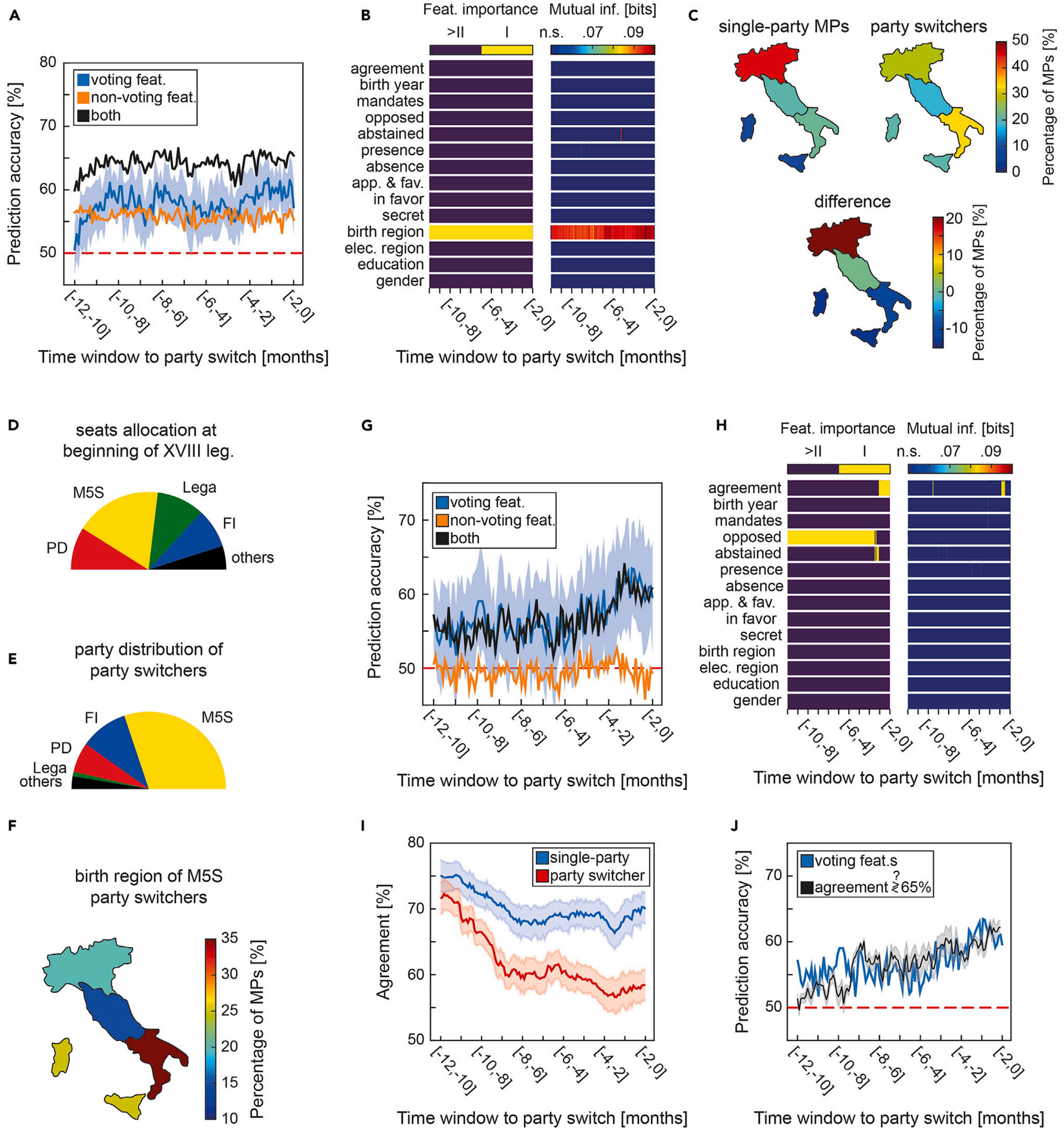


Figure 7. Prediction of party switching in the XVIII legislature of the Italian Chamber of Deputies

(A) Prediction accuracy of single-party and party switcher MPs (extracting voting information from 60-day windows) as a function of different times before the party switch. The algorithm was trained using voting (blue line), non-voting (orange line), and both (black line) sets of features. The dashed red line indicates the random guess level. Shaded regions indicate standard deviation of 5-fold cross-validation.

(B) (left) Most relevant feature (yellow) in the prediction of party switcher and single-party MPs identified by the *mrmr* algorithm. (right) Mutual information of the features in bits. "n.s." stands for non-significant mutual information (see STAR Methods).

(C) Distribution of the Italian birth region (grouped into north, center, south, and islands) in the XVIII legislature of single-party Deputies (upper-left map), party switchers (upper-right map), and their difference (bottom map).

(D) Distribution of seat allocation at the beginning of the XVIII legislature of the Italian Chamber of Deputies. From left to right: "Partito Democratico" (PD, red); "Movimento 5 Stelle" (M5S, yellow); Lega (green); FI ("Forza Italia", blue).

Figure 7. *Continued*

(E) Distribution of party switchers' original political party.

(F) Distribution of the Italian birth region (grouped into north, center, south, and islands) in the XVIII legislature of M5S party switcher Deputies.

(G) Same as in A) but excluding all the Deputies elected with the M5S.

(H) Same as in B) but excluding all the Deputies elected with the M5S.

(I) Median value of the fraction of votes in agreement with the party majority of single-party (blue line) and party switcher (red line) MPs as a function of the translation backward in time to the party switch. The shaded region indicates the confidence interval. Please note that the Deputies elected with the M5S were discarded from this analysis.

(J) Prediction accuracy of single-party and party switcher MPs (excluding all the Deputies elected with the M5S). The prediction was computed with two algorithms: (i) random forest as in G) (blue line); (ii) a simple binary yes-no question: "is the fraction of votes in agreement with party majority greater than 65%? If yes: single-party MP; else: party switcher" (black line). Shaded region indicates the standard deviation of prediction accuracy.

islands (Figure 7F). In a nutshell, therefore, the prediction algorithm employed the birth region as a proxy for identifying M5S Deputies.

We tested whether the M5S unbalanced contribution to the phenomenon of party switching was idiosyncratic to the XVIII legislature. We repeated the prediction workflow outlined before without considering all the Deputies elected with M5S. Similar to the previous legislature, we found the prediction accuracy to reach a level of (61.9 $\pm$ 2.4)% and then decreased as the window moved further away from the party switch (Figure 7G). Interestingly, we found the prediction accuracy obtained by exclusively using non-voting features to be indistinguishable from chance level (49.8 $\pm$ 1.4%, $p < 0.01$ t-test): this was indicative that eliminating M5S Deputies was able to dismiss the contribution of the birth region in the discrimination of single-party and party switcher MPs.

After having discarded M5S MPs, the *mrmr* algorithm found that the level of agreement with the party majority was the most relevant feature for the prediction until the 60-day window was moved to 1 month prior to the party switch (Figure 7H left). When the window was moved further back in time, the *mrmr* algorithm relied on other voting features, such as the fraction of opposed votes and the fraction of abstentions. Mutual information revealed, however, that these two features retained non-significant decoding roles (Figure 7H, right). The fraction of votes in agreement with the party majority (only within 1 month prior to party switching) was the only feature retaining significant information, in agreement with the *mrmr* algorithm.

We consequently tested to what extent the information contained in the fraction of votes in agreement with the party majority was sufficient to predict future switches. We found that as party change drew nearer, the median of votes in agreement cast by PS diverged below the single-party MPs (Figure 7I). Accordingly, we found a simple classification rule ("Is the fraction of votes in agreement on a 60-day window lower than 65%? Then the MP is a party switcher") performed similarly to the more sophisticated random forest algorithm trained on the whole set of voting features, reaching a prediction accuracy of [63.8 $\pm$ 0.3]% (Figure 7J, difference of accuracy [2.1 $\pm$ 1.5]%, mean $\pm$ standard deviation).

DISCUSSION

We presented a simple machine learning approach able to predict MPs' voting dynamics based on the availability of past votes data. We found that the information about future switches could be extracted within two to six months earlier than the actual switch. In the XVII legislature, the likelihood of predicting the switch was 20% above chance considering all the available features. The strongest determinant was the agreement level between the votes expressed by the MPs with the majority of their political group. Considering only this feature led to an estimate of the likelihood of party switch 15% more accurate than chance. This dynamic held (to a lesser extent) also for the XVIII legislature, although superimposed on the dynamics of the breakdown of a party, which also led to several party switches. These results suggest that relevant information about future party switches of individual MPs can be extracted by a simple analysis of the temporal pattern of their votes.

Crucially, previous studies[1,3,4,7,10,14–19] did not address the prediction of future switches but focused only on identifying features that distinguish party switchers from single-party MPs (as we did in Figures 1, 2, 3, and 6). We designed a prediction algorithm both because we aimed at a statistical tool potentially useful for political practices, and also because we hypothesized that the stronger constraints of this kind of algorithm (e.g., focusing on the time immediately preceding the switch) could also provide more information on the actual causes of the switch. Another major difference with the aforementioned works is that we followed a "hypothesis-free" approach, developing algorithms able to classify/predict switches without any assumption about the reasons leading to the switch. Coherently, the dataset collected about the Deputies in our work is made up of features that form the direct readout of their parliamentary activity, without any heuristic consideration of the political, institutional, or economic background. In this sense, this work did not rely on any composite variable highly dependent on some underlying hypothesis. This "hypothesis-free" non-parametric approach has two very important features: (i) it might be also applicable to other parliamentary democracies; (ii) it allowed us to highlight the (unexpected) relevance of participation in secret ballots (see Methods S1). Moreover, previous studies did not identify these features with a machine learning approach but with a linear proportional hazards regression model,[26,27] which requires the independence of survival times between the distinct individual in the sample, a very strong (and questionable) assumption when applied to the party switching of different MPs.

The possibility of predicting switches suggests that selected voting features are able to capture the dynamics leading to party switching. In particular, the slow but steady decrease in the level of voting in agreement with the party majority suggests that policy motivation is the most prominent determinant behind party switching, coherently with previous studies.[28,29] Indeed, it would be counter-productive for office-seeking MPs to express their disagreement months before deciding to leave it. The presence of the drop in agreement level might display instead an index of increasing ideological divergence between the PS and their (soon-to-be-ex) parties. Accordingly, once it drops below a certain threshold, its natural consequence is a change in political affiliation. Interestingly, also the prominence of secret votes in party switcher can be more easily interpreted as a way to push a political agenda, even when in disagreement with the MP's party, rather than as an office-seeking strategy, as only the former benefits from anonymity.

The efficacy of the proposed algorithms relies on the selection of the relevant features (and time window) for the classification and prediction of party switches rather than algorithmic optimization, as we heuristically selected a single machine learning algorithm, i.e., random forest. Despite being characterized by a simple approach, random forest is well suited for our particular datasets due to its ability to deal simultaneously with categorical and numerical variables and to its robustness against lack of balance and outliers. Accordingly, a survey of classification accuracy levels on the XVII legislature across different classification algorithms (specifically K Neighbors; Support Vector Classification; Gaussian Process; Decision Tree; Random Forest; Multi-Layer Perceptron; AdaBoost; Gaussian Naive Bayes; Quadratic Discriminant Analysis. See Figure S2C) revealed the random forest as being one of the best in terms of accuracy levels (Figure S2C). We want however to stress that we implemented this comparative test by using the default hyperparameters provided by the scikit-learn library implementation.[30] It is therefore plausible that hyperparameter optimization might yield better classification performance. This algorithmic analysis goes however beyond the scope of the current work, and it would be in our opinion best suited for follow-up methodological studies. We indeed reiterate that goal of this work is to prove the possibility of employing machine learning tools to shed light upon Parliamentarians' voting dynamics, regardless of the specific algorithm employed or the accuracy levels obtained. Note, furthermore, that the assessment of the features' importance relied also on information theory, a non-parametric approach that gives an upper bound to decoding performances and, most importantly, does not depend on the choice of the algorithm.[31]

The robustness of our approach could be further tested by extending it to other datasets as the other Chamber of the Italian Parliament, i.e., the Senate. Moreover, our algorithm might be extended by adding to the information about the individual MP information about the parties' dynamics (heterogeneity of each party's votes, numerosity of the party, the fraction of MPs that have already left the party, and so on). Future works might also involve extending (and possibly adapting) the presented algorithms to predict party switching in other Parliaments (such as the European Parliament or the Japanese Parliament).

Finally, we would like to highlight that the presented approach could be used to tackle different issues in parliaments in which party switching is not a common event, for instance, to predict whether MPs will vote coherently with their party on a specific topic.

Limitations of the study

Although the methodological framework outlined in this study has the potential to be applied as is to other parliamentary democracies, it is worth mentioning that our findings are solely focused on the Italian Parliament. Therefore, results and conclusions concerning the nature of party switching obtained in this study may not be generalizable to other parliamentary democracies.

We decided to investigate the determinants of party switching with a ''party-free'' approach. This means that we opted not to include any information concerning parliamentary party system. This choice allowed for greater generalizability of our results since we did not rely on any parametric political model that might not be applicable outside the Italian Parliament. However, it is worth noting that including party information (such as party labels before and after the switch or the ideological heterogeneity of the origin/destination party) could improve the accuracy of classification/prediction and provide a better understanding of how party organization influences party switching.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- METHOD DETAILS
  - Data collection
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Classification of party switchers and single-party MPs
  - Feature importance
  - Prediction

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2023.107098.

ACKNOWLEDGMENTS

N.M. and A.M. were supported by the Italian Ministry of University and Research (MIUR) PRIN2017, PROTECTION, Project 20178L7WRS. S.M. was supported by the Bertarelli Foundation and by the Italian Ministry of University and Research (MUR) PRIN-2017 INSPECT 2017JPMW4F. We thank the three anonymous reviewers for their helpful comments that have improved this paper. The authors would also like to extend their gratitude to Professor Nicola Giocoli for his valuable advice.

AUTHOR CONTRIBUTIONS

Conceptualization, F.P., F.B., E.R., A.M., S.M.; Methodology, N.M., M.C., A.M., S.M.; Investigation, N.M., M.C., A.M.; Writing – Original Draft, N.M., A.M.; Writing – Review & Editing, N.M., F.P., F.B., M.C., E.R., A.M., S.M., Supervision, E.R., A.M., S.M.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: December 13, 2022
Revised: May 16, 2023
Accepted: June 7, 2023
Published: June 14, 2023

REFERENCES

1. Di Virgilio, A., Giannetti, D., and Pinto, L. (2012). Patterns of party switching in the Italian Chamber of Deputies 2008-2011. Riv. Ital. Sci. Polit. 29–58. https://doi.org/10.1426/36739.
2. Bengtson, A. (2022). The Morality of Party Switching. J. Appl. Philos. *39*, 616–630. https://doi.org/10.1111/japp.12584.
3. Mershon, C., and Shvetsova, O. (2008). Parliamentary Cycles and Party Switching in Legislatures. Comp. Polit. Stud. *41*, 99–127. https://doi.org/10.1177/0010414007303651.
4. Heller, W.B., and Mershon, C. (2005). Party Switching in the Italian Chamber of Deputies, 1996–2001. J. Polit. *67*, 536–559. https://doi.org/10.1111/j.1468-2508.2005.00328.x.
5. Knott, C. (2017). The Cross-National Determinants of Legislative Party Switching. https://doi.org/10.31390/gradschool_theses.4501.
6. Rakner, L., and SvÂsand, L. (2004). From Dominant to Competitive Party System: The Zambian Experience 1991–2001. Party Polit. *10*, 49–68. https://doi.org/10.1177/1354068804041316.
7. Desposato, S.W. (2006). Parties for Rent? Ambition, Ideology, and Party Switching in

Brazil's Chamber of Deputies. Am. J. Pol. Sci. *50*, 62–80. https://doi.org/10.1111/j.1540-5907.2006.00170.x.

8. Grose, C.R., and Yoshinaka, A. (2003). The Electoral Consequences of Party Switching by Incumbent Members of Congress, 1947-2000. Legis. Stud. Q. *28*, 55–75. https://doi.org/10.3162/036298003X200809.

9. Kerevel, Y.P. (2017). The Costs and Benefits of Party Switching in Mexico. Lat. Am. Polit. Soc. *59*, 28–51. https://doi.org/10.1111/laps.12009.

10. Giannetti, D., and Laver, M. (2001). Party system dynamics and the making and breaking of Italian governments. Elect. Stud. *20*, 529–553. https://doi.org/10.1016/S0261-3794(00)00048-2.

11. Laver, M., and Kato, J. (2001). Dynamic approaches to government formation and the generic instability of decisive structures in Japan. Elect. Stud. *20*, 509–527. https://doi.org/10.1016/S0261-3794(00)00047-0.

12. McElroy, G., and Benoit, K. (2010). Party Policy and Group Affiliation in the European Parliament. Brit. J. Polit. Sci. *40*, 377–398. https://doi.org/10.1017/S0007123409990469.

13. Kato, J., and Yamamoto, K. (2009). Competition for Power: Party Switching and Party System Change in Japan. In Political Parties and Legislative Party Switching, W.B. Heller and C. Mershon, eds. (Palgrave Macmillan US), pp. 233–263. https://doi.org/10.1057/9780230622555_9.

14. Mcmenamin, I., and Gwiazda, A. (2011). Three roads to institutionalisation: Vote-office- and policy-seeking explanations of party switching in Poland: three roads to institutionalisation. Eur. J. Polit. Res. *50*, 838–866. https://doi.org/10.1111/j.1475-6765.2010.01985.x.

15. Desposato, S. (2009). Party Switching in Brazil: Causes, Effects, and Representation. In Political Parties and Legislative Party Switching, W.B. Heller and C. Mershon, eds. (Palgrave Macmillan US), pp. 109–144. https://doi.org/10.1057/9780230622555_5.

16. Heller, W.B., and Mershon, C. (2008). Dealing in Discipline: Party Switching and Legislative Voting in the Italian Chamber of Deputies, 1988-2000. Am. J. Pol. Sci. *52*, 910–925. https://doi.org/10.1111/j.1540-5907.2008.00350.x.

17. Heller, W.B., and Mershon, C. (2009). Political Parties and Legislative Party Switching, 1st ed. (Palgrave Macmillan).

18. Laver, M., and Benoit, K. (2003). The Evolution of Party Systems between Elections. Am. J. Polit. Sci. *47*, 215–233. https://doi.org/10.1111/1540-5907.00015.

19. Müller, W.C., and Strøm, K. (1999). Policy, Office, or Votes?: How Political Parties in Western Europe Make Hard Decisions 1st ed. (Cambridge University Press). https://doi.org/10.1017/CBO9780511625695.

20. Downs, A. (1957). An Economic Theory of Democracy Nachdr (Addison Wesley).

21. Curini, L., Marangoni, F., and Tronconi, F. (2011). Rebels With a Cause - But Which One? Defections from Legislative Party Unity in Italy and Their Individual and Institutional Determinants. Riv. Ital. Sci. Polit. 385–410. https://doi.org/10.1426/36012.

22. Poole, K.T., and Rosenthal, H. (2000). Congress: A Political-Economic History of Roll Call Voting 1st paperback ed. (Oxford Univ Pr).

23. Poole, K., Lewis, J., Lo, J., and Carroll, R. (2011). Scaling Roll Call Votes with **wnominate** in *R*. J. Stat. Soft. *42*. https://doi.org/10.18637/jss.v042.i14.

24. Hix, S., and Noury, A. (2016). Government-Opposition or Left-Right? The Institutional Determinants of Voting in Legislatures. PSRM *4*, 249–273. https://doi.org/10.1017/psrm.2015.9.

25. Ding, C., and Peng, H. (2003). Minimum redundancy feature selection from microarray gene expression data. In Computational Systems Bioinformatics. CSB2003. Proceedings of the 2003 IEEE Bioinformatics Conference. CSB2003 (IEEE Comput. Soc), pp. 523–528. https://doi.org/10.1109/CSB.2003.1227396.

26. Cleves, M.A. (2010). An Introduction to Survival Analysis Using Stata 3rd ed. (Stata Press).

27. Cox, D.R. (1972). Regression Models and Life-Tables. J. Roy. Stat. Soc. B *34*, 187–202. https://doi.org/10.1111/j.2517-6161.1972.tb00899.x.

28. Morgan, J.,W. (2014). The latent path model for dynamic social networks with an application to party switching in Poland. In New Faces in Political Methodology VII Conference Papers.

29. Hare, C., and Kutsuris, M. (2022). Measuring Swing Voters with a Supervised Machine Learning Ensemble. Polit. Anal. 1–17. https://doi.org/10.1017/pan.2022.24.

30. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Müller, A., Nothman, J., Louppe, G., et al. (2012). Scikit-learn: Machine Learning in Python. https://doi.org/10.48550/ARXIV.1201.0490.

31. Jumarie, G. (1990). A Critical Review of Shannon Information Theory. In Relative Information Springer Series in Synergetics (Springer Berlin Heidelberg), pp. 44–65. https://doi.org/10.1007/978-3-642-84017-3_3.

32. Breiman, L. (1996). Bagging predictors. Mach. Learn. *24*, 123–140. https://doi.org/10.1007/BF00058655.

33. Breiman, L. (2001). Random Forests. Mach. Learn. *45*, 5–32. https://doi.org/10.1023/A:1010933404324.

34. Shannon, C.E. (1948). A Mathematical Theory of Communication. Bell Syst. Tech. J. *27*, 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x.

35. Magri, C., Whittingstall, K., Singh, V., Logothetis, N.K., and Panzeri, S. (2009). A toolbox for the fast information analysis of multiple-site LFP, EEG and spike train recordings. BMC Neurosci. *10*, 81. https://doi.org/10.1186/1471-2202-10-81.

36. Panzeri, S., and Treves, A. (1996). Analytical estimates of limited sampling biases in different information measures. Network *7*, 87–107. https://doi.org/10.1080/0954898X.1996.11978656.

37. Meneghetti, N., Cerri, C., Tantillo, E., Vannini, E., Caleo, M., and Mazzoni, A. (2020). Thalamic inputs determine functionally distinct gamma bands in mouse primary visual cortex. Neuroscience. https://doi.org/10.1101/2020.07.09.194811.

38. Meneghetti, N., Cerri, C., Tantillo, E., Vannini, E., Caleo, M., and Mazzoni, A. (2021). Narrow and Broad γ Bands Process Complementary Visual Information in Mouse Primary Visual Cortex. eNeuro *8*. ENEURO.0106-21.2021. https://doi.org/10.1523/ENEURO.0106-21.2021.

39. Meneghetti, N., Cerri, C., Vannini, E., Tantillo, E., Tottene, A., Pietrobon, D., Caleo, M., and Mazzoni, A. (2022). Synaptic alterations in visual cortex reshape contrast-dependent gamma oscillations and inhibition-excitation ratio in a genetic mouse model of migraine. J. Headache Pain *23*, 125. https://doi.org/10.1186/s10194-022-01495-9.

40. Cohen, M.X. (2014). Analyzing Neural Time Series Data: Theory and Practice (The MIT Press).

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
Matlab	Mathworks	https://www.mathworks.com/
Python	Python Software Foundation	https://www.python.org/
Information theory toolbox	Publication	https://10.1186/1471-2202-10-81
Other		
Italian Chamber of Deputies platform for data sharing	Linked Open Data	https://dati.camera.it/it/
Custom code for classification and prediction of party switcher Deputies		https://github.com/nicolomeneghetti/Italian_Chamber_Deputies_Analysis

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Prof. Silvestro Micera (silvestro.micera@epfl.ch).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- This paper analyzes existing, publicly available data. The data are available at http://dati.camera.it/. This link leads to a publishing and sharing platform for Linked Open Data on the activities and bodies of the Chamber of Deputies, available for free download or querying.
- The original code for classification and prediction of party switcher Deputies has been deposited at https://github.com/nicolomeneghetti/Italian_Chamber_Deputies_Analysis and is publicly available as of the date of publication.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact Prof. Silvestro Micera (silvestro.micera@epfl.ch) upon request.

METHOD DETAILS

Data collection

We extracted the votes and the non-voting features of the MPs of the XVII and XVIII legislatures of the Italian Chamber of Deputies. The datasets were extracted from the official open data platform of the Italian Chamber of Deputies (http://dati.camera.it/). The query language employed for the dataset collection was Sparql. We extracted every voting session available from the Italian Chamber of Deputies for both legislatures. Specifically, we collected for the XVII [XVIII] legislature: Articles ('Articolo'): 2929 [1430] voting sessions; Closing of discussion ('Chiusura della discussione'): 10 [6] voting sessions; Resignations ('Dimissioni'): 13 [5] voting sessions; Documents ('Documenti'): 103 [23] voting sessions; Amendment ('Emendamento'): 15860 [5228] voting sessions; Final Act Chamber ('Finale atto Camera'): 466 [368] voting sessions; Motion ('Mozione'): 1561 [607] voting sessions; Note of variation ('Nota di variazione'): 5 [1] voting sessions; Agenda ('Ordine del Giorno'): 3489 [2546] voting sessions; Preliminary ('Pregiudiziale'): 102 [31] voting sessions; Suspensive matter ('Questione sospensiva'): 9 [1] voting sessions; Resolution ('Risoluzione'): 372 [440] voting sessions; Others ('Altre'): 9 [21] voting sessions.

There were 671 Deputies in the XVII legislature and 660 in the XVIII: in both cases, the maximum in office at the same time was 630 Deputies. For each one of them, we collected the following non-voting features: name and family name; date of birth; place of birth (divided into 5 categorical labels: north, center, south, islands, abroad); education level (divided into 5 categorical labels: low-level, middle-level, diploma, degree, Ph.D.); gender; beginning and ending date of the mandate; electoral district(divided into 5 categorical labels: north, center, south, islands, abroad); political party affiliation(s); the number of served terms. Please note that these features consisted of the whole set of personal data available in the repository we employed in our analysis - no personal information was discarded or taken from external resources. We furthermore extracted the preferences expressed in every voting session (for a total of 24928 in the XVII and 10707 in the XVIII throughout the whole legislatures) by every legislator (if present). Consequently, we computed, for every voting session and every Deputy: (i) present (or absent): whether the MP is present in a given vote; (ii) agreement: whether the MPs have voted alike the majority of their party; (iii) abstained: whether the MP abstained on the vote; (iv) in favor: whether the MP voted in favor of the proposed piece of legislation; (v) opposed: whether the MP voted in order to reject the proposed piece of legislation; (vi) approved and in favor: whether the MP voted in favor of the piece of legislation that was then approved by the Chamber; (vii) secret: whether the vote was a secret ballot. Hence, when investigating the difference between single-party and party switchers MPs in a given temporal window, the fraction of these voting features was computed (for each MP separately) and then fed to the classifier. The probability density functions of these features (considering the whole legislature, see Figure 1) in the Deputies groups were estimated via a kernel smoothing function estimate (*ksdensity* in Matlab).

QUANTIFICATION AND STATISTICAL ANALYSIS

Classification of party switchers and single-party MPs

The classification estimator was chosen to be random forests[32,33] due to the relative simplicity, robustness against dataset imbalance and outliers, and the possibility to handle categorial other than numerical variables.

A stratified (according to target distribution) 5-fold random search was conducted to find the best hyperparameters for the model, optimizing the accuracy on the test set. The tested hyperparameters were the number of estimators, the depth of each estimator, the minimal number of samples required to split a node, the minimal number of samples required to be at a leaf node, the number of features to consider at every split, and the maximum number of levels in the tree. We computed the improvement on the accuracy of the randomly searched hyperparameters as compared to the default hyperparameters based on the implementation of the scikit-learn library in Python.[30] As no significant improvements were found in 50 iterations of the random search ([0.6 ± 1.3] %, mean $\pm$ standard deviation) we kept for the analysis here presented the default hyperparameters values.

We applied the classification pipeline to our dataset in a time-varying fashion (Figure 2A). Briefly, for every discrete step *t* of 15 days across the whole legislature, the MPs were split into PS(t) (i.e., those MPs that changed their political affiliation prior to *t*) and non-PS(t) (i.e., those MPs that have not yet, or never will, change their political affiliation prior to *t*). For each time step, a stratified 5-fold random forest classifier estimation was performed on the resulting dataset.

Empirical chance levels were computed using a dummy classifier, i.e., a classifier fit on data with randomly shuffled labels.

Feature importance

To assess the contribution of every feature to the classification of party switchers and single-party Deputies, we used two different approaches. The first one is the maximum relevance minimum redundancy algorithm (*mrmr*).[25] Briefly, the *mrmr* is a minimal-optimal feature selection algorithm. This means it is designed to iteratively rank the features' relevance for a given machine learning task while minimizing the redundancy with previously identified features. In our case, we adopted the *mrmr* algorithm to sort the importance of every feature in the classification/prediction of single-party and party switcher MPs. For the sake of clarity, we decided to report only the three most important features in Figures 3 and 6, and the most important feature in Figures 5 and 7.

The second approach is mutual information.[34] All information quantities were computed in Matlab with Information Breakdown Toolbox.[35] Probabilities density estimates were computed by discretizing the numerical features (i.e., the birth year and the fraction of absence, presence, votes in agreement, abstained, in favor, opposed, approved and in favor, and of secret votes) into 5 equi-populated bins. Limited dataset bias was accounted for by applying the Panzeri-Treves correction.[36] The level of significant information was estimated by applying a bootstrap procedure (with 500 iterations) following the methods proposed in.[37–39] The bootstrap procedure consisted of randomly pairing the MPs' features with the labels (i.e., single-party or party switcher).

The alpha level was set to 0.05 and corrected with the extreme pixel-based multiple comparison method.[40]

Prediction

To test the prediction task, we adopted the procedure graphically described in Figure 4A. We paired each party switcher with a randomly (without replacement) selected Deputy belonging to the single-party group during the whole legislature. Please note how the resulting dataset is now balanced. For both components of each pair, we considered the votes over the same 60 days prior to the switch of the party switcher component of the pair. Every pair was consequently assigned to the train or the test set with a probability of 0.5. The resulting dataset was fed to a random forest algorithm. This procedure was repeated 5 times and the prediction results reported in the manuscript are the average across such repetitions. Afterward, this approach was repeated by sliding the 60-day window back towards the past in discrete steps of 3 days (i.e., moving away from the day of party switching).