

Politicians go social. Estimating intra-party heterogeneity (and its effects) through the analysis of social media.

Andrea Ceron and Alessandra Caterina Cremonesi – University of Milan

NYU La Pietra Dialogues on Social Media and Political Participation, Florence, 10-11 May 2013

Do not quote without authors' permission

April 24, 2013

Abstract

Scholars have emphasized the need to deepen investigation of intra-party politics in order to assess the effect of internal heterogeneity on the party and the party system. However, estimating the preferences of intra-party actors (e.g., party factions) is often a difficult task. Recent works shown that quantitative text analysis of documents drafted by intra-party subgroups can be a solution, but this strategy relies on the existence of such documentation. The striking rise of social media provides a substitute for that. In fact, politicians belonging to different party factions feel free to express their sincere preferences on social media (blogs) or social network sites like Facebook and Twitter. Taking advantage of these publicly released statements, we are able to assess the policy position of factions within the Italian Democratic Party (PD). Focusing on the 2012 centre-left primary election, we show that factional heterogeneity measured through content analysis of social media helps explaining the politician's choice to endorse or criticize one or the other of the two PD candidates running for the nomination.

Introduction

Although some scholars claim that unity is a source of party's strength, we hardly ever observe perfect cohesion within political parties. The party is by no means a monolithic structure as it is composed by politicians retaining similar but non-identical preferences. Those sharing the most similar views often cluster together and form party factions, in order to shape the party strategy and maximize their own share of benefits (Ceron 2012, 2013b). Given this premise, the fact that many parties all over the world are factionalized comes as no surprise, particularly since factionalism is not necessarily damaging to the party's fortune (Boucek 2012).

Scholars started to investigate the impact of intra-party politics proving that factional preferences affect party platform, parliamentary policy-making, coalition formation, and portfolio allocation (e.g. Giannetti and Benoit 2009). Furthermore factional affiliation helps explaining the voting behavior of MPs (Bernauer and Braüninger 2009; Ceron 2013a; Giannetti and Laver 2009) and could be crucial to enhance a politician's career (Cox et al. 2000). Despite the relevance of this topic, the research on factional politics has remained underdeveloped until recent year, mainly because intra-party politics is a kind of invisible politics that takes place in the shadow (Sartori 1976). It is hence difficult to assess the preferences of politicians and party factions. Nevertheless, latest improvements in the field of quantitative text analysis help to fill this gap and allow evaluating the degree of ideological heterogeneity by focusing on parliamentary speeches (Bäck et al. 2011; Bernauer and Braüninger 2009; Proksch and Slapin 2010), debates held at party congresses or documents drafted by intra-party subgroups (Ceron 2012, 2013a; Debus and Braüninger 2009; Giannetti and Laver 2009). These are valuable sources of information even though they present some limitation. Parliamentary speeches may overestimate internal cohesion while the analyses of congress debates and factional motions fail to track changes between two consecutive congresses and to discriminate between the ideological positions of intra-party actors that contest the congress together. The rise of social media and social network sites (SNS) provides an opportunity to overcome these drawbacks and allows assessing the preferences of party factions. After the successfully 2008 Obama campaign (Cogburn and Espinoza-Vasquez 2011) and the widespread growth of internet users, Italian politicians discovered the power of the web and started to make use of it. They usually express their opinions on-line following a top-down approach (Formenti 2012; Vaccari 2008). The statements released by politicians belonging to different party factions on social media (personal weblogs), micro-blogging platforms (Twitter) or social network sites (e.g., Facebook and Twitter) have been exploited to assess the different policy position of intra-party actors within the highly factionalized Italian Democratic Party (PD).

In the wake of 2012 centre-left primary election, the estimates of factional heterogeneity have been employed to explain the politician's choice to endorse or criticize one or the other of the two PD candidates running for the nomination. The analysis shows that politicians are likely to endorse the leader closer to their faction's policy position. Fur-

thermore, as the policy distance between a politician and the candidate increases, the politician feels free to openly criticize his/her party colleague. These results confirm that party factions are driven by policy motives and, in a highly factionalized party, politicians are prone to publicly criticize their intra-party opponents without fearing retaliation for having damaged the faade of unity. Furthermore, this paper show that public statements released on social media and SNS can be used to assess the policy position of political actors and to measure ideological heterogeneity within a party. Since heterogeneity increases the propensity for criticism, this work also suggests that political actors would rather express their sincere views on social media compared to more institutionalized contexts like parliamentary arenas.

1 The growing usage of social media in the wake of 2012 centre-left primary election

Nowadays, the usage of internet and social media is growing at very fast rates. In 2012 more than one third of the world population got access to the web and most of them were also active on SNS. Facebook surpassed one billion active accounts in 2012 while Twitter reached 200 million users in February 2013.¹ The growth of the internet audience made the web attractive to parties and candidates, mainly as a device to mobilize support (e.g., Cardenal 2011).

The rise of social media has involved Italy too. In the last five years a growing number of citizens joined the two most popular SNS. In 2013 over 22 million Italian citizens are active on Facebook (almost the 40% of the population) while around 5 millions are Twitter users (7.8% of the Italian population). For the sake of spreading their political views to a large number of people, Italian politicians, exactly like anyone else, subscribed and became users as well. Their activity rate is similar to that of the whole population: in February 2011 the 35.9% of Italian MPs had a profile on Facebook while the 9.7% of them was active on Twitter. SNS are used for personal communication but also to discuss about politics. The link between social media and politics has been unveiled during the campaign for the 2012 centre-left primary election, that started in October 2012. This election generated a heated debate on-line where hundreds of thousands messages have been posted by politicians, party activists, or common citizens. Episodes like the acquittal of Nicky Vendola, the third ranked candidate, or the two television debates held between the candidates produced a huge amounts of tweets and posts on the web. Nevertheless, citizens and politicians commented any event of the campaign and any statement issued by candidates and their staff. People expressed their policy preferences up to the point that most of them revealed their voting choices on SNS. This wide amount of data has been exploited by scholars (Ceron et al. 2012) that analyzed over 500,000 tweets to suc-

¹See <http://mashable.com/2012/12/26/facebook-2012/> and <http://blog.twitter.com/2013/03/celebrating-twitter7.html>

cessfully predict both the results of the first ballot (25th November 2012) and the run-off (December 2nd).

The primary election was called to select the leader of the centre-left coalition 'Italia Bene Comune' (IBC) in view of the 2013 general elections. Five candidates ran for the nomination but only two had a real chance to win the race: the frontrunner Pierluigi Bersani, leader of the PD, was challenged by the young mayor of Florence, Matteo Renzi, belonging to the same party. Both of them are active on SNS and they use to write, on average, two or three tweets per day. These comments are shared with a wide audience. Few weeks before the beginning of the campaign, in September 2012, Bersani had 80,000 fans on Facebook (nowadays: 140,000) and 135,000 followers on Twitter (number that eventually increased until 330,000). Renzi was even more popular on the web with 155,000 fans and 152,000 followers (few months later, in April 2013, Renzi reached the peak of 380,000 fans and 490,000 followers). These numbers give an idea about the relevance of the content published on-line. From 2008 on, a growing number of Italian politicians has been debating and releasing statements on weblogs or SNS and they publicly expressed their opinion on a variety of political topics. These sources of information can be exploited to estimate the policy view and the ideological position of political parties but they could be even more suitable to evaluate the degree of intra-party heterogeneity and to assess the ideal points of rival party factions.

2 Estimating policy positions from social media

Several scholars claim that the analysis of social media allows measuring the policy position of political actors. (Conover et al. 2010; King et al. 2011; Boutet et al. 2012). Taking the cue from these works, I contend that the information available on the web is particularly suitable to estimate the preferences of 'hidden actors', like formal and informal intra-party subgroups whose ideological viewpoints may be not formally recorded in official documents or publicly displayed through observable behavior.

In fact, estimating the policy positions of party factions is a challenging task. Several case studies indicate that parties are internally divided. However, party competition creates pressure to display cohesion in the eyes of the voters as unity may enhance a party's electoral fortune (Alesina and Cukierman 1990: 847; McGann 2002, Snyder and Ting 2002).

As a consequence, members supporting contrasting views about the party line and strategy should try to work out their differences on their own (Evans 2001; Messmer 2003) so that 'internal disagreements are usually resolved before party positions are defined formally, as in party manifestos, or behaviorally, as in legislative votes and speeches' (Heller 2008: 2, *italic added*). Whether the solution stems from consensual bargaining and compromise or from loyalty and enforced discipline does not really matter for our purpose, as in both cases we would fail to observe disunity and we lack the information necessary to estimate the internal heterogeneity of policy views. The display of party

unity may hide internal division and it does not imply perfect cohesion. But even when we observe conflict and splits, for instance in roll call votes, the extent of disagreement inside the party could be misestimated.

While some scholars have been estimating factional preferences by scaling roll call votes (e.g., Spirling and Quinn 2010), others argued that this technique only provides a description of the 'revealed behavioural space' (Hix and Jun 2009) and a measure of ex-post behaviour instead of ex-ante preferences. Furthermore the actual level of division can be underestimated due to potential selection bias in the use of roll call (Carrubba et al. 2006, 2008). For these reasons Giannetti and Benoit (2009) suggest measuring factions' positions relying on what intra-party actors say (the declared preferences) instead of on what they do (the actual behavior). Since talk is cheap heterogeneous declarations are less damaging to the party compared to the cost of incohesive behavior. This is even more true in the internet age when politicians can take advantage of the new media to spread their ideas and to comment over any political event, virtually in real time. Then, 'politicians may often toe the party line while at the same time generating texts that show far less subservience to the mechanisms of party discipline' (Giannetti and Benoit 2009: 233). The analysis of political texts allows discriminating contrasting preferences even when actors behave in the same manner (e.g., cast the same vote or endorse the same candidate) and therefore it is well suited to study intra-party politics.

Accordingly, several scholars measured the degree of intra-party heterogeneity by analyzing parliamentary speeches (Bäck et al. 2011; Bernauer and Braüninger 2009; Proksch and Slapin 2010) and debates held at party congresses or documents drafted by intra-party subgroups (Ceron 2012, 2013a; Debus and Braüninger 2009; Giannetti and Laver 2009). These approaches, however, are somehow problematic.

Under the assumption that speeches released during legislative debates are the outcome of an interplay between the party leader and backbenchers, Proksch and Slapin (2012: 16) analyzed MPs' discourses showing that they tend to misrepresent ideological polarization, hence are not the best source to catch intra-party divisions. In fact, speeches delivered in public and highly institutionalized arenas (like national parliaments) are easily observable and therefore subject to party whip. Different electoral systems alter the leader's propensity to employ the whip in order to impose discipline and affect MPs' incentives to express their sincere positions during the debate. Since the leader can decide whether to leave the floor to MPs or not, in competitive political systems where the value of party unity is higher (e.g., closed-list PR) she will be more likely to deliver the speech on his/her own or to give way to one of his/her followers. As a consequence, parliamentary speeches are subject to selection effects and 'may not reflect the true distribution of preferences' (Proksch and Slapin, 2012: 3) so that the analysis will overestimate party cohesion.

During intra-party debates the whip should only slightly bind the sincere expression of preferences, compared to the discussions held in the parliamentary arena. Investigating these debates by means of content analysis on texts drafted by intra-party groups could be useful to identify their preferences (Benoit et al. 2009). Through documents like

factional motions, i.e. omni-comprehensive policy documents issued by factions during party congresses, any internal subgroup is (almost) completely free to present its idea about how party position and strategy ought to be. Given that their content should be minimally affected by leaders' control, some authors analyzed these programmatic documents that express 'opposing views on the ideological direction of the party' (Giannetti and Laver 2009: 154) to map the distribution of preferences within the party.

Beside some advantages, using this kind of documents presents also some limitations. First of all, compared to analogous texts like party manifestos, subgroups' programmatic documents could be seldom available and not issued on a regular basis. This generates troubles in treating missing information and tracking changes in the position of subgroups over time. By the same token, it is impossible to analyze motions when, for the sake of displaying strong party unity, both mainstream and minority factions find an agreement in advance and submit to the congress a common text expressing a shared position that is approved by all party members. The same problem arises when the motion presented by the mainstream faction is unchallenged because the minority feels too weak to defy the leadership and refuse to contest the congress even though in this case the unity of the party is a faade. Similarly, we could not be able to discriminate the different positions of factions if they sign an alliance and present the same motion. Finally, as far as party congresses display the subgroups' preferences in a single point in time, the analysis of motions and congress debates would not get changes happened between one congress and the following one (for instance because a faction changes its mind or splits in two subgroups). This aspect could be especially tricky in low-institutionalized contexts and inside young or fluid parties where subgroups are not steady and party members often make and break factions, shaping and reshaping the intra-party structure.

To overcome these concerns we suggest using data gathered from social media and social network sites. Some specific features make this information interesting and valuable. Social media penetrate into people's lives. From the previous section, we know that users tend to react and comment on everyday political events and they express their attitudes on a variety of topics. In turn, politicians have incentives to open a website or a social networking profile and make use of social media to inform, mobilize, and cultivate personal votes (Vergeer et al. 2012). To do that they express their opinions and release public comments in order to influence the agenda or react to the statements released by others. Weblogs, SNS or micro-blogging devices like Twitter share a unique feature that is convenient for our purpose. They are unmediated and self-expression oriented tools (Formenti 2012) whose users release unsolicited statements.² This increases the likelihood that public declarations posted on-line reflect the true preferences of political actors. Although some statements could be instrumental, the extent of strategic behavior on-line should

²It has been argued that text analysis could be affected by 'content injection' (Conover et al. 2010) related to the misuse of Twitter hashtags. Other potential drawbacks concern 'retweets' (RT) made for ironic purposes, and the direct replies made by actors that are debating together (Barberá 2012). However, we discard RT and replies, retaining only original and unsolicited content generated by the single user. Furthermore, by focusing on several social media beside Twitter (see later) we limit the extent of any possible bias.

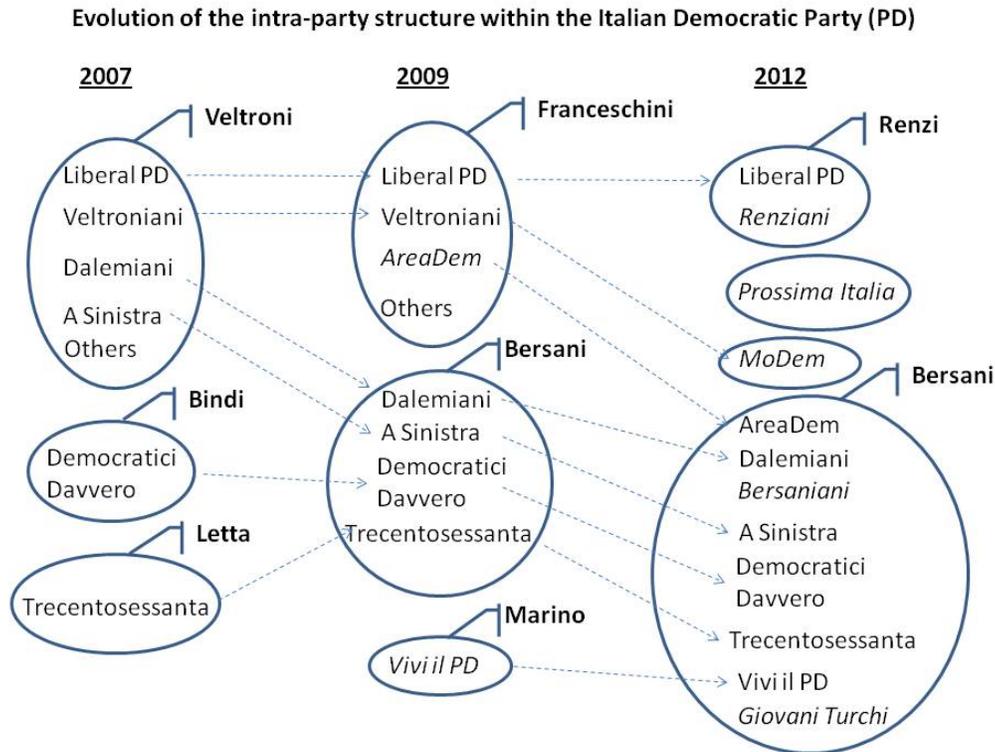
be lower if compared to what happens off-line in more formal environments like parliamentary arenas or party conferences. Texts written on-line are also more spontaneous compared to the content of interviews released to the media where politicians face direct (and sometimes unwanted) questions to which they must answer (this point will be retrieved later). Finally since mobile devices make social media accessible at any time, they will record reasoned behavior as well non-strategic impulsive activities. For all these reasons we argue that the analysis of content published on-line could avoid misestimating the true degree of heterogeneity and internal dissent within the party, particularly when these statements are perceived as being free preliminary personal opinions not subjected to party whip and not damaging for party unity.

3 Assessing the position of party factions through content analysis of social media

This section describes the methodology adopted to measure the differences between rival party factions. Later, these estimates will be used to analyze factional strife within the Italian Democratic Party (PD) in the wake of 2012 primary election held by the centre-left coalition. The PD was created few years ago, in October 2007, after the merger of the Democrats of the Left (DS) and the centre-left party named 'The Daisy' (DL). The party has been highly factionalized since its birth. Members gathered together and created intra-party associations supporting contrasting views on a variety of policy dimensions, ranging from economics to social policy. Most of these subgroups are steady and institutionalized factions while others are blur. In 2007 and 2009 the party held direct elections open to party members and non-members in order to select the leadership. Several factions grouped together presenting a unique programmatic document and supporting a common candidate. In 2007 the party leader Walter Veltroni was sustained by a large number of factions including long-standing rivals tied to Massimo D'Alema ('Dalemini'). Only two proposed alternative candidates: Roby Bindi (supported by the faction 'Democratici Davvero') and Enrico Letta who led his followers ('Associazione Trecentosessanta'). In 2009 three candidates fought for the election. Pierluigi Bersani, who became the party leader, proposed his own platform that was supported by several subgroups including the factions tied to Bindi and Letta. His main opponent, Dario Franceschini, was the incumbent party leader, selected in February 2009 after the resignation of Veltroni. Franceschini was supported by the faction 'AreaDem' as well as by the liberals ('Liberal PD') and the followers of Veltroni. The third candidate was Ignazio Marino ('Vivi il PD'). Bersani won the election with a narrow margin of 53.2% votes. For the sake of preserving unity both Marino and Franceschini rejoined the party mainstream soon after the 2009 congress and decided to sustain Bersani in his attempt to rule the party through consensual internal dynamics. Veltroni strongly disagreed with the drift made by Franceschini and decided to form a new minority faction, 'MoDem', to oppose the party mainstream (Ceron 2011a, 2011b). Beside this split, between 2009 and 2012 many

other reshuffles altered the factional structure of the Democratic Party.³ New subgroups like ‘Rottamatori’, ‘Giovani Turchi’, and ‘Prossima Italia’ were created by young party members (respectively Matteo Renzi, Stefano Fassina, and Giuseppe Civati) in order to promote leadership turnover and policy renewal (albeit in different directions). These changes are illustrated in Figure 1.

Figure 1: Party factions within the PD and candidates in primary elections



Given this fluidity, sticking to the intra-party structure portrayed in the 2009 congress would not be the best solution to understand factional politics within the PD. Following the definition of faction proposed by Boucek (2009) we classify 12 distinct subgroups and identify their leaders: A Sinistra (Livia Turco and Vincenzo Vita), AreaDem (Franceschini), Bersaniani (Bersani), D’Alemiani (D’Alema), Democratici Davvero (Bindi), Giovani Turchi (Fassina), Liberal PD (Enrico Morando), MoDem (Walter Veltroni and Paolo Gentiloni), Prossima Italia (Giuseppe Civati and Debora Serrachiani), Rottamatori (Renzi),

³The fractionalization of the PD yields effects on the party system. For instance, in the election of the Head of State, held in April 2013, factions supported different names and the party split during the vote when left-wings and right-wings coalesced to impose their veto on the two candidates supported by the party mainstream, Franco Marini and Romani Prodi. Note that several rebels publicly declared their dissent on the web. Indeed both politicians and citizen expressed their view on the preferred candidate on social media and social network sites up to the point that several journalists argued about the impact of the public opinion (expressed on-line) on the voting behavior of party MPs.

Trecentosessanta (Letta), Vivi il PD (Marino). The policy position of each faction has been measured by means of text analysis applied to blogs, Facebook accounts and Twitter profiles of their factional leaders. We downloaded the comments published during the XVI Legislature, between April 2008 and December 2012, and retained only the posts with a political content. When there was no web content available for a leader or when the amount of information was insufficient to carry out a reliable analysis we compensate using data related to other prominent politicians belonging to the same faction.⁴ Table 1 shows the details of the documentation related to each faction.

Table 1: Details on data gathered.

% Faction	N. of sources	Words
A Sinistra	2	27324
AreaDem	4	49176
Bersaniani	3	78911
Dalemiani	5	22539
Democratici Davvero - Bindiani	2	47417
Giovani Turchi	1	5074
Liberal PD	5	79422
MoDem	4	5805
Prossima Italia	5	143807
Rottamatori	5	53841
Vivi il PD	4	36033
Trecentosessanta - Lettiani	4	21754

These documents have been analyzed through Wordfish, an automated scaling model that performs quantitative text analysis (Proksch and Slapin 2009a; Slapin and Proksch 2008).⁵ Under the assumption that words usage is informative of the policy position of textual documents, Wordfish compares the relative frequencies of all the words contained in a text and assigns a value to each.⁶ This value represents the discriminating power. A higher absolute value of indicates that the word is more useful to discriminate among documents. Overall, values of words that appear only in a few documents will be greater and these words will be located on the extremes of the scale. On the contrary, words that

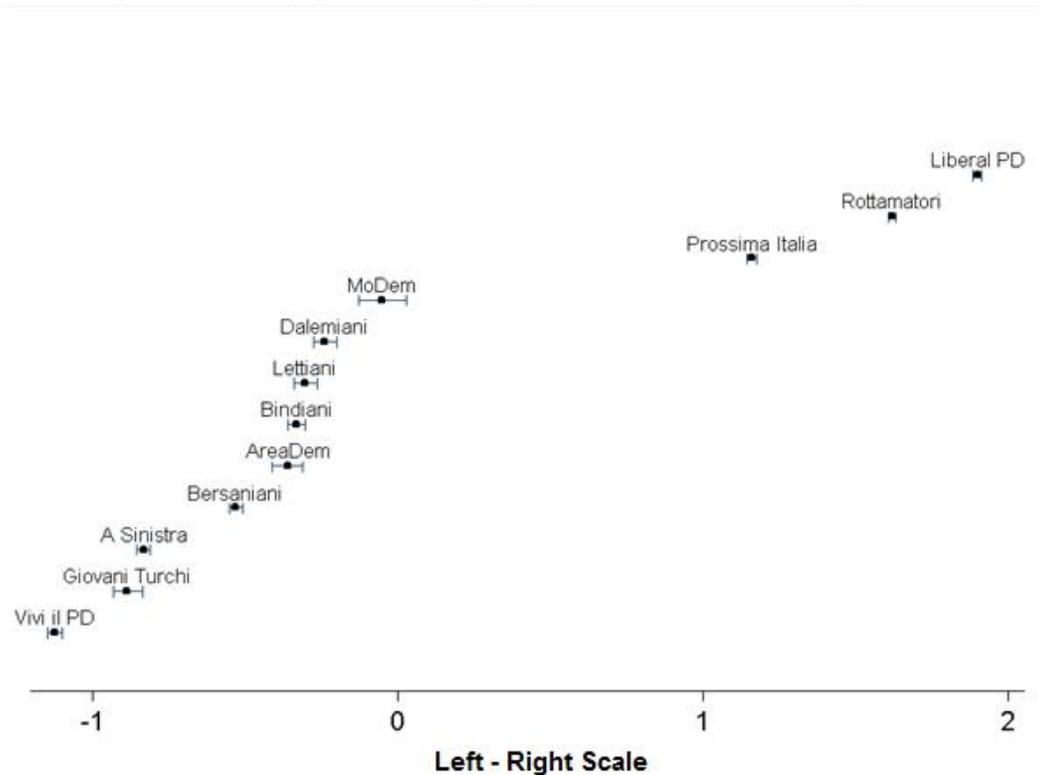
⁴For instance, instead of Morando we analyzed posts written by Luigi Bobba and Pietro Ichino, while we used comments made by David Sassoli and Marina Sereni to supplement Franceschini.

⁵Wordfish has been already employed to study the policy positions of political actors within Germany, Italy and the European Union, providing reliable estimates of actors' positions (Ceron 2011a; Klüver 2009; Proksch and Slapin 2009b, 2010). Wordfish estimates are also robust to text selection and reliable to other techniques of content analysis (Klüver 2009; Slapin and Proksch 2008)

⁶Wordfish allows avoiding any concern with common or rare words. Common words not associated with any political meaning (i.e., articles, prepositions, etc.) appear with higher frequency but they are evenly spread across documents. Despite their strong fixed effect (frequency), their discriminating power is close to zero. Conversely the discriminating power of rare words (for instance, words included in only one document, whose fixed effect is very low) is systematically lowered by the software to solve any trouble with uncommon words.

appear frequently in all documents have a discriminating power close to zero and will be placed at the centre. Then, according to the words' discriminating power, Wordfish arrays documents along the latent dimension whose meaning depends on the content of the texts submitted to the analysis and has to be interpreted ex-post. When documents encompass a wide variety of topics, like in this case, Wordfish extracts policy positions on a latent ideological dimension that should be interpreted as a left-right scale. Figure 2 displays the placement of each faction along the scale.

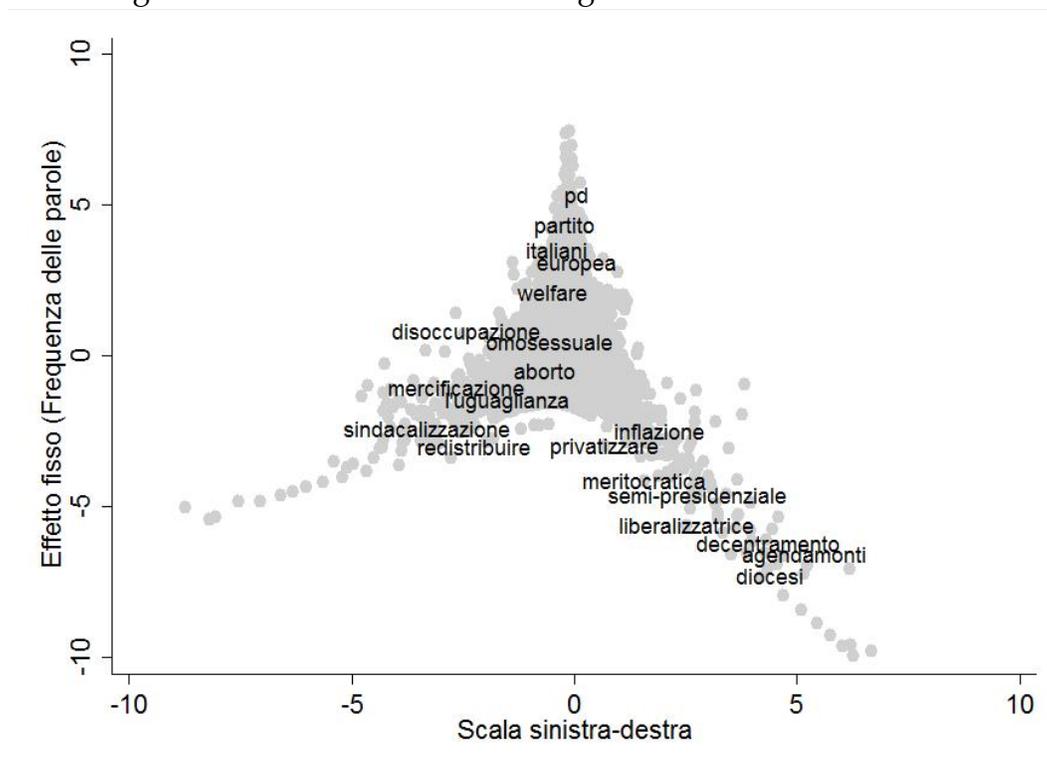
Figure 2: Policy position of party factions on the left-right scale



The policy positions of party factions are in line with the expectations. On the left side we find factions like 'Vivi il PD', 'Giovani Turchi', and 'A Sinistra' that usually express left-wing positions. The 'Bersaniani', followers of the party leader, are still on the left though on a more moderate position and all the other factions that supported Bersani (AreaDem, Bindiani, Lettiani, and Dalemiani) are quite close as well. The position of the MoDem is more centrist and statistically different from that of the mainstream factions rallied behind Bersani. Finally, liberals and reformist subgroups such as the 'Liberal PD', 'Rottamatori', and 'Prossima Italia' (a splinter group of the latter) are on the right wing. The positioning of Bersaniani and Rottamatori is similar to that based on the analysis of policy platforms presented by Bersani and Renzi during the 2012 primary election (Ceron et al. 2012). The placement of words on the latent dimension is in line with their substan-

tial meaning in the Italian political language and this confirms that the analysis catches the actual meaning of those words. In Figure 3 we reports the coefficients and highlight some words in example. Terms like ‘redistribuire’ (redistributing) and ‘l’uguaglianza’ (equality) are located on the left as well ‘disoccupazione’ (unemployment). Concerns about ‘inflazione’ (inflation) are instead typical of the right of the party likewise support for the ‘agendamonti’ (a platform of reforms proposed by the former premier Mario Monti).

Figure 3: Beta coefficients and diagnostic of words’ estimates.



4 Endorsement or criticism of party leaders: Hypotheses and analysis

In this section we employ these estimates of factions’ position to test whether factional membership and intra-party polarization have an effect on the internal debate of the Italian Democratic Party, shaping the behavior of its members and the outcomes of factional strife.

We do that focusing on two aspects that are closely linked to each other and relate to the idea of publicly supporting or opposing a party leader. The first aspect is about the choice to openly endorse one of the two main party candidates running for the nomina-

tion in the 2012 primary election, Pierluigi Bersani and Matteo Renzi. The second element refers to the propensity to overtly criticize and attack these two leaders.

We will analyze the behavior of PD politicians taking into consideration the effect of policy motives. Since party members divide along policy lines and build factions accordingly (Bernauer and Braüninger 2009; Ceron 2011a), we would expect that ideological preferences affect the decision to endorse and/or criticize Bersani or Renzi. These two party leaders contested the primary election presenting contrasting views about the party strategy. Bersani strongly supported the idea of a centre-left coalition including the PD, the small sized Italian Socialist Party (PSI), and the radical 'Left, Ecology, and Freedom' party (SEL) while Renzi was skeptic about this alliance and claimed that the PD should have ran the general election alone. Their struggle was also fought on the choice of party platform as their policy preferences were divergent indeed (Ceron et al. 2012): Bersani ran for the nomination proposing a left-leaning policy manifesto while Renzi presented a more centrist platform and aimed to attract moderate voters. We raise the following hypotheses.

Hypothesis 1: Politicians endorse the candidate closer to the policy position of their party faction

Hypothesis 2: Politicians are more prone to publicly criticize a party leader as his policy preferences diverge from that of their party faction.

By focusing on the Italian Democratic Party we hold intra-party rules and party system features constant. Furthermore, in each analysis we will check the impact of policy motives net of reelection bid and career rewards, and we will include personal traits and socio-demographic attributes of each politician as control variables.

The first analysis deals with the choice to endorse a candidate. We focus on a list of endorsements made by PD politicians in favor of Bersani or Renzi (Seddone 2012). The list has been gathered considering all the official declarations publicly released in national newspapers in the last weeks before the primary election. We assess the factional tie of each politician based on personal biography and membership in one of the rival intra-party associations.

The dependent variable is Endorsement, which is equal to 1 when the politician supports the party leader Bersani, and takes the value 0 when not.⁷ Our main independent variable is Distance, which records the absolute distance between the policy position of Bersani and that of politician i 's faction, according to the estimates shown in the previous section. Notably, politicians are also interested in office payoffs linked to career rewards. As a consequence, incumbent MPs could have an incentive in supporting the party mainstream and standing for Bersani, while fringe politicians may side with the opponent for the sake of taking advantage from a leadership turnover. To evaluate this point we include the variable Incumbent, a dummy that assumes the value of 1 when the politician

⁷Note that a few politicians explicitly declared their non-endorsement for both candidates. Considering the endorsement in favor of Renzi leads to similar results.

was an incumbent MP in the 16th Legislature. The fight between Bersani and Renzi was usually presented as a clash of two generations and Renzi itself often urged the need to scrap the leadership and promote a party renewal. As a consequence, we hypothesize that younger politicians are more likely to support Renzi instead of Bersani and test this claim through the variable Age. Finally we include two other control variables: Female, a dummy to detect female politicians, and Experience, which accounts for the number of years spent in parliament.

Data have been analyzed through logit regression.⁸ The results are shown in Table 2. Model 1, the baseline model, tests the impact of policy heterogeneity. Model 2 includes some controls.

Table 2: Logit regression of Endorsement

Variables	(I)	(II)
Distance	-2.393*** (0.723)	-2.414*** (0.911)
Incumbent		0.657 (0.544)
Age		0.056** (0.023)
Female		2.368** (0.918)
Experience		-0.097** (0.042)
Constant	2.505*** (0.410)	-0.741 (1.275)
Observations	87	87
Log pseudolikelihood	-31.702	-27.473

Standard errors clustered by faction in parentheses

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

The results confirm that policy preferences estimated through the analysis of posts published on social media affect politicians' endorsements in the primary election. Politicians belonging to factions whose policy position is closer to that of the candidate are more likely to support him and publicly declare their endorsement in the media. The likelihood of an endorsement decreases as Distance increases.⁹ Females tend to support Bersani rather than Renzi, while younger politicians endorse the latter confirming the idea of a clash of generations. Keeping Age constant, we observe that more experienced politicians are less likely to support Bersani even though this does not automatically im-

⁸Since we have repeated observations (politicians) within the same party subgroup we provide standard errors clustered on factions.

⁹Furthermore, when measuring intra-party heterogeneity through text analysis of factional motions presented in the 2009 party congress (Ceron 2012) we observe that the distance between factions and candidates is less informative about the probability to endorse the party leader or not, if compare to the more updated measure provided in this paper

plies that they endorse the opponent.¹⁰ Model 2 also shows that being an incumbent MP does not have any effect on the politician's choice. Since the electoral system is a closed list PR we would expect that incumbent MPs toe the line and support the party leader for the sake of being included in the party list. However, the Democratic Party selects its candidates through party primaries. This rule loosens the link between the MPs and the leadership, which is not the solely determinant of a politician's career.

The second analysis considers the public expression of dissent. Focusing on the content published on social media, we count the number of negative mentions made by factional leaders against the two candidates and measure a politician's propensity to explicitly criticize Bersani and Renzi. Figure 4 displays the distribution of negative mentions across time. Almost all the direct attacks have been made in 2012 between June, when the primary election was called, and November, when it was held. In view of this, we can consider these statements as negative campaign (e.g., Ansolabehere and Snyder 2000; Curini and Martelli 2010).¹¹ The source of negative attacks is particularly intriguing since declarations made on the web are voluntary and unsolicited statements.

Attacks, our dependent variable, records the number of attacks made against Bersani and those made against Renzi. The main independent variable, Distance, measures the absolute distance between the policy position of each factional leader and that of Bersani and Renzi. Since the dependent variable is a count model we analyzed data through the negative binomial regression. Model 1 tests the impact of policy heterogeneity alone, while in model 2 we added the control variables used before (Incumbent, Age, Female, and Experience). Standard errors have been clustered by faction. Controlling for the length of the words does not alter the results.

Policy motives affect the propensity to criticize and attack another leader. Within the party, politicians are more likely to attack each other as their Distance increases.¹² Overall, Incumbent MPs are less prone to attack and criticize members of the same party. Due to their status, they could have a higher interest in party unity and want to tone down the debate. After controlling for the degree of Experience, we observe a higher number of attacks made by older politicians while there is no difference in male's and female's propensity. Despite the interest in toeing the line provided by a closed list PR, politicians within the Democratic Party are inclined to wash the dirty linen in public. Since the party is highly factionalized and selects candidates through open primaries, however, it is

¹⁰Experience does not convey any statistically significant effect on the likelihood to endorse Renzi.

¹¹Examples of direct attacks found in social networks are as follows: *'We can't travel to the future with Bersani who has never written a single page talking about the future in the last years'* (Matteo Renzi against Pierluigi Bersani); *'I have never heard a single proposition made by Renzi about the economic crisis, unemployment, development of Southern regions or the Europe'* (Rosy Bindi against Matteo Renzi)

¹²This aspect however seems to concern only issue based attacks, while we do not observe any effect of Distance on valence-related attacks.

Figure 4: Frequency of negative statements against party leaders across time

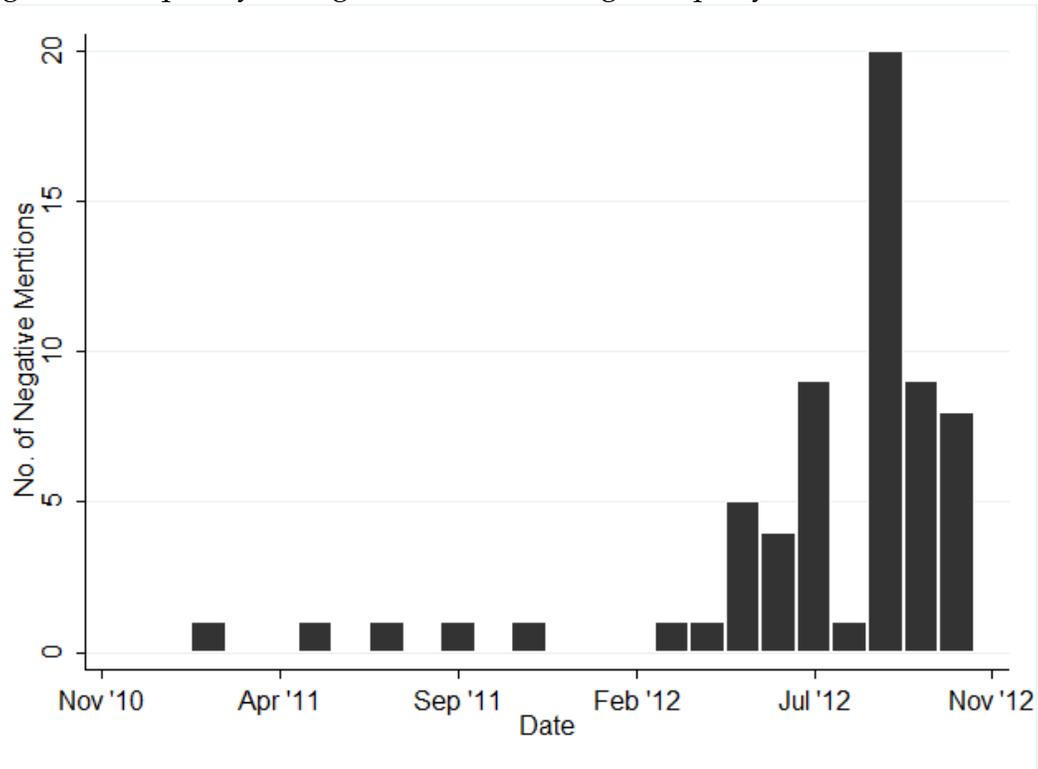


Table 3: Negative binomial regression of Attacks

Variables	(I)	(II)
Distance	0.719*** (0.161)	0.670*** (0.151)
Incumbent		-1.620** (0.640)
Age		0.046* (0.025)
Female		-1.148 (0.897)
Experience		0.040 (0.057)
Constant	-10.849*** (0.352)	-12.243*** (0.942)
Observations	54	54
Log pseudolikelihood	-56.958	-53.674
Alpha	1.599 (1.004)	0.913 (0.693)

Standard errors clustered by faction in parentheses
 * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

not surprising that politicians prioritize factional loyalty instead of pursuing party unity. Furthermore, the incentive to cultivate a personal vote (Kam 2009: 128) encourages them to release public statements on social network to uphold their preferred policy views or to overtly attack and criticize intra-party rivals. In turn, the declarations published on-line are a reliable source of information to assess the policy preferences of intra-party actors and they help to understand politicians decision to positively or 'negatively' endorse a party leader. Both analyses show that factional affiliation and policy heterogeneity have a huge impact on the behavior of political actors and affect intra-party equilibriums.

5 Conclusion

The present paper deals with the question of estimating policy preferences of intra-party actors. Due to the strong need for party unity traditional methods such as the analysis of roll call votes or parliamentary debates tend to misestimate the actual level of heterogeneity within the party. Taking advantage of the increasing use of social media and social network sites we analyze posts published on-line by politicians belonging to rival party faction through a modern technique of quantitative text analysis. Since on-line declarations are unsolicited statements released in an informal environment, they seem to be a good source to assess the sincere preferences of political actors and should be less affected by strategic behavior, at least compared to the behavior (speeches or votes) observable in formal arenas or to the content of official documents.

In view of this, we analyzed the nature and the effects of factionalism within the Italian Democratic Party (PD), by means of content analysis of posts published by factional leaders on social media. By doing this, we assessed the policy position of factions within the PD. Then, we apply our estimates to understand the impact of ideological heterogeneity on the behavior of intra-party actors in the wake of 2012 centre-left primary election. The empirical analyses show that factional affiliation and divergent policy preferences affect the likelihood to endorse one or the other of the two PD candidates running for the nomination. These elements also affect the propensity to adopt negative campaign against rival politicians.

This paper proved that in a highly factionalized party that adopts primary elections to select the leadership, the content of posts published on social media seems the best source to assess the sincere preferences of intra-party actors. Given that the usage of social network can dramatically increase in the next few years, this approach can be extended to study internal polarization across parties and countries, in a comparative perspective. This technique can also be used to estimate the position of citizens and influent opinion makers, such as bloggers and journalists media, and provides a way to better assess the congruence of preferences between parties, activists, and voters.

References

- Alesina, A. and Cukierman, A. (1990). The Politics of Ambiguity. *The Quarterly Journal of Economics* 105(4): 829-850.
- Ansolabehere, S. and Snyder, Jr. J.M. (2000). Valence politics and equilibrium in spatial election models. *Public Choice* 103: 327-336.
- Bäck, H., Debus M. and Müller, W. C. (2011). The Ideological Cohesion of Political Parties. An Evaluation of the Method of Deriving MPs' Policy Positions from Parliamentary Speeches. Paper presented at the workshop Cohesion, Dissent And Partisan Politics in European Legislatures, ECPR Joint Sessions of Workshops, St. Gallen, Switzerland, April 12-16, 2011.
- Barber, P. (2012). Birds of the Same Feather Tweet Together. Bayesian Ideal Point Estimation Using Twitter Data, Draft.
- Bernauer, J. and Bräuninger, T. (2009). Intra-party Preference Heterogeneity and Faction Membership in the 15th German Bundestag: A Computational Text Analysis of Parliamentary Speeches. *German Politics* 18(3): 385-402.
- Boucek, F. (2012). *Factional Politics. How Dominant Parties Implode or Stabilize.* London: Palgrave MacMillan
- Boutet, A., H. Kim, E. Yoneki, et al. (2012). Whats in your tweets? I know who you supported in the uk 2010 general election. In Sixth International AAAI Conference on Weblogs and Social Media.
- Cardenal, A. S. (2011). Why mobilize support online? The paradox of party behaviour online. *Party Politics* 19(1): 83103
- Carrubba, C.J., Gabel, M., Murrain, L., Clough R., Montgomery, E., and Schambach R. (2006). Off the record: Unrecorded legislative votes selection bias and roll-call vote analysis. *British Journal of Political Science* 36(4): 691704.
- Carrubba, C.J., Gabel, M., and Hug S. (2008). Legislative voting behavior, seen and unseen: A theory of roll-call vote selection. *Legislative Studies Quarterly* 33(4): 543572.
- Ceron, A. (2011a). Correnti e frazionismo nei partiti politici italiani (1946-2010): una analisi quantitativa delle mozioni congressuali. *Rivista Italiana di Scienza Politica* 41(2): 237264.
- Ceron, A. (2011b). Il governo Berlusconi alla prova delle due fiducie: coesione e divisione tra gruppi parlamentari a fine 2010. *Polena* 3/2011.
- Ceron, A. (2012). Bounded Oligarchy: How and When Factions Constrain Leaders in Party Position-Taking. *Electoral Studies* 31: 689-701.
- Ceron, A. (2013a). Brave Rebels Stay Home: Assessing the Effect of Intra-Party Ideological Heterogeneity and Party Whip on Roll-Call Votes. *Party Politics* doi:10.1177/1354068812472581.
- Ceron, A. (2013b) Gamson Rule not for All. Patterns of Portfolio Allocation among Italian Party Factions. *European Journal of Political Research* doi:10.1111/1475-6765.12020.

- Ceron A., Curini L., Iacus S. M., and Giuseppe P. (2013). Sentimeter e primarie previsioni e programmi al tempo di Twitter. *Questioni Primarie* 5: 8-10. Accessible from: <http://www.candidateandleadersselection.eu/content/qp-5-sentimeter-e-primarie-previsioni-e-programmi-al-tempo-di-twitter>
- Cogburn, D. L., and Espinoza-Vasquez, F. K. (2011). From Networked Nominee to Networked Nation: Examining the Impact of Web 2.0 and Social Media on Political Participation and Civic Engagement in the 2008 Obama Campaign. *Journal of Political Marketing* 10(1-2): 189-213
- Conover, M., Goncalves, B., Ratkiewicz, J., Flammini, A., and Menczer, F. (2010). Predicting the political alignment of twitter users. Unpublished manuscript.
- Curini L., and Martelli, P. (2010). Ideological proximity and valence competition. Negative campaigning through allegation of corruption in the Italian legislative arena from 1946 to 1994. *Electoral Studies* 29(4): 636-647.
- Cox, G. W., Rosenbluth, F. M., and Thies, M. F. (2000). Electoral Rules, Career Ambitions, and Party Structure: Comparing Factions in Japan's Upper and Lower Houses. *American Journal of Political Science* 44(1): 115-122
- Debus, M. and Bräuninger, T. (2009). Intra-Party Factions and Coalition Bargaining, in Giannetti, Daniela and Benoit, Kenneth (edited by) *Intra-Party Politics and Coalition Government*, Routledge.
- Evans, C. L. (2001). Committees, Leaders, and Message Politics, in *Congress Reconsidered*, ed. Lawrence C. Dodd, and Bruce I. Oppenheimer. Washington, D.C.: CQ Press.
- Formenti, L. N. (2012). Italian Politics and Social Networks: Modern Attempts to Close the Distance. Universit de Fribourg, draft
- Giannetti, D. and Benoit, K., eds. (2009). *Intra-Party Politics and Coalition Government*. New York, NY: Routledge.
- Giannetti, D. and Laver, M. (2009). Party Cohesion, Party Discipline, Party Factions in Italy. Pp. 146-68 in *Intra-Party Politics and Coalition Government*, edited by Daniela Giannetti and Kenneth Benoit. New York, NY: Routledge.
- Heller, W. B. (2008). Legislative Organization, Party Factions, and Party Policy Position: Endogenous Adaptation to Exogenous Institutions. Unpublished manuscript.
- Hix, S., and Jun, H.W. (2009). Party behaviour in the parliamentary arena. *Party Politics* 15(6): 667-694.
- King, A., F. Orlando, and D. Sparks (2011). Ideological extremity and primary success: A social network approach. Paper presented at the 2011 MPSA Conference.
- Klüver, H. (2009). Measuring Interest Group Influence Using Quantitative Text Analysis. *European Union Politics* 10: 535-549.
- McGann, A. J. (2002). The Advantages of Ideological Cohesion: A Model of Constituency Representation and Electoral Competition in Multi-Party Democracies. *Journal of Theoretical Politics* 14: 37-70.
- Messner, W. B. (2003). Taming Labour's MEPs. *Party Politics* 9(2): 201-218.
- Proksch, S-O, and Slapin J.B. (2009a) WORDFISH: Scaling Software for Estimating Political Positions from Texts. Version 1.3 (22 January 2009), <http://www.wordfish.org>.

- Proksch, S-O, and Slapin J.B. (2009b) How to avoid pitfalls in statistical analysis of political texts: The case of Germany. *German Politics* 18(3): 323-344.
- Proksch, Sven-Oliver and Slapin, Jonathan B. (2010). Position Taking in European Parliament Speeches. *British Journal of Political Science* 40: 587-611.
- Proksch, S-O, and Slapin J.B. (2012) Institutional foundations of legislative speeches. *American Journal of Political Science* DOI: 10.1111/j.15405907.2011.00565.x.
- Sartori, G. (1976). *Parties and Party System: A Framework for Analysis*. Cambridge University Press, New York.
- Seddone, A. (2013). Per chi suona la campana. Uno sguardo sugli endorsement. *Questioni Primarie* 4: 7-8 Accessible from: <http://www.candidateandleaderselection.eu/content/qp-4-chi-suona-la-campana-uno-sguardo-sugli-endorsement>
- Slapin, J. B. and Proksch, S-O (2008). A scaling model for estimating time-series party positions from texts. *American Journal of Political Science* 52(3):705-722.
- Snyder, Jr J. M. and Ting, M. M. (2002). An Informational Rationale for Political Parties. *American Journal of Political Science* 46(1): 90-110.
- Spirling, A. and Quinn, K. M. (2010). Identifying Intra-Party Voting Blocs in the UK House of Commons. *Journal of the American Statistical Association* 490: 447-457.
- Vaccari, C. (2008). Research Note: Italian Parties' Websites in the 2006 Elections. *European Journal of Communication* 23(1): 69-77
- Vergeer, M., Hermans, L., and Cunha, C. (2012) Web campaigning in the 2009 European Parliament elections: A cross-national comparative analysis. *New Media Society* doi:10.1177/1461444812457337

Politicians Go Social

Estimating intra-party heterogeneity (and its effects)
through the analysis of social media

Andrea Ceron & Alessandra C. Cremonesi

University of Milan

May 10, 2013

Introduction

Politics & Social Media

Text Analysis of SM

Estimating policy positions from SM

Why Social Media

From Words to Fact(ion)s

Counting Factions

Text Analysis

Applications

Endorsement

Negative Campaign

Conclusion

Content Analysis of SM

What's Next?

Politics & Social Media

Rise of social media (SM & SNS)

- ▶ Italians: 40% active on FB; 8% on TW (2013)
- ▶ attractive to parties and candidates
- ▶ mobilize support

Politicians make use of it

- ▶ Italian MPs: 36% active on FB; 10% on TW (2011)
- ▶ politicians v/ population: FB usage is equal; TW usage seems higher
- ▶ express opinions
- ▶ debate day-by-day over agenda
- ▶ top-down approach

People talk about politics (e.g., 2012 Primary Elections)

- ▶ lot of data to be exploited

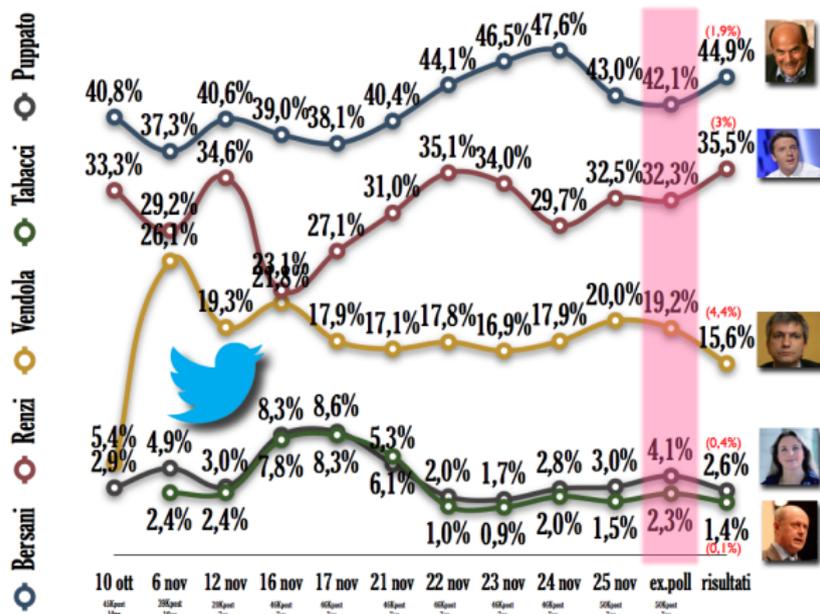
2012 Primary Elections

Centre-left primary elections (PD, SEL, PSI)

- ▶ November/December 2012
- ▶ 5 candidates
- ▶ frontrunner: Pierluigi Bersani, PD (80k friends; 135k followers)
- ▶ runnerup: Matteo Renzi, PD (155k friends; 152k followers)
- ▶ heated debate (over 500k tweets)

2012 Primary Elections

Analysis of opinions expressed on-line useful to predict the results
cf. Ceron et al. 2012 (Stefano Iacus will talk about!)



How to assess the preferences of 'hidden actors'?

Analysis of SM helps assessing policy position of political actors
(Conover et al. 2010; King et al. 2011; Boutet et al. 2012)

Can solve a challenging task: Estimating positions of 'hidden actors'

- ▶ party \neq unitary actor
- ▶ formal and informal intra-party subgroups (they matter!)
- ▶ policy views not always formally recorded or observable
- ▶ pressure to party unity

Display of party unity may hide internal division

- ▶ due to whip, loyalty, or consensual bargaining

Even when splits are observable, heterogeneity could be misestimated

Estimating positions of intra-party actors

Roll call

- ▶ ex-post behaviour instead of ex-ante preferences (Hix & Jun 2009)
- ▶ selection bias (Carrubba et al. 2006,2008)

Parliamentary speeches

- ▶ speeches in highly institutionalized arenas are subject to party whip
- ▶ overestimate party cohesion (Proksch & Slapin 2012)

Documents drafted by intra-party groups

- ▶ intra-party debates less affected by party whip (still)
- ▶ seldom available and not issued on a regular basis
- ▶ misestimate cohesion when the minority feels too weak
- ▶ troubles when subgroups are not steady

Estimating positions of hidden actors

Content Analysis of SM

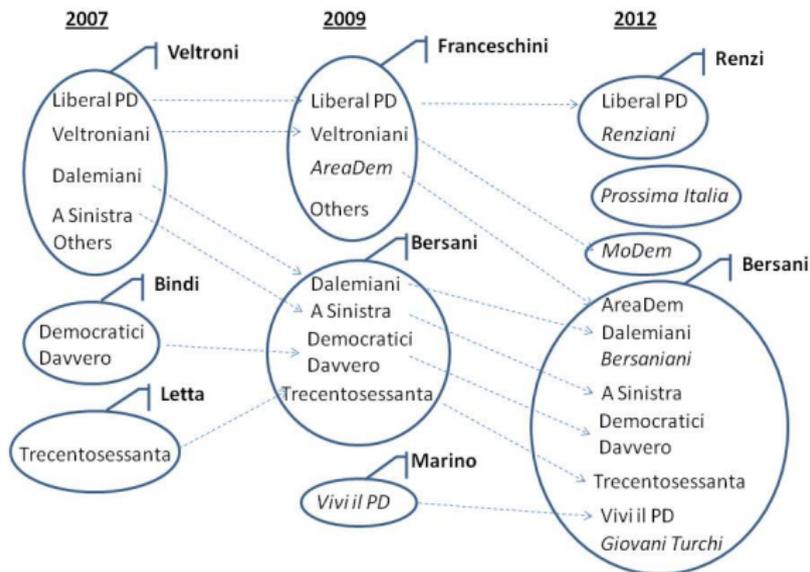
- ▶ express people's attitudes covering all the topics of political agenda
- ▶ unmediated, unsolicited, and self-expression oriented statements
- ▶ strategic behavior lower in on-line informal environment
- ▶ could be more spontaneous and impulsive
- ▶ perceived as free preliminary personal opinions not subjected to party whip

SM declarations are more likely to reflect the sincere preferences

Factions within the Italian Democratic Party (2007-2012)

Splits, reshuffles and emergence of new factions

Evolution of the intra-party structure within the Italian Democratic Party (PD)



Data collection

12 Factions identified ex-ante

Data: on-line texts written by factional leaders

- ▶ several sources: Facebook, Twitter, blogs
- ▶ comments posted during XVI Leg. (apr2008-dec2012)
- ▶ posts with political content

Policy positions based on content analysis (Wordfish)

- ▶ avg. 48k words per document (5k,144k)

Wordfish

Words Frequencies/Policy Positions

$$y_{ijt} \sim \text{Poisson}(\lambda_{ijt}) \quad (1)$$

$$\lambda_{ijt} = \exp(\alpha_{it} + \psi_j + \beta_j \times \omega_{it}) \quad (2)$$

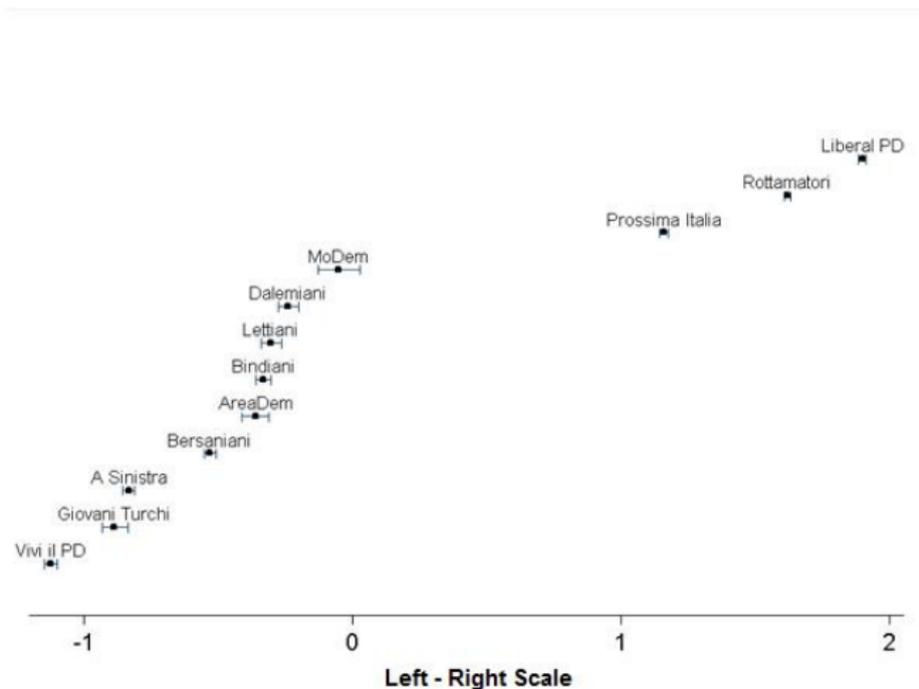
Documents cover several topics

- ▶ arrayed on ideological Left-Right scale

Valid & reliable results

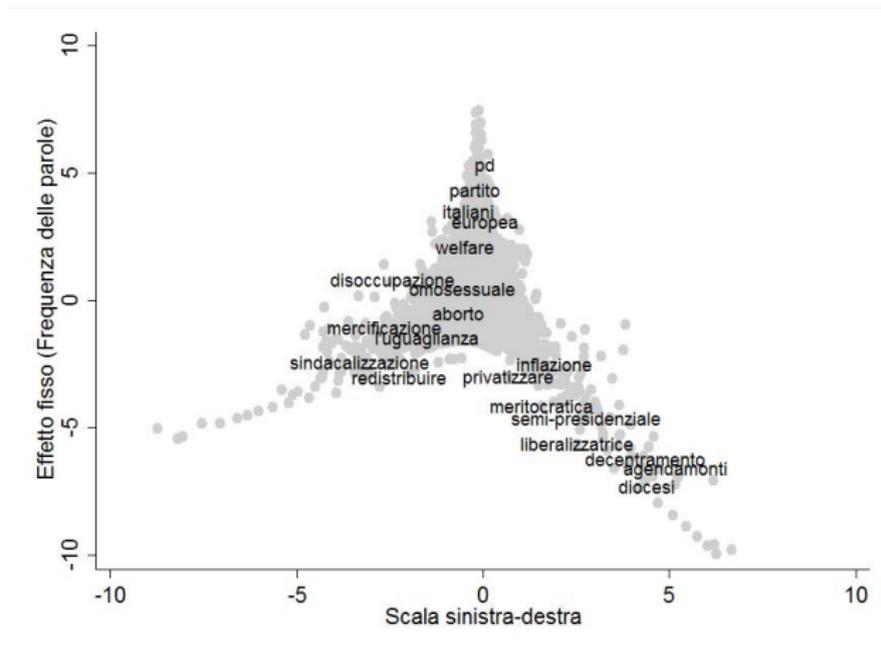
From Words to Fact(ion)s

Positions of party factions



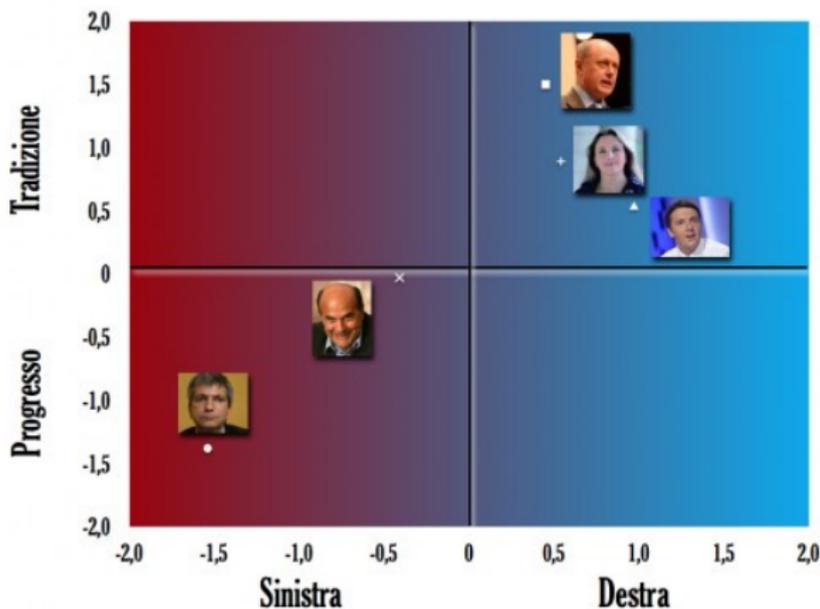
Diagnostic

Words' Discriminating Power



Reliability

Placement of Bersani and Renzi based on 2012 primary election manifestos



Endorsement

List of official endorsements made by PD politicians (Seddone 2012)

- ▶ on three main national newspapers (Corriere, Repubblica, Unità)

Assess factional affiliation of each

- ▶ based on membership or personal biography

Likelihood of endorsing Bersani

Control variables

- ▶ clash of two generations?

Endorsement

Table: Logit regression of Endorsement

Variables	(I)	(II)
Distance	-2.393*** (0.723)	-2.414*** (0.911)
Incumbent		0.657 (0.544)
Age		0.056** (0.023)
Female		2.368** (0.918)
Experience		-0.097** (0.042)
Constant	2.505*** (0.410)	-0.741 (1.275)
Observations	87	87
Log pseudolikelihood	-31.702	-27.473

Standard errors clustered by faction in parentheses

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

Negative Campaign

Attacks: #direct mentions against Bersani or Renzi

- ▶ made on-line by factional leaders
- ▶ 80% after primary election was announced
- ▶ negative campaign

Propensity to criticize and attack a party leader/candidate

Control variables

- ▶ insiders want to tone down the debate?

Negative Campaign

Table: Negative binomial regression of Attacks

Variables	(I)	(II)
Distance	0.719*** (0.161)	0.670*** (0.151)
Incumbent		-1.620** (0.640)
Age		0.046* (0.025)
Female		-1.148 (0.897)
Experience		0.040 (0.057)
Constant	-10.849*** (0.352)	-12.243*** (0.942)
Observations	54	54
Log pseudolikelihood	-56.958	-53.674
Alpha	1.599 (1.004)	0.913 (0.693)

Standard errors clustered by faction in parentheses

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

Content Analysis of SM

Politicians inclined to wash the dirty linen on SM/SNS

Express policy opinions on-line

Unsolicited and personal sincere preferences

- ▶ less affected by loyalty & party whip
- ▶ compared to official documents or behavior in formal arenas (speeches/votes)

Can be used to assess preferences of 'hidden actors' (e.g., factions)

- ▶ help understanding intra-party politics
- ▶ policy motives drive endorsement & negative campaign

What's Next?

What happens when

- ▶ incentive to cultivate personal votes are lower?
- ▶ emphasis on party unity is higher?
- ▶ dissent is not allowed?

Comparison across parties and countries

This technique can help estimating positions of influent opinion makers

- ▶ e.g., bloggers and journalists