

Mining Democracy

Matthias Grossglauser, EPFL

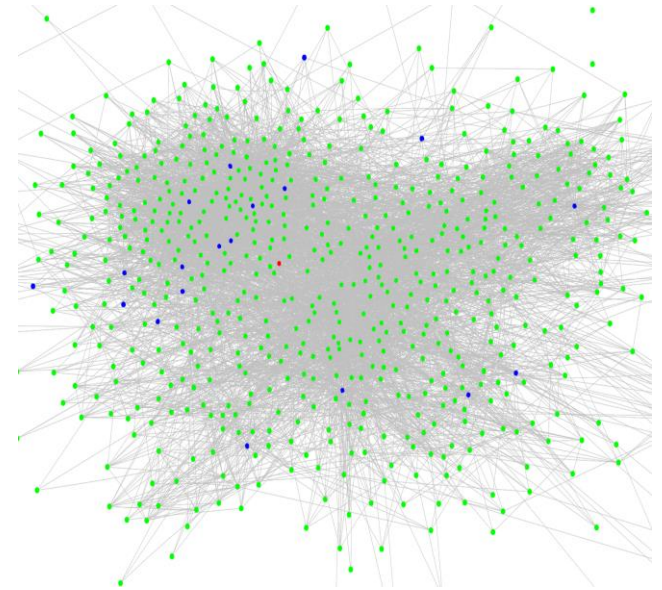
Collaborators:

Vincent Etter, Julien Herzen, Emtiyaz Khan, Victor Kristof,
Patrick Thiran

EIT ICT Labs Summer School
August 2015

Overview: Information and Network Dynamics Group

- Team goals: modeling large systems of social interactions
 - Online social networks
 - Mobility
 - Epidemics
 - Crowdsourcing
 - ...



- Democracy = a rich & complex social system
 - Switzerland: sophisticated political system, direct democracy
 - Data: Open Government initiatives

Our laboratory: Switzerland

- Diversified party landscape
- Four official languages
- smartvote: available since 2003
- Direct democracy with frequent issue votes on various subjects
 - at both parliamentary and citizen levels



Switzerland: direct democracy



LEERE KASSEN?



NEIN ZU DEN KAMPFJET-
MILLIARDEN!

Data sources

- 1: Smartvote
 - 32 questions covering different societal & political themes
 - Answers from candidates and citizens before the election of the Nationalrat in 2011
 - ~3'000 candidates (82.4% of all candidates)
 - ~220'000 citizens (9% of active voters)
- 2: Parliament votes
 - ~2'500 votes (2011 - 2013) of the 181 candidates elected in 2011 (with smartvote profile)
- 3: Federal initiatives (plebiscites)
 - Result of 245 federal votes (1981 - 2011) for ~2'400 municipalities

Dimensionality reduction: 2-D is easy

- Example:
 - Chocolate consumption vs Nobel laureates
- Visual detection of relationships and trends

principal direction («component»)

[F. H. Messerli, New Engl J Med 2012]

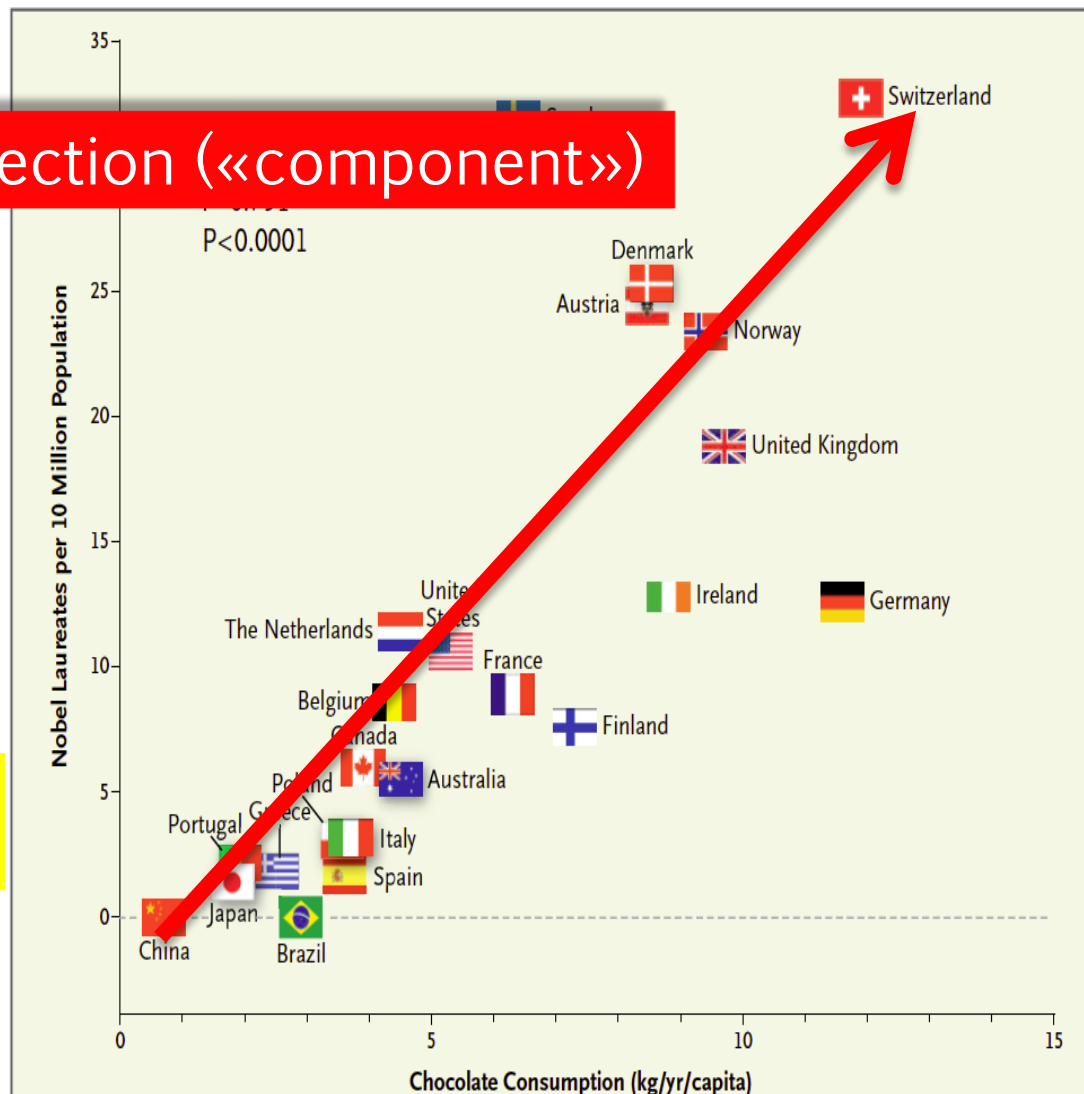
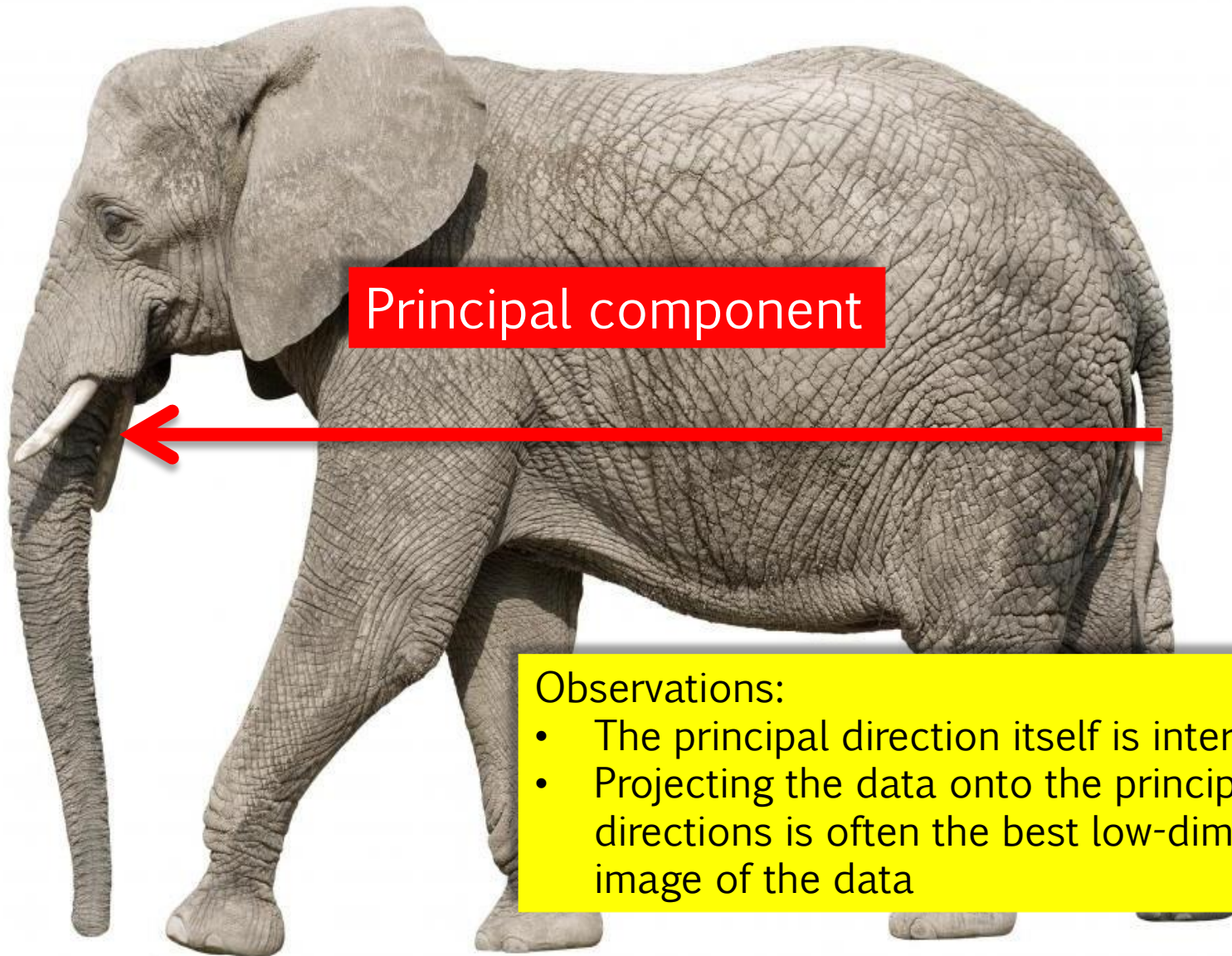


Figure 1. Correlation between Countries' Annual Per Capita Chocolate Consumption and the Number of Nobel Laureates per 10 Million Population.

Which is the best «direction of projection»?



Which is the best «direction of projection»?



Principal component

Observations:

- The principal direction itself is interesting
- Projecting the data onto the principal directions is often the best low-dimensional image of the data

High-dimensional data is hard

- First example: «chocolate» vs «Nobel awards»
 - 2 → 1 dimension
- Second example: elephant in 3-D
 - 3 → 2 dimensions
- Typical data: very high dimension
 - Smartvote data: 32 questions (point=candidate)
 - Municipality data: 245 votes (point=municipality)
 - Parliament: data 2500 votes (point=legislator)
- Questions:
 - What are the principal components?
 - What does the data projected into the «ideological space» look like?

smartvote

Smartvote dataset


Questionnaire > Modifier la recommandation de vote > Recommandation de vote

Questionnaire

1. Etat social & famille (0/4)

Catégories

Question par question

1. Êtes-vous favorables à une hausse de l'âge de la retraite pour les hommes et les femmes (p. ex. à 67 ans)? 

Oui

Plutôt oui


Plutôt non

Non

Pas de réponse

Pondération



2. Approuveriez-vous l'introduction d'indemnités journalières dégressives dans le cadre de l'assurance chômage (c'est-à-dire que le montant de l'indemnité journalière diminue au fur et à mesure que la durée du chômage augmente)? 

Soutenir smartvote!

J'aimerais soutenir smartvote:

 Flattr

254

Pour les candidat-e-s

Vous trouverez plus d'informations sur notre portail candidat.

Remarque

- Afin de pouvoir établir une recommandation de vote, vous devez répondre à au moins une question.
- La recommandation de vote sera plus précise si vous répondez à un grand nombre de questions.
- Si vous désirez en savoir plus sur une question, cliquez sur information. Des

Smartvote dataset

- smartvote pre-electoral opinions of the 2011 parliamentary elections
 - 2,985 candidates (82.4% of all candidates)
 - 229,133 citizens (~9% of total turnout)
- Examples of questions:
 - “Should Switzerland embark on negotiations in the next four years to join the EU?”
 - “How much should the public transport budget be?”
- Possible answers
 - strongly disagree - disagree - agree - strongly agree
 - less - no change - more

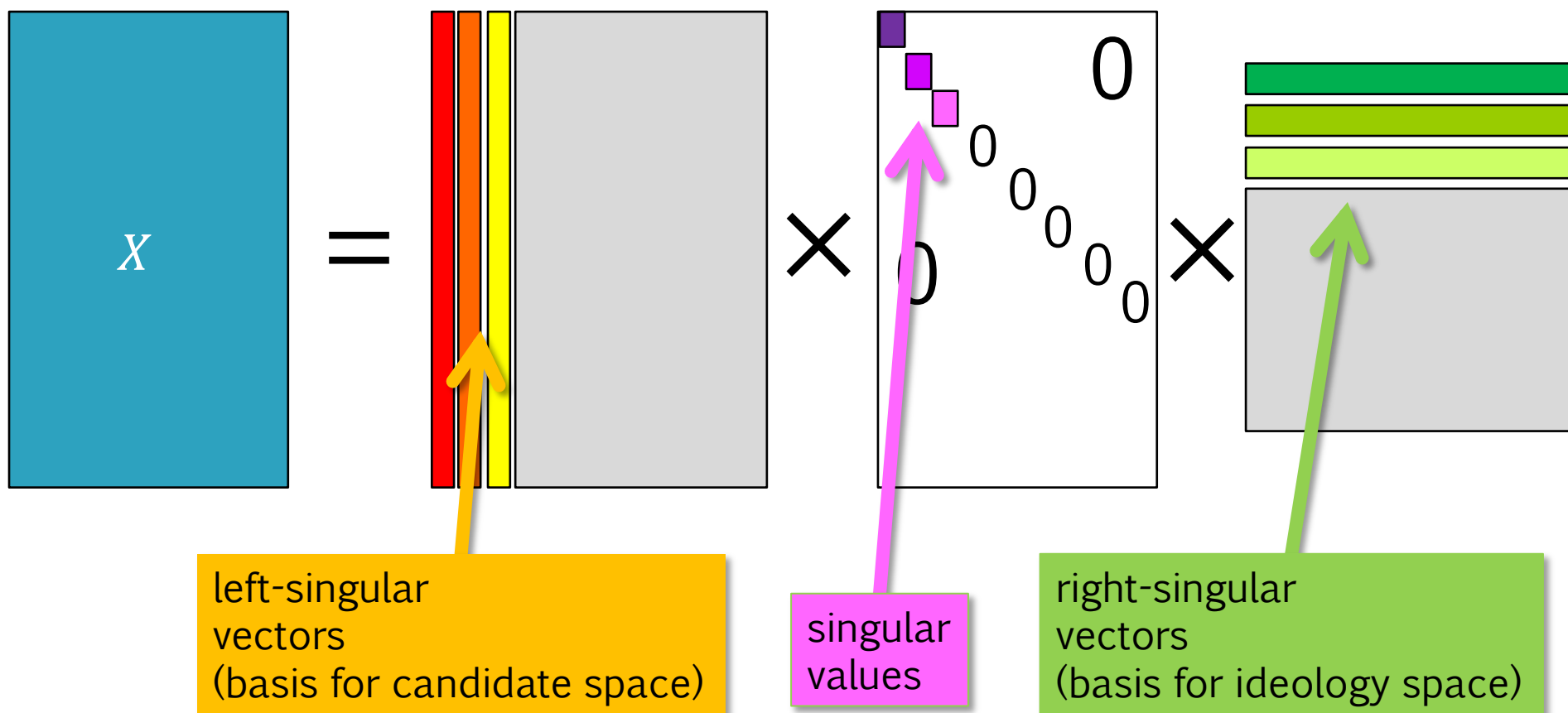
Discriminative questions

- What questions discriminate best the opinion of candidates?
 - Is the traditional left/right view meaningful?
- Use dimensionality reduction to find out
- Use SVD on the matrix X of candidates' responses

$$C \text{ candidates} \left\{ \begin{array}{c} \overbrace{\left(\begin{array}{cccc} 0.5 & 0.25 & \dots & 0.0 \\ 0.75 & 0.5 & \dots & 1.0 \\ \vdots & \vdots & \ddots & \vdots \\ 1.0 & 0.25 & \dots & 0.75 \end{array} \right)}^{n \text{ questions}} \end{array} \right.$$

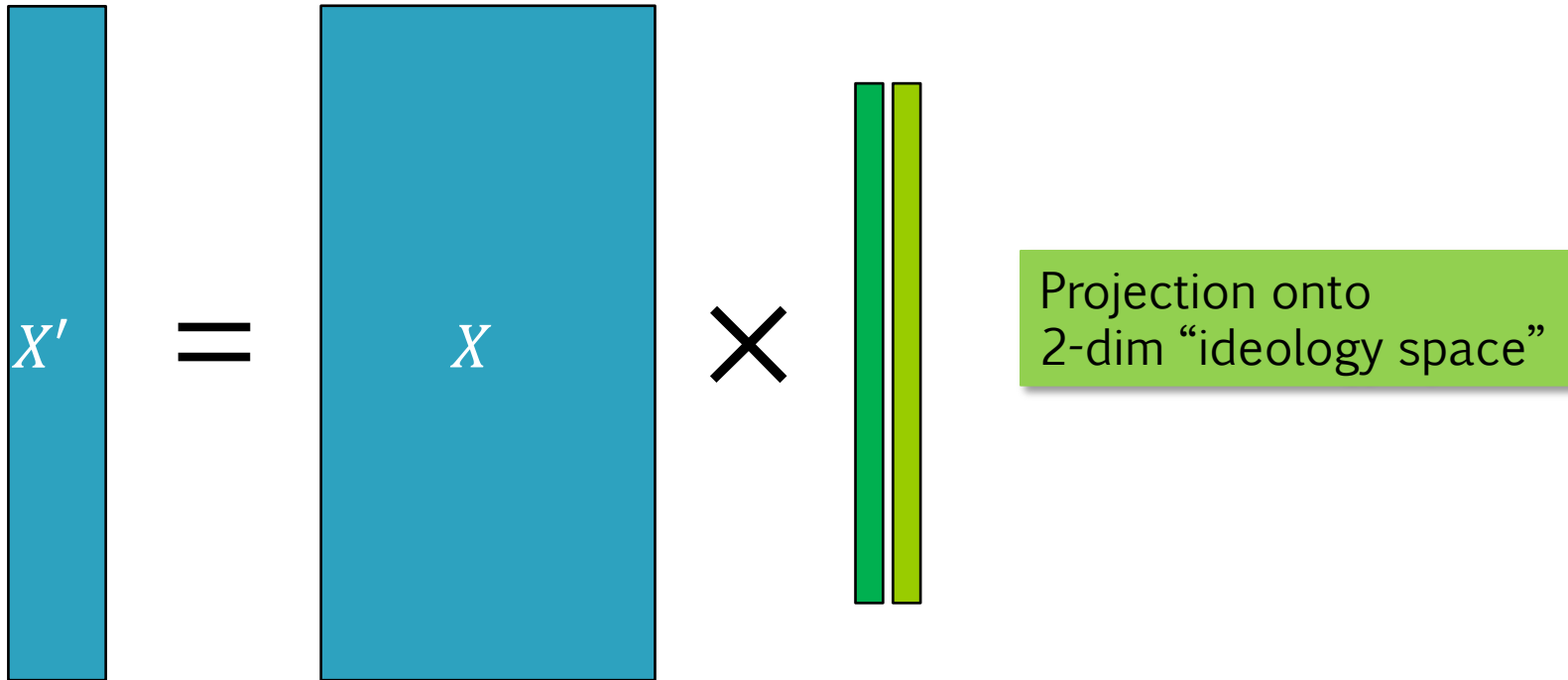
Dimensionality reduction

- 1: Center X
- 2: Singular Value Decomposition: $X = U\Sigma V^T$

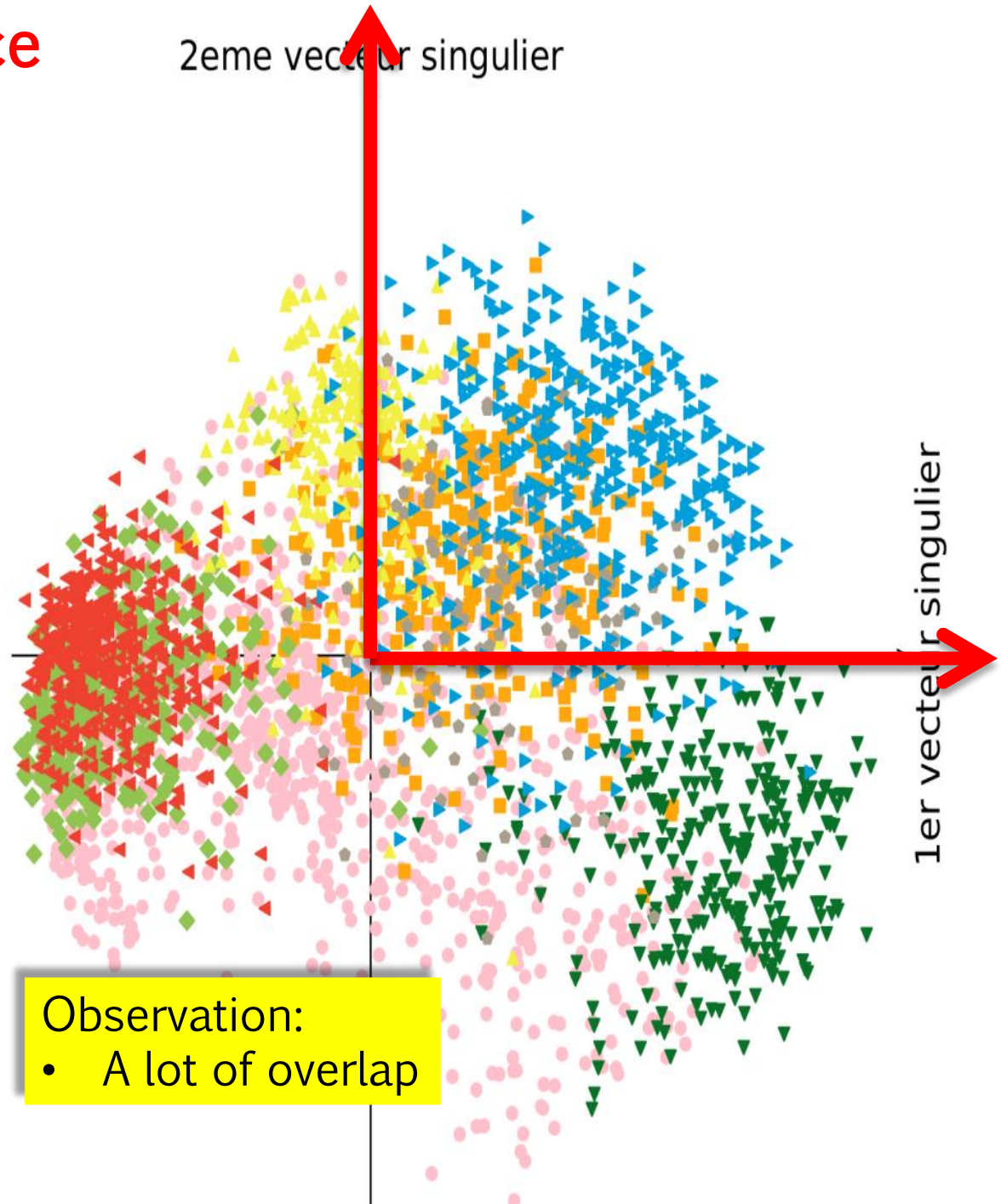
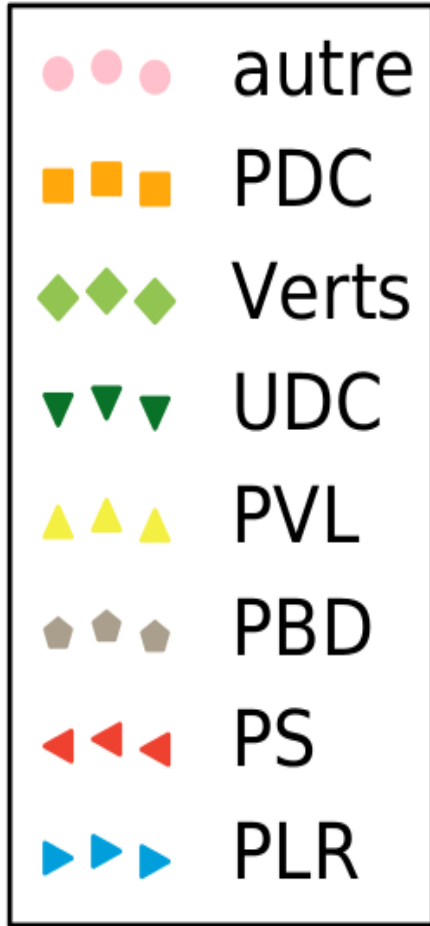


Dimensionality reduction: projection

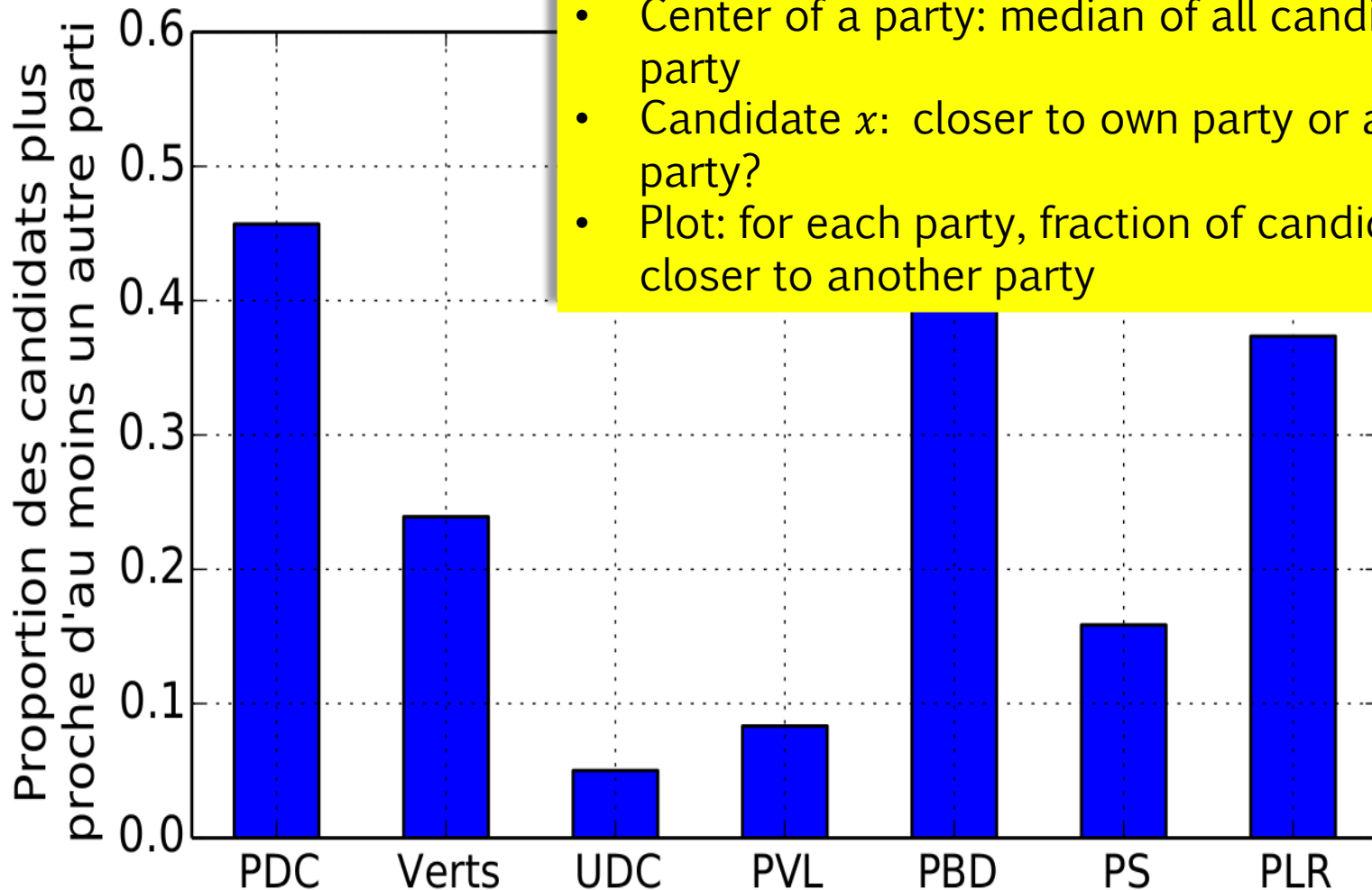
- Projecting X onto first two (right) singular vectors
 - $X' = X V_{\{1,2\}}$
- X' is $C \times 2$



Ideological space



Party overlap



Observation:

- Center of a party: median of all candidates of party
- Candidate x : closer to own party or another party?
- Plot: for each party, fraction of candidates closer to another party

Principal components

1st axis

- Seriez-vous favorable à ce que le **droit de vote** au niveau communal soit instauré pour les **étrangers** qui vivent en Suisse depuis au moins dix ans et ce, dans toute la Suisse?
- Approuveriez-vous que la **concurrence fiscale** entre les **cantons** soit plus limitée?
- Soutenez-vous l'initiative populaire qui souhaite que le **salaire** le plus élevé au sein d'une **entreprise** ne puisse pas être plus de douze fois supérieur au salaire le plus bas versé par la même entreprise. (initiative 1:12)?
- Une initiative populaire souhaite instaurer une **caisse maladie** unique et publique pour l'assurance de base. Êtes-vous favorable à ce projet?

Social questions («égalité»)

Principal components

2nd axis

- Approuvez-vous des engagements de soldats armés (pour l'autoprotection) de l'**armée** suisse à l'**étranger** dans le cadre de missions de maintien de la paix de l'ONU ou de l'OSCE?
- Êtes-vous en faveur d'un accord de **libre-échange** agricole avec l'**UE** ?
- Êtes-vous favorable à l'accord sur la **libre circulation** des personnes existant avec l'UE?
- Une imposition centrale sur les quantités dans la production laitière doit-elle être réinstaurée en Suisse à la place du **libre marché** laitier?

Economics, globalisation («liberté»)

Principal components

3rd axis

- Seriez-vous favorables à ce que l'**euthanasie** active directe soit légalement possible par le biais d'un médecin en Suisse?
- Les couples **homosexuels** sous le régime du partenariat enregistrés devraient-ils pouvoir adopter des enfants?
- La Suisse possède des règles relativement strictes concernant la **procréation** médicalement assistée. Celles-ci devraient-elles être assouplies?
- La consommation ainsi que la consommation personnelle de d... doivent-elles être légalisées?

Society, ethics («fraternité»)

Observation:

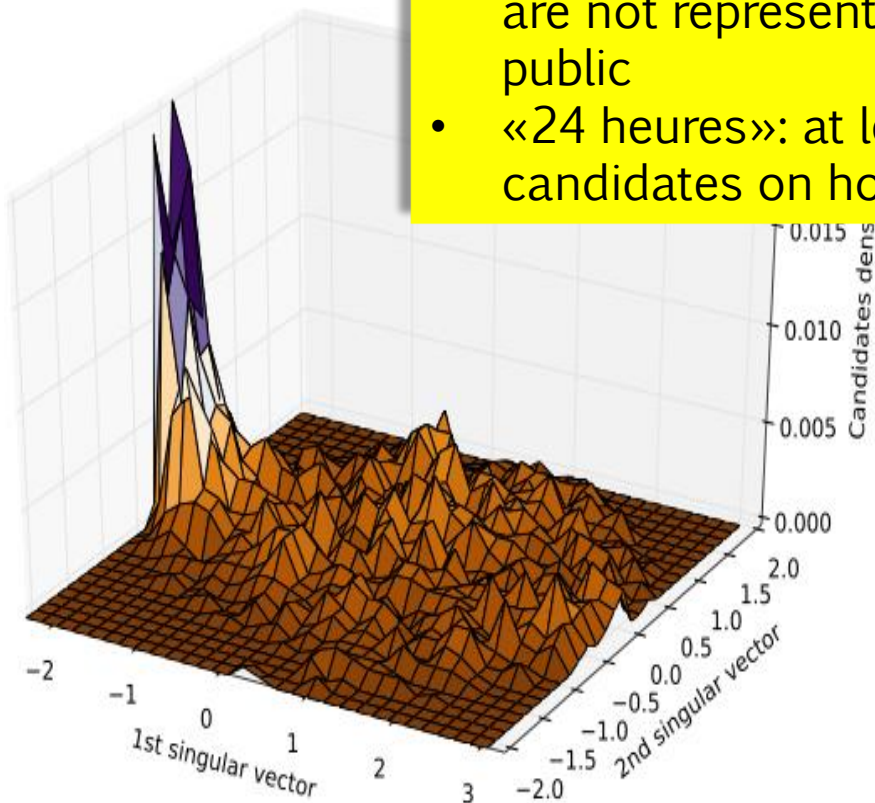
- Principal components correspond to clearly interpretable political and ideological dimensions

Smartvote: density over ideology space

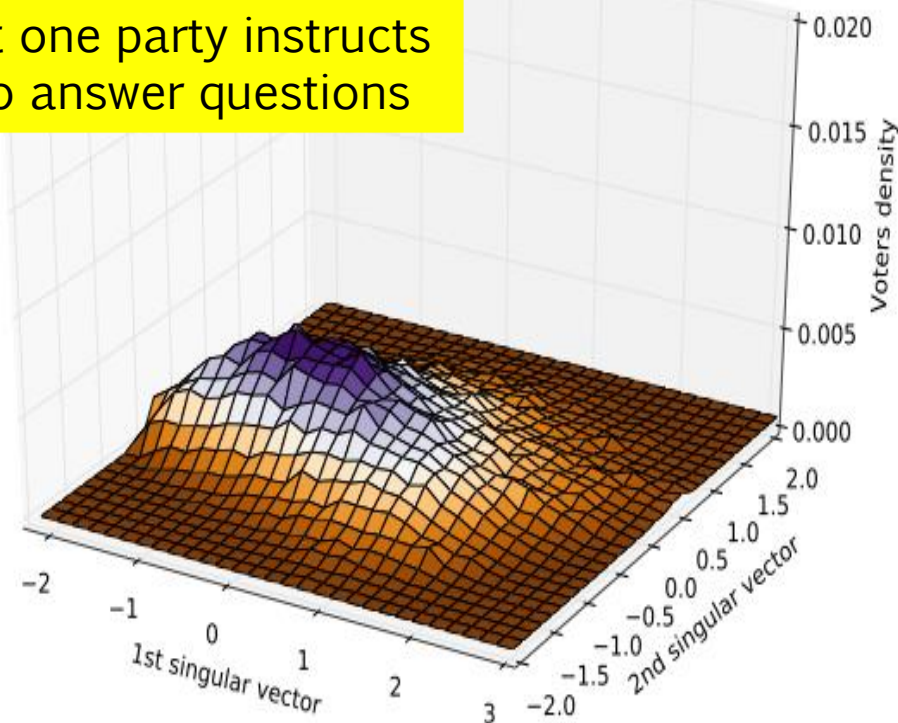
- The density profiles of politicians and citizens are very different

Observation:

- The positions of politicians on smartvote are not representative of the general public
- «24 heures»: at least one party instructs candidates on how to answer questions



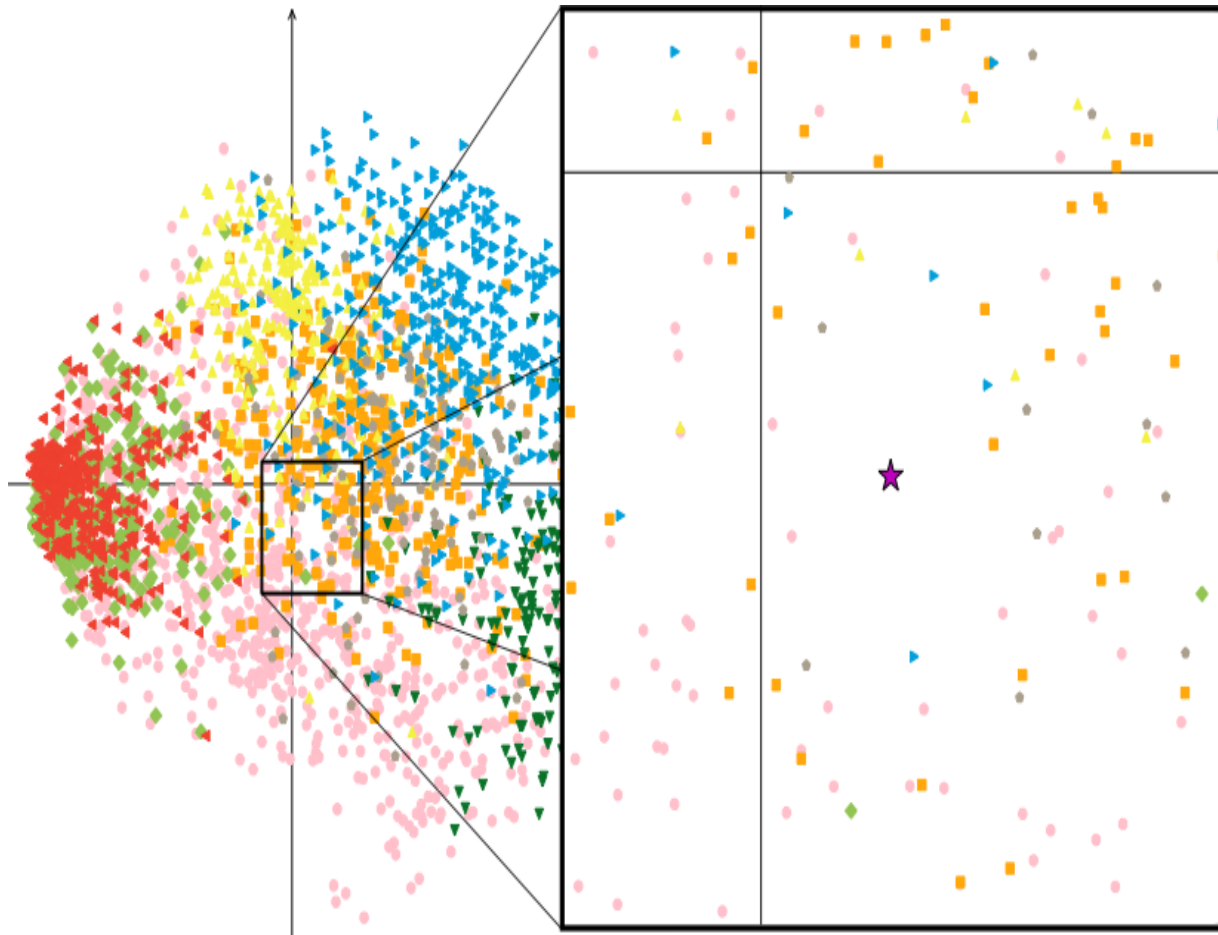
Candidates



Citizens

Political marketing?

- Thought experiment: could a candidate identify a position in ideology space which is (a) close to many voters, but (b) far from other candidates?

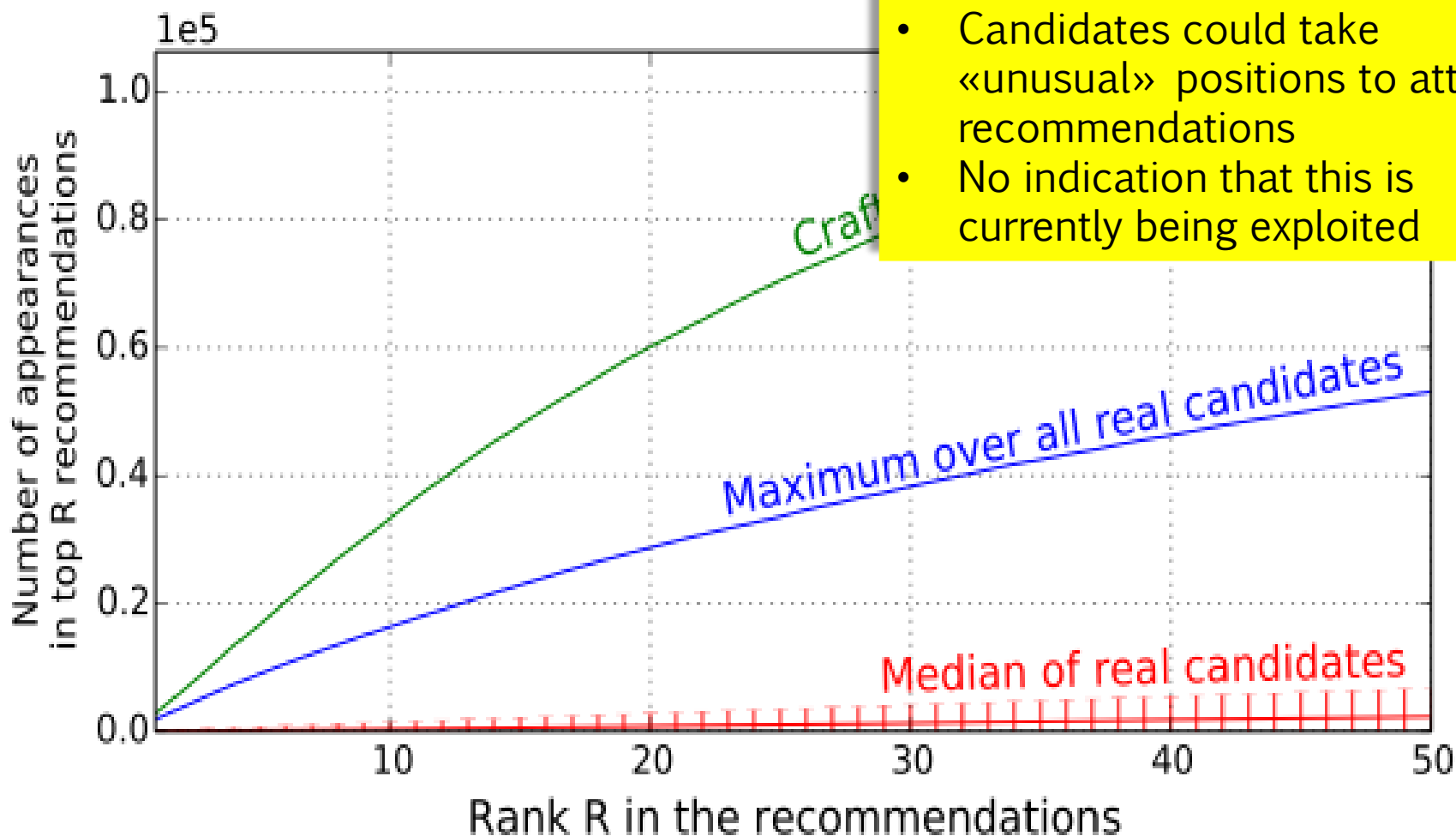


Impact of an crafted smartvote profile

- We compute optimal set of answers according to (a) and (b)
- Recomputed recommendations for the 229133 smartvote users
- Computed # recommendations by our «fake» candidate relative to genuine candidates

Results: fake vs genuine candidate

- The «fake» candidate would be recommended to almost $\frac{1}{2}$ of all smartvote users!



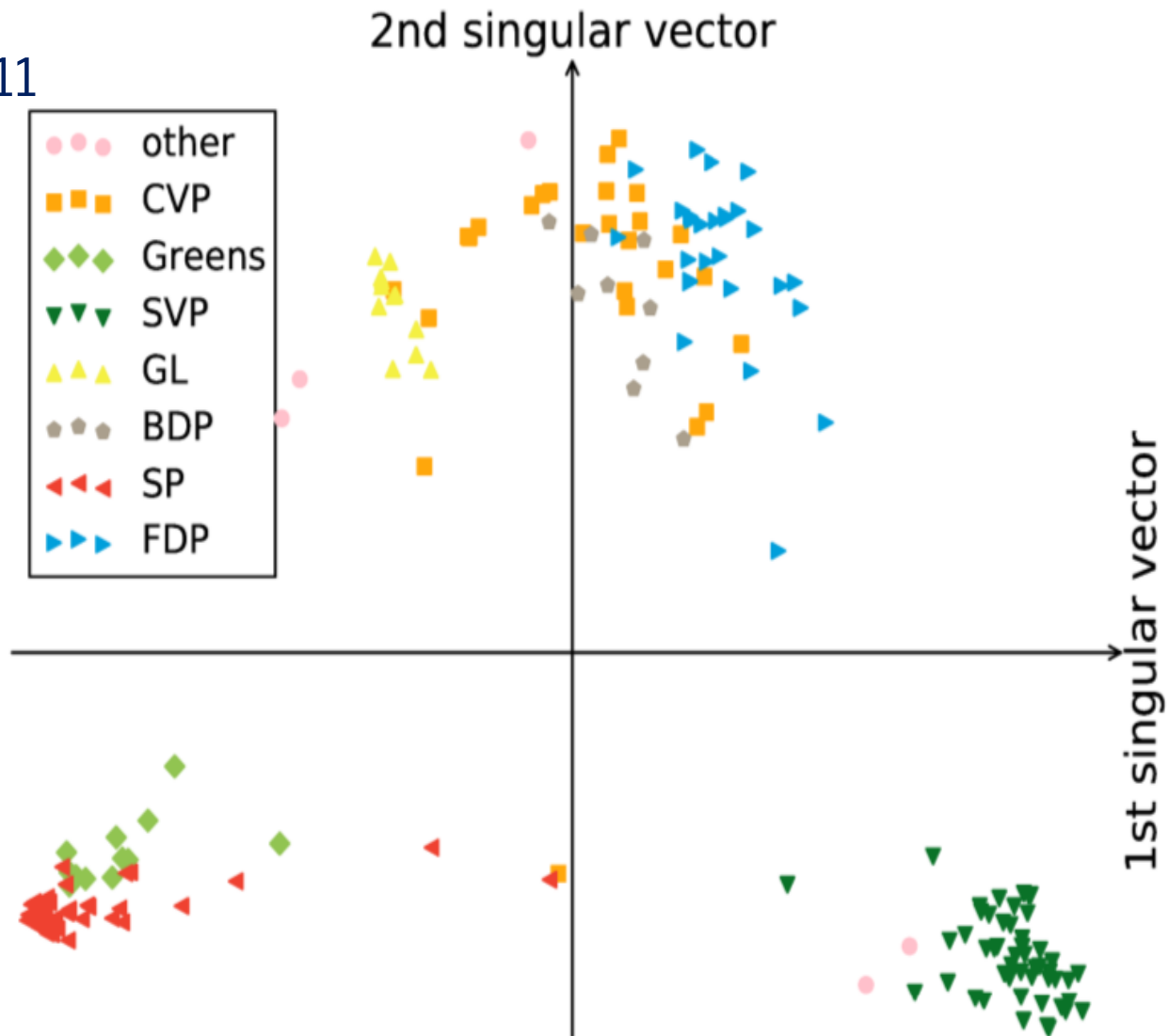
Observation:

- Candidates could take «unusual» positions to attract recommendations
- No indication that this is currently being exploited

parliamentary votes

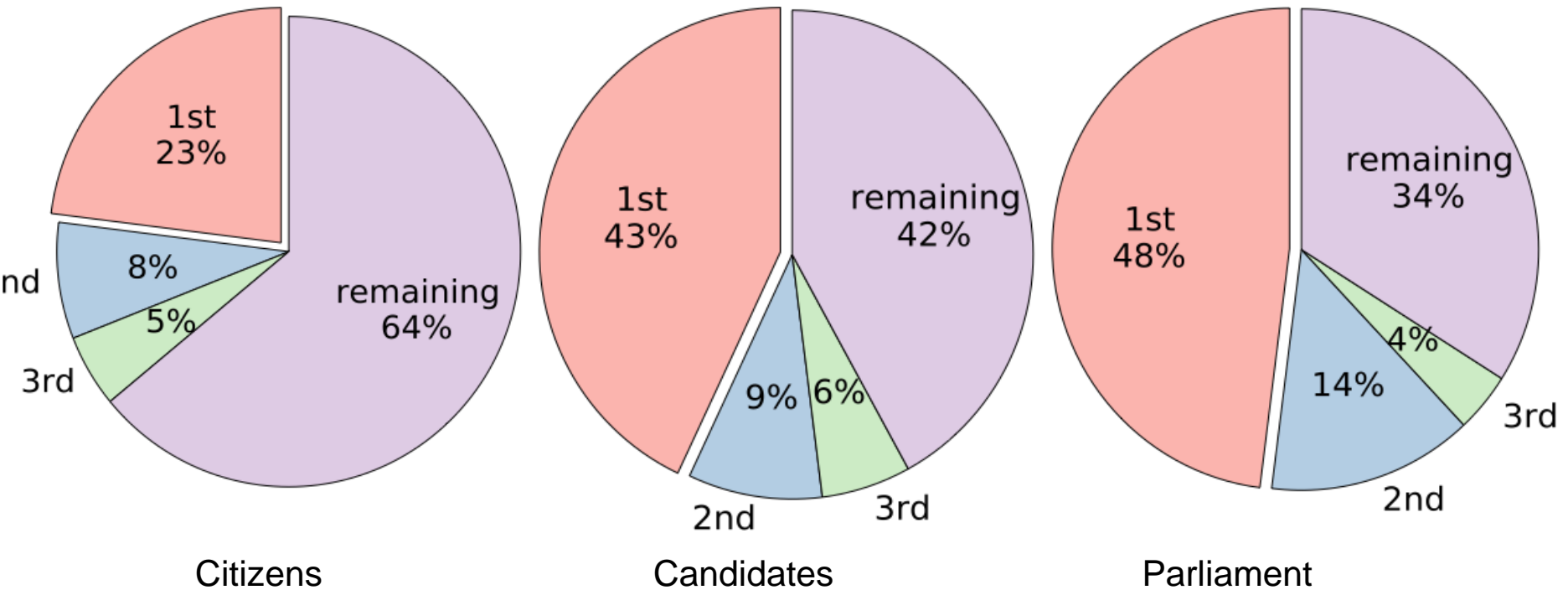
Parliament votes

- Public
 - 2,494 since the 2011 elections



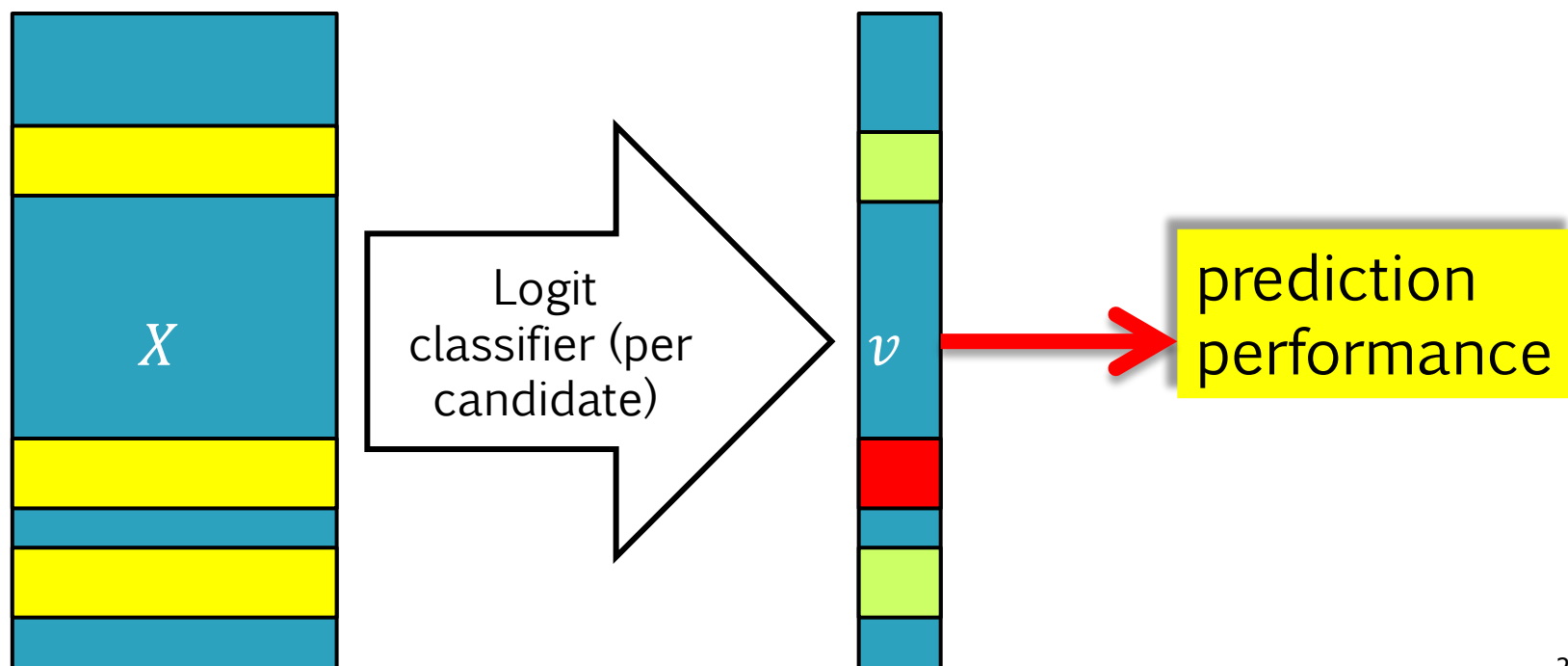
Polarization

- Citizens are more spread on the ideological plane, while candidates are more polarized
- This can be seen in the proportion of the variance captured by each singular vector



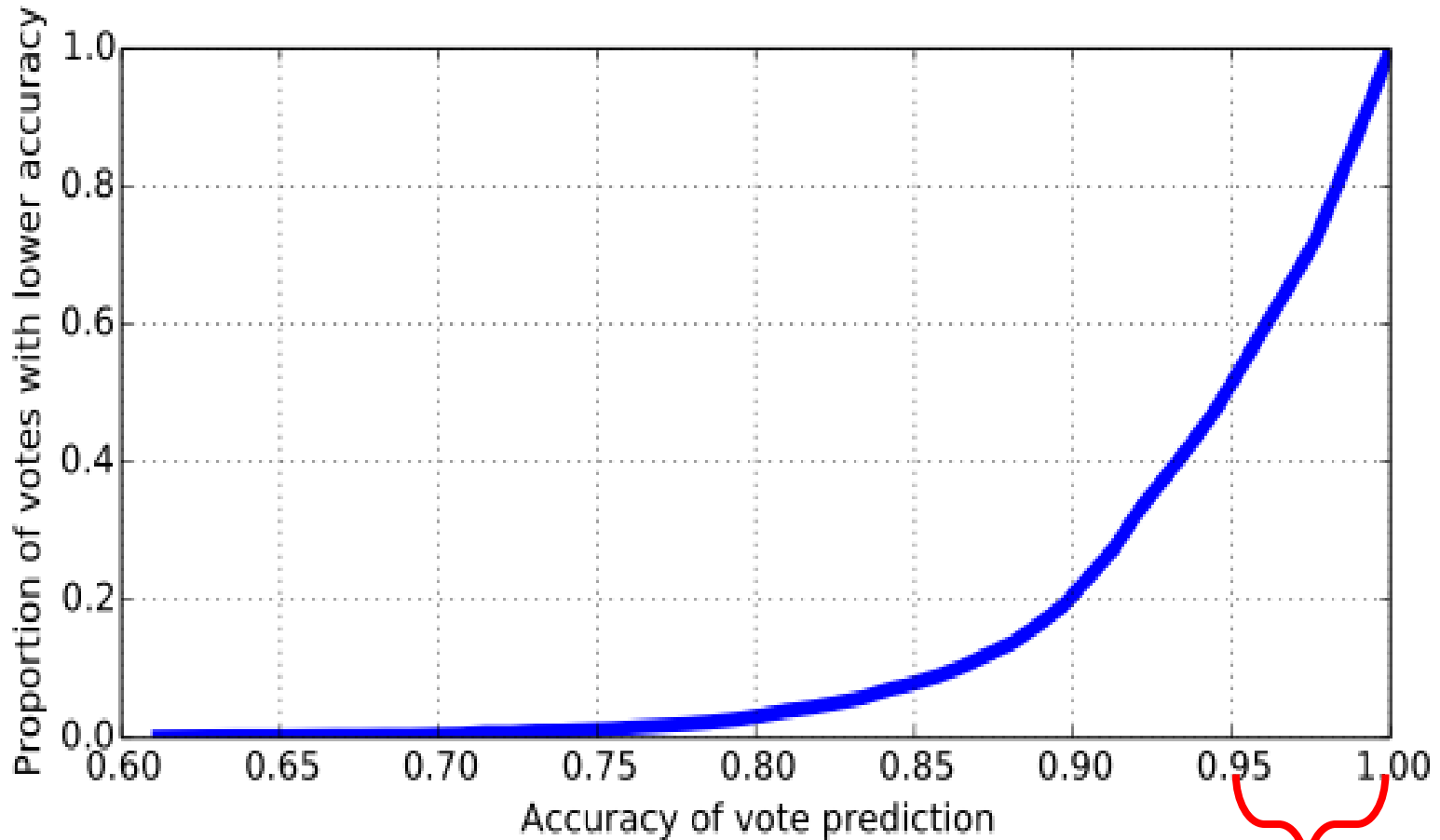
Do politicians flip-flop after being elected?

- Is it possible to detect whether a politician crafted his/her profile, given the way he/she votes once elected?
- Approach: learn a predictor for parliament vote v from smartvote answers
 - Then compare predicted and actual vote



Smartvote responses → predict parliament votes

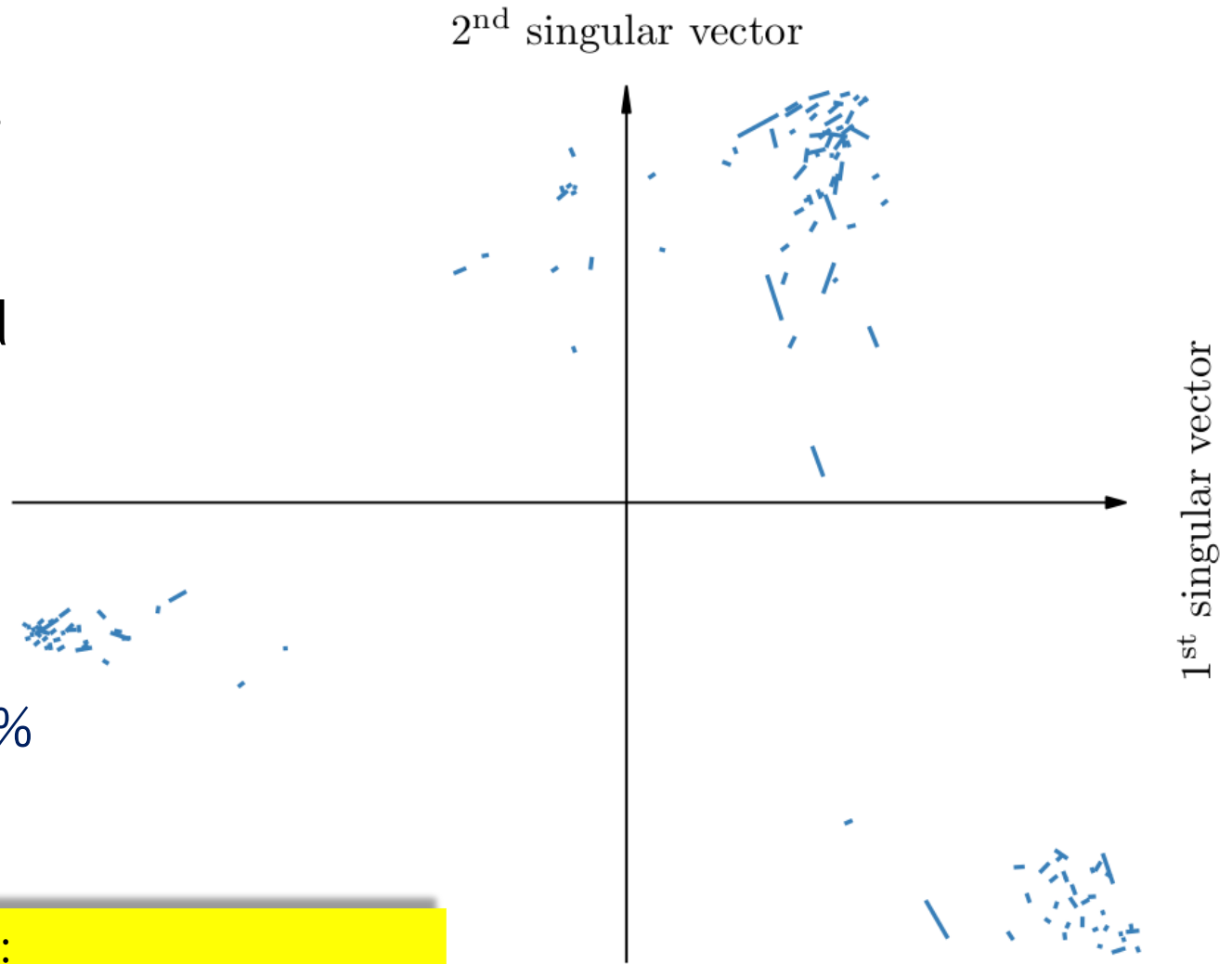
- Logit classifier predicts $\geq 50\%$ of the votes with $\geq 95\%$ accuracy



Predictable votes

Opinion shifts from campaign to parliament?

- Comparison between votes expected from smartvote responses and actual votes cast in parliament
- Using votes predicted with accuracy > 95%



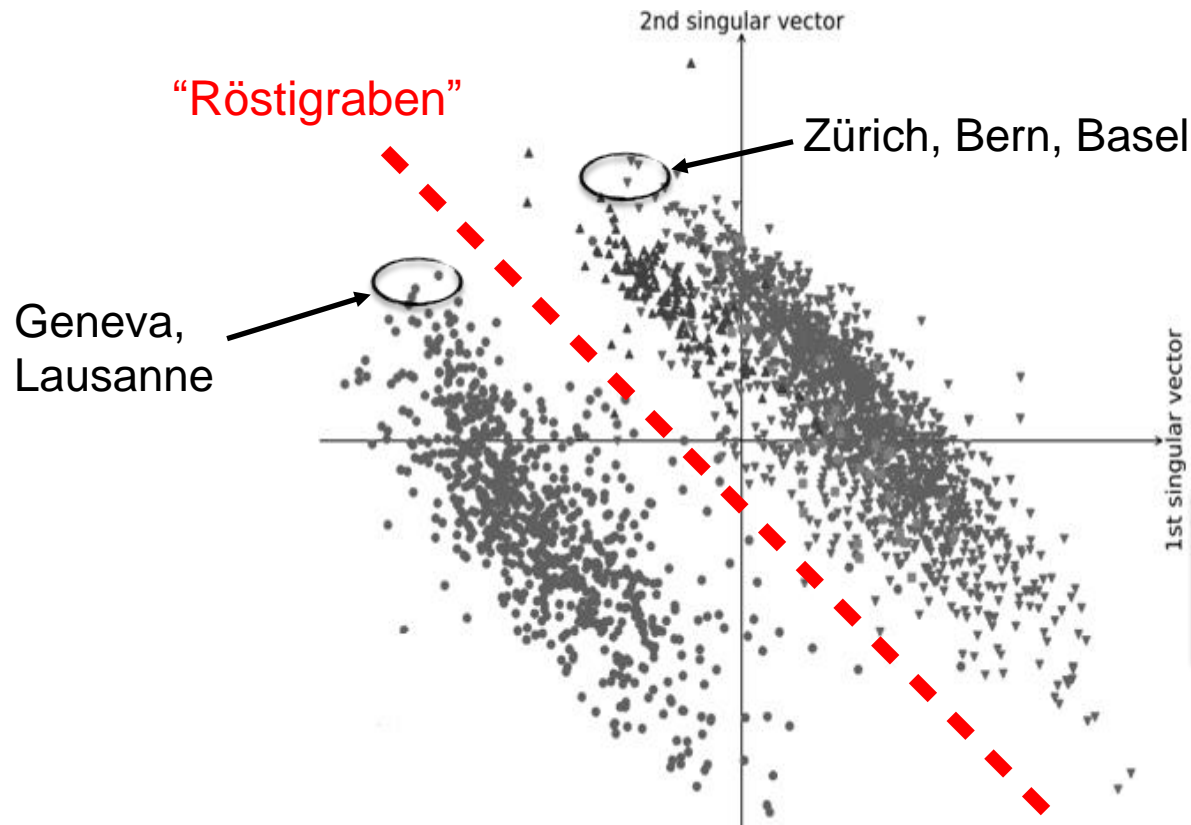
Observation:

- No significant flip-flopping effect

people votes

Voting patterns in Switzerland

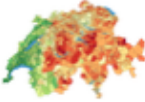









- Dataset: outcome (% yes) of 245 votes since 1981 in 2,398 municipalities
- Dimensionality reduction highlights linguistic / cultural contrasts



Explore the data:
www.predikon.ch











First principal component

Composition of the horizontal axis:

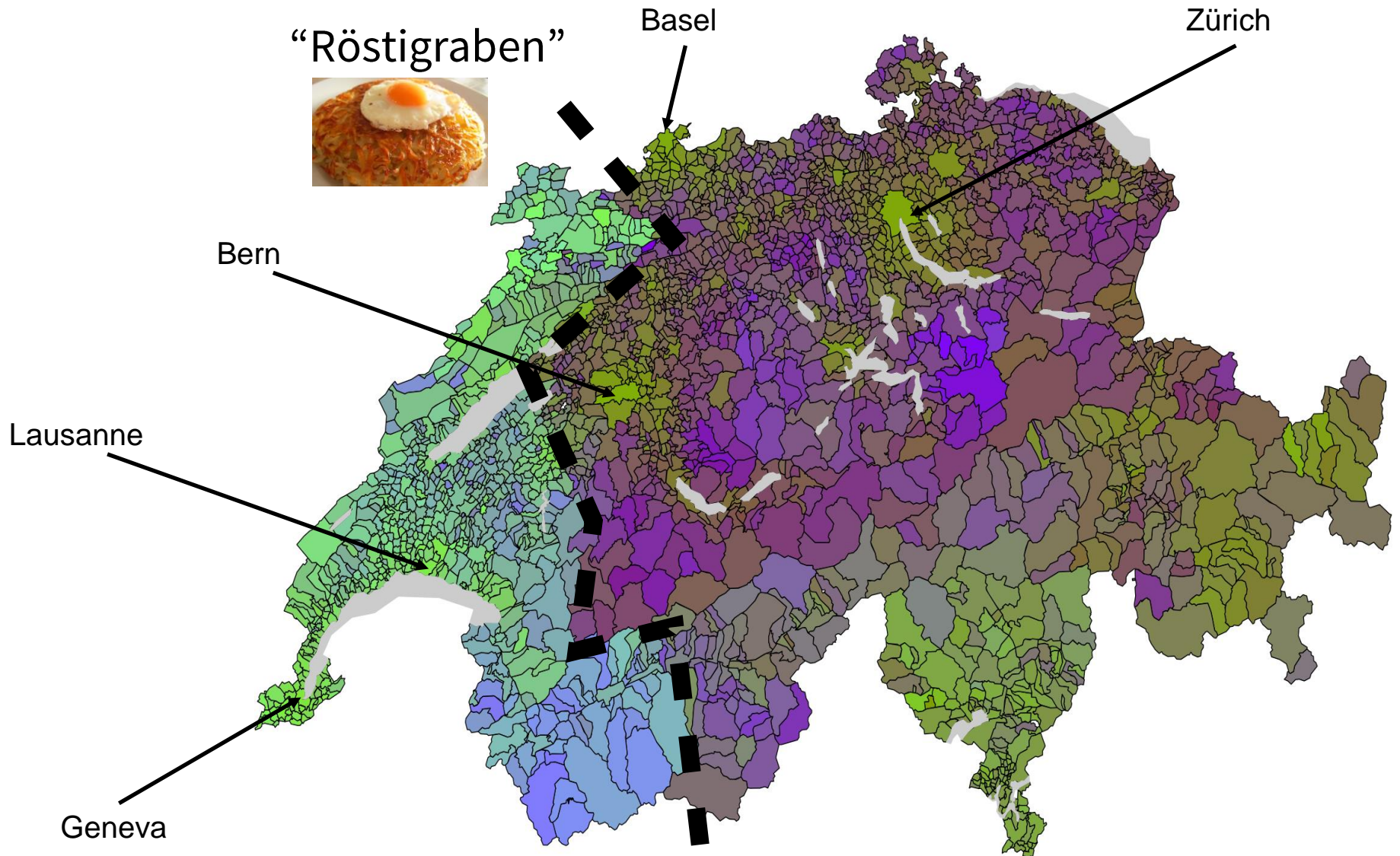
Results	Vote date	Vote name	Relative importance
	December 6, 1992	Federal decree on the European Economic Area (EEA) of October 9, 1992	100.0%
	September 22, 1985	Federal decree proposed October 5, 1984 on the popular initiative "calling for the harmonization of the beginning of the school year in all cantons" (counter-proposal)	98.13%
	June 13, 1999	Federal law on maternity insurance	97.33%
	September 26, 2004	Change proposed October 03, 2003 of the federal law on the income loss allowances for people serving in the military, civil service or civil protection (Allowances act for income loss)	92.94%
	September 26, 2004	Federal decree proposed October 03, 2003 on the acquisition of nationality by third-generation foreigners	79.88%
	April 18, 1999	Federal decree on an update of the Federal Constitution	79.73%
	September 26, 2004	Federal decree proposed October 03, 2003 on the regular naturalization and the facilitated naturalization of second-generation young foreigners	78.18%
	June 1, 2008	Popular initiative proposed November 18, 2005 « For democratic naturalizations »	69.89%
	March 3, 2013	Federal decree proposed June 15, 2012 on family policies	68.69%
	December 4, 1994	Federal act on health insurance of March 18, 1994	64.86%

Second principal component

Composition of the vertical axis:

Results	Vote date	Vote name	Relative importance
	February 20, 1994	Federal decree on the continuation of the fee on heavy vehicles traffic from June 18, 1993	100.0%
	February 20, 1994	Federal law regarding the introduction of a fee on the traffic of heavy vehicles, related either to services or to the consumption, of June 18, 1993	98.64%
	May 17, 1992	Federal law on the protection of waters (Water Protection Act) of January 24, 1991	91.54%
	September 28, 1986	Federal decree on the domestic sugar economy, change of June 21, 1985	87.96%
	February 20, 1994	Federal decree on the continuation of the national road tax	86.57%
	June 4, 1989	Popular initiative "for the protection of farms and against animal factories (Initiative for small farmers)"	86.44%
	September 27, 1998	Federal law regarding a fee on the traffic of heavy vehicle related to services (Law related to a fee on the traffic of heavy vehicles)	84.64%
	February 26, 1984	Federal decree proposed June 24, 1983 concerning the collection of a fee on the traffic of heavyweight trucks	78.36%
	June 12, 1988	Federal decree proposed March 20, 1987 concerning the amendment of the federal constitution for a coordinated transport policy	77.37%
	March 12, 1995	Federal Decree on the popular initiative 'for an environmentally sound and efficient peasant agriculture' (counter-proposal)	71.7%

Voting patterns at municipal level



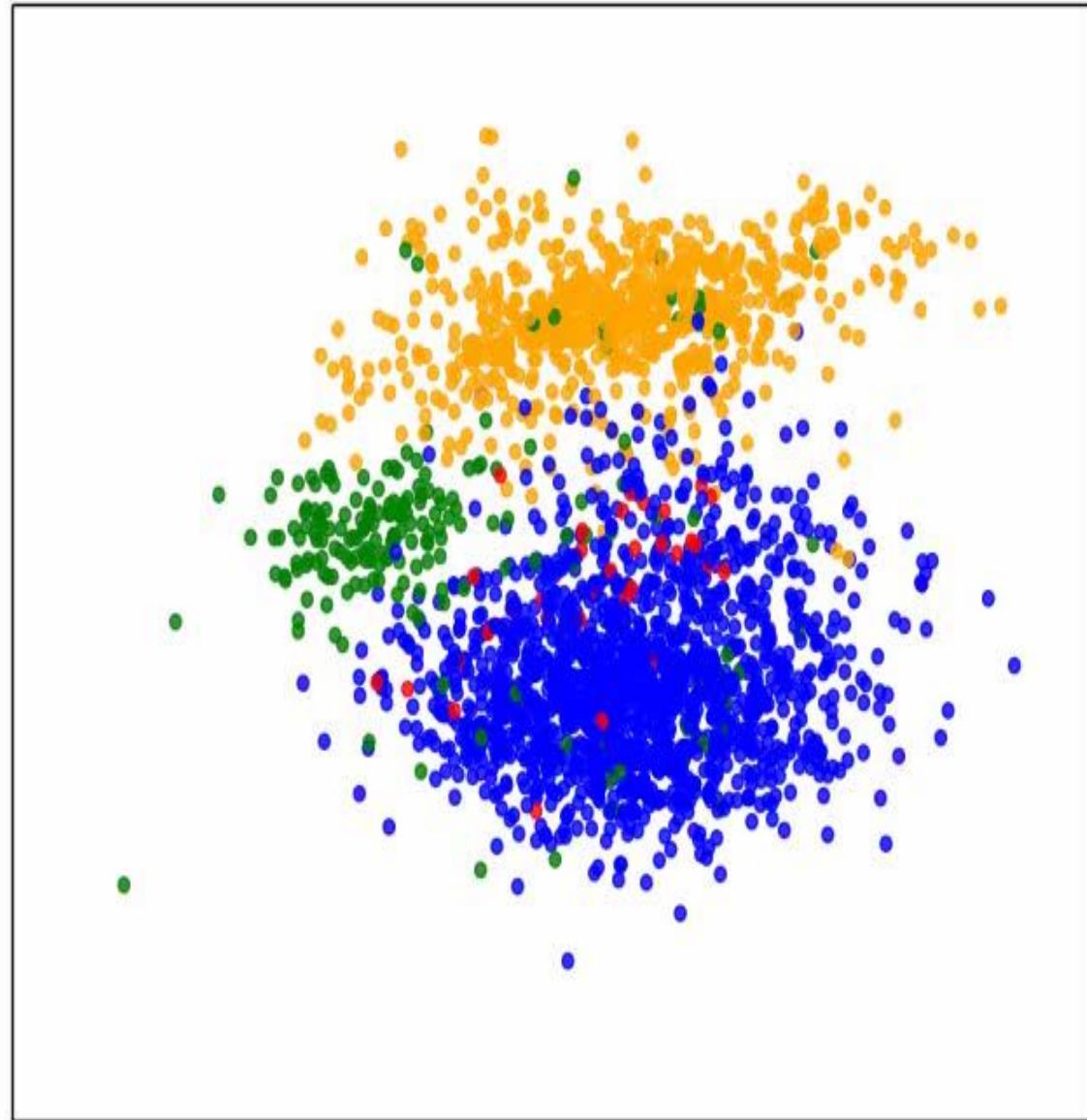
Evolution

11.2014 - 05.1992

french

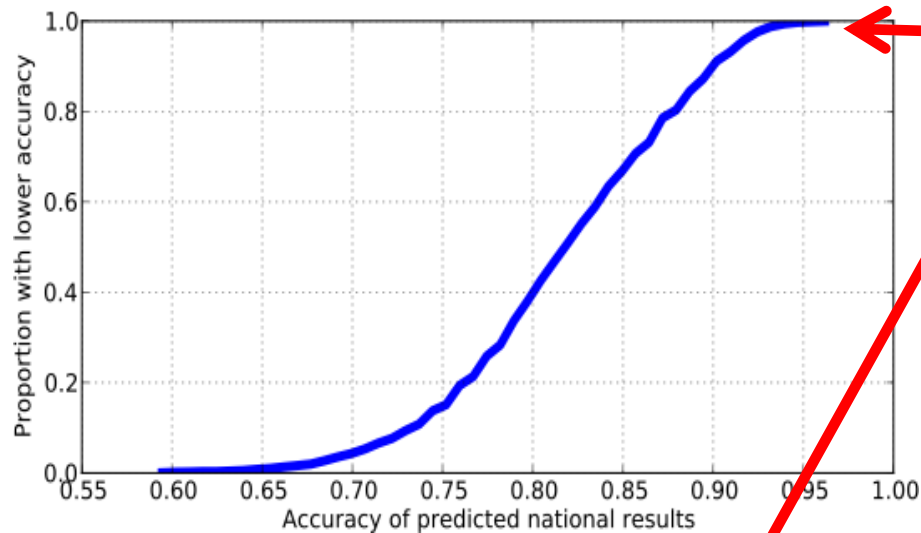
german

italian

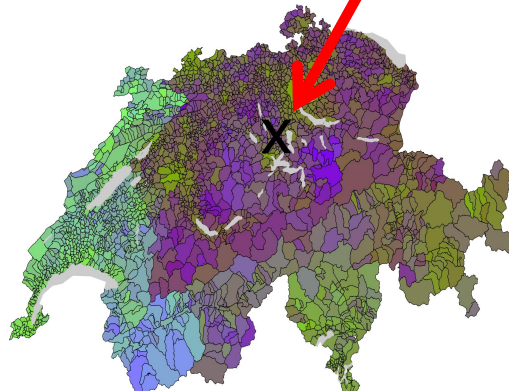


Prediction of national results

- Knowing the result of one municipality in advance (e.g., from polling/survey), can we predict the final result?
- Answer: Yes, but it depends on which municipality!



Ebikon (accuracy 95.9% on test set)



Der Traum von «Predikon»

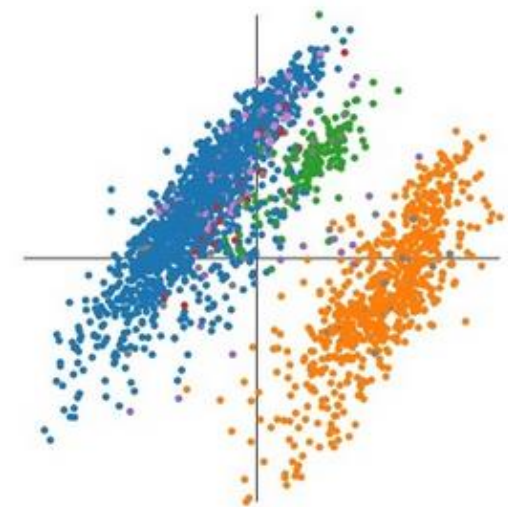
Forscher haben berechnet, dass sich der Erfolg einer Abstimmungsvorlage ziemlich genau am Resultat einer Innerschweizer Gemeinde ablesen lässt. Doch lässt sich damit künftig eine präzise Prognose erstellen?



Malerische Lage: Ebikon am Rotsee. Bild: Gemeinde Ebikon

Von Martin Wilhelm
Redaktor Inland

Die Luzerner Vorortsgemeinde Ebikon ist jene Gemeinde, aus deren Abstimmungsergebnissen sich am



Für Ihre Studie untersuchten die ETH-Forscher Vincent Etter und Julien Herzen die Abstimmungsergebnisse von 245 nationalen Abstimmungen von 1981 bis 2011. Zur «Prognose» des nationalen Ergebnisses aufgrund der Gemeindergebnisse fütterten sie ihre Computer zunächst mit den Werten von 196 Abstimmungen. Das so «trainierte» Programm liessen sie dann für die restlichen 49 Abstimmungen das nationale Ergebnis voraussagen. In 47 Fällen lag das Programm

Daniel Casser, le président FDC de la commune, perdut un triple non ce dimanche lors des votations (Dominic Büttner)

Ongoing work: vote prediction from partial results

- How to predict national outcomes from partial results?
- State of the art: average of municipalities reporting so far
- Might be biased:
 - Small municipalities might complete earlier
 - Different rules in different cantons/municipalities
 - ...
- Can we do better? Learn latent-factor model:
 - $Y_{in} = V_i^T U_n + \beta_i^T x_i + \beta_n^T x_n$
 - Y_{in} : fraction “yes” in municipality i for vote n
 - i : municipality, n : vote, ϵ_{in} : noise, l : latent dim = lengths of V_i, U_n
 - Interaction between municipality and vote through $V_i^T U_n$

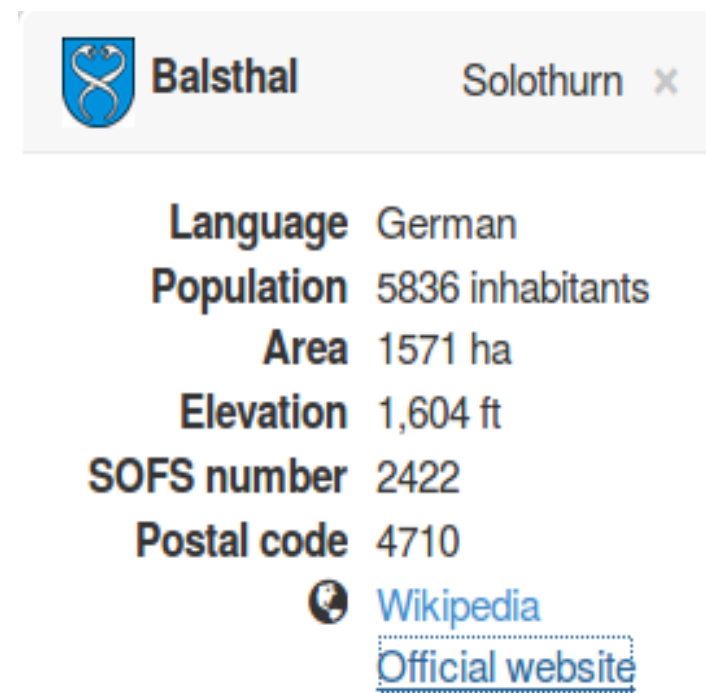
Going further: richer models with features


- Vote features β_n : environment, social, economic, Europe, ...
 - Most objective: party recommendations
- Municipality β_i : size, elevation, density, ...
 - Easy to obtain from official sources & wikipedia




	FDP	CVP	BDP	SP	Grüne
Erbschaftssteuer-Initiative	NEIN	NEIN	NEIN	JA	JA
Präimplantationsdiagnostik	JA	JA	JA	Frei	JA
Änderung Radio-Fernsehgesetz	NEIN	JA	JA	JA	JA
Stipendien-Initiative	NEIN	NEIN	NEIN	JA	JA

ABSTIMMUNG VOM 14. JUNI 2015 

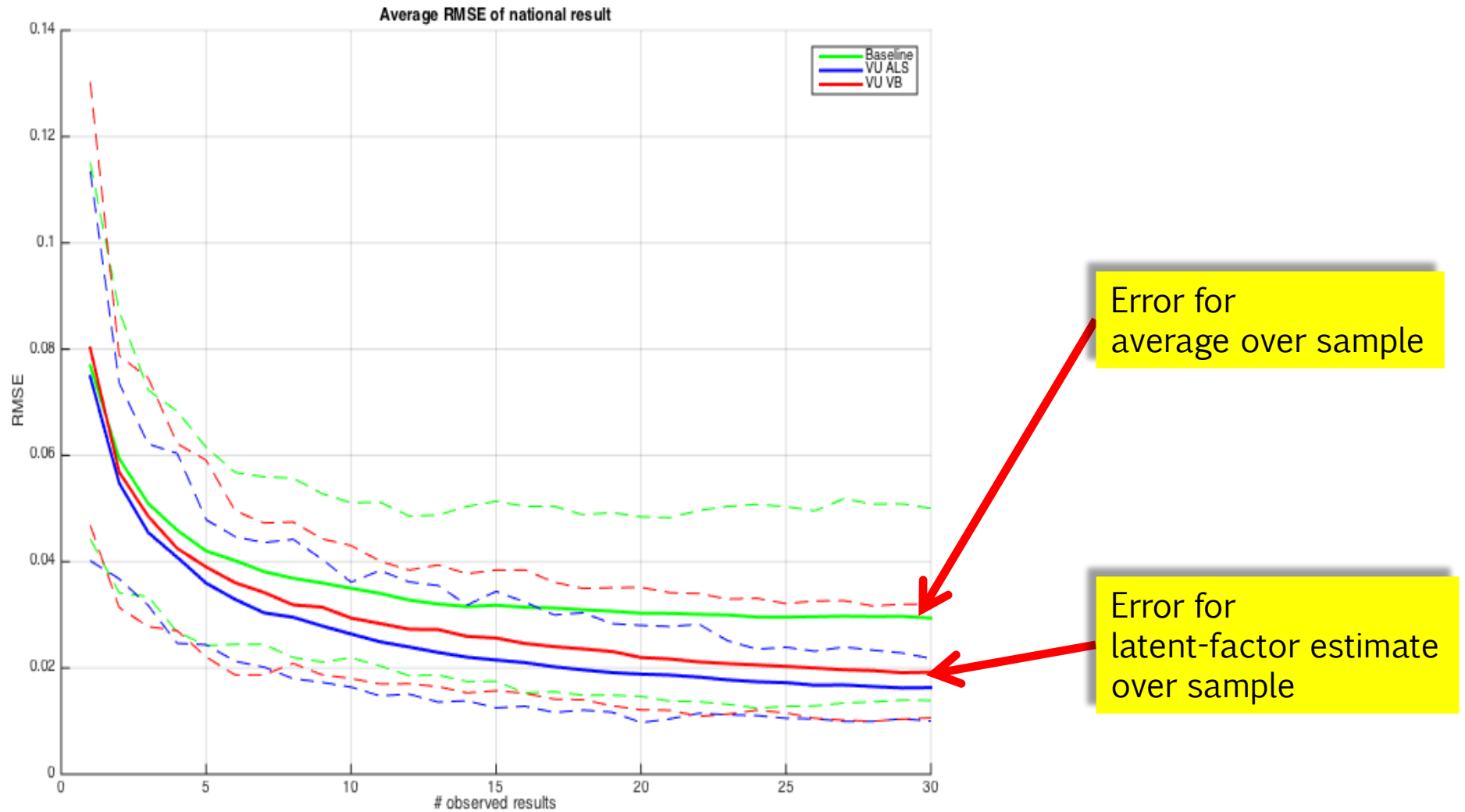


 **Balsthal** Solothurn 

Language German
Population 5836 inhabitants
Area 1571 ha
Elevation 1,604 ft
SOFS number 2422
Postal code 4710

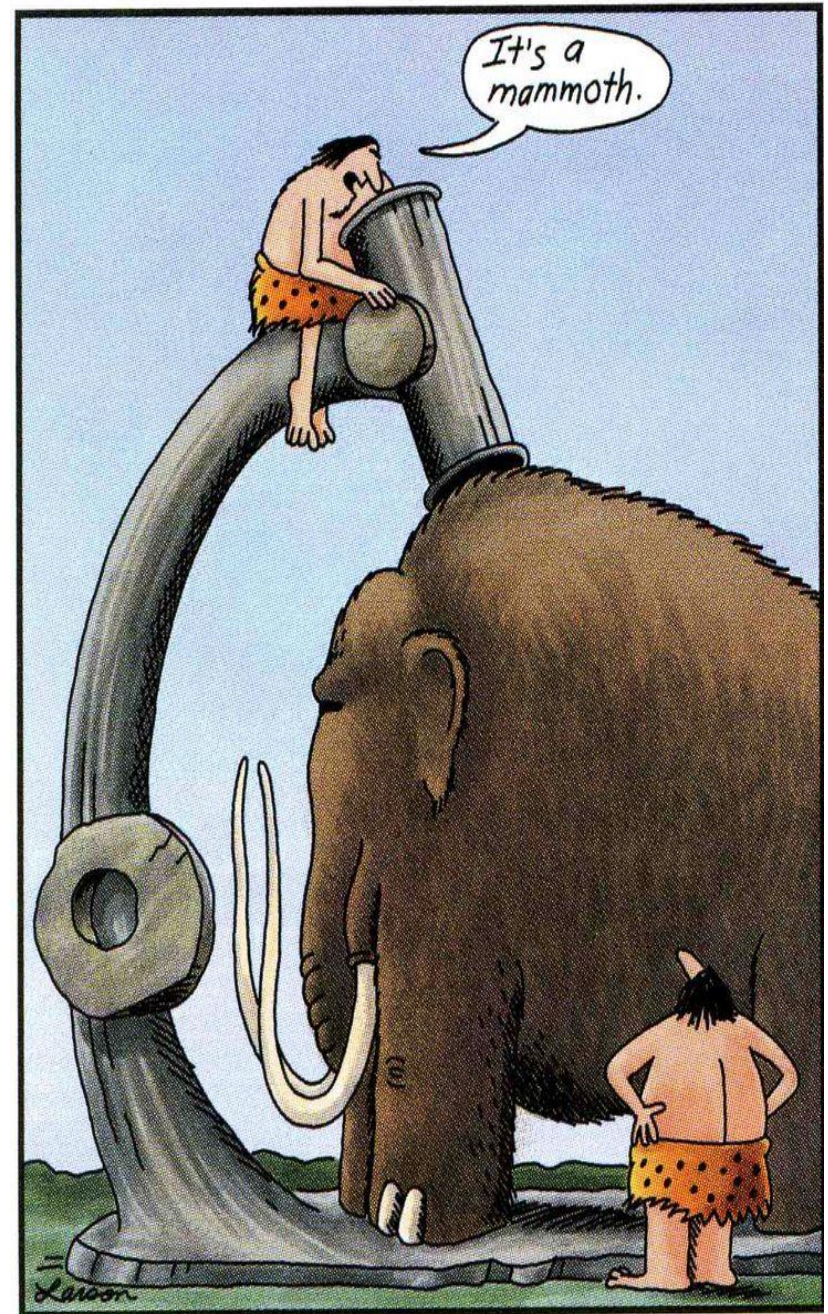
 [Wikipedia](#)
[Official website](#)

Predicting vote outcomes from a few municipalities



Summary

- Goals:
 - Scientific: improve methods and tools for data analysis
 - Societal: enrich our democracy
- Motto:
 - Mine first, ask questions later!
- Explore yourself:
 - predikon.ch, [COSN2014]



Early microscope

Thanks! Questions?

Matthias Grossglauser, EPFL

Collaborators:

Vincent Etter, Julien Herzen, Emtiyaz Khan, Victor Kristof,
Patrick Thiran

EIT ICT Labs Summer School
August 2015