

Online Appendices:

Candidate Awareness in Mixed-Member Electoral Systems: A Data-Driven Approach

Oliver Rittmann
University of Mannheim

Marie-Lou Sohnus
University of Oxford

Thomas Gschwend
University of Mannheim

A Bootstrap Procedure	1
B GLES Question Wording and Variable Codes	5

A Bootstrap Procedure

Figure A1 visualizes the bootstrap procedure we implement to account for the clustered structure of our data.

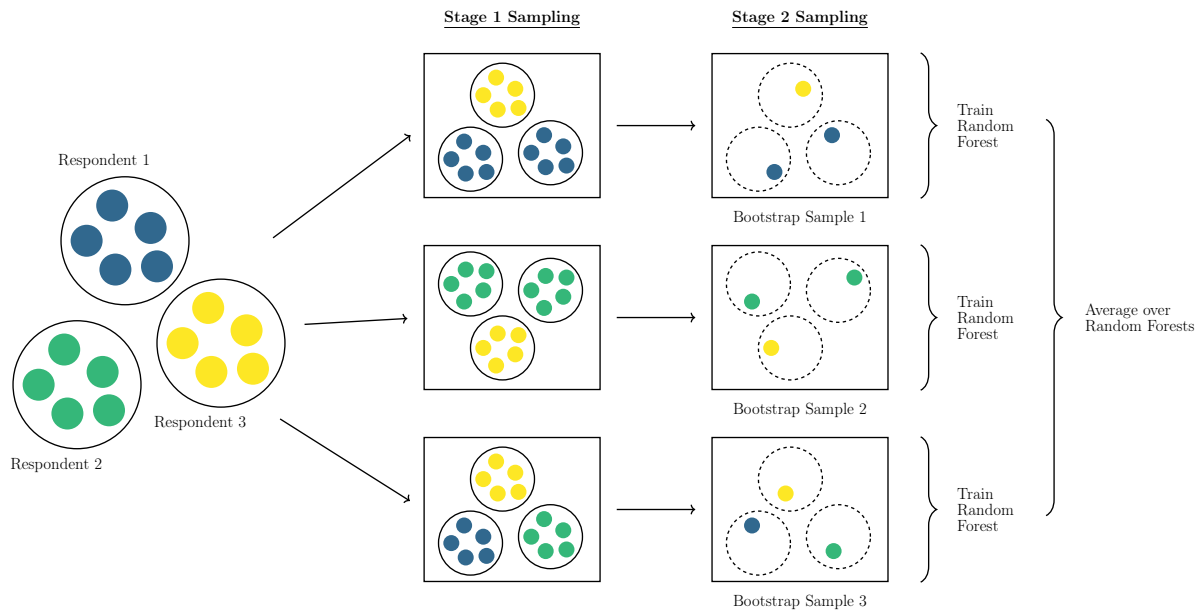


Figure A1: Illustration of the two-stage subject-level bootstrapping procedure.

Methodological Literature on Random Forests and Clustered Data

There is an active field of scholarship that develops extensions of the random forest model to account for clustered data structures. This literature identifies a series of weaknesses of standard random forest models when applied to data with a clustered structure. Here, we summarize the main contributions within this literature in previous years and explain how our approach to dealing with clustered data relates to the solutions proposed in this field of research.

Random Forest Models, such as all common machine learning models, rely on the assumption that individual observations are independent and identically distributed (Breiman, 2001, p. 6). This assumption is violated for clustered data where one subject is observed multiple times. Observations about the same subject are correlated due to subject-level characteristics. This applies twice in our case. Our data set consists of respondent-candidate dyads. This means that each survey respondent and most candidates enter the data set multiple times. Respondents enter the data set once for each relevant candidate

in their electoral district, with the outcome variable of each row indicating whether the respondent correctly recalled the candidate or not. We consider candidates of all parties who ultimately entered the Bundestag. Based on this rule, each respondent occurs five to six times in our data set, depending on the election year. District candidates enter the data set once for each survey respondent who resides in their electoral district. This number ranges between one and 37 and averages at eleven. Neither observations of the same respondent nor observations of the same candidate are independent of one another. Yet, we perceive the challenge posed by dependencies on the respondent level as more pressing, because the observations of respondents are likely to be more strongly correlated than the observations that belong to one candidate.

In light of the popularity of random forest models for prediction tasks and the ubiquity of clustered data, approaching the challenges posed by the correlated structure of clustered data became an active field of research in recent years (Hu and Szymczak, 2023). The solutions proposed in this emerging literature fall into two categories: One strand of methodological work proposes to adapt the bootstrapping algorithm built into the random forest estimation algorithm to reflect the sampling process of clustered data (Karpievitch et al., 2009; Adler, Potapov and Lausen, 2011). The second strand proposes to integrate random forest models with classical statistical models for clustered data, namely (generalized) linear mixed effects models (Hajjem, Bellavance and Larocque, 2014; Pellagatti et al., 2021; Speiser et al., 2019).

Solution 1: Subject-level Bootstrapping

Classical random forests are a collection of decision trees, where each decision tree is built upon a bootstrapped sample of the training data set. Each decision tree results in one outcome prediction. The prediction of the forest aggregates the predictions of all its decision trees (Breiman, 2001, p. 6). The bootstrap sampling of the original random forest algorithm treats observations independently. If this independence holds, then each bootstrapped sample contains, on average, around two-thirds of the observations in the training data. It is key to the performance of random forests that each decision tree does not get exposed to the information of the remaining one-third of observations that are not bootstrapped. If observations are not independent, this implies that the observations in a bootstrapped sample are correlated with observations that do not end up in the bootstrapped sample. In other words, the observations in a bootstrapped sample contain information about the observations that do not end up in the bootstrapped sample. Each decision tree is thus exposed to more information than it is supposed to be. This leads to a higher correlation between trees, which can be expected to impair the performance of the forests as a whole (Adler, Potapov and Lausen, 2011, p. 357).

To address this issue, Karpievitch et al. (2009) propose to replace the original bootstrap algorithm with a subject-level bootstrap algorithm that reflects the sampling process of clustered data. Instead of sampling observations directly, their bootstrap algorithm first samples subjects (in our case respondents), and, in the second stage, bootstraps observations (in our case knowledge about individual candidates) within the sampled subject clusters. This ensures that each bootstrapped sample, on average, leaves out one-third of the subjects. Since subjects are independent, this makes sure that the observations selected within one bootstrapped sample are not correlated with the observations by subjects who are not sampled. Adler, Potapov and Lausen (2011) extend this approach. They propose to sample only one observation per subject in the second stage of the subject-level bootstrap algorithm. For each decision tree, this breaks the clustered data structure of the learning data. Adler, Potapov and Lausen (2011) show that this approach further improves the performance of random forests in the presence of clustered data.

Solution 2: Integrating Random Forests with Mixed Models

More recent methodological work proposes to integrate random forest models with (generalized) linear mixed-effects models. Mixed effects models—also called multilevel models, random effects models, or hierarchical models—belong to standard statistical approaches to deal with the problems caused by the dependency structures in clustered data. The solution of mixed models is to partition unexplained variation into two levels: Unexplained variation on the subject level and unexplained variation within subjects (Bell and Jones, 2015). Mixed models incorporate two types of effects. Random effects explain variation on the subject level, and fixed effects explain variation within subjects. All integrations of random forests and mixed models have in common that they replace the fixed effects component of mixed models with random forests, and keep the random effects component to model the dependencies in the data caused by its clustered structure (Hu and Szymczak, 2023).

Different models have been proposed in this strand of research. Hajjem, Bellavance and Larocque (2014) first introduced Mixed Effects Random Forests (MERF) for continuous outcomes. Pellagatti et al. (2021) extends this work by proposing a generalized version of this model applicable to all outcome variables of the exponential family, including categorical outcomes. Speiser et al. (2019) add a Bayesian solution that is applicable to binary outcomes. While Generalized Mixed Effects Random Forests are generally appealing for the purpose of our study, there is currently no software or code available that allows researchers to implement proposed models.¹

¹A software implementation for Mixed Effects Random Forests as proposed by Hajjem, Bellavance and Larocque (2014) is implemented in the R-package `LongituRF` (Capitaine, 2020), but this is only

The approach of our study

Our own approach builds on the solutions that adjust the bootstrap algorithm of random forests. By and large, we adapt the bootstrap algorithm proposed by Adler, Potapov and Lausen (2011). This means that we implement a two-stage subject-level bootstrap procedure where we first bootstrap subject in the first stage, and sample one observation for each sampled subject in the second stage (see figure A1). The difference between our approach and the approach by Adler, Potapov and Lausen (2011) is that instead of using one bootstrap sample to fit one decision tree, we fit a random forest on each bootstrapped sample. This is more computationally demanding but allows us to fine-tune the random forests.² The key idea behind the approach by Adler, Potapov and Lausen (2011) is not affected by that: Due to the two-stage subject-level bootstrapping, each random forest gets trained on a reduced data set that is free of within-subject dependencies.

Limitation of our approach

Our approach accounts for dependencies caused by the respondent clusters in our data. However, it is limited by the fact that it does not account for dependencies caused by candidate clusters in our data. Based on our reading of the literature, there is no solution that adjusts the bootstrap procedure of the random forest with respect to cross-cutting clusters. At the same time, we are convinced that candidate clusters pose a significantly smaller problem than respondent clusters. This is because observations that belong to the same respondents are likely to be more strongly correlated than observations that belong to the same candidate. Nevertheless, observations that ask different respondents about the same candidate are potentially correlated, and our approach does not account for such dependencies.

applicable to continuous outcomes. Our hope is that software implementations of the generalizations of this mode will follow in the future.

²To the best of our knowledge, there is no software implementation of the algorithm proposed by Adler, Potapov and Lausen (2011) that allows automatic fine-tuning.

B GLES Question Wording and Variable Codes

Here, we list all items from the German Longitudinal Election Studies in 2009, 2013, and 2017 that were used in the analysis. For variables with multiple answer categories that were recoded to a binary variable, underlined numbers (e.g. (02)) indicate the category coded as 1. Categories with italic numbers are coded as missing values (e.g. *(98)*).

Candidate Knowledge

2009: q81m1 - q81m5:

Kennen Sie den Namen von einem oder vielleicht sogar mehreren der hiesigen Wahlkreiskandidaten oder -kandidatinnen und können Sie mir sagen, für welche Partei diese bei der Bundestagswahl am 27. September 2009 antreten? Bitte nennen Sie mir den Namen und die Partei der Kandidatinnen bzw. Kandidaten.

English version: *Do you know one or several of the local district candidates and can you tell me for which party they are running in the federal election on 27th September 2009? Please tell me the names and the corresponding parties of the candidates.*

2013: q82a - q82e:

Kennen Sie den Namen von einem oder vielleicht sogar mehreren der hiesigen Wahlkreiskandidaten und können Sie mir sagen, für welche Partei diese bei der Bundestagswahl am 22. September 2013 antreten? Bitte nennen Sie mir den Namen und die Partei der Kandidaten.

English version: *Do you know one or several of the local district candidates and can you tell me for which party they are running in the federal election on 22nd September 2013? Please tell me the names and the corresponding parties of the candidates*

2017: q77a1 - q77f1:

Kennen Sie den Namen von einem oder vielleicht sogar mehreren der hiesigen Wahlkreiskandidaten und können Sie mir sagen, für welche Partei diese bei der Bundestagswahl am 24. September 2017 antreten? Bitte nennen Sie mir den Namen und die Partei der Kandidaten.

English version: *Do you know one or several of the local district candidates and can you tell me for which party they are running in the federal election on 24th September 2017? Please tell me the names and the corresponding parties of the candidates.*

⇒ Coded 1 if respondent named the candidate and the party, 0 otherwise.

Party Identification

2009: q139a:

Und jetzt noch einmal kurz zu den politischen Parteien. In Deutschland neigen viele Leute längere Zeit einer bestimmten politischen Partei zu, obwohl sie auch ab und zu eine andere Partei wählen. Wie ist das bei Ihnen: Neigen Sie - ganz allgemein gesprochen - einer bestimmten Partei zu? Und wenn ja, welcher?

English version: *Now, let's look at the political parties. In Germany, many people lean toward a particular party for a long time, although occasionally, they vote for another party. How about you, do you lean toward a particular political party? If yes, which party is that?*

2013: q119a:

Und nun noch einmal kurz zu den politischen Parteien. In Deutschland neigen viele Leute längere Zeit einer bestimmten politischen Partei zu, obwohl sie auch ab und zu eine andere Partei wählen. Wie ist das bei Ihnen: Neigen Sie - ganz allgemein gesprochen - einer bestimmten Partei zu? Und wenn ja, welcher?

English version: *Now, let's look at the political parties. In Germany, many people lean toward a particular political party for a long time, although occasionally, they vote for another party. How about you, do you lean toward a particular political party? If yes, which party is that?*

2017: q99a:

Und nun noch einmal kurz zu den politischen Parteien. In Deutschland neigen viele Leute längere Zeit einer bestimmten politischen Partei zu, obwohl sie auch ab und zu eine andere Partei wählen. Wie ist das bei Ihnen: Neigen Sie - ganz allgemein gesprochen - einer bestimmten Partei zu? Und wenn ja, welcher?

English version: *Now, let's look at the political parties. In Germany, many people lean towards a particular party for a long time, although they may occasionally vote for a different party. How about you, do you in general lean towards a particular party? If so, which one?*

⇒ Coded 1 if respondent indicated to identify with the party of the candidate, 0 otherwise.

Voted for Candidate

2009: q11aa:

Bei der Bundestagswahl können Sie ja zwei Stimmen vergeben. Die Erststimme für einen Kandidaten aus Ihrem Wahlkreis und die Zweitstimme für eine Partei. Hier ist ein Musterstimmzettel, ähnlich wie Sie ihn bei der Bundestagswahl erhalten.

Was werden Sie auf Ihrem Stimmzettel ankreuzen? Bitte nennen Sie mir jeweils die Kennziffer für Ihre Erst- und Zweitstimme.

(A) jetzt bitte für die Erststimme

English version: *You have two votes in the federal election. The first vote is for a candidate in your local district, the second vote is for a party. Here is a ballot paper similar to the one you receive for federal elections. How will you mark your ballot? Please just tell me the identification code for your first and second vote.*

(A) now for your first vote (aka candidate vote) please.

2013: q11aa:

Bei der Bundestagswahl können Sie ja zwei Stimmen vergeben. Die Erststimme für einen Kandidaten aus Ihrem Wahlkreis und die Zweitstimme für eine Partei. Hier ist ein Musterstimmzettel, ähnlich wie Sie ihn bei der Bundestagswahl erhalten. Was werden Sie auf Ihrem Stimmzettel ankreuzen? Bitte nennen Sie mir jeweils die Kennziffer für Ihre Erst- und Zweitstimme.

(A) Erststimme

English version: *You have two votes in the federal election. The first vote is for a candidate in your local district, the second vote is for a party. Here is a ballot paper similar to the one you receive for federal elections. How will you mark your ballot? Please just tell me the identification code for your first and second vote*

(A) First vote (aka candidate vote)

2017: q11a:

Bei der Bundestagswahl können Sie ja zwei Stimmen vergeben. Die Erststimme für einen Kandidaten aus Ihrem Wahlkreis und die Zweitstimme für eine Partei. Hier ist ein Musterstimmzettel, ähnlich wie Sie ihn bei der Bundestagswahl erhalten. Was werden Sie auf Ihrem Stimmzettel ankreuzen? Bitte nennen Sie mir jeweils die Kennziffer für Ihre Erst- und Zweitstimme.

(A) jetzt bitte für die Erststimme

English version: *You have two votes in the federal election. The first vote is for a candidate in your local district, the second vote is for a party. Here is a ballot paper similar to the one you receive for federal elections. How will you mark your ballot? Please just tell me the identification code for your first and second vote*

(A) First vote (aka candidate vote)

⇒ Coded 1 if respondent indicated to vote for the candidate, 0 otherwise. If respondent already voted via mail, the mail vote was taken instead of intended vote choice.

Party Contact

2009: Und nun zum Bundestagswahlkampf 2009 und den Aktivitäten der Parteien und Kandidaten. Lassen Sie bitte alle anderen Wahlen, die in diesem Jahr stattgefunden haben, außer Acht. Haben Sie...

q45m Wahlveranstaltungen oder Kundgebungen von Parteien besucht? ...von welchen Parteien waren die?

q46m Emails oder SMS von Parteien erhalten? ...von welchen Parteien waren die?

q48m Informationsmaterial von Parteien wie z.B. Flugblätter, Handzettel, Broschüren oder Postwurfsendungen gelesen? ...von welchen Parteien waren die?

q50m Wahlanzeigen von Parteien in Zeitungen oder Zeitschriften gelesen? ...von welchen Parteien waren die?

q51m Wahlplakate von Parteien oder Kandidaten gesehen? ...von welchen Parteien waren die?

q53m einen oder mehrere Wahlkampfstände hier im Ort besucht? ...von welchen Parteien waren die?

q54m Werbesendungen von Parteien im Fernsehen gesehen? ...von welchen Parteien waren die?

English version: *Now let's have a look at the 2009 federal election campaign and the activities of the parties and candidates. Please leave aside all the other elections that have taken place this year. Did you...*

q45m *go to any of the election meetings or rallies organised by the political parties? Which parties organised the election meetings or rallies you attended?*

q46m *receive any e-mails or text messages from political parties? From which parties did you receive e-mails or text messages?*

q48m *read any information material produced by political parties, such as campaign flyers, handbills, leaflets or posted material? From which parties did you read information material, such as campaign flyers, leaflets and similar material?*

q50m *read any campaign adverts placed by political parties in newspapers or magazines? From which parties did you read campaign adverts in newspapers or magazines?*

q51m *see any campaign posters for political parties or candidates? Which party election posters did you see?*

q53m *go to one or several election campaign booths here locally? Which party election booths did you visit?*

q54m *see any party political broadcasts on television? From which political parties were the political broadcasts you have seen?*

2013: Und nun zum Bundestagswahlkampf 2013 und den Aktivitäten der Parteien und Kandidaten. Lassen Sie bitte alle anderen Wahlen, die in diesem Jahr stattgefunden haben, außer Acht. Haben Sie...

q48a-i Wahlveranstaltungen oder Kundgebungen von Parteien besucht? ...von welchen Parteien waren die?

q49a-i E-Mails oder SMS von Parteien erhalten? ...von welchen Parteien waren die?

q51a-i Informationsmaterial von Parteien wie z.B. Flugblätter, Handzettel, Broschüren oder Postwurfsendungen gelesen? ...von welchen Parteien war das?

q53a-i Wahlanzeigen von Parteien in Zeitungen oder Zeitschriften gelesen? ...von welchen Parteien waren die?

q54a-i Wahlplakate von Parteien oder Kandidaten gesehen? ...von welchen Parteien waren die?

q56a-i einen oder mehrere Wahlkampfstände hier im Ort besucht? ...von welchen Parteien waren die?

q57a-i Werbesendungen von Parteien im Fernsehen gesehen? ...von welchen Parteien waren die?

q224a-i Wurden Sie direkt von einer Person angesprochen, z.B. auf der Straße oder an Ihrer Wohnungs- bzw. Haustür? ...von welcher Partei war diese?

q225a-i Wurden Sie telefonisch von Parteien kontaktiert? Von welchen Parteien wurden Sie telefonisch kontaktiert?

q226a-i Wurden Sie über ein soziales Netzwerk im Internet, wie z.B. Facebook, von Parteien kontaktiert? Von welchen Parteien wurden Sie über ein soziales Netzwerk im Internet wie Facebook kontaktiert?

English version: *Now let's have a look at the 2013 federal election campaign and the activities of the parties and candidates. Please leave aside all the other elections that have taken place this year. Did you...*

q48a-i *go to any of the election meetings or rallies organised by the political parties? Which parties organised the election meetings or rallies you attended?*

q49a-i *receive any e-mails or text messages from political parties? From which parties did you receive e-mails or text messages?*

- q51a-i *read any information material produced by political parties, such as campaign flyers, handbills, leaflets or posted material? From which parties did you read information material, such as campaign flyers, leaflets and similar material?*
- q53a-i *read any campaign adverts placed by political parties in newspapers or magazines? From which parties did you read campaign adverts in newspapers or magazines?*
- q54a-i *see any campaign posters for political parties or candidates? Which party election posters did you see?*
- q56a-i *go to one or several election campaign booths here locally? Which party election booths did you visit?*
- q57a-i *see any party political broadcasts on television? From which political parties were the political broadcasts you have seen?*
- q224a-i *Have you had direct contact with a person at an election campaign booth, at an election meeting, on your own doorstep or elsewhere? Which party did he/she belong to?*
- q225a-i *Have you been contacted by phone? Which parties contacted you by telephone?*
- q226a-i *Have you been contacted via social networks on the Internet, such as Facebook? Which parties contacted you via social networks on the Internet (such as Facebook etc.)?*

2017: Und nun zum Bundestagswahlkampf 2017 und den Aktivitäten der Parteien und Kandidaten. Lassen Sie bitte alle anderen Wahlen, die in diesem Jahr stattgefunden haben, außer Acht. Haben Sie...

- q40a-i *Wahlveranstaltungen oder Kundgebungen von Parteien besucht? ...von welchen Parteien waren die?*
- q41a-i *E-Mails oder SMS von Parteien erhalten? ...von welchen Parteien waren die?*
- q43a-1 *Informationsmaterial von Parteien wie z.B. Flugblätter, Handzettel, Broschüren oder Postwurfsendungen gelesen? ...von welchen Parteien war das?*
- q45a-i *Wahlanzeigen von Parteien in Zeitungen oder Zeitschriften gelesen? ...von welchen Parteien waren die?*
- q46a-i *Wahlplakate von Parteien oder Kandidaten gesehen? ...von welchen Parteien waren die?*
- q48a-i *einen oder mehrere Wahlkampfstände hier im Ort besucht? ...von welchen Parteien waren die?*
- q49a-i *Werbesendungen von Parteien im Fernsehen gesehen? ...von welchen Parteien waren die?*

q50a-i Wurden Sie direkt von einer Person angesprochen, z.B. auf der Straße oder an Ihrer Wohnungs- bzw. Haustür? ...von welcher Partei war diese?

q51a-i Wurden Sie über ein soziales Netzwerk im Internet, wie z.B. Facebook, von Parteien kontaktiert? Von welchen Parteien wurden Sie über ein soziales Netzwerk im Internet wie Facebook kontaktiert?

English version: *Now let's have a look at the 2017 federal election campaign and the activities of the parties and candidates. Please leave aside all the other elections that have taken place this year. Did you...*

q40a-i *go to any of the election meetings or rallies organised by the political parties? Which parties organised the election meetings or rallies you attended?*

q41a-i *receive any e-mails or text messages from political parties? From which parties did you receive e-mails or text messages?*

q43a-1 *read any information material produced by political parties, such as campaign flyers, handbills, leaflets or posted material? From which parties did you read information material, such as campaign flyers, leaflets and similar material?*

q45a-i *read any campaign adverts placed by political parties in newspapers or magazines? From which parties did you read campaign adverts in newspapers or magazines?*

q46a-i *see any campaign posters for political parties or candidates? Which party election posters did you see?*

q48a-i *go to one or several election campaign booths here locally? Which party campaign booths did you visit?*

q49a-i *see any party political broadcasts on television? From which political parties were the political broadcasts you have seen?*

q50a-i *Have you had direct contact with a person on the street, on your own doorstep or elsewhere? Which party did he/she belong to?*

q51a-i *Have you been contacted via social networks on the Internet, such as Facebook? Which parties contacted you via social networks on the internet (such as Facebook etc.)?*

⇒ Coded 1 if respondent mentioned a party for at least one item, and zero otherwise.

Candidate Contact

2009: Und nun zum Bundestagswahlkampf 2009 und den Aktivitäten der Parteien und Kandidaten. Lassen Sie bitte alle anderen Wahlen, die in diesem Jahr stattgefunden haben, außer Acht. Haben Sie...

- q47a-d Emails oder SMS von Parteien erhalten? Kam die Email oder SMS von dem Wahlkreiskandidaten oder der Wahlkreiskandidatin der [Partei] persönlich?
- q49a-d Informationsmaterial von Parteien wie z.B. Flugblätter, Handzettel, Broschüren oder Postwurfsendungen gelesen? War das Informationsmaterial, wie z.B. Flugblätter oder Broschüren, von dem Wahlkreiskandidaten oder der Wahlkreiskandidatin der Partei [Partei] persönlich?
- q52a-d Wahlplakate von Parteien oder Kandidaten gesehen? Waren das Plakate, die Werbung für die Person des Wahlkreiskandidaten oder der Wahlkreiskandidatin der Partei [Partei] machen?
- q56m Haben Sie direkten persönlichen Kontakt zu einer Wahlkreiskandidatin oder einem Wahlkreiskandidaten, z.B. an einem Wahlkampfstand, im Rahmen einer Wahlveranstaltung oder durch Hausbesuche gehabt? Von welchen Parteien war das?

English version: *Now let's have a look at the 2009 federal election campaign and the activities of the parties and candidates. Please leave aside all the other elections that have taken place this year. Did you...*

- q47a-d *receive any e-mails or text messages from political parties? Was the e-mail or text messages sent by the district candidate for the [party] personally?*
- q49a-d *read any information material produced by political parties, such as campaign flyers, handbills, leaflets or posted material? Was the information material, such as campaign flyers or leaflets sent by the party's district candidate [party] personally?*
- q52a-d *see any campaign posters for political parties or candidates? Were the posters advertising the actual district candidate for the [party]?*
- q56m *Have you had direct contact with a local candidate at an election campaign booth, at an election meeting, on your own doorstep or elsewhere? Which party did he/she belong to?*

2013: Und nun zum Bundestagswahlkampf 2013 und den Aktivitäten der Parteien und Kandidaten. Lassen Sie bitte alle anderen Wahlen, die in diesem Jahr stattgefunden haben, außer Acht. Haben Sie...

- q50a-h E-Mails oder SMS von Parteien erhalten? Kam die E-Mail oder SMS von dem Wahlkreiskandidaten der [Partei] persönlich?
- q52a-h Informationsmaterial von Parteien wie z.B. Flugblätter, Handzettel, Broschüren oder Postwurfsendungen gelesen? War das Material von dem Wahlkreiskandidaten der [Partei] persönlich?

q54a-i Wahlplakate von Parteien oder Kandidaten gesehen? Waren das Plakate, die Werbung für die Person des Wahlkreiskandidaten der [Partei] gemacht haben?

q59a-i Haben Sie direkten persönlichen Kontakt zu Wahlkreiskandidaten im Rahmen einer Wahlveranstaltung oder durch Hausbesuche gehabt? Von welchen Parteien war das?

English version: *Now let's have a look at the 2013 federal election campaign and the activities of the parties and candidates. Please leave aside all the other elections that have taken place this year. Did you...*

q50a-h *Did you receive any e-mails or text messages from political parties? Was the e-mail or text message sent by the district candidate of the [party] personally?*

q52a-h *Did you read any information material produced by political parties, such as campaign flyers, handbills, leaflets or posted material? Was the material provided by the district candidate of the [party] personally?*

q54a-i *Did you see any campaign posters for political parties or candidates? Did these posters promote the district candidate of the [party]?*

q59a-i *Did you have direct contact with district candidates, e.g. at an election meeting or a house call? Which party did he/she belong to?*

2017: Und nun zum Bundestagswahlkampf 2017 und den Aktivitäten der Parteien und Kandidaten. Lassen Sie bitte alle anderen Wahlen, die in diesem Jahr stattgefunden haben, außer Acht. Haben Sie...

q42a-i E-Mails oder SMS von Parteien erhalten? Kam die E-Mail oder SMS von dem Wahlkreiskandidaten der [Partei] persönlich?

q44a-1 Informationsmaterial von Parteien wie z.B. Flugblätter, Handzettel, Broschüren oder Postwurfsendungen gelesen? War das Material von dem Wahlkreiskandidaten der [Partei] persönlich?

q47a-i Wahlplakate von Parteien oder Kandidaten gesehen? Waren das Plakate, die Werbung für die Person des Wahlkreiskandidaten der [Partei] gemacht haben?

English version: *Now let's have a look at the 2017 federal election campaign and the activities of the parties and candidates. Please leave aside all the other elections that have taken place this year.. Did you...*

q42a-i *receive any e-mails or text messages from political parties? Was the e-mail or text message sent by the district candidate of the [party] personally?*

q44a-1 *read any information material produced by political parties, such as campaign flyers, handbills, leaflets or posted material? Was the material provided by the district candidate of the [party] personally?*

q47a-i *see any campaign posters for political parties or candidates? Did these posters promote the district candidate of the [party]?*

⇒ Coded 1 if respondent mentioned a party for at least one item, and zero otherwise.

Political Knowledge

2009: q6:

Bei der Bundestagswahl haben Sie ja zwei Stimmen, eine Erststimme und eine Zweitstimme. Wie ist das eigentlich, welche der beiden Stimmen ist ausschlaggebend für die Sitzverteilung im Bundestag?

(01) die Erststimme

(02) die Zweitstimme

(03) beide sind gleich wichtig

(98) weiß nicht

(99) keine Angabe

English version: *In the federal elections you have two votes, the first vote and the second vote. What do you think: Which vote decides how many seats each party will have in parliament?*

(01) the first vote (aka candidate vote)

(02) the second vote (aka party-list vote)

(03) both are equally important

(98) don't know

(99) No answer

2013: q7:

Bei der Bundestagswahl haben Sie ja zwei Stimmen, eine Erststimme und eine Zweitstimme. Wie ist das eigentlich, welche der beiden Stimmen ist ausschlaggebend für die Sitzverteilung im Bundestag?

(01) die Erststimme

(02) die Zweitstimme

(03) beide sind gleich wichtig

- (98) weiß nicht
- (99) keine Angabe

English version: *In the federal elections you have two votes, the first vote and the second vote. What do you think: Which vote decides how many seats each party will have in parliament?*

- (01) the first vote (aka candidate vote)
- (02) the second vote (aka party-list vote)
- (03) both are equally important
- (-98) don't know
- (-99) No answer

2017: q7:

Bei der Bundestagswahl haben Sie ja zwei Stimmen, eine Erststimme und eine Zweitstimme. Wie ist das eigentlich, welche der beiden Stimmen ist ausschlaggebend für die Sitzverteilung im Bundestag?

- (01) die Erststimme
- (02) die Zweitstimme
- (03) beide sind gleich wichtig
- (98) weiß nicht
- (99) keine Angabe

English version: *In the federal elections you have two votes, the first vote and the second vote. What do you think: Which vote decides how many seats each party will have in parliament?*

- (01) the first vote (aka candidate vote)
- (02) the second vote (aka party-list vote)
- (03) both are equally important
- (-98) don't know
- (-99) No answer

Local Representation Important

2009: q84b:

Es gibt unterschiedliche Auffassungen darüber, wen ein Abgeordneter repräsentieren soll. Wie wichtig ist Ihnen das Folgende. Bitte sagen Sie mir den zutreffenden Wert auf dieser Skala. Der Abgeordnete sollte alle Bürger im Wahlkreis repräsentieren.

- (01) überhaupt nicht wichtig
- (02)
- (03)
- (04)
- (05) sehr wichtig
- (98) weiß nicht
- (99) keine Angabe

English version: *There are different views about who a delegate should represent. Please tell me, using this scale how important the following is to you: the delegate should represent all citizens living in the district.*

- (01) not at all important
- (02)
- (03)
- (04)
- (05) very important
- (98) don't know
- (99) no answer

⇒ Recoded to match the scale in 2013 and 2017.

2013: q95b:

Es gibt unterschiedliche Auffassungen darüber, wen ein Abgeordneter repräsentieren soll. Wie wichtig ist Ihnen das Folgende. Bitte sagen Sie mir den zutreffenden Wert auf dieser Skala. Der Abgeordnete sollte die Wähler des Abgeordneten im Wahlkreis repräsentieren.

- (1) sehr wichtig
- (2) wichtig
- (3) mittelmäßig
- (4) nicht so wichtig
- (5) überhaupt nicht wichtig
- (-98) weiß nicht
- (-99) keine Angabe

English version: *There are different views about who a delegate should represent. Please tell me, using this scale how important the following is to you: the delegate should represent his/her voters in the local district.*

- (1) very important
- (2) somewhat important
- (3) in between
- (4) not very important
- (5) not important at all
- (-98) don't know
- (-99) no answer

2017: q90b:

Es gibt unterschiedliche Auffassungen darüber, wen ein Abgeordneter repräsentieren soll. Wie wichtig ist Ihnen das Folgende. Bitte sagen Sie mir den zutreffenden Wert auf dieser Skala. Der Abgeordnete sollte die Wähler des Abgeordneten im Wahlkreis repräsentieren.

- (1) überhaupt nicht wichtig
- (2)
- (3)
- (4)
- (5) sehr wichtig
- (98) weiß nicht
- (99) keine Angabe

English version: *There are different views about who a delegate should represent. Please tell me, using this scale how important the following is to you: the delegate should represent his/her voters in the local district.*

- (1) very important
- (2) somewhat important
- (3) in between
- (4) not very important
- (5) not important at all
- (-98) don't know
- (-99) no answer

Political Interest

2009: q2:

Einmal ganz allgemein gesprochen: Wie stark interessieren Sie sich für Politik: sehr stark, ziemlich stark, mittelmäßig, weniger stark oder überhaupt nicht?

- (01) sehr stark
- (02) ziemlich stark
- (03) mittelmäßig
- (04) weniger stark
- (05) überhaupt nicht
- (98) weiß nicht
- (99) keine Angabe

English version: *Quite generally, how interested are you in politics: very interested, somewhat interested, in between, not very interested, or not at all interested?*

- (01) very interested
- (02) somewhat interested
- (03) in between
- (04) not very interested
- (05) not at all interested
- (98) don't know
- (99) no answer

2013: q3:

Einmal ganz allgemein gesprochen: Wie stark interessieren Sie sich für Politik: sehr stark, ziemlich stark, mittelmäßig, weniger stark oder überhaupt nicht?

- (01) sehr stark
- (02) ziemlich stark
- (03) mittelmäßig
- (04) weniger stark
- (05) überhaupt nicht
- (98) weiß nicht
- (99) keine Angabe

English version: *Quite generally, how interested are you in politics: very interested, somewhat interested, in between, not very interested, or, not at all interested?*

- (01) very interested
- (02) somewhat interested
- (03) in between
- (04) not very interested

- (05) not at all interested
- (-98) don't know
- (-99) no answer

2017: q3:

Einmal ganz allgemein gesprochen: Wie stark interessieren Sie sich für Politik: sehr stark, ziemlich stark, mittelmäßig, weniger stark oder überhaupt nicht?

- (01) sehr stark
- (02) ziemlich stark
- (03) mittelmäßig
- (04) weniger stark
- (05) überhaupt nicht
- (98) weiß nicht
- (99) keine Angabe

English version: *Quite generally, how interested are you in politics: very interested, somewhat interested, in between, not very interested, or, not at all interested?*

- (01) very interested
- (02) somewhat interested
- (03) in between
- (04) not very interested
- (05) not at all interested
- (-98) don't know
- (-99) no answer

Age

2009: q1a

Sagen Sie mir bitte, wie alt Sie sind.

English version: *What is your age?*

2013: q2a

Würden Sie mir bitte Ihr Geburtsdatum nennen?

English version: *Would you please tell me your date of birth?*

2017: q2a

Würden Sie mir bitte sagen, in welchem Jahr Sie geboren wurden? Und in welchem Monat? Und an welchem Tag?

English version: *What year were you born in? What month were you born in?
On what day were you born?*

Subjective Economic Situation

2009: q18:

Und nun zu Ihrer wirtschaftlichen Lage. Wie beurteilen Sie Ihre derzeitige eigene wirtschaftliche Lage? Bitte sagen Sie es mir anhand dieser Liste.

- (01) sehr gut
- (02) gut
- (03) teils/teils
- (04) schlecht
- (05) sehr schlecht
- (98) weiß nicht
- (99) keine Angabe

English version: *Now, let us shift attention to the economic situation. How would you evaluate your own current economic situation? Please tell me using this list.*

- (01) very good
- (02) good
- (03) neither good nor bad
- (04) bad
- (05) very bad
- (98) don't know
- (99) no answer

2013: q17:

Und nun zu Ihrer wirtschaftlichen Lage. Wie beurteilen Sie Ihre derzeitige eigene wirtschaftliche Lage? Bitte sagen Sie es mir anhand dieser Liste.

- (01) sehr gut
- (02) gut
- (03) teils/teils
- (04) schlecht
- (05) sehr schlecht
- (98) weiß nicht

(99) keine Angabe

English version: *Now, let us shift attention to the economic situation. How would you evaluate your own current economic situation? Please tell me using this list.*

(01) very good

(02) good

(03) neither good nor bad

(04) bad

(05) very bad

(-98) don't know

(-99) no answer

2017: q15:

Und nun zu Ihrer wirtschaftlichen Lage. Wie beurteilen Sie Ihre derzeitige eigene wirtschaftliche Lage? Bitte sagen Sie es mir anhand dieser Liste.

(01) sehr gut

(02) gut

(03) teils/teils

(04) schlecht

(05) sehr schlecht

(98) weiß nicht

(99) keine Angabe

English version: *Now, let us shift attention to the economic situation. How would you evaluate your own current economic situation? Please tell me using this list.*

(01) very good

(02) good

(03) neither good nor bad

(04) bad

(05) very bad

(-98) don't know

(-99) no answer

Female

2009: q1-1:

Ist die Zielperson männlich oder weiblich?

(01) männlich

(02) weiblich

English version: *Is the respondent male or female?*

(01) male

(02) female

2013: q1-1:

Ist die Zielperson männlich oder weiblich?

(01) männlich

(02) weiblich

English version: *Is the respondent male or female?*

(01) male

(02) female

2017: q1:

Ist die Zielperson männlich oder weiblich?

(01) männlich

(02) weiblich

English version: *Is the respondent male or female?*

(01) male

(02) female

Turnout Intention

2009: q9:

Wenn Wahlen stattfinden, geben viele Leute ihre Stimme ab, andere kommen nicht dazu, ihre Stimme abzugeben oder nehmen aus anderen Gründen nicht an der Wahl teil. Einmal angenommen, Sie wären schon wahlberechtigt: Wie wahrscheinlich würden Sie dann am 27. September an der Bundestagswahl teilnehmen?

(01) bestimmt zur Wahl gehen

(02) wahrscheinlich zur Wahl gehen

(03) vielleicht zur Wahl gehen

(04) wahrscheinlich nicht zur Wahl gehen

- (05) bestimmt nicht zur Wahl gehen
- (98) weiß nicht
- (99) keine Angabe
- (100) trifft nicht zu

English version: *When elections are held a lot of people vote, others do not manage to vote or do not participate in elections for other reasons. How likely is it that you will vote in the federal election on 27th September? Please tell me using this list.*

- (01) certain to vote
- (02) likely to vote
- (03) might vote
- (04) not likely to vote
- (05) certain not to vote
- (98) don't know
- (99) no answer
- (100) not applicable

2013: q10:

Wenn Wahlen stattfinden, geben viele Leute ihre Stimme ab, andere kommen nicht dazu, ihre Stimme abzugeben oder nehmen aus anderen Gründen nicht an der Wahl teil. Einmal angenommen, Sie wären schon wahlberechtigt: Wie wahrscheinlich würden Sie dann am 22. September zur Bundestagwahl gehen.

- (1) bestimmt zur Wahl gehen
- (2) wahrscheinlich zur Wahl gehen
- (3) vielleicht zur Wahl gehen
- (4) wahrscheinlich nicht zur Wahl gehen
- (5) bestimmt nicht zur Wahl gehen
- (6) habe bereits per Briefwahl meine Stimme abgegeben
- (-97) trifft nicht zu
- (-98) weiß nicht
- (-99) keine Angabe

English version: *When elections are held a lot of people vote, others do not manage to vote or do not participate in elections for other reasons. How likely is it that you will vote in the federal election on 22nd September? Please tell me using this list.*

- (1) certain to vote

- (2) likely to vote
- (3) might vote
- (4) not likely to vote
- (5) certain not to vote
- (6) have already made a postal vote
- (-97) not applicable
- (-98) don't know
- (-99) no answer

2017: q10:

Wenn Wahlen stattfinden, geben viele Leute ihre Stimme ab, andere kommen nicht dazu, ihre Stimme abzugeben oder nehmen aus anderen Gründen nicht an der Wahl teil. Einmal angenommen, Sie wären schon wahlberechtigt: Wie wahrscheinlich würden Sie dann am 24. September an der Bundestagswahl teilnehmen?

- (1) bestimmt zur Wahl gehen
- (2) wahrscheinlich zur Wahl gehen
- (3) vielleicht zur Wahl gehen
- (4) wahrscheinlich nicht zur Wahl gehen
- (5) bestimmt nicht zur Wahl gehen
- (5) habe bereits per Briefwahl meine Stimme abgegeben
- (5) bin nicht wahlberechtigt
- (-97) trifft nicht zu
- (-98) weiß nicht
- (-99) keine Angabe

English version: *When elections are held a lot of people vote, others do not manage to vote or do not participate in elections for other reasons. How likely is it that you will vote in the federal election on 24th September? Please tell me using this list.*

- (1) certain to vote
- (2) likely to vote
- (3) might vote
- (4) not likely to vote
- (5) certain not to vote
- (6) have already made a postal vote
- (7) not eligible to vote

- (-97) not applicable
- (-98) don't know
- (-99) no answer

⇒ Vote intention was also coded as 1 if respondent indicated to have already vote via mail.

Highschool

2009: d206:

Welchen allgemeinbildenden Schulabschluss haben Sie?

- (01) Schule beendet ohne Abschluss
- (02) Hauptschulabschluss, Volksschulabschluss, Abschluss der polytechnischen Oberschule 8. oder 9. Klasse
- (03) Realschulabschluss, Mittlere Reife, Fachschulreife oder Abschluss der polytechnischen Oberschule 10. Klasse
- (04) Fachhochschulreife (Abschluss einer Fachoberschule etc.)
- (05) Abitur bzw. erweiterte Oberschule mit Abschluss 12. Klasse (Hochschulreife)
- (06) anderen Schulabschluss, und zwar:
- (09) bin noch Schüler
- (98) weiß nicht
- (99) keine Angabe

English version: *What's your highest level of general education?*

- (01) Finished school without school leaving certificate
- (02) Lowest formal qualification of Germany's tripartite secondary school system, after 8 or 9 years of schooling
- (03) Intermediary secondary qualification, after 10 years of schooling
- (04) Certificate fulfilling requirements to study at a polytechnical college
- (05) Higher qualification, entitling holders to study at a university
- (06) other school leaving certificate, please enter:
- (09) still at school
- (98) don't know
- (99) no answer

2013: q163:

Welchen allgemeinbildenden Schulabschluss haben Sie?

- (01) Schule beendet ohne Abschluss
- (02) Hauptschulabschluss, Volksschulabschluss, Abschluss der polytechnischen Oberschule 8. oder 9. Klasse
- (03) Realschulabschluss, Mittlere Reife, Fachschulreife oder Abschluss der polytechnischen Oberschule 10. Klasse
- (04) Fachhochschulreife (Abschluss einer Fachoberschule etc.)
- (05) Abitur bzw. erweiterte Oberschule mit Abschluss 12. Klasse (Hochschulreife)
- (06) anderen Schulabschluss, und zwar:
- (09) bin noch Schüler
- (-98) weiß nicht
- (-99) keine Angabe

English version: *What's your highest level of general education?*

- (01) Finished school without school leaving certificate
- (02) Lowest formal qualification of Germany's tripartite secondary school system, after 8 or 9 years of schooling
- (03) Intermediary secondary qualification, after 10 years of schooling
- (04) Certificate fulfilling requirements to study at a polytechnical college
- (05) Higher qualification, entitling holders to study at a university
- (06) other school leaving certificate, please enter:
- (09) still at school
- (-98) don't know
- (-99) no answer

2017: q136:

Welchen allgemeinbildenden Schulabschluss haben Sie?

- (01) Schule beendet ohne Abschluss
- (02) Hauptschulabschluss, Volksschulabschluss, Abschluss der polytechnischen Oberschule 8. oder 9. Klasse
- (03) Realschulabschluss, Mittlere Reife, Fachschulreife oder Abschluss der polytechnischen Oberschule 10. Klasse
- (04) Fachhochschulreife (Abschluss einer Fachoberschule etc.)
- (05) Abitur bzw. erweiterte Oberschule mit Abschluss 12. Klasse (Hochschulreife)

(06) anderen Schulabschluss, und zwar:

(09) bin noch Schüler

(98) weiß nicht

(99) keine Angabe

English version: *What's your highest level of general education?*

(01) Finished school without school leaving certificate

(02) Lowest formal qualification of Germany's tripartite secondary school system, after 8 or 9 years of schooling

(03) Intermediary secondary qualification, after 10 years of schooling

(04) Certificate fulfilling requirements to study at a polytechnical college

(05) Higher qualification, entitling holders to study at a university

(06) other school leaving certificate, please enter:

(09) still at school

(-98) don't know

(-99) no answer

Satisfaction with Democracy

2009: q5:

Wie zufrieden oder unzufrieden sind Sie - alles in allem - mit der Demokratie, so wie sie in Deutschland besteht? Sind Sie ... sehr zufrieden, ziemlich zufrieden, teils/teils, ziemlich unzufrieden oder sehr unzufrieden?

(01) sehr zufrieden

(02) ziemlich zufrieden

(03) teils/teils

(04) ziemlich unzufrieden

(05) sehr unzufrieden

(98) weiß nicht

(99) keine Angabe

English version: *On the whole, are you very satisfied, fairly satisfied, neither satisfied nor dissatisfied, not very satisfied, or not at all satisfied with the way democracy works in Germany?*

(01) very satisfied

- (02) fairly satisfied
- (03) neither satisfied nor dissatisfied
- (04) not very satisfied
- (05) not at all satisfied
- (98) don't know
- (99) no answer

2013: q6:

Wie zufrieden oder unzufrieden sind Sie - alles in allem - mit der Demokratie, so wie sie in Deutschland besteht? Sind Sie ... sehr zufrieden, ziemlich zufrieden, teils/teils, ziemlich unzufrieden oder sehr unzufrieden?

- (01) sehr zufrieden
- (02) ziemlich zufrieden
- (03) teils/teils
- (04) ziemlich unzufrieden
- (05) sehr unzufrieden
- (-98) weiß nicht
- (-99) keine Angabe

English version: *On the whole, are you very satisfied, fairly satisfied, neither satisfied nor dissatisfied, not very satisfied, or not at all satisfied with the way democracy works in Germany?*

- (01) very satisfied
- (02) fairly satisfied
- (03) neither satisfied nor dissatisfied
- (04) not very satisfied
- (05) not at all satisfied
- (-98) don't know
- (-99) no answer

2017: q6:

Wie zufrieden oder unzufrieden sind Sie - alles in allem - mit der Demokratie, so wie sie in Deutschland besteht? Sind Sie ... sehr zufrieden, ziemlich zufrieden, teils/teils, ziemlich unzufrieden oder sehr unzufrieden?

- (01) sehr zufrieden
- (02) ziemlich zufrieden

- (03) teils/teils
- (04) ziemlich unzufrieden
- (05) sehr unzufrieden
- (-98) weiß nicht
- (-99) keine Angabe

English version: *On the whole, are you very satisfied, fairly satisfied, neither satisfied nor dissatisfied, not very satisfied, or not at all satisfied with the way democracy works in Germany?*

- (01) very satisfied
- (02) fairly satisfied
- (03) neither satisfied nor dissatisfied
- (04) not very satisfied
- (05) not at all satisfied
- (-98) don't know
- (-99) no answer

References

- Adler, Werner, Sergej Potapov and Berthold Lausen. 2011. "Classification of repeated measurements data using tree-based ensemble methods." *Computational Statistics* 26(2):355–369.
- Bell, Andrew and Kelvyn Jones. 2015. "Explaining Fixed Effects: Random Effects Modeling of Time-Series Cross-Sectional and Panel Data." *Political Science Research and Methods* 3(1):133–153.
- Breiman, Leo. 2001. "Random Forests." *Machine Learning* 45:5–32.
- Capitaine, Louis. 2020. "LongituRF: Random Forests for Longitudinal Data." **URL:** <https://cran.r-project.org/package=LongituRF>
- Hajjem, Ahlem, François Bellavance and Denis Larocque. 2014. "Mixed-effects random forest for clustered data." *Journal of Statistical Computation and Simulation* 84(6):1313–1328.
- Hu, Jianchang and Silke Szymczak. 2023. "A review on longitudinal data analysis with random forest." *Briefings in Bioinformatics* pp. 1–11.

- Karpiévitch, Yuliya V., Elizabeth G. Hill, Anthony P. Leclerc, Alan R. Dabney and Jonas S. Almeida. 2009. “An Introspective Comparison of Random Forest-Based Classifiers for the Analysis of Cluster-Correlated Data by Way of RF++.” *PLoS ONE* 4(9).
- Pellagatti, Massimo, Chiara Masci, Francesca Ieva and Anna M. Paganoni. 2021. “Generalized mixed-effects random forest: A flexible approach to predict university student dropout.” *Statistical Analysis and Data Mining* 14(3):241–257.
- Speiser, Jaime Lynn, Bethany J. Wolf, Dongjun Chung, Constantine J. Karvellas, David G. Koch and Valerie L. Durkalski. 2019. “BiMM forest: A random forest method for modeling clustered and longitudinal binary outcomes.” *Chemometrics and Intelligent Laboratory Systems* 185(January):122–134.
URL: <https://doi.org/10.1016/j.chemolab.2019.01.002>

Declaration of Interest Statement

Competing interests: The authors declare none.

Acknowledgements

We are grateful to Franziska Quosß, two anonymous reviewers, and the Editor for their constructive feedback on earlier versions of this manuscript.