Supplement material for

Analyzing Nonresponse in Longitudinal Surveys Using Bayesian Additive Regression Trees:

A Nonparametric Event History Analysis

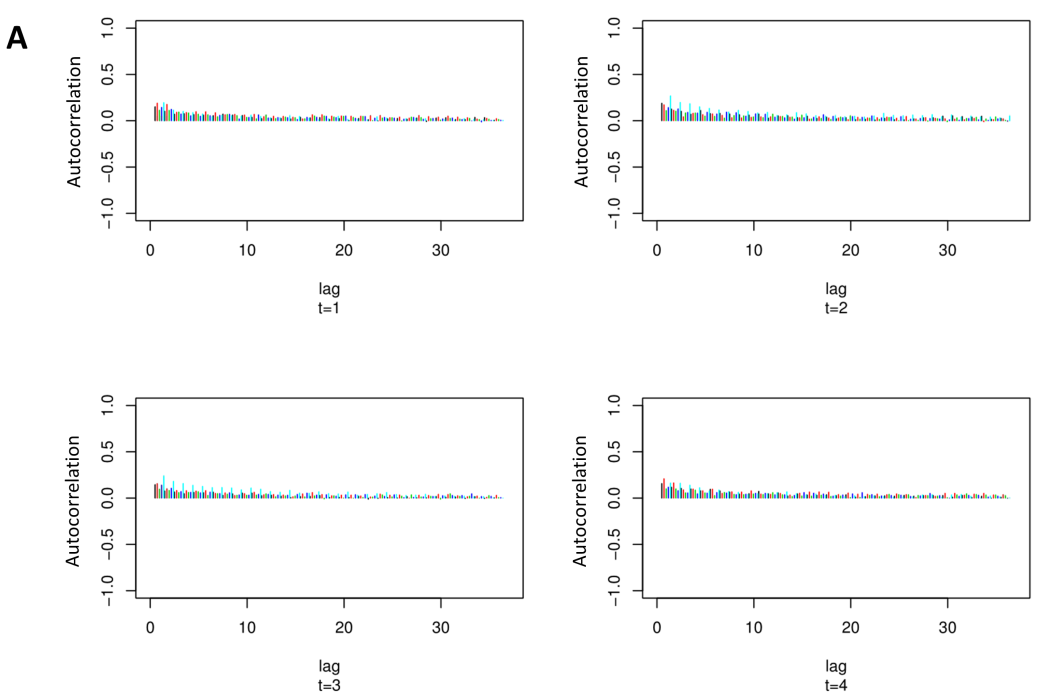
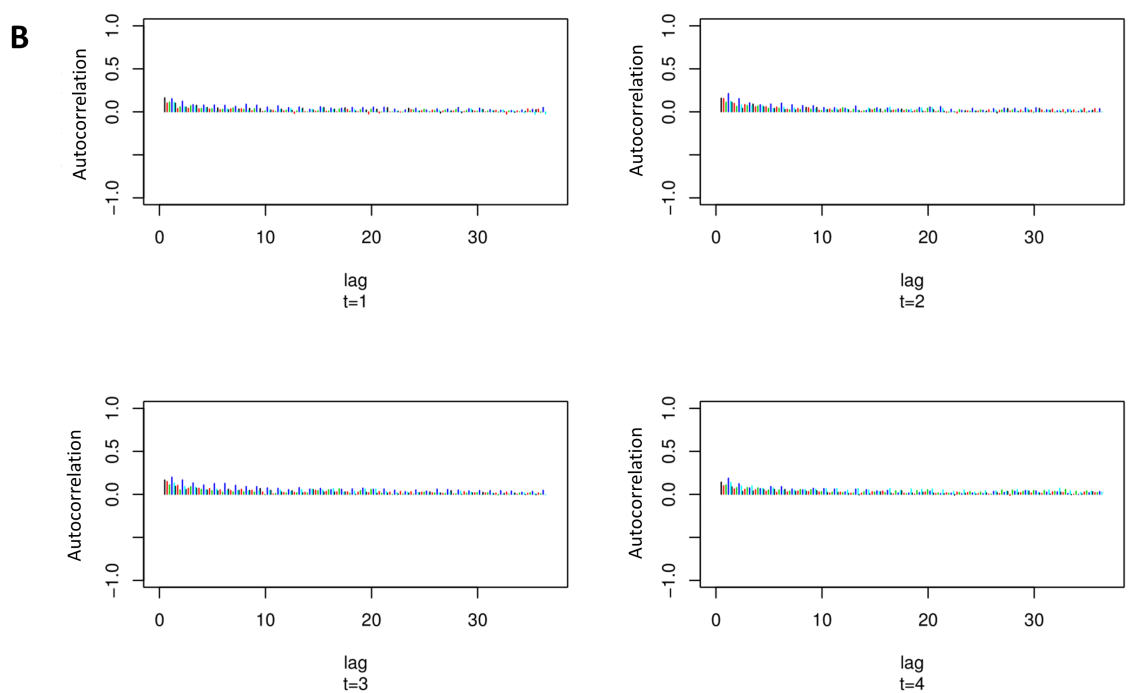
S1. Derivation of Decomposition of Formula for Survey Participation at Wave

In the stochastic process , represents the survey participation resulting from a non-recurrent event (i.e., permanent dropout) and the respective respondent outcome for a recurrent event (i.e., temporary dropout) at time point *t*. The joint transition probabilities for and can be derived from two submodels by splitting the respective probabilities into conditional probabilities for the non-recurrent event and the recurrent event in the following way:

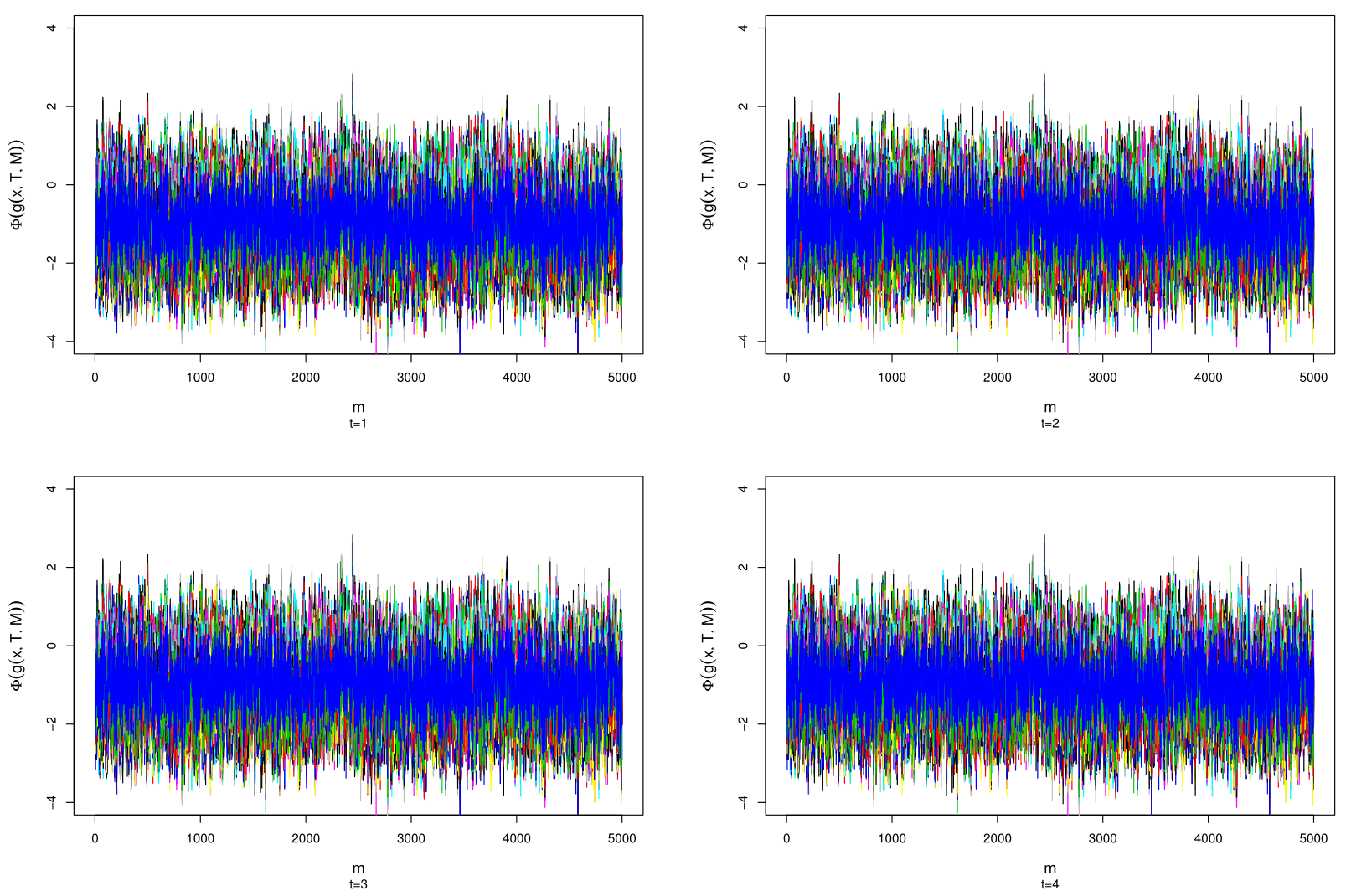
Table S1.

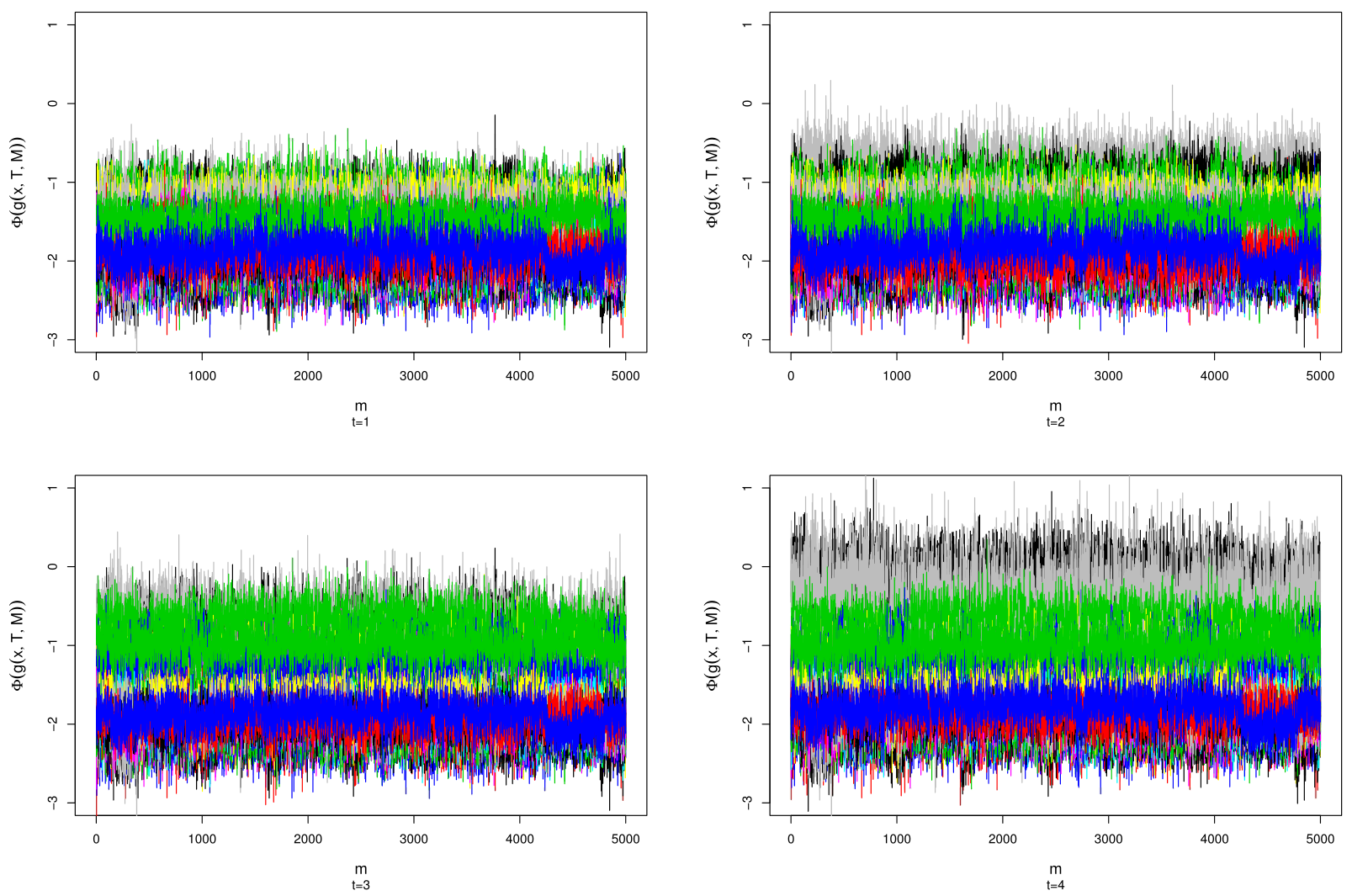
*Conditioning Variables for Nonresponse Analyses*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | *M* | *SD* | Range | MV |
|  | *Respondent characteristics* |  |  |  |  |
| 1. | Sex | 0.5 | 0.5 | [0, 1] | 0% |
| 2. | Age | 15.1 | 0.5 | [12.7, 18.2] | 0% |
| 3. | Mother tongue | 0.1 | 0.3 | [0, 1] | 0% |
| 4. | Migration background | 0.2 | 0.4 | [0, 1] | 0% |
| 5. | Household size | 4.5 | 1.7 | [2, 35] | 5% |
| 6. | Number of books at home | 4.0 | 1.4 | [0, 6] | 1% |
|  | *Student characteristics* |  |  |  |  |
| 7. | Repeated school | 0.1 | 0.3 | [0, 1] | 2% |
| 8. | Number of days sick | 2.0 | 3.8 | [0, 50] | 21% |
| 9. | Grade in German | 2.3 | 0.9 | [1, 6] | 7% |
| 10. | Grade in mathematics | 2.3 | 0.9 | [1, 6] | 7% |
|  | *Psychological characteristics* |  |  |  |  |
| 11. | Satisfaction with life | 8.2 | 2.3 | [0, 10] | 5% |
| 12. | Satisfaction with current living standards | 8.7 | 2.2 | [0, 10] | 4% |
| 13. | Satisfaction with health | 8.8 | 2.2 | [0, 10] | 3% |
| 14. | Satisfaction with family | 9.1 | 2.0 | [0, 10] | 4% |
| 15. | Satisfaction with friends | 8.9 | 2.0 | [0, 10] | 3% |
| 16. | Satisfaction with school | 7.8 | 2.5 | [0, 10] | 3% |
| 17. | Subjective health | 1.7 | 0.7 | [1, 5] | 1% |
| 18. | Self-esteem | 4.0 | 0.7 | [1.0, 5.0] | 15% |
| 19. | German self-concept | 3.0 | 0.6 | [1, 4] | 7% |
| 20. | Mathematical self-concept | 2.9 | 0.8 | [1, 4] | 7% |
| 21. | School self-concept | 3.2 | 0.6 | [1, 4] | 7% |
| 22. | Perceptual speed | 44.0 | 13.0 | [1, 93] | 0% |
| 23. | Reading speed | 21.5 | 6.9 | [0, 51] | 0% |
| 24. | Reasoning | 7.0 | 2.6 | [0, 12] | 0% |
| 25. | Orthography | 0.0 | 1.3 | [-7.2, 4.6] | 0% |
| 26. | Mathematical competence | 0.1 | 1.2 | [-4.4, 4.0] | 0% |
| 27. | Reading competence | 0.0 | 1.2 | [-4.2, 4.1] | 0% |
|  | *School characteristics* |  |  |  |  |
| 28. | School type: intermediate secondary | 0.2 | 0.4 | [0, 1] | 0% |
| 29. | School type: other | 0.3 | 0.4 | [0, 1] | 0% |
| 30. | Number of students in grade 8 | 102.8 | 50.7 | [12, 346] | 2% |
| 31. | Number of classes in grade 8 | 3.9 | 1.7 | [1, 12] | 2% |
| 32. | Institution type | 0.1 | 0.3 | [0, 1] | 0% |
| 33. | School location: part urban / part rural | 0.4 | 0.5 | [0, 1] | 0% |
| 34. | School location: rural | 0.1 | 0.3 | [0, 1] | 0% |
| *Note*. Respondent, student, and psychological characteristics aggregated to the school level and dummy-indicators for federal states and missing values are not included. | | | | | |



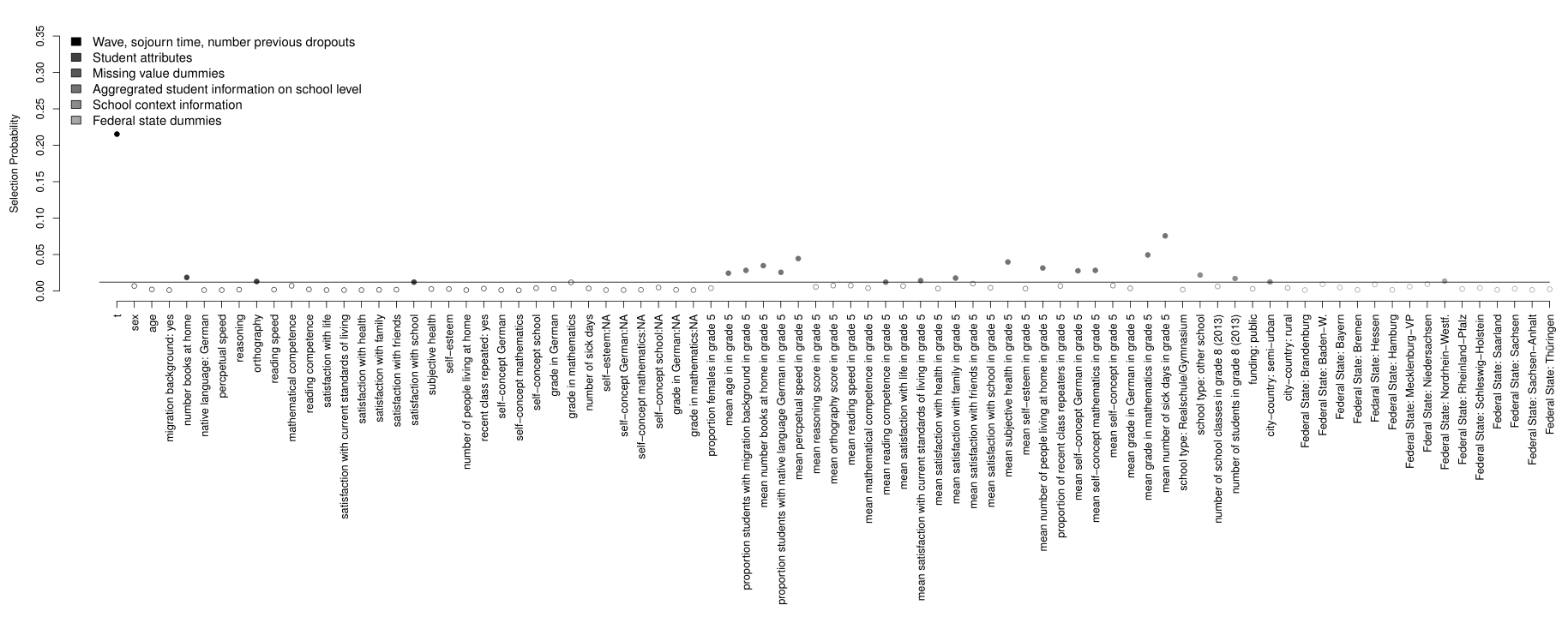
*Figure S1*. Autocorrelation plots for BART event history model. A = Permanent dropout model, B = Temporary dropout model.

A.

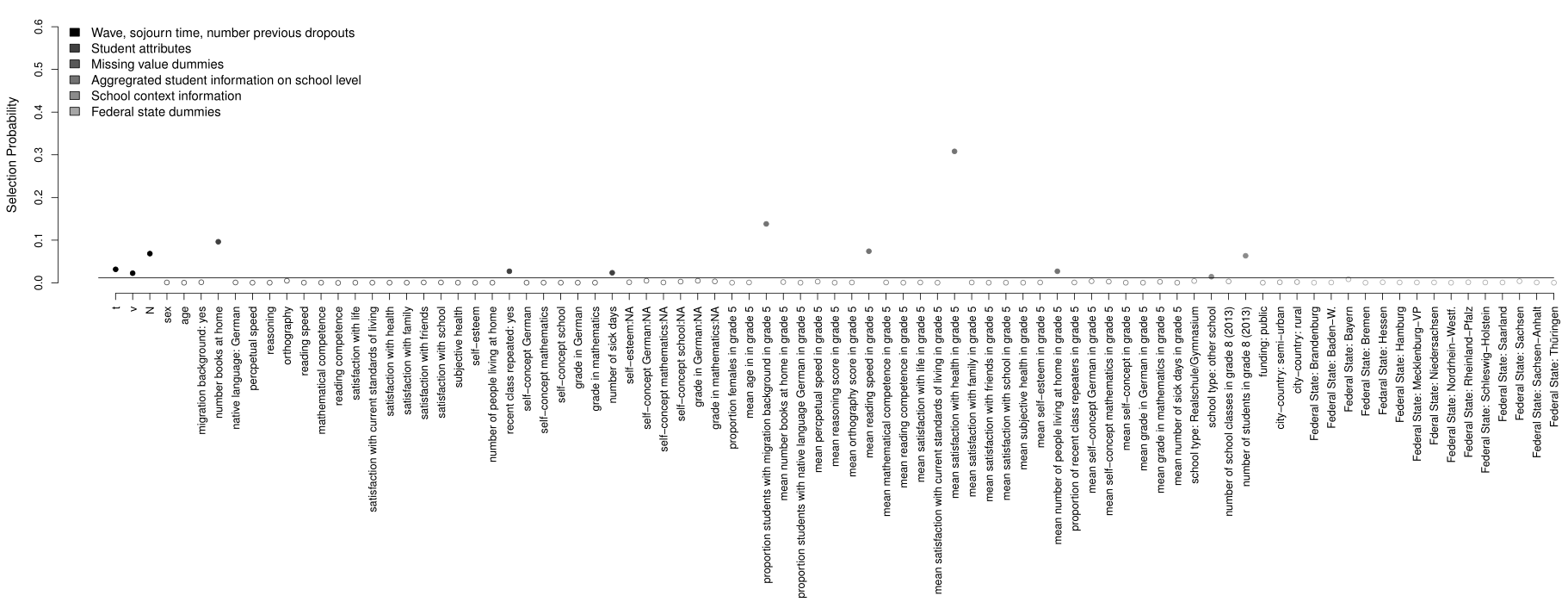


B.

*Figure S2*. Trace plots for BART event history model. A = Permanent dropout model, B = Temporary dropout model.



*Figure S3*. Relative covariate importance for predicting permanent dropout with *t* as survey wave, *v* as the event time, and *N* as the number of previous dropouts for BART. The solid line represents the threshold for nonignorable importance, filled dots mark variables of nonignorable importance, and empty dots mark variables of ignorable importance.



*Figure S4*. Relative covariate importance for predicting temporary dropout with *t* as survey wave, *v* as the event time, and *N* as the number of previous dropouts for BART. The solid line represents the threshold for nonignorable importance, filled dots mark variables of nonignorable importance, and empty dots mark variables of ignorable importance.