Predicting GPA at Age 15 in the Fragile Families and Child Wellbeing Study

Louis Raes

Abstract
In this paper, we describe in detail the different approaches we used to predict the GPA of children at the age of 15 in the context of the Fragile Families Challenge. Our best prediction improved about 18 percent in terms of mean squared error over a naive baseline prediction and performed less than 5 percent worse than the best prediction in the Fragile Families Challenge. After discussing the different predictions we made, we also discuss the predictors that tend to be robustly associated with GPA. One remarkable predictor is related to teacher observations at the age of nine. We end with a reflection on our participation in the Fragile Families Challenge and provide some suggestions for follow-up work.

Keywords
Fragile Families, predicting GPA, computational social science

Introduction
This paper describes our effort in predicting GPA of children in the context of the Fragile Families Challenge (FFC). The FFC was a mass collaboration effort to study six important outcomes in the lives of children and families when the child was 15. Participants in the FFC were diverse in terms of experience (from undergraduate students to tenured faculty) and fields of expertise (with backgrounds in sociology, physics, economics, etc.). Each participant was required to make predictions of the outcome variables using a large data set consisting of earlier waves of the Fragile Families and Child Wellbeing Study. Otherwise participants were completely free to use any approach they deemed appropriate. The participants were blind to the actual outcome but could submit predictions to a leaderboard that provided feedback in terms of predictive accuracy.

In this paper, we describe our approach to predict one outcome, GPA. This score reflects school performance and ranges from 1 (worst) to 4 (best).

The approach taken in this paper reflects our background as a social scientist. We browsed the relevant literature, read through the variable descriptions, and experimented with various methods. We started with a rather compact data set. Then, we created new data sets by adding (or removing) variables to this initial data set. The changes to this data set were based on our reading of the literature and the results of our predictions. This led to a total of 12 data sets we explored using different prediction models.

The approach we took grew organically throughout our participation in the FFC. As the goal of this paper is to explicitly describe what we did, we restrict ourselves to the predictions we made during the contest.

We found that a simple approach using only constructed variables already led to a substantial improvement in terms of predictive accuracy over a baseline prediction. Further improvements were modest. Our best prediction had a mean squared error (MSE) that was about 18 percent lower than a baseline prediction but was more than 4 percent off the best prediction in the FFC. Key variables are related to parental demographics and background and teacher observations at the age of nine.

This manuscript is structured as follows. We describe the creation of various data sets, present the prediction models, discuss the performance of our predictions, review the key covariates, and end with a conclusion.

Data Cleaning and Handling
We created 12 different data sets to predict GPA. To start, we restricted ourselves to the constructed variables. These

1Details can be found in Salganik et al. (2019).

Creative Commons Non Commercial CC BY-NC. This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (http://www.creativecommons.org/licenses/by-nc/4.0/) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).
variables are constructed by the Fragile Families research staff. This is a convenient starting point because these variables have often little missingness and combine variables into meaningful constructs.

During our participation in the FFC, we created different data sets by adding (and sometimes removing) variables. Both removing and adding variables can improve performance (Kuhn and Johnson 2013). One example is splitting up a categorical variable into a series of dummy variables.

We cleaned the data set, consisting of constructed variables, by removing variables that we deemed not to have a meaningful interpretation such as indicators of sample characteristics. We also removed variables with no variability. In case of missing observations, we used median imputation.

Next, we reduced the number of variables related to poverty and household income. The data contain 40 of such closely related variables. We reduced these to three variables via principal component analysis. Then we transformed categorical variables containing four or less categories to a series of dummy variables (one hot encoding). We applied the same transformation to the categorical variables capturing the relationship status of the parents. This data set we call ConstrTeacher.

To this data set, we added different sets of variables or transformed the same set of variables differently.

We started by adding variables coming from the teacher survey at age nine, the most recent information available. We believe that information from teachers is important to consider because children spend a large amount of time in school (Hofferth and Sandberg 2001). We focused on questions related to classroom behavior and social skills, and we dealt with these questions in different ways. First, we used a narrow set of these questions focusing on social skills. These were added to the data set previously described to create the data set ConstrSocial. Then, we added all the questions on classroom behavior to ConstrTeacher to create data set ConstrTeacher.

The answers on some of these questions are highly correlated as they relate to the same underlying concept. For example, teachers had to rate children on questions like: “Controls temper in conflict situations . . . ” or “Responds appropriately to teasing . . . ” We aimed at capturing the underlying construct using four different strategies. First, we created cumulative scores by summing the answers to related questions. These indices were added to ConstrTeacherPC. Second, we added categorical variables of five to seven levels based on these indices. This led to ConstrTeacherIndex. Third, we create a dummy variable for each level of the corresponding categorical variable to allow for nonlinear effects. The resulting data set is called ConstrTeacherDummy. Finally, we extracted principal components from these sets of related questions. The resulting data set is called ConstrTeacherPr.

We explored these seven data sets by assessing the predictive performance of various elastic net models (see the following) through cross-validation. From this, we learned that the data set ConstrTeacherPr tends to yield the lowest prediction errors.

Next, we searched for existing research using the Fragile Families and Child Wellbeing Study. We used Google Scholar to look for papers using survey data and studying some of the outcome variables in the FFC. We considered the first 25 papers in the search results. Two were irrelevant, and 2 could papers could not be accessed. We read the remaining 21 papers and listed relevant variables and constructs. We

---

2 The raw data set contains 591 constructed variables.

3 With the term missing observations, we refer to all observations that are not available (coded with a negative value in the original data set).

4 Note that some prediction models are insensitive to having many (highly correlated) predictors and lack a strong theoretical foundation. Our aim here is to provide an accurate description of our predictions.

---

Table 1. Data Sets Constructed for the Fragile Families Challenge.

| Name of Data Set          | Submitted Predictions | Number of Variables | Elements                                                                 |
|---------------------------|-----------------------|---------------------|--------------------------------------------------------------------------|
| ConstrTeacher             | 1                     | 328                 | Based on the constructed variables in the Fragile Families survey         |
| ConstrTeacherPC           | 2, 3                  | 310                 | ConstrTeacher+ t5b variables                                             |
| ConstrTeacherIndex        | 302                   |                     | ConstrTeacher+ t5b variables as index variables                          |
| ConstrTeacherCat          | 302                   |                     | ConstrTeacher+ t5b variables as categorical variables                    |
| ConstrTeacherPr           | 351                   |                     | ConstrTeacher+ t5b variables as principal components                     |
| ConstrTeacherDummy        | 321                   |                     | ConstrTeacher+ t5b variables as dummy variables                         |
| ConstrTeacherPCAge9       | 4, 5, 6               | 351                 | ConstrPC + items from interviewer and mother at age 9                    |
| ConstrTeacherPCAge9Age5   | 396                   |                     | ConstrTeacherPCAge9 + variables at age 5                                 |
| ConstrTeacherPCAge9Age5PC | 362                   |                     | ConstrTeacherPCAge9 + variables at age 5 transformed                     |
| Interaction               | 7, 8                  | 25                  | Interactions between relevant variables                                   |
| InteractionPCAge9         | 9, 10, 11             | 80                  | Interaction + items from interviewer and mother at age nine               |

Socius: Sociological Research for a Dynamic World
then browsed various questionnaires to add variables to our data set. We added variables coming from the impressions the interviewer got when making the interviews at age nine. These questions concern the childhood environment, hygiene of the child, and so on. We also added questions from the mother related to her perception of the father, any hardship she might face, as well as the mother’s well-being (e.g., religious attendance, life satisfaction, smoking). This is ConstrTeacherPCAge9.

Next we considered questions asked to the parents when the child was five. These questions were related to how the mother raises the child (e.g., spanking, kindergarten) and how she feels about raising a child. The answers from the father concern any financial aid the father was receiving as well as the father’s involvement in raising the child. These variables were added to ConstrTeacherPCAge9 to yield ConstrTeacherPCAge9Age5. We also made a variation of this data set, ConstrTeacherPCAge9Age5PC, where we added principal components from subsets of related variables instead.

The variables we chose to add to the constructed variables thus come from the waves at ages five and nine. This does not mean that we disregard the importance of early childhood, but we thought that the early life conditions would be well captured by the constructed variables.

We guessed that more recent information, such as from a teacher in elementary school, would be more helpful in predicting outcomes at age 15 than earlier data.

Finally, we constructed two additional data sets focusing on interactions between variables we considered relevant (based on the predictions we already made). We used the lasso (see the following) to select the variables. We fitted a lasso model on ConstrTeacherPCAge9Age5. We took all variables that were not set to zero and created manually a range of interaction variables. This formed the data set Interaction. The final data set is InteractionPCAge9, which combines Interaction with answers from the mother and father at age nine we earlier used. Table 1 provides an overview.

In Table 1, we show the different data sets we constructed and explored in the FFC. The table shows which data sets were ultimately used in predictions we submitted. The third column indicates the number of variables in these data sets. The last column shows the relation between the different data sets.

**Methods**

In our prediction exercise, we used four types of prediction models.

---

5There is an extensive literature on the importance of early childhood. The overview paper by Heckman (2008) contains many references. There are government-sponsored programs aimed at helping disadvantaged children in their early years. Garecs, Thomas, and Currie (2002) evaluate the longer term effects of one program.

6We only provide here a short description. For a succinct introduction, see Kuhn and Johnson (2013).

---

The elastic net (Zou and Hastie 2005) is a (penalized) linear model that has two penalties. First, there is regularization, which means that large coefficients are shrunk toward zero. Second, there is feature selection that means that some features are set to zero. These penalties are governed by two tuning parameters: \( \lambda \) (the ridge parameter) and \( \alpha \) (the lasso parameter). Setting \( \alpha = 0 \) yields ridge regression (no feature selection), whereas \( \alpha = 1 \) yields lasso regression (feature selection). Intermediate values of \( \alpha \) mix both.

Multivariate adaptive regression splines (Friedman 1991), or MARS, uses surrogate features instead of the original predictors. It is an inherently nonlinear model where nonlinear relationships between predictors do not need to be specified in advance. MARS creates contrasted features of a predictor to enter a model. These new features are added to a basic linear regression to create a piecewise linear model where each feature models an isolated portion of the data (Kuhn and Johnson 2013). The MARS model sequentially searches for prediction/cut point combinations that achieve the smallest error. After an initial selection, the model searches for the next set of features that, given the initial set, provides the best model fit. Once the full set of features is determined, a pruning procedure removes superfluous features.

We also use two tree-based models. These are models that build on regression trees where the idea is to partition the data into smaller groups that are homogenous with respect to the response.

We consider generalized boosted modeling, or gbm (Ridgeway 2017). The underlying idea is to combine individual regression trees that may not have strong predictive performance to improve predictive performance. There are different approaches to construct an ensemble of models that yields the final prediction. Boosting is one approach (Friedman 2002) that is known to work well (Kuhn and Johnson 2013).

We also consider Cubist, a rule-based model based on Quinlan (1987, 1992, 1993). Rules are distinct paths through a tree. Rule-based models try to simplify rules generated from trees, cutting down model complexity and hopefully also improving predictive performance. Cubist draws on other approaches (e.g., boosting) to improve the predictive performance. We chose this method because of its strong performance in horse race presented by Kuhn and Johnson (2013).

All models require tuning parameters that were chosen using cross-validation.

**Performance**

We submitted 11 predictions to the leaderboard. An overview of the models and data sets we used can be found in Table 2. This table provides the performance in terms of MSE given by the leaderboard and the performance on the holdout data. We notice that our predictions always did better on the holdout data than on the leaderboard data. This could be coincidental but could also be due to the fact that the leaderboard
data contained randomly imputed data (Salganik et al. 2019). The performance on the leaderboard is also more variable than on the holdout data.

We see that our best submission, based on the performance on the holdout data, is only 2.5 percent better than our worst in MSE terms. To see this in perspective, consider a naive baseline that uses the mean observed value as a prediction. Such a model yields an MSE of .4251. Our best prediction yields an MSE that is more than 18 percent lower. Consider on the other hand the best submission in the FFC. This submission obtained a score of slightly less than .344, which is an improvement of less than 5 percent over our best prediction.

From this, we conclude that the model we initially submitted already yielded a considerable improvement over the baseline model. We were then able to improve slightly, but we were still a bit removed from the best prediction in the FFC.

Table 2 shows the data sets and predictive models used in the different predictions we submitted to the Fragile Families Challenge. The MSE scores are rounded to four decimals. \( \lambda_{\text{min}} \) is the lambda for which the mean cross-validated error is minimized. The lowest MSE on the holdout data is shown in bold. Additional information on our predictions can be found in the supporting online materials.

### Interpretation

The discussion so far centered around predictive performance as this was the explicit goal of the prediction contest. Here we look at the variables that play a role in our predictions.

We do not discuss all submitted models. For the elastic net models, we focus on the models with \( \alpha = 1 \) (lasso) because the other models have more predictors (with very small values) that encompass the predictors of the lasso models. We also do not show the variables in the models based on the data sets with interactions because these are by construction mostly based on the variables we report here (see the previous discussion on the construction of the data sets) and importantly because these models do not seem to outperform the ones we report here.

This leaves us with Models 2, 4, 5, and 6 (see also Table 2). For the elastic net models, we show which variables were not set to zero. For Model 5, we show the variables that are used to make a prediction (which are left after the pruning procedure; see our previous description of MARS). For Model 6, we show the 10 variables with the largest relative influence. These are the variables that contributed the most in reducing the loss function (Friedman 2001).

We see four variables playing a role in the four models. Model 5 only uses these four variables to predict GPA (after pruning).

We notice that all variables are either related to the socioeconomic background of the parents or capture classroom behavior at age nine. The variable with the largest relative influence in our best model (Model 6) is \( \text{schoolpr1} \). This is the first principal component extracted from a set of 25 items in the teacher survey at age nine. This set contains items that reflect how much the child is participating in educational activities, for example, whether homework is finished in time, the child attends teacher instructions, keeps the desk clean, and so on. Table 3 shows which variables played a role in different predictions.

### Conclusion

Our first prediction proved to be a substantial improvement over a baseline prediction that simply predicts the mean value of all observations. Subsequently, we explored different data sets using different types of prediction models. This resulted in only a modest improvement.

Looking back on our participation in the Fragile Families Challenge, we have the following takeaways.
First, variables related to teacher observations are relevant to consider when making predictions about GPA at the age of 15. In our best performing model (Model 6, see “Interpretation” section), these variables played a major role (as measured by relative influence).

Second, in our predictions, we relied solely on median imputation. It is unlikely that this is the best approach (Kuhn and Johnson 2013), and some improvement could be expected by investigating alternative imputation schemes.

Third, the effort invested in creating new data sets based on our reading of the literature and the codebook as well as exploring new algorithms only led to modest improvements. This suggests that a simple prediction approach, using an elastic net on the constructed variables with median imputation, captures most of the low hanging fruit. After that, improvements were increasingly difficult to achieve.

We tried to improve by exploring different models and reading through the literature. Implementing different machine learning models was more alien to us and took us more efforts without substantial benefits—in hindsight. For this reason, we would recommend researchers improve predictions by putting effort in what they are familiar with.

### Appendix

**A. Predicting GPA at Age 15 in the Fragile Families and Child Wellbeing Study: Supporting Information**

In this appendix, we provide some additional information on the predictions we submitted to the leaderboard in the Fragile Families Challenge. To know exactly what we did, one needs to go through the R code.

Here we show some additional results, and we provide some verbal explanation on what we did to create our submitted predictions.

#### A1. Submitted Predictions 1–3

We created several data sets, as explained in the paper. We subsequently estimated various elastic net models. Cv.glmnet, the implementation we used, finds an optimal \( \lambda \) parameter for a given \( \alpha \) parameter. Hence, we estimated the glmnet repeatedly and got a sense of the fit through the reported deviance. We considered initially seven data sets and four different values of the \( \alpha \) parameter. We considered each time two different sets of folds (for cross-validation). So we see for each choice of \( \alpha \) each time two sets of results. One should therefore compare even columns with even columns and odd columns with odd columns. In the following, we show the outcome of this exercise.

| Variable                                                                 | Model 2 | Model 4 | Model 5 | Model 6 | Relative Influence |
|--------------------------------------------------------------------------|---------|---------|---------|---------|--------------------|
| Mother’s age at her 1st birth \( cm2fbir \)                               | Y       | Y       | Y       | Y       | 3.74               |
| Father has spent time in jail between birth and age nine \( cm5fevjail \)| Y       | Y       |         |         |                    |
| First PC of items on poverty ratio and household income \( povertycomp \)| Y       | Y       | Y       | Y       | 7.58               |
| Mother did not finish high school \( dummy cm1edu 1 \)                    | Y       | Y       | Y       | Y       | 5.54               |
| Mother is black \( dummy cm1ethrace 2 \)                                 | Y       | Y       | Y       | Y       | 2.04               |
| Father has college or graduate degree \( dummy cf1edu 4 \)               | Y       | Y       | Y       | Y       | 11.48              |
| First PC of items on class behavior (age nine) \( schoolpr1 \)          | Y       | Y       | Y       |         | 25.77              |
| Second PC of items on poverty ratio and household income \( povertycomp2 \)| Y       |         |         |         |                    |
| Third PC of items on social behavior (age nine) \( socialpr3 \)          |         |         |         | Y       | 2.88               |
| Third PC of items on mental well-being (age nine) \( mentalhealthpr3 \)  |         |         |         | Y       | 2.22               |
| First PC of items on child’s self control (age nine) \( selfcontrolpr1 \)|         |         |         | Y       | 2.04               |

Note: PC = principal component.
The lowest loss we obtained is in the vicinity of .36xx. We submitted three predictions. The first submitted prediction was based on cvgpa1 and $\alpha = .2$, the second submitted prediction was based on cvgpaPr and $\alpha = .5$, and the third submitted prediction was based on cvgpaPr and $\alpha = .2$.

The results suggest (roughly) that a high $\alpha$ tends to do better than a low $\alpha$. We also see that we tend to make the most accurate predictions with cvgpaPr.

A2. Submitted Predictions 4. For this prediction, we added a selection of items from the survey at age nine to construct ConstrTeacherPC to create ConstrTeacherPCAge9. First we explored different elastic net models. We used cross-validation on a single set of folds and compared $\alpha$ parameters equal to 1, .9, .8, .5, .2, .1, 0: We chose to submit predictions based on the elastic net with an $\alpha$ equal to 1 (and optimal $\lambda$ parameter as before).

A3. Submitted Predictions 5. We now explored multivariate adaptive regression splines. To do so, we set up a grid of tuning parameters. The grid is formed by considering one and two degrees of interaction and three to nine terms in the pruned model. For a discussion, see the vignette of the Earth package. We applied this model to four data sets: Constructed, ConstrTeacherPC, ConstrSocial, and ConstrTeacherPCAge9. Because our grid forms $2 \times 7 = 14$ combinations of tuning parameters, we explored 14 models on each of these data sets. From all these models, we chose the mars model with degree = 1 and nprune = 5, applied to ConstrTeacherPCAge9.

A4. Submitted Predictions 6. We now explored the generalized boosted modeling (gbm) algorithm. Given the success of the ConstrTeacherPCAge9, we only explored different parameterizations of gbm on this data set. We again set up a tuning grid. We chose to consider an interaction depth of 1, 3, 5, 7; a number of trees equal to 100, 250, 400, 550, 700, 850, 1,000; a shrinkage parameter of .01 and .1, and a minimum number of observations per node equal to 8 and 14. This implies that our grid comprises $4 \times 7 \times 2 \times 2 = 56$ different models that are explored using cross-validation. The choice of tuning parameters was based on some initial experimentation. The optimal model is chosen on the basis of root mean squared error (RMSE). The tuning parameters for this model are an interaction depth = 1, a number of trees = 550, a shrinkage parameter of 1, and a minimum of 8 observations per node.

A5. Submitted Predictions 7. We continued with exploring new data sets. As explained in the main text, we added to ConstrTeacherPCAge9 items taken from the survey at the age of five. We did this in two ways. Once we used the items as they were, and once we use transformed variables (e.g., by extracting principal components from related variables). This led to ConstrTeacherPCAge9Age5 and ConstrTeacherPCAge9Age5PC. We looked at the elastic net model with $\alpha$ equal to 1 and the optimal lambda, applied to ConstrTeacherPCAge9Age5PC. We took all predictors that were not constrained to be zero and used these variables to create a number of interaction variables, 25 in total. We call this data set Interaction and used it to make predictions as well. On each of these three data sets, we explored elastic net models with $\alpha$ parameters equal to 1, .9, .8, .5, .2, .1, 0. The best performance was achieved with an $\alpha$ parameter of .2, and so we used that $\alpha$ (and optimal lambda) for our predictions.

| $\alpha$ | Performance |
|----------|-------------|
| 1        | .3699623    |
| .9       | .3700333    |
| .8       | .3701268    |
| .5       | .3706004    |
| .2       | .3721780    |
| .1       | .3746903    |
| 0        | .3699623    |

A6. Submitted Predictions 8. We now considered only the data set interaction and explored gbm and Cubist, respectively. For gbm, we set up a tuning grid with an interaction depth of 1, 3, 5; with 100, 300, 500, 700, 900 trees; with a shrinkage parameter of .01, .05, 1; and with a minimum of observations per node of 20, 26, 32, 38, 44. This leads to $3 \times 5 \times 5 = 225$ models. For our Cubist models, we set up a tuning grid of 10, 20, 30, 40, 50, 60, 70, 80, 90, 100 committees and 0, 1, 5, 9 neighbors, leading to 40 models. The best among the gbm models had an RMSE that was slightly lower than the
best among the Cubist models. We opted to make predictions based on that gbm model, with tuning parameters: 100 trees, an interaction depth of 1, a shrinkage parameter of .05, and at least 38 observations per node.

A7. Submitted Predictions 9 and 10. We now took the Interaction data set and added the answers from the interviewer and the mother at age nine. We then fitted different elastic net models. We used cross-validation on a single set of folds and compared α parameters equal to 1, .9, .8, .5, .2, .1, 0. We submitted the model with an α equal to 1 as well as the model with the best RMSE.

A8. Submitted Predictions 11. For our last predictions, we once again compared different gbm and Cubist models. For gbm, we set up a tuning grid with an interaction depth of 1, 3, 5, 7; with 100, 200, 300, 400, 500, 600, 700, 800, 900, 1,000 trees; with a shrinkage parameter of .01, .05, 1; and with a minimum of observations per node of 4, 8, 12, 16, 20, 24, 28, 32. This leads to $4 \times 10 \times 3 \times 8 = 660$ models. For our Cubist models, we set up a tuning grid of 10, 20, 30, 40, 50, 60, 70, 80, 90, 100 committees and 0, 1, 5, 9 neighbors, leading to 40 models.

B. Literature Used

We used the following search terms with Google Scholar in July 2017: fragile families, eviction, and material hardship. The first 25 hits, contained 23 unique hits:

- Kalil, Ariel, and Rebecca M. Ryan. 2010. “Mothers’ Economic Conditions and Sources of Support in Fragile Families.” The Future of Children 20:39–61.
- Heflin, Colleen M., and John Iceland. 2009. “Poverty, Material Hardship, and Depression.” Social Science Quarterly 90:1051–71.
- Reichman, Nancy E., Julien O. Teitler, and Marah A. Curtis. 2005. “TANF Sanctioning and Hardship.” Social Service Review 79:215–36.
- Heflin, Colleen, John Sandberg, and Patrick Rafail. 2009. “The Structure of Material Hardship in US Households: An Examination of the Coherence behind Common Measures of Well-being.” Social Problems 56:746–64.
- Nepomnyaschy, Lenna, and Irwin Garfinkel. 2011. “Fathers Involvement with Their Nonresident Children and Material Hardship.” Social Service Review 85:3–38.
- Desmond, Matthew, and Rachel Tolbert Kimbro. 2015. “Eviction’s Fallout: Housing, Hardship, and Health.” Social Forces 94:295–324.
- Slade, Alexander N. 2013. “The Relationship between Nonresident Father Involvement and Maternal Depression in Fragile Families.” Social Service Review 87:3–39.
- Kang, Jiyoun. 2013. “Instrumental Social Support, Material Hardship, Personal Control and Neglectful Parenting.” Children and Youth Services Review 35:1366–73.
- Warren, Emily J., and Sarah A. Font. 2015. “Housing Insecurity, Maternal Stress, and Child Maltreatment: An Application of the Family Stress Model.” Social Service Review 89:9–39.
- Reichman, Nancy E., Julien O. Teitler, Irwin Garfinkel, and Sandra Garcia 2004. “Variations in Maternal and Child Well-being among Financially Eligible Mothers by TANF Participation Status.” Eastern Economic Journal 30:101–18.
- Turney, Kristin. 2013. “Perceived Instrumental Support and Children’s Health across the Early Life Course.” Social Science & Medicine 95:34–42.
- Neckerman, Kathryn M., Irwin Garfinkel, Julien O. Teitler, Jane Waldfogel, and Christopher Wimer. 2016. “Beyond Income Poverty: Measuring Disadvantage in Terms of Material Hardship and Health.” Academic Pediatrics 16:552–59.
- King, Christian. 2016. “Food Insecurity and Housing Instability in Vulnerable Families.” Review of Economics of the Household.
- Fremstad, Shawn. 2004. “Recent Welfare Reform Research Findings: Implications for TANF Reauthorization and State TANF Policies.” Center on Budget and Policy Priorities.
- Geller, Amanda, and Allyson Walker Franklin. 2014. “Paternal Incarceration and the Housing Security of Urban Mothers.” Journal of Marriage and Family 76:411–27.
- Teitler, Julien O., Nancy E. Reichman, and Lenna Nepomnyaschy. 2002. “A Balancing Act: Sources of Support, Child Care, and Hardship among Unwed Mothers.” Indicators survey center working Paper, no. 0201. Columbia University School of Social Work, New York.
- Donnelly, Louis, Sara McLanahan, and Jeanne Brooks-Gunn. “The Protective Effects of Housing Assistance Programs on Eviction.”
- Desmond, Matthew. 2015. “Fast Focus.”
- Pilkauskas, Natasha V., Colin Campbell, and Christopher Wimer. 2017. “Giving unto Others: Private Financial Transfers and Hardship among Families with Children.” Journal of Marriage and Family 79:705–22.
- Donnelly, Louis, Sara McLanahan, and Jeanne Brooks-Gunn. “The Protective Effects of Housing Assistance Programs on Eviction.”
- Corman, Hope, et al. 2016. “Maternal Depression as a Risk Factor for Children’s Inadequate Housing Conditions.” Social Science & Medicine 149:76–83.
- Wildeman, Christopher. 2014. “Parental Incarceration, Child Homelessness, and the Invisible Consequences of Mass Imprisonment.” The ANNALS of the American Academy of Political and Social Science 651:74–96.
Emory, Allison. 2017. “Explaining the Consequences of Paternal Incarceration for Children’s Behavioral Problems.” No. wp17-01-ff.
Desmond, Matthew, et al. 2013. “Evicting Children.” Social Forces 92:303–27.

Acknowledgments
This paper benefited from the feedback of the editors and three anonymous reviewers. The results in this paper were created with software written in R 3.4.1. (R Core Team, 2017) using the following packages: Caret (Kuhn 2017), Cubist (Kuhn et al. 2016), dplyr (Wickham et al. 2017), factoextra (Kassambara and Mundt 2017), gbm (Ridgeway 2017), Glimnet (Friedman, Hastie, and Tibshirani 2010), and Hmisc (Harrell 2017). Funding for the Fragile Families and ChildWellbeing Study was provided by the Eunice Kennedy Shriver National Institute of Child Health and Human Development through grants R01HD36916, R01HD39135, and R01HD40421 and by a consortium of private foundations, including the Robert Wood Johnson Foundation. Funding for the Fragile Families Challenge was provided by the Russell Sage Foundation. Errors remain our own.

ORCID iD
Louis Raes https://orcid.org/0000-0003-2640-7493

Supplemental Material
Supplemental material for this article is available with the manuscript on the Socius website.

References
Friedman, Jerome H. 1991. “Multivariate Adaptive Regression Splines.” Annals of Statistics 19(1):1–67.
Friedman, Jerome H. 2001. “Greedy Function Approximation: A Gradient Boosting Machine.” Annals of Statistics 29(5): 1189–232.
Friedman, Jerome H. 2002. “Stochastic Gradient Boosting.” Computational Statistics & Data Analysis 38(4):367–78.
Friedman, Jerome, Trevor Hastie, and Rob Tibshirani. 2010. “Regularization Paths for Generalized Linear Models via Coordinate Descent.” Journal of Statistical Software 33(1): 1–22.
Garces, Eliana, Duncan Thomas, and Janet Currie. 2002. “Longer-term Effects of Head Start.” American Economic Review 92(4):999–1012.
Harrell Jr., Frank E. 2017. Hmisc: Harrell Miscellaneous (R package version 4.0-3). Vienna, Austria: R Foundation for Statistical Computing.
Heckman, James J. 2008. “Role of Income and Family Influence on Child Outcomes.” Annals of the New York Academy of Sciences 1136(1):307–23.
Hofferth, Sandra L., and John. F. Sandberg. 2001. “How American Children Spend Their Time.” Journal of Marriage and Family 63(2):295–308.
Kassambara, Aalboukadel, and Fabian Mundt. 2017. factoextra: Extract and Visualize the Results of Multivariate Data Analyses (R package version 1.0.4). Vienna, Austria: R Foundation for Statistical Computing.
Kuhn, Max. 2017. caret: Classification and Regression Training. R package version 6.0-76.
Kuhn, Max, and Kjell Johnson. 2013. Applied Predictive Modeling. Vol. 810. New York: Springer.
Kuhn, Max, Steve Weston, Chris Keefer, and Nathan Coulter. 2016. Cubist: Rule- and Instance-Based Regression Modeling (R package version 0.0.19). Vienna, Austria: R Foundation for Statistical Computing.
Milborrow, Stephen. 2017. earth: Multivariate Adaptive Regression Splines (R package version 4.5.1). Vienna, Austria: R Foundation for Statistical Computing.
Quinlan, J. Ross. 1987. “Simplifying Decision Trees.” International Journal of Man-Machine Studies 27(3):221–34.
Quinlan, J. Ross. 1992. “Learning with Continuous Classes.” 5th Australian Joint Conference on Artificial Intelligence 92: 343–48.
Quinlan, J. Ross. 1993. “Combining Instance-based and Model-based Learning.” Pp. 236–43 in Proceedings of the Tenth International Conference on Machine Learning. Retrieved January 10, 2019 (http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.34.6358&rep=rep1&type=pdf).
R Core Team. 2017. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.
Ridgeway, Greg. 2017. gbm: Generalized Boosted Regression Models (R package version 2.1.3). Vienna, Austria: R Foundation for Statistical Computing.
Salganik, Matthew J., Ian Lundberg, Alexander T. Kindel, and Sara McLanahan. 2019. “Introduction to the Special Collection on the Fragile Families Challenge.” Socius 5. doi:10.1177/237802319871580.
Wickham, Hadley, Romain Francois, Lionel Henry, and Kirill Miller. 2017. dplyr: A Grammar of Data Manipulation (R package version 0.7.4). Vienna, Austria: R Foundation for Statistical Computing.
Zou, Hui, and Trevor Hastie. 2005. “Regularization and Variable Selection via the Elastic Net.” Journal of the Royal Statistical Society: Series B (Statistical Methodology) 67(2):301–20.

Author Biography
Louis Raes is an assistant professor of economics at Tilburg University (the Netherlands).