UCLA

Title

Predicting Layoff among Fragile Families

Permalink

https://escholarship.org/uc/item/9vg1k6d7

Authors

Ahearn, Caitlin E
Brand, Jennie E

Publication Date

2019

DOI

10.1177/2378023118809757

Peer reviewed

Predicting Layoff among Fragile Families

Caitlin E. Ahearn[1], **Jennie E. Brand**[1]

[1]University of California–Los Angeles, Los Angeles, CA, USA

Abstract

The loss of a job is the loss of a major social and economic role and is associated with long-term negative economic and psychological consequences for workers and families. Modeling the causal effects of a social process like layoff with observational data depends crucially on the degree to which the model accounts for the characteristics that predict loss. We report analyses predicting layoff in the Fragile Families data as part of the Fragile Families Challenge. Our model, grounded in empirical social science research on layoff, did not perform substantially worse than the best-performing model using data science techniques. This result is not fully unforeseen, given that layoff functions as a relatively exogenous shock. Future work using the results of the Challenge should attend to whether small improvements in prediction models, like those we observe across models of layoff, nevertheless significantly increase the validity of subsequent models for causal inference.

Keywords

job loss; Fragile Families Challenge; propensity scores; disadvantaged mothers

Introduction

Job loss is a disruptive event to individuals and families with far-reaching implications. Sociologists and economists have persuasively shown that displaced workers experience lengthy periods of unemployment, earnings losses, job quality declines, higher levels of depressive symptoms, social withdrawal, and poorer physical health (see Brand 2015 for a review of the literature). Studies have also documented a notable intergenerational effect; that is, parental job loss affects the well-being and attainment of children (Johnson, Kalil, and Dunifon 2012; Kalil and Ziol-Guest 2005, 2008; Peter 2016; Stevens and Schaller 2011). Brand and Simon Thomas (2014) show that children of single mothers, who may be particularly vulnerable to economic shocks, experience lower social-psychological well-being and educational attainment in young adulthood after a parental job loss.

Corresponding Author: Caitlin E. Ahearn, Department of Sociology, University of California–Los Angeles, 264 Haines Hall, Los Angeles, CA 90095-1551, USA. cahearn3@ucla.edu.

With the Brand and Simon Thomas (2014) analysis as motivation, we joined the Fragile Families Challenge (hereafter referred to as the Challenge) to consider alternative models predicting layoff among disadvantaged primary caregivers and to ultimately improve causal estimates of the effects of parental layoff on children. The Challenge asked scholars from social and data science to develop models predicting six outcomes using the Fragile Families data: layoff of the child's primary caregiver (the focus of this article), family material hardship, eviction, job training of the primary caregiver, academic achievement of the child, and grit of the child. Participants of the Challenge used information from the first five waves of the Fragile Families data (through age 9 of the child) to predict outcomes in the sixth wave (age 15). Participants developed models using training data, which accounted for one half of the total data in wave 6, and predictions were scored based on two sets of holdout data. The Challenge was designed to improve our knowledge of how techniques from social and data science can be used collectively to understand social processes and to inform policy to help mitigate some of the disadvantages facing low-income children. For more information about the details of the Challenge, see the Introduction to this special collection (Salganik et al. 2019).

In this article, we describe models that draw on social science research to predict layoff of the primary caregiver. We find that these simple models performed only moderately worse than the best-performing submissions that use more complex data science techniques. We emphasize that no model performed particularly well. As a relatively less self-selective social process among those considered in the Challenge, layoff presents some distinct lessons concerning our ability to generate predictive models. The article begins with a discussion of three key types of heterogeneity in social processes, focusing on that which is most relevant to the Challenge: pretreatment heterogeneity, that is, individual characteristics that vary systematically with the likelihood of experiencing a treatment.[1] We then describe our analytic approach and present our results from the leaderboard and the final holdout sample. We end with a discussion of our ability to predict an event like layoff and the generalizability and utility of models predicting layoff from the Challenge for future research.

Heterogeneity and Layoff

Three types of heterogeneity warrant discussion in considering a model of layoff and the degree to which we can infer causality in the effects of layoff on subsequent outcomes. First, and most relevant to the Challenge, is *pretreatment heterogeneity*. Modeling the causal effects of layoff with observational data, as opposed to randomized controlled trials, depends crucially on the degree to which the model accounts for these pretreatment characteristics. Research on layoff suggests that it functions as a relatively exogenous, or random, shock to individuals (Brand 2015). That is, while many social processes involve a high degree of self-selection (e.g., divorce or migration), layoff is typically the result of macroeconomic and business conditions beyond the control of individual workers. As a widespread event affecting workers with varied sociodemographic characteristics, attempts by social scientists

---

[1]We adopt the conventions of the potential outcome approach to causal inference and thus refer to layoff as a "treatment." Layoff is of course not randomly assigned and is not analogous to a treatment in a randomized trial.

to model layoff have produced generally weak predictions with low levels of variation explained (see Fallick 1996 for a review). It is thus unsurprising that different strategies for adjusting for pretreatment characteristics have produced substantively similar results (Brand 2006; Coelli 2011; Stevens and Schaller 2011). If layoff involved more self-selection, then strategies that differ in the degree to which they adjust for pretreatment characteristics used to predict layoff may yield more meaningfully different effect estimates. Outcomes in the Challenge like material hardship and academic achievement, by contrast, are social processes known to be strongly associated with a range of sociodemographic characteristics (Edin 2015; Reardon 2011; Sirin 2005).

Two other types of heterogeneity deserve mention: treatment heterogeneity and treatment effect heterogeneity. *Treatment heterogeneity* refers to variation in the key treatment in a causal process, in this case layoff. In prior research, respondents are typically asked open-ended questions pertaining to job separation, which then are categorized as layoff, downsizing, and plant closing. Researchers generally combine these categories to form an indicator of job loss. The Challenge focuses on self-reported layoff of the primary caregiver, which may differ from the combined measure of job loss characteristic of prior studies. *Treatment effect heterogeneity* indicates the degree to which the impact of layoff varies by population characteristics and contextual factors. For example, workers who are accustomed to socioeconomic stability and have few peers undergoing economic distress may be more psychologically harmed by a layoff event (Brand 2015). The submissions and results of the Challenge deal with modeling pretreatment heterogeneity, but these two additional types of heterogeneity have important implications for the generalizability of the Challenge and future work using its results. We return to the discussion of these implications in the Conclusion.

Modeling Approach

To construct a model predicting layoff, we draw on prior social science research that has demonstrated some association between job loss and sociodemographic, employment, family, psychosocial, and family background characteristics (Brand and Simon Thomas 2014).[2] To improve prediction of layoff, we use the most proximate information on the primary caregiver's education and employment characteristics from the wave 5 interviews. Table 1 provides descriptive statistics of our selected variables. Sociodemographic and socioeconomic variables include race, sex, citizenship status, economic hardship, and education of the primary caregiver. Family variables include primary caregiver's age, household size, and marriage/cohabitation status. We include family background of the primary caregiver using highest education of either parent (grandparent of the child in the Fragile Families data) and whether the caregiver reported having lived with both parents when she or he was age 15. We also include key employment variables. Workers in the manufacturing sector and other blue-collar industries are generally more likely to experience layoff (Brand 2015; Farber 2005). Other employment measures include employment status, self-employment status, full-time work, and union membership. Finally, models account for

[2]These variables have been theorized and shown empirically to have some impact on the likelihood of layoff. The decisions, however, governing the precise specification of models predicting layoff are seldom well articulated.

the possibility that individuals who exhibit poor mental health and delinquent behaviors are at higher risk of layoff (Brand and Simon Thomas 2014). We include variables indicating a major depressive episode, drinking, and impulsive or delinquent behavior.[3]

We estimate logistic regression models to predict layoff of the primary caregiver; results are reported in odds ratios. Our submissions included a sequence of attempts to improve our predictions of layoff both by tailoring the models to the specific population in the Fragile Families data and by developing more parsimonious models to avoid overfitting. We present three of these specifications in Table 2. Model 1 was our first Challenge submission. It includes a set of covariates we hypothesized captured pretreatment heterogeneity in layoff based on prior literature. As 90 percent of the primary caregivers are mothers, a model used to predict job loss among single mothers in Brand and Simon Thomas (2014) guided our specification in model 2. Models 1 and 2 are quite similar, however, and include most of the same variables with slightly different specifications. We then specified a third, more parsimonious model. The motivation underlying this final model was to determine whether a more parsimonious model performed similarly on the leaderboard. Missing data presented an analytic challenge. We tested models that did and did not use imputation to handle missing data and tested various imputation strategies. The appendix provides more information about missing data and our imputation strategy. Below we discuss our interpretation of these models, their success on the leaderboard and in the final holdout samples, and how results changed with missing data imputation.

Results

Figure 1 presents our results from the leaderboard and final holdout samples. Success in the Challenge is indicated by mean-squared errors, which measure the accuracy of predicted probabilities of submitted models in the training and final holdout data. Despite our varying specifications and imputation approaches, we found minimal variation in success in predicting layoff in the Challenge. Our first model produced the best (lowest) scores in both the leaderboard and the final holdout samples. Using multiple imputation did not improve our score on the leaderboard and only slightly improved our score in the final holdout sample. Model 2 produced our worst results in the leaderboard and final holdout samples. Model 3 produced midrange leaderboard and final holdout scores.

We were surprised to find that our first model was most successful in the Challenge. Model 1, although grounded in prior work on job loss, less closely followed research on layoff among disadvantaged mothers. However, we emphasize that the margin of improved performance between this and our other specifications was quite small, as shown in Figure 1. Notably, there was little variation across our scores. Moreover, we observe little variation between our models and the best prediction. The mean-squared error of our best model (.164) is only about 1 percent worse than the mean-squared error produced by the winning model in the Challenge (.162). Likewise, the winning model scored only about 3 percent

---

[3]Our models do not include two potentially key predictors of job loss identified by prior research: tenure of current employment and mental ability. Tenure of employment is not available in the data set. Although Fragile Families administered a verbal aptitude test, almost 40 percent of the primary caregivers were missing information on this variable (i.e., more than any other variable for which we chose to impute missing data). We tested models that included this variable, but those models did not improve fit.

better than the baseline (.167), which predicted layoff in the leaderboard and final holdout samples using the mean of layoff in the training sample.

The difference between the scores from the best model and those from the baseline is smaller than for the other outcomes, as shown in Figure 2. While layoff has a difference of only about 3 percent between the baseline and the best model, material hardship has a difference of about 30 percent, and GPA has a difference of about 24 percent. The winning model predicting eviction, for which individual characteristics also have limited predictive power (Desmond 2016), performed 6 percent better than the baseline. Although the best submissions did improve the prediction of layoff, the relatively limited predictive power of even the top scores likely reflects the relatively exogenous process governing the risk of layoff. This result aligns with our theoretical prior.

We present our regression results for models 1 through 3 in Table 2. A few of our regression results are worth noting. As expected, conditional on all other covariates, respondents who are black, unemployed, and blue-collar or manufacturing workers were more likely to experience layoff six years later. Education is not associated with layoff in this sample after adjusting for other primary caregiver characteristics. We tested many measures of education and consistently found this result. This is not unexpected given the lower educational attainment, and thus less variation, among the disadvantaged Fragile Families sample. It is, however, a departure from prior findings (Brand and Simon Thomas 2014). In addition, citizenship was one of the most consistently statistically significant predictors of layoff across the models we tested, indicating the precarious position of noncitizen women in the urban labor market.

Conclusion

The Challenge offers insights into the utility of mass scientific collaboration for understanding social processes and ultimately improving social science research. We focus on a particular social process, parental layoff, that has been used to study the causal impact of socioeconomic mobility on a range of life outcomes. Our best-performing model of parental layoff, grounded in empirical social science research on job loss, did not produce substantially more error than the best-performing model using data science techniques. Indeed, we find notably small differences across the results of our own models and between ours and better-scoring submissions to the Challenge. Given prior work emphasizing the relatively less self-selective processes governing job loss, these findings are not unexpected. They also do not speak to the utility of data science techniques and collaborative models of more self-selective social processes.

We caution that results predicting layoff in the Challenge, and future work using these models to estimate the effects of layoff, pertain to a particular type of job loss and a particular population. First, the effects of job loss vary according to how we characterize loss (i.e., treatment heterogeneity). Researchers typically include layoff, downsizing, and plant closing in studies of job loss, while the Challenge used only a measure of whether the primary caregiver experienced a self-reported layoff. Plant closings are a more exogenous form of loss than (more individualized) layoffs and involve different economic and

psychological consequences (Brand 2015). In addition, layoff in the Challenge pertains to disadvantaged primary caregivers, 90 percent of whom are mothers and almost half of whom are single parents. Results estimating effects of layoff on this group of individuals will likely differ from the results pertaining to the full population of workers experiencing a layoff (i.e., treatment effect heterogeneity). Brand and Simon Thomas (2014) show that children of the most disadvantaged single mothers experienced smaller reductions in social-psychological well-being and educational attainment. The effects of experiencing a layoff in this sample on subsequent outcomes will thus be influenced by both treatment and treatment effect heterogeneity and will not necessarily provide insight into the processes governing job loss among the majority of workers who experience loss.

Despite the limitations to drawing broader conclusions, future work on job loss will benefit from the knowledge advanced in this article and the Challenge more broadly. As data scientists develop increasingly powerful techniques, it will be important to marry those with social science knowledge about the particulars of social processes under consideration. Future research also should attend to whether small gains in predictive performance, like those we observe across models of layoff in the Challenge, significantly reduce selection bias and improve estimates of causal effects on subsequent outcomes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Appendix:: Missing Data

We imputed missing data for several variables from the Fragile Families data. First, layoff is measured at the level of primary caregiver, but our covariates predicting layoff come from individual surveys—either the biological mother, the biological father, or a nonparental primary caregiver. The primary caregiver's status (mother, father, nonparent) was asked in wave 5, for which 14 percent of the sample is missing. For these cases, we could not match covariates with the appropriate survey response and could not impute from prior waves. More than 90 percent of the primary caregivers of known status are mothers. We took advantage of the high likelihood that primary caregivers are mothers and used mothers' information for missing primary caregivers. In other words, we assumed that the missing primary caregivers were the biological mothers of the children and used the mother's data to

predict primary caregiver layoff in wave 6. We did this for all models reported in this article. We suspect this did not significantly affect our results because a case missing primary caregiver status in wave 5 was also likely to be missing in wave 6. Second, some data were missing in the manner that is typical of survey analysis. To handle these missing data, we used two different types of imputation: multiple imputation and single imputation. We first tested multiple imputation with model 1. Using Stata's mi suite of commands, we used chained imputation with 10 imputations to impute variables that had 3 percent or more missing (see Table 1). We then tested whether results differed when we used a simpler, single imputation process. Results were virtually identical using both types of imputation, and the single imputation process is less computationally intensive. We therefore used the single imputation process in models 2 and 3. In sum, all models presented in Table 2 replace missing primary caregiver data with the biological mothers' data. The imputation version of model 1 uses Stata's mi suite of commands to conduct multiple imputation, and the imputation versions of models 2 and 3 use Stata's single imputation command to impute missing data. Our imputed models nevertheless resulted in some missing predictions because we imputed only variables with at least 3 percent missing. After the submission window had closed for the Challenge, we reran these models using fully imputed predictors and found no substantive differences in our results. Finally, the Fragile Families Challenge required submissions to include predictions for all observations, and to comply with this requirement, we replaced remaining missing predicted probabilities with the mean of the layoff outcome variable in the training data (about 0.21).

Author Biographies

Caitlin E. Ahearn is a doctoral student of sociology at the University of California–Los Angeles. She is broadly interested in causal analysis and social stratification and mobility. Her current research interests are in inequality in higher education and differential outcomes to college attendance and graduation, particularly different degrees and college levels. She also has studied vocational education and is currently working on a project related to career development in high schools.

Jennie E. Brand is a professor of sociology and statistics at the University of California–Los Angeles (UCLA). She is also a codirector of the Center for Social Statistics and the associate director of the California Center for Population Research at UCLA. Her research centers on inequality and its implications for various outcomes that indicate life chances. Current research projects include evaluation of heterogeneity in the effects of education on socioeconomic and civic outcomes as well as the socioeconomic and social-psychological consequences of disruptive events such as job displacement and parental divorce. Her research also focuses on causal inference and the application and innovation of statistical models for panel data.

References

Brand Jennie E.2006. "The Effects of Job Displacement on Job Quality: Findings from the Wisconsin Longitudinal Study." Research in Social Stratification and Mobility 24(3):275–98.

Brand Jennie E.2015. "The Far-reaching Impact of Job Loss and Unemployment." Annual Review of Sociology 41:359–75.

Brand Jennie E., and Juli Simon Thomas. 2014. "Job Displacement among Single Mothers: Effects on Children's Outcomes in Young Adulthood." American Journal of Sociology 119(4):955–1001.

Coelli Micheal B.2011. "Parental Job Loss and the Education Enrollment of Youth." Labour Economics 18:25–35.

Desmond Matthew. 2016. Evicted: Poverty and Profit in the American City. New York: Crown.

Edin Kathryn. 2015. $2.00 a Day. New York: First Mariner Books.

Fallick Bruce C.1996. "A Review of the Recent Empirical Literature on Displaced Workers." Industrial and Labor Relations Review 50(1):5–16.

Farber Henry S.2005. "What Do We Know about Job Loss in the United States? Evidence from the Displaced Workers Survey, 1984–2004." Economic Perspectives 2Q:13–28.

Jann Ben. 2005. "Making Regression Tables from Stored Estimates." The Stata Journal 5(3):288–308.

Jann Ben. 2007. "Making Regression Tables Simplified." The Stata Journal 7(2):227–44.

Johnson Rucker C., Kalil Ariel, and Dunifon Racehl E.. 2012. "Employment Patterns of Less-skilled Workers: Links to children's Behavior and Academic Progress." Demography 49:747–72. [PubMed: 22246798]

Kalil Ariel, and Ziol-Guest Kathleen M.. 2005. "Single Mothers' Employment Dynamics and Adolescent Well-being." Child Development 76(10):196–211. [PubMed: 15693767]

Kalil Ariel, and Ziol-Guest Kathleen M.. 2008. "Parental Employment Circumstances and Children's Academic Progress." Social Science Research 37:500–15.

Peter Frauke. 2016. "The Effect of Involuntary Maternal Job Loss on Children's Behaviour and Non-Cognitive Skills." Labour Economics 42:43–63.

Reardon Sean. 2011. "The Widening Academic Achievement Gap between the Rich and the Poor: New Evidence and Possible Explanations." Pp. 91–116 in Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances, edited by Duncan GJ and Murnane RJ. New York: Russell Sage.

Salganik Matthew J., Lundberg Ian, Kindel Alexander T., and McLanahan Sara. 2019. "Introduction to the Special Collection on the Fragile Families Challenge." Socius 5. doi:10.1177/2378023119871580.

Sirin Selcuk R.2005. "Socioeconomic Status and Academic Achievement: A Meta-analytic Review of Research." Review of Educational Research 75(3):417–53.

Stevens Ann Huff, and Schaller Jessamyn . 2011. "Short-run Effects of Parental Job Loss on Children's Academic Achievement." Economics of Education Review 30(2):289–99.

Wolfe Fred. 2002. "FSUM: Stata Module to Generate and Format Summary Statistics." Statistical Software Components S426501. Boston College Department of Economics, revised May 06, 2014.
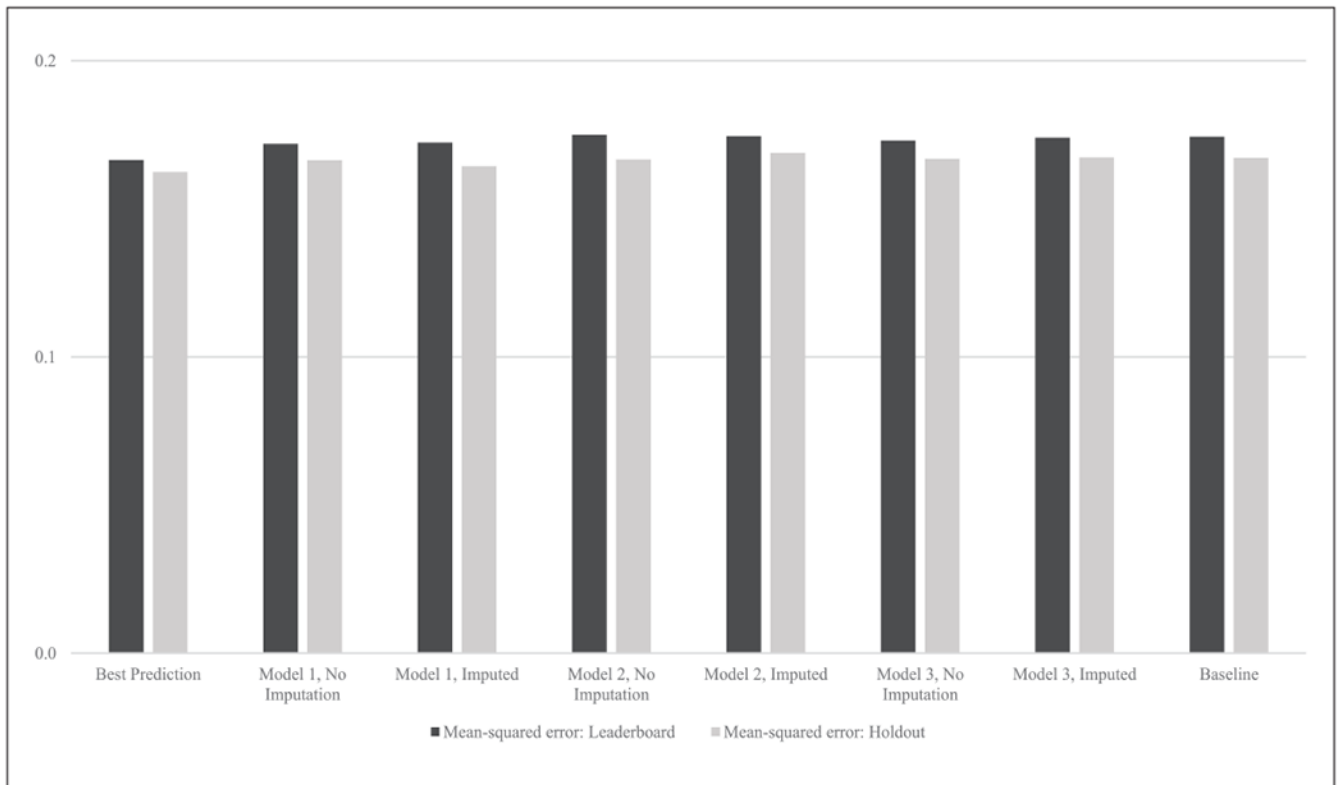
Figure 1.

Mean-squared error from leaderboard and holdout samples, models 1 through 3, with and without imputation, with best prediction and baseline.

Note: Baseline is the mean-squared error produced by using the mean value of layoff in the training sample to predict layoff for all observations in the leaderboard and final holdout sample. Best prediction is the highest-scoring model, or the model that produced the lowest mean-squared error with layoff in the leaderboard and final holdout sample.
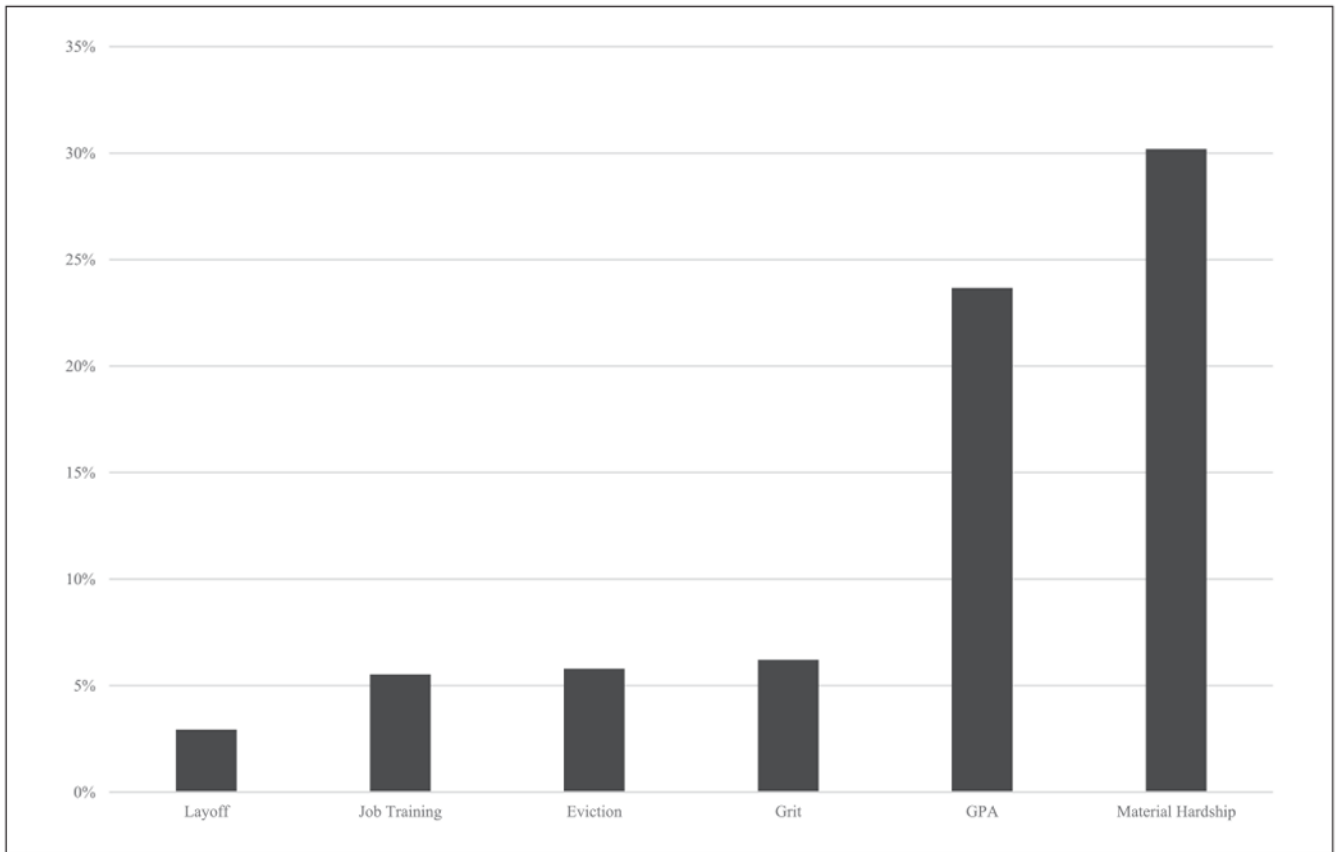
Figure 2.
Difference in predictive power between the baseline and the winning submissions for six outcomes in the Fragile Families Challenge.

Note: Difference in predicted power is the extent to which the best prediction, or the model that produced the mean-squared error for any given outcome, differs from the baseline prediction. The baseline is the mean-squared error produced by using the mean value of the outcome in the training sample to predict layoff for all observations in the leaderboard and final holdout sample.

Table 1.

Descriptive Statistics of Variables Used to Predict Layoff.

Primary Caregiver Characteristic	Mean	Standard Deviation	Minimum	Maximum	% Missing	n
Dependent variable						
Experienced a layoff in last 12 months	0.21		0	1	69.9	1,277
Independent variables					70	
Sociodemographic factors						
Female	0.95	—	0	1	0	4,242
Race/ethnicity						
White	0.22	—	0	1	2	4,165
Black	0.49	—	0	1	2	4,165
Hispanic	0.25	—	0	1	2	4,165
Other race/ethnicity	0.04	—	0	1	2	4,165
U.S. citizenship status	0.84	—	0	1	2	4,160
Socioeconomic factors						
Highest level of education						
Less than high school	0.29	—	0	1	1	4,184
High school graduate	0.29	—	0	1	1	4,184
2-year degree or some college	0.27	—	0	1	1	4,184
4-year degree or graduate degree	0.15	—	0	1	1	4,184
Economic hardship indicator	0.50	—	0	1	8	3,908
Employment factors						
Employed	0.61	—	0	1	5	4,031
Self-employed	0.07	—	0	1	9	3,861
Industry						
Manufacturing industry	0.03	—	0	1	9	3,854
Blue-collar job	0.07	—	0	1	9	3,854
Full-time work	0.70	—	0	1	6	3,970
Union membership status	0.10	—	0	1	12	3,723
Family factors						
Age at child's birth	25.89	6.98	15	68	3	4,117

Primary Caregiver Characteristic	Mean	Standard Deviation	Minimum	Maximum	% Missing	n
Married or cohabiting	0.49	—	0	1	0	4,242
Household size	4.56	1.62	1	13	7	3,935
Number of children in household	2.60	1.32	1	10	7	3,953
Psychosocial factors						
Acute depression	0.15	—	0	1	5	4,033
Highest number of drinks in one day	0.76	0.74	0	3	10	3,818
Delinquency indicator	0.33	—	0	1	18	3,463
Family background factors						
Parents' highest education	2.33	1.01	1	4	14	3,652
Lived with both parents age 15	0.43	—	0	1	5	4,022

Note: Dashes indicate that the standard error is not reported because the variable is dichotomous with a range of 0–1.

Author Manuscript   Author Manuscript   Author Manuscript   Author Manuscript

Table 2.

Logistic Regression Models 1 through 3 of Layoff Experienced in the Past 12 Months.

	Model 1[a]		Model 2[b]		Model 3[b]	
	No Imputation	Imputed	No Imputation	Imputed	No Imputation	Imputed
Sociodemographic factors						
Female	1.494 (0.893)	1.594 (0.720)	1.258 (0.849)	1.325 (0.794)		
White	0.509 *** (0.113)	0.561 *** (0.118)	0.460 *** (0.111)	0.555 *** (0.119)	0.523 *** (0.107)	0.544 *** (0.106)
Hispanic	0.628 ** (0.131)	0.689 * (0.136)	0.716 (0.161)	0.693 * (0.141)	0.673 * (0.138)	0.647 ** (0.125)
Other race	0.296 ** (0.152)	0.299 ** (0.152)	0.268 ** (0.154)	0.282 ** (0.146)	0.338 ** (0.174)	0.274 ** (0.140)
Citizenship status	0.486 *** (0.126)	0.492 *** (0.123)	0.523 ** (0.146)	0.508 *** (0.128)	0.601 ** (0.154)	0.564 ** (0.136)
Socioeconomic factors						
Highest level of education						
At least some college	0.930 (0.150)	0.922 (0.142)				
Less than high school			0.909 (0.196)	0.870 (0.167)		
2-year degree or some college			0.903 (0.192)	0.791 (0.154)		
4-year or graduate degree			0.942 (0.247)	0.859 (0.208)		
Economic hardship indicator[c]	1.153 (0.182)	1.175 (0.178)				
Employment factors						
Employed[c]	0.735 * (0.115)	0.731 ** (0.110)	0.637 *** (0.108)	0.724 ** (0.111)	0.716 ** (0.111)	0.735 ** (0.107)
Self-employed[c]	0.580 (0.208)	0.776 (0.244)				
Industry						
Blue collar or manufacturing[c]	1.520 * (0.371)	1.425 (0.337)				
Blue collar[c]			0.973 (0.335)	1.082 (0.315)		
Manufacturing[c]			1.726 (0.893)	2.590 * (1.052)	2.080 * (0.907)	2.554 ** (0.981)
Full-time employment	0.995 (0.164)	0.965 (0.152)	0.973 (0.172)	0.962 (0.152)		
Union membership status[c]	0.730 (0.180)	0.747 (0.182)				
Family factors						
Age at child's birth (mean centered)	1.000 (0.013)	1.001 (0.011)	0.996 (0.015)	0.990 (0.013)		

	Model 1[a]		Model 2[b]		Model 3[b]	
	No Imputation	Imputed	No Imputation	Imputed	No Imputation	Imputed
Married or cohabiting	0.891 (0.146)	0.833 (0.130)	0.930 (0.158)	0.831 (0.127)		
Household size (mean centered)[c]	0.931 (0.046)	0.933 (0.043)				
Number of children in household[c]			0.912 (0.058)	0.927 (0.052)		
Psychosocial factors						
Acute depression[c]	0.923 (0.195)	0.910 (0.185)	0.870 (0.200)	0.949 (0.192)		
Highest number of drinks[c]	1.107 (0.115)	1.132 (0.114)				
Delinquency[c]			1.261 (0.211)	1.230 (0.196)	1.218 (0.190)	1.226 (0.190)
Family background factors						
Parents' highest education[c]			1.134 (0.099)	1.117 (0.092)		
Lived with both parents at age 15[c]	0.865 (0.139)	0.884 (0.137)	0.764 (0.135)	0.845 (0.137)		
Constant	0.544 (0.377)	0.510 (0.296)	0.738 (0.678)	0.809 (0.664)	0.581[*] (0.169)	0.606[*] (0.167)
Observations	1,192	1,270	1,070	1,246	1,146	1,264

Note: Standard errors are in parentheses. Coefficients are reported in log odds. Both models impute mother's data for cases with a missing primary caregiver.

[a] Imputation was done with mi suite (chained imputation) in Stata.

[b] Imputation was done with "impute," Stata's command for single imputation using regression.

[c] The variable was imputed. Missing indicator variables were included where applicable.

[*] $p < .1$.

[**] $p < .05$.

[***] $p < .01$ (two-tailed tests).