Housing, economic hardship, and maternal depression: a subgroup analysis

Seleeke Flingai[a], Adam Porton[b], Jinyuan Qi[c]

[a]Woodrow Wilson School of Public and International Affairs, Princeton University, Princeton, NJ 08540
[b]The Eviction Lab, Princeton University, Princeton, NJ 08540
[c]Office of Population Research, Princeton University, Princeton, NJ 08540

Significance Statement
Which mothers are most at risk of experiencing adverse mental health outcomes? We examine a data set of 2,103 mothers to identify what characteristics best predict the likelihood that a mother has experienced depression in the previous two years. Important predictors considered include race, age, marital status, intimate partner violence, neighborhood demographic characteristics, economic hardship, parents' mental health history and housing instability. After using an innovative statistical model called LASSOplus, we find that economic hardship and whether the mother's own mother experienced depression are the strongest predictors of depression. These findings may help inform programs designed to screen for mental health difficulties among mothers.

Abstract
Previous literature has established a relationship between socioeconomic status and adverse mental health outcomes. In particular, a 2011 study by Suglia et al. found that economic hardship and housing instability predicts the likelihood of maternal depression. We build on this research by analyzing a group of 2,103 mothers from the Fragile Families and Child Wellbeing Study (FFCWS). We first confirm the findings of Suglia et al with a logistic regression model that shows that a mother who experiences economic hardship/housing instability is indeed more likely to report symptoms of depression net of demographic and certain behavioral characteristics. We next employ a Bayesian model LASSOplus to select the variables *with large effects* in predicting maternal depression and compare the results with LASSO regression. After incorporating dozens of covariates including age, race, marital status, neighborhood characteristics and parental mental history, the LASSOplus model indicates that economic hardship and a history of depression by the mother's own mother are in fact the strongest predictors of the mother's mental health status. These findings suggest that mothers facing economic hardship or those with a family history of maternal depression are most at risk of experiencing depression themselves.

Introduction

Depression is one of the most common forms of mental illness in the United States, with nearly 16 million Americans (approximately 7% of the U.S. population) reporting at least one major depressive episode in the last year (SAMHSA, 2013). Prevalence of depression is higher among women and the poor, and the onset of depression for women peaks during childbearing years (Heflin and Iceland, 2009; Whiteford, Ferrari, and Degenhardt, 2016; Eaton et al., 1997). Furthermore, depression can be transmitted across generations due to genetic factors, environmental stressors, or both (Lau and Eley, 2008; Rice, 2010). For example, children of mothers with a history of depression are more likely to be diagnosed with depression than children of mothers without a diagnosis (Tompson et al., 2010).

Many studies have reported on the association between poor mental health and economic hardship (e.g. Dohrenwend et al., 1992; Mirowsky and Ross, 2001). Economic hardship, which describes the inadequate access to goods and services due to adverse economic conditions, may affect mental health by inducing stress and depressive symptoms, as proposed in the stress process model (Pearlin et al., 2005). Alternatively, the health selection hypothesis posits that mental health issues decrease access to economic opportunities, contributing to economic hardship (Butterworth et al., 2012; Whooley et al., 2002). Prior research has affirmed this bidirectionality (e.g., Williams and Cheadle, 2016), though the associations depend on the mental health and economic measures used and the populations studied (Dohrenwend et al., 1992; Muntaner et al., 2004).

Mental health and housing also share a bidirectional association. Inadequate housing conditions, such as a lack of heat, peeling paint, mold, and structural deterioration have been associated with mental distress and depression (Shenassa et al., 2007; Corman et al., 2016). Housing instability, marked by frequent moves or the inability to make rent payments, has also been associated with poor mental health (Magdol, 2002; Davey-Rothwell, German, and Latkin, 2008; Suglia, Duarte, and Sandel, 2011). Considering the aforementioned relationship between mental health and economic status, it follows that poor mental health not only follows housing instability, but can also serve as a precursor to housing instability and inadequate housing conditions (Phinney et al., 2007; Corman et al., 2016).

Assessing the strength of the relationships between economic hardship, housing instability, and other variables of interest with maternal depression for different subgroups of a study population is of great interest to social scientists, as the results can lead to more nuanced research takeaways and policy interventions. Yet determining which variables are of most relevance among hundreds of potential options is a non-trivial task. Our study aims to address these issues using various statistical methods.

We use data from the Fragile Families and Child Wellbeing Study (FFCWS) to examine the association between housing instability, economic hardship, and maternal depression using a subgroup analysis approach. Building on previous research, particularly the work of Suglia et al. (2011), we use logistic regression to assess these associations while accounting for housing

conditions and exposure to intimate partner violence (IPV), individual covariates such as age, race, and marital status, data on the mental health and risk behavior of the mother's parents, and neighborhood characteristics. We then employ a powerful sparse modeling approach called LASSOplus, a Bayesian method for variable selection in high dimensional settings (e.g., large number of variables). This method allows the researcher to estimate and select the most relevant variables associated with maternal depression without explicitly subsetting data, which can prove to be a relatively arbitrary process that may reduce statistical power. LASSOplus has several theoretical and practical advantages over other variable selection and subgroup analysis methods (Ratkovic and Tingley, 2017). Our results indicate that experiencing two or more economic hardships and having a mother without a diagnosis of depression are the variables with the strongest relationships to maternal depression, suggesting that mothers with economic hardship or a maternal history of depression may be the subgroups at greatest risk of reporting depression.

Results
Of the sample of 2,103 mothers who completed the year 3 in-home assessment, 55% identified as non-Hispanic Black, 27% as Hispanic, and 19% as non-Hispanic White; 58% obtained at least a high school diploma or equivalent; and 28% were married to the baby's father at the time of the year 3 assessment. Twenty percent of mothers reported experiencing IPV at any time point, while 17% percent of mothers had experienced one economic hardship within the past year and 34% had experienced two or more hardships. Fourteen percent of mothers were living in deteriorated housing, 27% reported housing disarray, and 15% had experienced housing stability in the past two years. Eighteen percent of mothers were classified as having probable depression at year 3.

Logistic regression. We first conducted a series of logistic regressions to estimate the association between maternal depression, housing conditions/instability, economic hardship, and intimate partner violence, adjusting for mother's demographic covariates (data not shown). We found maternal depression to be positively associated with housing instability (OR, 1.4 [95% CI, 1.0, 1.8]), one or more instances of economic hardship (OR, 1.6 [95% CI, 1.1, 2.2] and 2.8 [95% CI, 2.2, 3.7] for one instance or two or more instances, respectively), IPV between 12 and 36 months (OR, 1.6 [95% CI, 1.1, 2.4]), and chronic IPV (OR, 2.0 [95% CI, 1.2, 3.1]), while negatively associated with Hispanic self-identification (OR, 0.5 [95% CI, 0.4, 0.8]). After further adjustment for neighborhood characteristics (percent below poverty level, percent foreign born, and percent black) and mother's nativity, housing instability was still positively associated with probable depression (OR, 1.3 [95% CI, 1.0, 1.8]), but this association was no longer statistically significant when adjusting for the respondent's parents' mental health histories and risk behaviors.

We assessed the effects of each of the selected variables on the predicted probability of probable depression. A respondent's mother's history of depression or father's history of anxiety were both found to have statistically significant effects on probable depression, as was having two or more economic hardships or chronic IPV (Figure 1). These data corroborate previous

analyses on the familiality of depression and mental illness, as well as the relationships between economic hardship, IPV, and mental distress. As shown in Figure 1, most variables including house characteristics and housing instability have big confidence intervals, while only economic hardship, father's anxiety and mother's depression history have non-overlapping confidence intervals across categories to indicate significant differences on predicting maternal mental health.

Relevant variable selection and subgroup analysis using Logistic LASSO regression. While logistic regression provides some insight into the associations between the variables selected for the model and probable depression, the technique becomes less robust when conducting subgroup analysis on datasets with a large number of potential variables. Specifically, subsetting data results in reduced statistical power as the number of variables in the model increases. An alternative approach to variable selection is the LASSO regression, in which redundant or irrelevant covariates with small effects included in a model are down-weighted in favor of relevant variables in the model, highlighting the covariates with larger effects on the outcome variable. The result is a less-arbitrary selection of covariates that may also improve the model's predictive accuracy.

LASSO selected 14 variables when using the lambda that gives the minimum mean binomial deviance (lambda.min), which typically generates an overfitted model (Table 1). Using the lambda that gives the most regularized (i.e. simpler) model with cross-validation error within one standard error of the minimum (lambda.1se), only five variables are returned. Notably, housing instability was selected in the overfit model (lambda.min), but not in the simpler, more stringent model (lambda.1se). The relevant variables from the simpler model mirror that of the logistic regression, with the addition of the respondent's mother's anxiety having a minor relevant effect.

Relevant variable selection using LASSOplus. LASSOplus selected only two variables -- economic hardship (two or more instances) and a respondent's mother having no history of depression -- of the 169 original variables (main terms and interaction terms) included in the model (Figure 2). One contribution of LASSOplus is the uncertainty estimates for the model parameter as approximate confidence intervals in the figure of main effects - zero was not included in the uncertainty intervals of the two parameters. Having either severe economic difficulties or a maternal history of depression are thus the strongest indicators for probable depression in mothers. These two effects had the largest coefficients and smallest p-values in the logistic regression analysis and were strong in the LASSO analyses, suggesting that these results are not likely to be false positives. All interacted variables included in the LASSOplus model had no effect.

Figure 3 plots the posterior distribution of the effects of the two selected variables. While most of the mass of the two distributions are positive or negative, some of the mass is centered at 0 for both distributions, suggesting that some mothers in the distribution may be less affected by economic hardship or maternal depression history than others, albeit this heterogeneity is not strong enough to be selected by LASSOplus.

Discussion

Our findings provide a multifaceted analysis of the relationship between housing, economic hardship, and maternal mental health. Economic hardship and the respondent's mother's depression history were strongly associated with maternal depression, and these results were the only consistently relevant variables across all three statistical methods. These robust results corroborate other studies that have found strong independent associations between each variable and maternal depression (Pearlin et al., 2005; Butterworth et al., 2012; Whooley et al., 2002; Williams and Cheadle, 2016; Tompson et al., 2010).

As our logistic regression analyses increased in model complexity with more additional variables -- particularly the mother's paternal mental health and risk behaviors -- the association between housing instability and maternal depression decreased greatly. One potential explanation is that the depression history of the respondent's mother, amongst other structural factors, may have reduced the economic opportunities afforded to the mother's mother. Other studies have described the intergenerational opportunity gap, in which Black Americans have substantially lower intergenerational income mobility than Whites, Hispanics, and Asians (Chetty et al., 2018). The depression history and associated economic hardship risk of the respondent's mother may thereby limit the economic opportunities of respondent, increasing their susceptibility to housing instability. Therefore, including the respondent's mother's depression history in our models may diminish the independent effect of housing instability on the respondent's own depression. Further studies are needed to corroborate these links.

Our use of LASSO and LASSOplus sought to address the issues of arbitrary variable selection and reductions in statistical power in subgroup analysis with logistic regression. Both techniques attempt to remove irrelevant effects from complex regressions, leading to simpler, more predictive models. LASSOplus in particular has a very low false discovery rate compared to other sparse estimators (Ratkovic and Tingley, 2017). A comparison between LASSOplus, LASSO, adaptive LASSO and horseshoe estimator found that LASSOplus's false positive rate for 2000 samples on 181 effects (similar dimensionality to our study) was almost zero, and the false negative rate was ~2%; LASSO had a higher false positive rate (>5%) and slightly lower false negative rate on the same combination of sample size (N) and number of effects (K).

In the LASSO model, both large effects and small effect are reduced by the same amount $\lambda/(N-1)$, which depends on the tuning parameter λ. With LASSOplus, all small and intermediate effects are reduced to zero using a thresholding rule. Thus, LASSOplus is a more conservative tool to select variables (i.e., identify non-zero effects) by aggressively shrinking the small and medium effects, and LASSO is more likely to detect intermediate effects. Housing conditions and instability, for example, may be a small/intermediate effect that LASSOplus considers irrelevant to maternal depression. However, this does not suggest that housing conditions are unimportant. Instead, the results of this study, and particularly the LASSOplus analyses, suggest that a familial history of depression and current economic hardship have incredibly strong associations with poor maternal mental health. Yet the complementary

techniques we've employed also suggest that housing instability and intimate partner violence may have small or intermediate effects, for which amelioration could lessen the stresses on mothers that increase the probability of depression. As such, policy interventions targeting improving maternal health should consider poverty reduction and poverty prevention strategies, while also providing mental health services for new parents.

Some of the limitations of our study reflect the limitations of the original analysis performed by Suglia et al. For starters, the survey instrument used to assess maternal depression was an interviewer-administered screening, not a medical diagnostic overseen by a physician. Further, the parents' mental health history was obtained retrospectively and multiple imputation would not address the missing data problem if the respondents who were not asked the relevant questions differed from respondents in other cities (that is, they were not missing at random). In addition, there were no variables that assessed housing or neighborhood quality for past residences. Especially for housing unstable families, including information about prior residences could potentially alter the results of the model.

Apart from the extended discussion about LASSOplus above, the limitation of selecting small and intermediate effects may also be an advantage for predicting purposes. However, we also acknowledge the limitation of the sparsereg package not allowing the exploration of the interaction effects between two categorical variables ("treatment covariates"). Finally, the analysis performed here was cross-sectional, and as such we are only able to make predictive findings. Future research examining longitudinal data sets (including Fragile Families) may be able to find causal relationships between variables like economic hardship/family history of depression and a mother's likelihood of depression.

Methods
Study sample. The Fragile Families and Child Wellbeing Study (FFCWS) is a prospective cohort study that follows 4,898 children born in one of 20 large U.S. cities between 1998 and 2000. The FFCWS consists largely of unmarried parents: nonmarital births were oversampled compared to marital births at a ratio of 3 to 1. Interviews with mothers, fathers, and/or primary caregivers took place at birth (baseline) and when the child was one, three, five, nine, and 15 years of age. A range of characteristics, including (but not limited to) demographic, economic, health, housing, and behavioral information was collected at each interview. Furthermore, in-home assessments of child behavior and living environment were also conducted. Details on the FFCWS design have been previously published (Reichman et al., 2001). The analytical sample used in the present study consisted of mothers who had completed the baseline, year 1, and year 3 interviews and year 3 in-home assessment (n = 2,103).

Covariates. Mothers were asked to report sociodemographic factors during baseline and subsequent interviews. The present study adjusts the analyses on mother's self-reported age, race/ethnicity, educational attainment, marital status, and nativity.

To ascertain the relative socioeconomic and demographic composition of a mother's neighborhood at the time of their child's birth, we included census tract-level indicators for the proportion of foreign born residents, proportion of residents who are Hispanic, proportion of residents who are Black, and share of incomes below the federal poverty line.

Housing conditions (deterioration and disarray), housing instability, probable depression, economic hardship, and intimate partner violence were coded under the same processes as described in Suglia, Duarte, and Sandel (2011).

Mothers were asked a series of questions about their biological father's and mother's mental health and risk behavior histories at the year 3 follow-up assessment. Mothers were asked: "Did biological father/mother ever have periods where depressed 2 weeks or more?" "Did biological father/mother have periods of month or more when he was nervous/anxious?" "Did biological father/mother ever have a problem with drinking?" "Did he/she abuse prescription drugs like valium/sleeping pills?" "Did he/she have a problem with illegal drugs?" "Did biological father/mother have problems with drinking or drugs?" and "Did biological father/mother attempt suicide?" A total of 14 questions (7 for each biological parent of the mother) were asked, and each were coded as 0 (No) or 1 (Yes).

Analyses. Logistic regression analyses were conducted to examine the relationship between housing instability, economic hardship, and maternal mental health. LASSO regression was conducted using the R package glmnet. Data were partitioned into training (50%) and test (50%) datasets, and regression models were trained using one of two lambda values: 1) the lambda that gives the minimum mean binomial deviance (lambda.min), which typically generates an overfitted model, and 2) the lambda that gives the most regularized (i.e. simpler) model with cross-validation error within one standard error of the minimum (lambda.1se). The LASSO regression predictive accuracy was found to be 81% when lambda.min was used (accuracy measurements could not be obtained for the model using lambda.1se).

Due to the need for complete data for LASSOplus analyses, we used multiple imputation (via the R package Amelia) to address missing data. Thirty percent of data on the respondent's parental mental health history were missing due to the questions not being asked in two cities. After creating five imputed datasets, we used the R package sparsereg to conduct our LASSOplus analyses and found similar results using Markov Chain Monte Carlo implementation and Expectation Maximization (EM) implementation (the latter of which also found the respondent's father's anxiety history to be a relevant variable).

References

Butterworth, P., Leach, L. S., Pirkis, J., & Kelaher, M. (2012) Poor mental health influences risk and duration of unemployment: a prospective study. *Social psychiatry and psychiatric epidemiology 47*(6):1013-1021.

Chetty, R., Hendren, N., Jones, M. R., & Porter, S. R. (2018) *Race and Economic Opportunity in the United States: An Intergenerational Perspective* (No. w24441). National Bureau of Economic Research.

Corman, H., Curtis, M. A., Noonan, K., & Reichman, N. E. (2016) Maternal depression as a risk factor for children's inadequate housing conditions. *Social Science & Medicine 149*:76-83.

Davey-Rothwell, M. A., German, D., & Latkin, C. A. (2008) Residential Transience and Depression: Does the Relationship Exist for Men and Women? *Journal of Urban Health : Bulletin of the New York Academy of Medicine 85*(5):707–716.

Dohrenwend, B. P., Levav, I., Shrout, P. E., Schwartz, S., Naveh, G., Link, B. G., ... & Stueve, A. (1992) Socioeconomic status and psychiatric disorders: the causation-selection issue. *Science 255*(5047):946-952.

Eaton, W. W., Anthony, J. C., Gallo, J., Cai, G., Tien, A., Romanoski, A., ... & Chen, L. S. (1997) Natural history of Diagnostic Interview Schedule/DSM-IV major depression: The Baltimore epidemiologic catchment area follow-up. *Archives of general psychiatry 54*(11):993-999.

Heflin, C. M., & Iceland, J. (2009) Poverty, material hardship, and depression. *Social science quarterly 90*(5): 1051-1071.

Lau, J. Y., & Eley, T. C. (2008) Disentangling gene-environment correlations and interactions on adolescent depressive symptoms. *Journal of Child Psychology and Psychiatry 49*(2):142-150.

Magdol, L. (2002) Is moving gendered? The effects of residential mobility on the psychological well-being of men and women. *Sex roles 47*(11-12):553-560.

Mirowsky, J., & Ross, C. E. (2001) Age and the effect of economic hardship on depression. *Journal of health and social behavior* 132-150.

Muntaner, C., Eaton, W. W., Miech, R., & O'campo, P. (2004) Socioeconomic position and major mental disorders. *Epidemiologic reviews 26*(1):53-62.

Pearlin, L. I., Schieman, S., Fazio, E. M., & Meersman, S. C. (2005) Stress, health, and the life course: Some conceptual perspectives. *Journal of health and Social Behavior 46*(2):205-219.

Phinney, R., Danziger, S., Pollack, H. A., & Seefeldt, K. (2007) Housing Instability Among Current and Former Welfare Recipients. *American Journal of Public Health 97*(5):832–837.

Ratkovic, M., & Tingley, D. (2017) Sparse estimation and uncertainty with application to subgroup analysis. *Political Analysis 25*(1):1-40.

Reichman, N. E., Teitler, J. O., Garfinkel, I., & McLanahan, S. S. (2001) Fragile families: Sample and design. *Children and Youth Services Review 23*(4-5):303-326.

Rice, F. (2010) Genetics of childhood and adolescent depression: insights into etiological heterogeneity and challenges for future genomic research. *Genome Medicine 2*(9):68.

Shenassa, E. D., Daskalakis, C., Liebhaber, A., Braubach, M., & Brown, M. (2007) Dampness and mold in the home and depression: an examination of mold-related illness and perceived control of one's home as possible depression pathways. *American Journal of Public Health 97*(10):1893-1899.

Suglia, S. F., Duarte, C. S., & Sandel, M. T. (2011) Housing quality, housing instability, and maternal mental health. *Journal of Urban Health 88*(6):1105-1116.

Tompson, M. C., Pierre, C. B., Boger, K. D., McKowen, J. W., Chan, P. T., & Freed, R. D. (2010) Maternal depression, maternal expressed emotion, and youth psychopathology. *Journal of abnormal child psychology 38*(1):105-117.

Whiteford, H., Ferrari, A., & Degenhardt, L. (2016) Global burden of disease studies: implications for mental and substance use disorders. *Health Affairs 35*(6): 1114-1120.

Whooley, M. A., Kiefe, C. I., Chesney, M. A., Markovitz, J. H., Matthews, K., & Hulley, S. B. (2002) Depressive symptoms, unemployment, and loss of income: The CARDIA Study. *Archives of Internal Medicine 162*(22):2614-2620.

Williams, D. T., & Cheadle, J. E. (2016) Economic hardship, parents' depression, and relationship distress among couples with young children. *Society and mental health 6*(2):73-89.
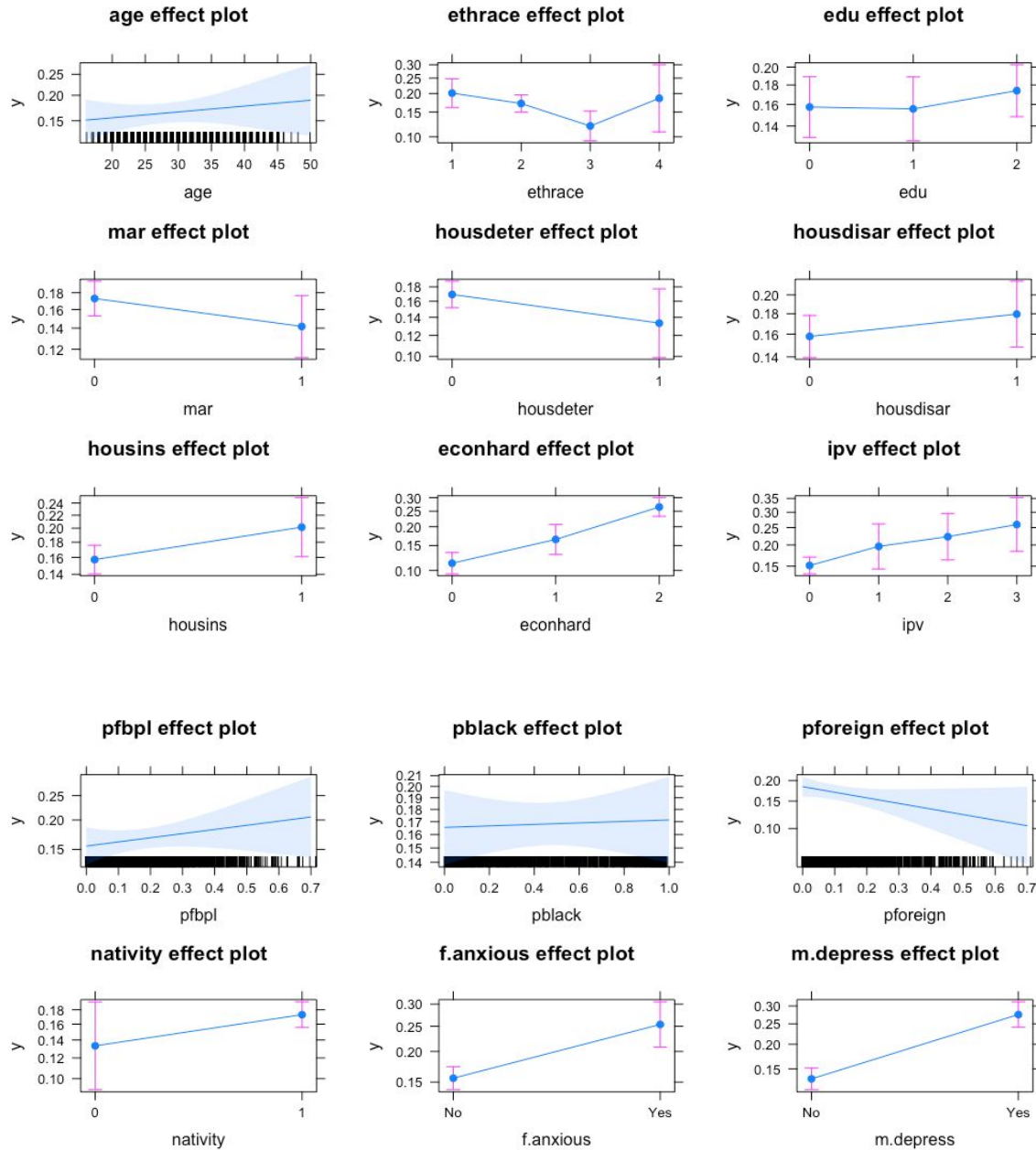
Figures

Figure 1. Logistic regression analysis: Mothers' parental poor mental health, economic hardship, and intimate partner violence associated with probable depression.
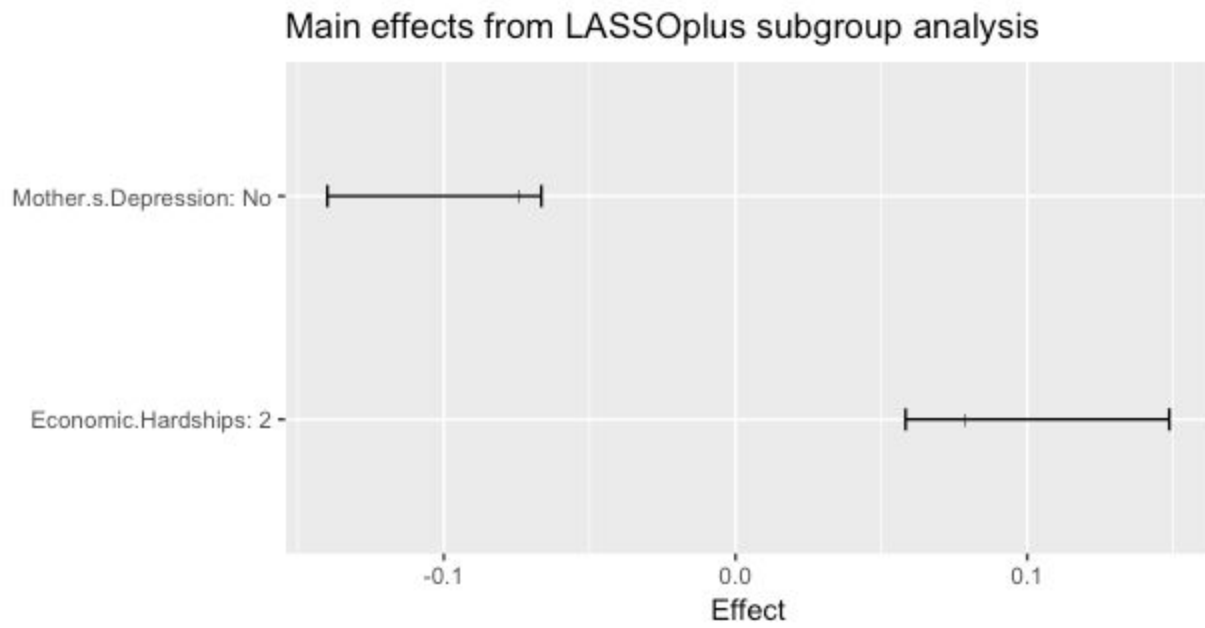
Figure 2. Estimated effects for non-zero coefficients for main effects (two or more instances of economic hardship and respondent's maternal depression history).
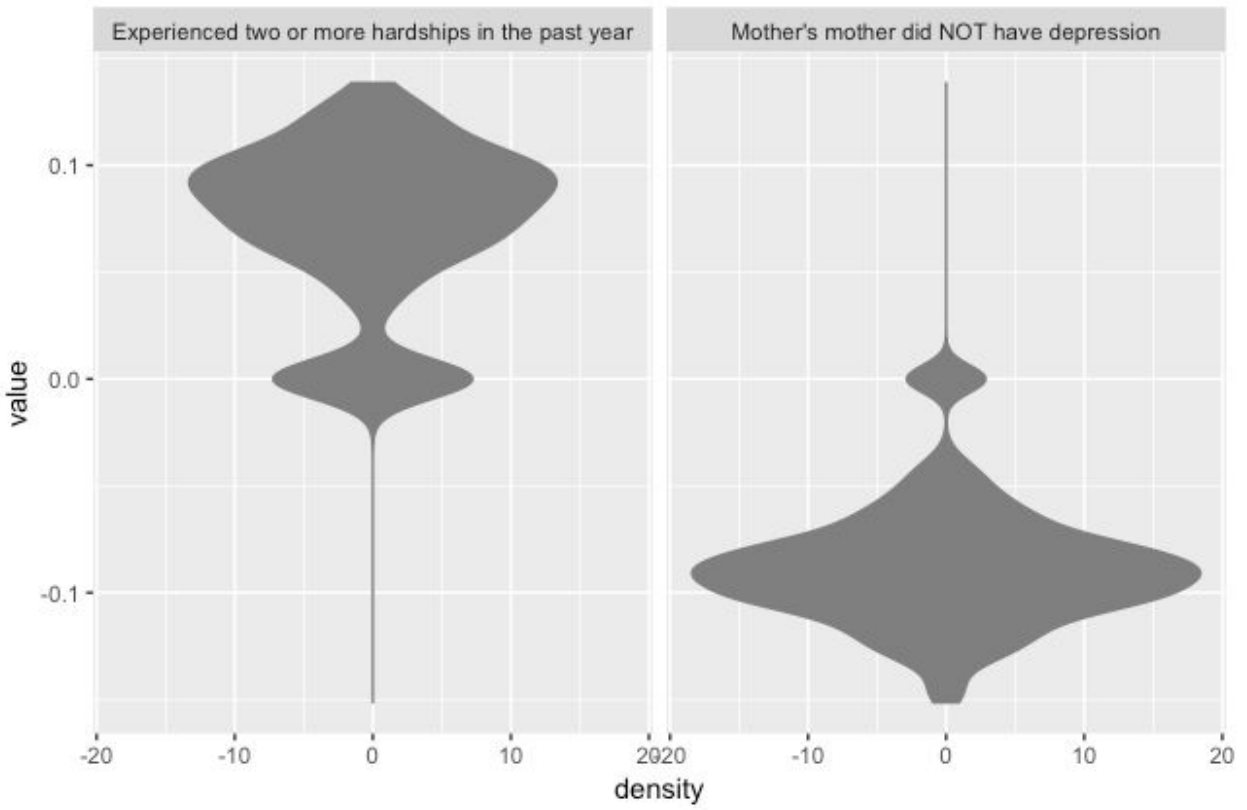
Figure 3. Posterior distribution for economic hardship (two or more instances) and respondent's maternal depression history (not having depression diagnosis).

Tables

Variable	Coefficient (Lambda.min)	Coefficient (Lambda.1se)
Race	-0.02	-
Marital status	-0.15	-
Housing instability	0.21	-
Economic Hardship	0.31	0.19
IPV	0.22	0.08
Percent foreign	-0.23	-
Respondent's father depressed?	0.05	-
Respondent's father anxious?	0.50	0.23
Respondent's father illegal drug use?	0.07	-
Respondent's father drinking/drug problem?	0.08	-
Respondent's mother depressed?	0.34	0.19
Respondent's mother anxious?	0.14	0.05
Respondent's mother drinking problem?	0.26	-
Respondent's mother attempt suicide?	0.07	-

Table 1. LASSO regression analysis: The estimated coefficients in the logistic LASSO regression at two different lambda levels.