Effects of Great Recession on Income Poverty

Subharati Ghosh, Subhrangshu Nandi, Susan Parish

Study Aim 1

Aim 1 of this study was to analyze whether households with working age adults with disability differed from households with no working age adults with disability, during the *great recession* [1], on 'Income Poverty' levels, controlling for demographic factors such as gender, marital status, education, race and ethnicity.

- Hypothesis 1: It was hypothesized that households with working age adults with disability would experience worse income poverty '*levels*' through the great recession, controlling for demographic factors.
- Hypothesis 2: It was hypothesized that households with working age adults with disability would experience worse income poverty '*trends*' during the great recession, controlling for demographic factors.

Study Aim 2

Aim 2 of this study was to analyze differences in income poverty profiles between demographic subgroups within households with working age adults with disability, during the *great recession*.

[1] Business Cycle Dating Committee, National Bureau of Economic Research (NBER)

DATA AND METHOD

DATA

Data was drawn from the Survey of Income and Program Participation (SIPP) 2008 panel. SIPP is administered by the US Census Bureau ([2]), and is representative of non-institutionalized US households. The SIPP 2008 panel started from July 2008 and lasted till June 2013, including a total of 13 waves. The waves overlapped with twelve of the eighteen months [3] of the Great recession and its long wake. Households selected, were followed through the entire panel, and were interviewed very fourth month on a set of core questions, which inquired on household demographics, labor-force participation, participation in the various safety-net programs, asset ownership over the last three months, etc. In addition to the core questions, the SIPP also administered specific modules or topical questions, asked only once during the entire study panel. The topical modules varied by the waves and included questions on marital history, disability, material hardships, assets-liabilities etc. The reference period for the modules varied. Of interest to this study was wave 6 of the 2008 panel, which specifically inquired on adult disability status.

To be included in the study sample, respondents had to meet a set of criteria. First, it was necessary for households to have at least participated in wave six of the study, the wave that had a specific module inquiring on adult disability. Second, the reference persons of the households had to be adults (18 years and older) throughout the household's participation in the study. And third, the households should have participated in at least one calendar year. A total of 33,547 households met the sample inclusion criteria.

MEASURES

Dependent variable

Income poverty, the ratio of average quarterly household income and average quarterly federal poverty level (100% FPL) was used as the dependent variable in our analysis, henceforth referred to as FPL100-ratio. We used quarterly averages because quarters are the most widely accepted time windows when analyzing economic trends. The four quarters that make up the year are: January,

[2]For more information on the SIPP 2008 panel schedule, please refer to this US Census Bureau website
[3]NBER Recession Cycles

February and March (Q1); April, May and June (Q2); July, August and September (Q3); and October, November and December (Q4). To calculate quarterly averages total monthly household incomes and monthly federal poverty levels were averaged [4] over three months in each quarter. An FPL100-ratio lower than one in any quarter indicated the household was below 100% of the federal poverty level in that quarter. In the sample, the quarterly income data ranged from -$27,180 to $108,900, the average being $5240 and median $3,874. The negative incomes were associated with households owning business that incurred losses in those quarters. The FPL100-ratio ranged from -17.95 to 89.48, with the average being 3.817 and the median 2.924.

Key Predictors

There were two key predictors in our analysis: *time* and *adult disability*. Time was measured in quarters and was treated as a continuous variable. Data between 2008-Q3 and 2013-Q1 were analyzed. Wave 6 of the 2008 panel included detailed questions to assess adult disability. Adult Disability was assessed by asking the household reference person, whether there was any adult in the household who had experienced difficulties with activities of daily living, or have been using assistive devices, or had mental retardation, learning disability, developmental disability, or Alzheimer's or any disease that impacted memory, resulting in loss of memory, forgetfulness. Once identified, only those households were selected where at least one person between ages 18 and 64 had a chronic illness/disability of the duration of at least one year, to capture the severity of the disease or illness. A total of 7,443 households (22.16%) met the inclusion criteria. A dichotomous variable indicated whether the household had a working age adult with disability or whether it did not.

Control variables

We controlled for four demographic factors: *gender*, *marital status*, *education*, *race/ethnicity* of the reference persons of the households in our analysis. Gender was a dichotomous variable (male & female), and so was marital status. Household reference persons who were divorced, widowed and never married were categorized as 'not married'. Education had three categories, 'high-school or less', 'some college, diploma, associated degrees' and 'bachelors or higher'. We also included a variable labeled race/ethnicity, which was based on two variables, ethnicity and racial origin. The SIPP assessed ethnicity using a dichotomous variable, which assessed whether the householder was

[4] These averages were weighted by monthly longitudinal survey weights

or was not of Hispanic or Latino origin. Racial origin was assessed by asking respondents to identify themselves as 'White alone', 'Black alone', 'Asian' and 'Others', including Native Hawaiian and Pacific Islanders. Asians and Others were collapsed into one category. Taking the two measures, racial origin and ethnicity, resulted in a measure race/ethnicity, which had four categories 'non-Hispanic White', 'non-Hispanic Black', 'Hispanic' and 'Others' who are henceforth referred to as 'White', 'Black', 'Hispanic' and 'Others'.

ANALYTIC STRATEGY

Descriptive statistics was used to identify the key demographic characteristics of the study sample. Longitudinal household weights provided by SIPP [5] were used for all analytic purposes.

For aim 1, a mixed (fixed and random) effects model was fit to analyze how households with a working age adult with disability differed from households with no working age adult with disability, on Income Poverty, controlling for demographic factors. Since this dataset was longitudinal in nature, to account for 'between household' differences a mixed effect model was used.

Suppose Y_{ij} denote the FPL100-ratio of household i at time period t_j, X_i denote the demographic factors (gender, marital status, education level, race/ethnicity and their interactions) associated with household i. Then, a simple mixed-effects model for the analysis could be written as in equation 1.

$$Y_{ij} = \beta_0 + \beta t_j + X_i\Theta + b_i + \epsilon_{ij} \tag{1}$$

The model in equation 1 would estimate the following parameters: (1) β_0, the overall intercept, (2) β, the fixed effect of time, (3) Θ, the vector of fixed effects of demographic factors, (4) b_i, the household level random effect (random intercept) and (5) ϵ_{ij}, the residual, or the 'within household' variability. In this model, the response from the ith household at time t_j is estimated to differ from the overall mean $\beta_0 + \beta t_j + X_i\Theta$ by a household effect b_i and a within household measurement error ϵ_{ij}. The within household and between household errors are assumed to be normal and independent [6].

The trough of the great recession was reached in the second quarter of 2009 (marking the

[5] https://www.census.gov/programs-surveys/sipp/methodology/weighting.html

[6] $(b_i \sim \mathcal{N}(0, \sigma_b^2), \quad \epsilon_{ij} \sim \mathcal{N}(0, \sigma^2), \quad b_i \perp\!\!\!\perp \epsilon_{ij}, \forall i, j)$

technical end of the recession, defined as at least two consecutive quarters of declining GDP) [7]. According to NBER, June 2009 was the final month of the recession. We checked if this was reflected in the FPL100-ratio as a downward trend in the initial quarters followed by an upward trend. A linear term in 'time' was insufficient to capture this effect. Therefore, we added a second order term time-squared (t^2) to test the change in direction of trend. The second order term was created after centering the original 'time' variable, to avoid introducing multicollinearity. An indicator variable $\mathbb{1}_{D_i}$ [8] was used to denote the presence of working age adult with disability in a household i. An interaction term between $\mathbb{1}_D$ and time was also included to estimate the difference in slopes between households with and without a working age adult with disability. Below is the final model that was fit for aim 1:

$$Y_{ij} = \beta_0 + \beta_1 t_j + \beta_2 t_j^2 + \beta_D \mathbb{1}_{D_i} + \beta_t(\mathbb{1}_{D_i} * t_j) + X_i\Theta + b_i + \epsilon_{ij} \tag{2}$$

The model in equation 2 estimated the following parameters: (1) β_0, the overall intercept, (2) β_1, the fixed effect of time, (3) β_2, the fixed effect of time-squared, (4) β_D, the effect of disability on income poverty (hypothesis 1) (5) β_t, the effect of interaction between disability and time on income poverty (hypothesis 2), (6) Θ, the vector of fixed effects of demographic factors and their interactions, (7) b_i, the household level random effect (random intercept) and (8) ϵ_{ij}, the residual, or the 'within household' variability. Significance of coefficients β_1 and β_2 were tested to analyze the overall trends of FPL100-ratio over the study period. In addition, interactions between demographic factors, and between disability and demographic factors were also tested. The demographic factors and their interactions were all considered as fixed effects.

Preliminary analysis revealed that over time the differences in FPL100-ratio trends were much smaller than the differences in FPL100-ratio levels between the demographic subgroups. Hence, in order to keep the final model simpler, we did not include interaction terms of time with demographic factors. This can be seen in figures 1 and 3.

The final model was selected using 'backward elimination'. Elimination of the fixed effects were done by the principle of marginality, that is: the highest order interactions are tested first: if they are significant, the lower order effects were included in the model without testing for significance. The p-values for the fixed effects are estimated from the F statistics, with 'Satterthwaite' approximation

[7] Business Cycle Dating Committee, National Bureau of Economic Research (NBER)

[8] $\mathbb{1}_{D_i} = 1$, if household i had a working age adult with disability, else $\mathbb{1}_{D_i} = 0$

(Satterthwaite (1946)) denominator degrees of freedom. The p-values for the random effect were computed from likelihood ratio tests (Morrell (1998)).

For aim 2, a separate mixed effects model was was fit, with FPL100-ratio as the dependent variable, time, time-squared and demographic factors as the predictors, only for households that had working age adults with disability. To carefully analyze differences in income poverty profiles between demographic subgroups within these households, interactions between the factors were included in the model. Post-hoc tests were conducted between categories of all demographic factors and their interactions, by calculating differences of 'Least Squares Means' using R package 'lmerTest' (Kuznetsova et al. (2015)), with 'Satterthwaite' approximation (Satterthwaite (1946)) of the denominator degrees of freedom. Since there were multiple categories in these factors the size of the tests could be inflated, hence inflating type I errors. Holm's sequentially rejective *Bonferroni procedure* (Holm (1979)) and *Benjamini-Hochberg procedure* (Benjamini and Hochberg (1995)) remain the two most popular multiple testing correction techniques to address type I error inflation. Holm's sequentially rejective Bonferroni procedure controls the family-wise type-I error rate (FWER) and is more powerful than the classical Bonferroni procedure. Benjamini-Hochberg controls the false discovery rate (FDR) which is the expected value of false discovery proportion. Controlling FWER usually proves to be too conservative. Hence, we used the Benjamini-Hochberg procedure, which is less conservative, but more powerful than Holm's sequentially rejective Bonferroni correction. All post-hoc test p-values reported were Benjamini-Hochberg corrected.

All analysis were conducted using the statistical software R (R Core Team (2016)), version 3.3.1. The mixed effects models were fit using the R-package 'lme4' (Bates et al. (2014)) and all hypothesis tests were done using the R package 'lmerTest' (Kuznetsova et al. (2015)).

RESULTS

Table 1 describes the sample from wave six of the 2008 panel. As seen in table 1, 22.19% of households who participated in wave six had at least one working age adults with disability. In 54.62% of the households with disability, the reference persons were females, 46.67% of the reference persons were 'not married', majority of the reference persons were White (66.71%), 40.23% of the

reference persons had high school or less education and 21.31% had bachelors or higher degrees. In 51.63% of the households with no disability, the reference persons were females, 48.07% of the reference persons were 'not married', majority of the reference persons were White (71.89%), 32.55% of the reference persons had high school or less education and 33.30% had bachelors or higher degrees.

Results for aim 1 of the study are presented in table 2. Households with working age adults with disability, on an average experienced significantly worse income poverty (FPL100-ratio, $\beta_D = -0.725, p < 0.001$) compared with households with no working age adults with disability, through the great recession, when controlled for demographic factors. This average difference between the two types of households is illustrated in figure 1. The findings supported hypothesis 1.

Results showed that the trends of income poverty were not significantly different between households with and without working age adults with disability, over the study period ($\beta_t = 0.015, p = 0.9994$). This is evident from the trendlines in figure 1, which shows that although income poverty worsened in the early stages of the great recession and experienced gradual recovery, the average difference in income poverty profiles between the two household types have remained similar. Hence, we could not substantiate hypothesis 2.

Table 2 also showed that FPL100-ratio, on an average, **decreased** by 0.054 every year ($\beta_1 = -0.054, p < 0.001$). The coefficient of the quadratic term of Time ($\beta_2 = 0.0073, p < 0.01$) indicated that the rate of change of slope was positive. In other words, although the FPL100-ratio decreased over time (as $\beta_1 < 0$), it flattened out and started increasing, in the latter parts of the study period (figure 1).

Of the main effects of demographic factors on the FPL100-ratio, households with 'female' reference persons had on an average 0.368 lower FPL100-ratio ($\beta = -0.368, p < 0.001$) compared to households with 'male' reference persons. Marital status was statistically significant; households with 'not-married' reference persons had on an average 0.611 lower FPL100-ratio ($\beta = -0.611, p < 0.001$) compared to households with 'married' reference persons. The impact of race/ethnicity on FPL100-ratio were statistically significant; households with Black reference persons had on an average 1.28 lower FPL100-ratio ($\beta = -1.284, p < 0.001$) than households with White reference persons, but the difference was greater between households with Hispanic and White reference persons

($\beta = -1.505, p < 0.001$). The impact of education on FPL100-ratio was also statistically significant; households with reference persons with education 'high school or less' had on an average 2.204 ($\beta = -2.204, p < 0.001$) lower FPL100-ratio than households with reference persons with education 'bachelors or higher'. In addition to these main effects, table 2 also shows the significant interactions between disability and demographic factors and among the demographic factors themselves.

Based on the interaction effects reported in table 2, figure 2 illustrates the composite effects of the various interacting factors on income poverty profiles. We present profiles of the two most contrasting household types on a combination of demographic factors that show the best show the best and the worst profiles on FPL100-ratios. We compare households where the reference persons are White, married, male, with education bachelors or higher and with no working age adults with disability (type 1) to households where the reference persons are Black, not married, female with education high school or less, with a working age adult with disability (type 2). As evident, households of type 2 had, on an average, income poverty worse than 200% FPL throughout the study period, starting at 1.52 and reaching lower than 1.3 FPL100-ratio. In comparison, household type 1 always had their FPL100-ratios higher than 6.3.

Aim 2

Aim 2 of this study was to analyze differences in income poverty profiles between demographic subgroups within households with working age adults with disability, during the *great recession*. Results from table 3 show that gender, marital status, education, race/ethnicity and some of their interactions have statistically significant associations with income poverty for households with disability during the great recession. The main effects of the demographic factors are illustrated in figures 3a for gender, 3b for marital status, 3c for race/ethnicity and 3d for education. The difference in associations of marital status with FPL100-ratios between tables 2 and 3 is worth highlighting. The association is almost double in households with disability ($\beta = -0.611$ in table 2 and $\beta = -1.119$ in table 3).

For example, 3a shows that households with 'female' reference persons had consistently worse income poverty compared to households with 'male' reference persons. Figure 3b shows that households with 'not married' reference persons had consistently worse income poverty compared to households with 'married' reference persons. Figure 3c shows that households with 'Hispanic'

reference persons had the worst income poverty throughout the study period. Also illustrated in figure 3c are the different shapes of the income poverty profiles of the four races. Households with 'White' reference persons had a gradual and steady improvement in their average income poverty after 2011. However, this behavior was not observed in households with 'Black', 'Hispanic' or 'others' reference persons. In figure 3d, households where the education levels of their reference persons were 'high school or less' experienced a decline in their FPL100-ratios, just like the other groups, but never experienced any improvement in the latter parts of the study. To conclude, subgroups of households headed by Hispanic, not married females, with education high school or less experienced consistently worst income poverty throughout the great recession.

Limitations

1. Although a linear mixed effects regression model discovered some conventional and some interesting patterns in the relationships between response and demographic factors, along with disability, the trajectory of income poverty over the study period for some households were not linear. This modeling approach did not capture trajectory shapes of individual households. A non-parametric fitting of the income poverty trajectories could be tried as a pre-processing step before testing for differences in behavior between different groups of households.

2. Some households in the sample did not participate over all the waves. Since households that participated in wave six were included there were some households that were first interviewed in wave six and some that were no longer interviewed after wave six. There were no means of determining the reasons for dropping out from the survey, nor the reasons for late inception into the survey. Since the *great recession* was a significant economic and social event, we included households without complete participation in order to maximize the sample size, and incorporate the effect of the recession on more households. If, however, the reasons for dropping out or late joining had an association with the outcome of the study (income poverty), including those households could increase bias in the estimates, in spite of the estimates being more stable (less variance). Chapter 2 in SIPP users guide [9] mentions that

[9] https://www2.census.gov/programs-surveys/sipp/guidance/SIPP_2008_USERS_Guide_Chapter2.pdf

the survey weights are adjusted to account for some types of household non-response with the objective of ameliorating the non-response bias.

1 Figures

Figure 1: Income poverty profiles of households, by disability status

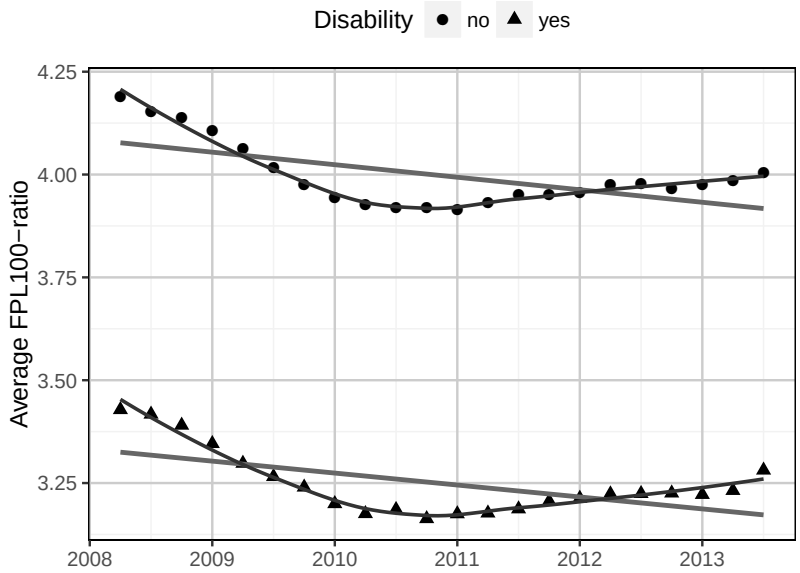
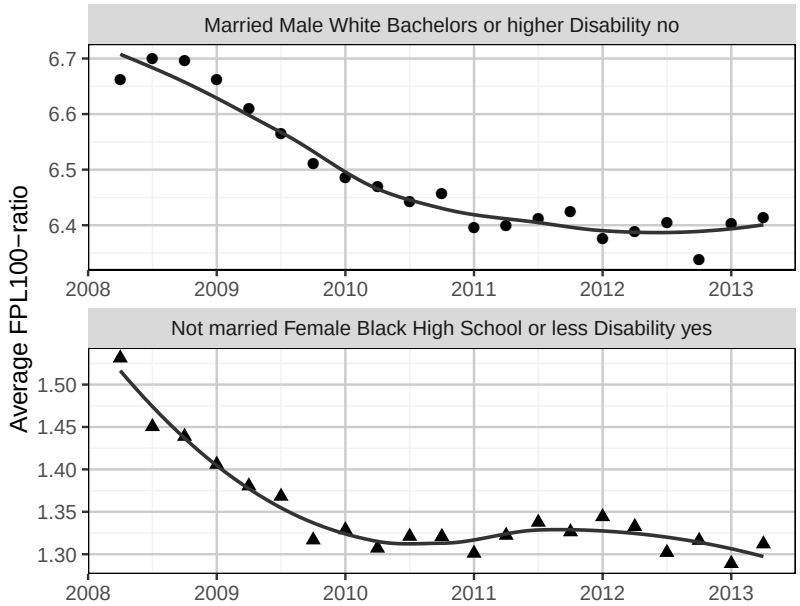
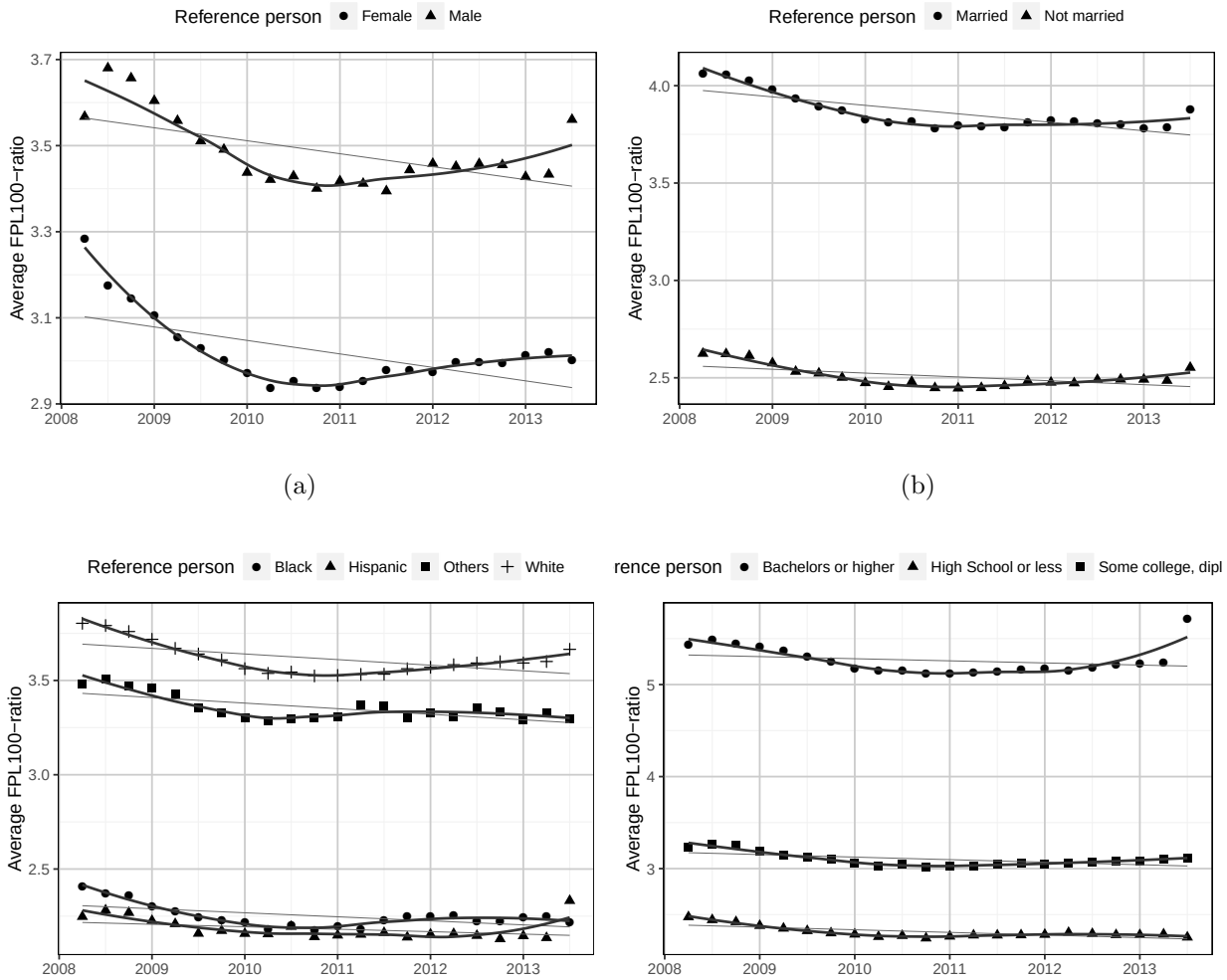


Figure 2: Income poverty profiles of two most contrasting household types



Note: The y-axes of the two subplots have different ranges.

Figure 3: Income poverty profiles for households with working age adults with disability (a) by gender (b) by marital status (c) by race and ethnicity (d) by education, of reference person



(a) (b) (c) (d)

2 Tables

Table 1: Description of the study sample in wave six of 2008 SIPP panel

	HH with No Disability		HH with Disability	
Demographic factors	**Number**	**Percentage**	**Number**	**Percentage**
Total	26,104	77.81	7,443	22.19
Gender:				
Male	12,626	48.37	3,378	45.38
Female	13,478	51.63	4,065	54.62
Marital status:				
Married	13,555	51.93	3,968	53.31
Not Married	12,549	48.07	3,474	46.67
Race and ethnicity:				
White	18,765	71.89	4,965	66.71
Hispanic	2,996	11.48	908	12.20
Black	2,912	11.16	1,092	14.67
Others	1,431	5.48	479	6.44
Education:				
High school or less	8,496	32.55	2,994	40.23
Some college, diploma, assoc	8,916	34.16	2,863	38.47
Bachelors or higher	8,692	33.30	1,586	21.31

Table 2: Mixed effects regression ouput testing the difference in income poverty between families with and without a working age adult with disability, over the *great recession* controlling for demographic factors. Only significant interactions (p-value < 0.05) have been reported.

Predictor and Control variables	β	**Std. Error**	**p-value**
Intercept	5.896	0.037	0.0000
Time (t)	-0.054	0.002	0.0000
Time-squared (t^2)	0.007	0.001	0.0000
Adult Disability	-0.725	0.068	0.0000[5]
Adult Disability x Time	0.015	0.004	0.9994[5]
Gender[1]: (Female)	-0.368	0.043	0.0000
Marital status[2]: (Not married)	-0.611	0.034	0.0000
Race[3]:			
Race2: (Black)	-1.284	0.086	0.0000
Race3: (Hispanic)	-1.505	0.088	0.0000
Race4: (Others)	-0.355	0.087	0.0000
Education[4]:			
Education2: (Some college, diploma, assoc)	-1.516	0.045	0.0000
Education3: (High School or less)	-2.204	0.048	0.0000
Adult Disability x Gender	0.195	0.050	0.0001
Adult Disability x Education:			
Adult Disability x Education2	0.112	0.068	0.1002
Adult Disability x Education3	0.219	0.072	0.0022
Gender x Marital status	-0.523	0.032	0.0000
Gender x Education:			
Gender x Education2	0.102	0.051	0.0448
Gender x Education3	0.255	0.052	0.0000
Marital status x Race:			
Marital status x Race2	0.209	0.051	0.0000
Marital status x Race3	0.529	0.046	0.0000
Marital status x Race4	0.182	0.069	0.0081
Marital status x Education:			
Marital status x Education2	0.125	0.037	0.0008
Marital status x Education3	0.234	0.038	0.0000
Race x Education:			
Race2 x Education2	0.337	0.087	0.0001
Race3 x Education2	0.401	0.092	0.0000
Race4 x Education2	-0.208	0.104	0.0460
Race2 x Education3	0.371	0.093	0.0001
Race3 x Education3	0.499	0.092	0.0000

[1] Base category of gender is "Male" [2] Base category of marital status is "Married"
[3] Base category of race is "White" [4] Base category of education is "Bachelors or higher"
[5] Note that these are p-values of one-sided tests based on hypotheses 1 and 2.

Table 3: Mixed effects regression ouput testing the associations of demographic factors with income poverty, for families with a working age adult with disability over the *great recession*. Only significant interactions (p-value < 0.05) have been reported.

Predictor and Control variables	β	**Std. Error**	**p-value**
Intercept	5.361	0.076	0.0000
Time (t)	-0.039	0.003	0.0000
Time-squared (t^2)	0.006	0.002	0.0127
Gender[1]: (Female)	-0.254	0.088	0.0039
Marital status[2]: (Not married)	-1.119	0.070	0.0000
Race[3]:			
Race2: (Black)	-1.121	0.167	0.0000
Race3: (Hispanic)	-1.295	0.159	0.0000
Race4: (Others)	-0.325	0.162	0.0442
Education[4]:			
Education2: (Some college, diploma, assoc)	-1.668	0.088	0.0000
Education3: (High School or less)	-2.155	0.089	0.0000
Gender x Marital status	-0.371	0.058	0.0000
Gender x Education:			
Gender x Education3	0.223	0.100	0.0263
Marital status x Race:			
Marital status x Race3	0.636	0.082	0.0000
Marital status x Race4	0.288	0.115	0.0124
Marital status x Education:			
Marital status x Education2	0.641	0.074	0.0000
Marital status x Education3	0.759	0.074	0.0000
Race x Education:			
Race2 x Education2	0.350	0.174	0.0438

[1] Base category of gender is "Male" [2] Base category of marital status is "Married"
[3] Base category of race is "White" [4] Base category of education is "Bachelors or higher"

References

Bates, D., M. Maechler, B. Bolker, S. Walker, et al. (2014). lme4: Linear mixed-effects models using eigen and s4. *R package version 1*(7).

Benjamini, Y. and Y. Hochberg (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the royal statistical society. Series B (Methodological)*, 289–300.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian journal of statistics*, 65–70.

Kuznetsova, A., P. B. Brockhoff, and R. H. B. Christensen (2015). Package lmertest. *R package version 2*.

Morrell, C. H. (1998). Likelihood ratio testing of variance components in the linear mixed-effects model using restricted maximum likelihood. *Biometrics*, 1560–1568.

R Core Team (2016). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Satterthwaite, F. E. (1946). An approximate distribution of estimates of variance components. *Biometrics bulletin 2*(6), 110–114.