**Unpaid Labor: How the habit of Home Cooked Meals could lower Women’s Earnings.**

*Ethan Norton*

*Economics Undergraduate Program*

*Valparaiso University*

*Abstract*

*Using data on households/individuals in America from the 2011 Panel Study of Income Dynamics(PSID), this article focuses on the motherhood penalty and specifically how unpaid work labor for females influences their pay. While previous studies have looked at this variable as well as using Ordinary Least Squares(OLS), there are no major articles that articulate it using slope dummy format. This article articulates the impact unpaid labor hours has on the earnings of a female. The results found demonstrate that the combination of high unpaid labor hours for females as well as the amount of meals eaten together as a family weekly. Also, the study will show that women are more heavily affected by meals prepared than men, through the use of a slope dummy. This study will study the relationship between frequency of unpaid housework hours and women and men’s wages separately.*

***Keywords:*** *unpaid-labor, wages, gender.*

**Introduction**

It is considered a social norm for a wife to perform more household chores than her husband. Society portrays women as being the cleaner, chef, and primary caretaker of the home. The wife puts a lot of effort into rearing a child and labor time, and it has a negative impact on her earnings and overall well-being. **Is there a relationship between the frequency of meals eaten together as a family and a difference in women’s and men’s wages separately?** This is a rational question and the findings will be explored throughout this paper. Throughout this study, the goal is to focus on the variable “MealW” and the theory behind how this variable influences the two parents.

This study focuses on the economic impacts of unpaid labor activities for females in the labor force, and primarily is structured around unpaid labor hours. There is a plethora of evidence that suggests an increase in unpaid labor leads to decrease earnings for women (Schneider 2011; Craig and Sawrikar 2009; Becker 2010; Raley, Bianchi, Wang 2010; Staff and Jeylan 2011). One study found that even men and women make the same amount of money(or the women make more), the women still performs more house duties(Killewald 2011). One of the goals of this study is to shine light on how the difference in unpaid labor hours is reflected on men's and women’s earnings separately.

This analysis contributes to the discussion of the relationship of the motherhood penalty reflected on earnings in multiple ways. First of all, the study targets the housework hours of the female relative to the housework hours of the male. The variable looked at here has not specifically been looked at this way in the context of other research. Also, it targets a specific variable of unpaid labor hours- the amount of times the family has a meal together during a given week. The range from this is zero times to seven times(every day). The association with amount of meals eaten together as a family and earnings is analyzed through: MealW, a slope dummy, of female and the amount of meals the family eats together weekly.

In the subsequent section, the current literature related to experience and earnings, and religion and earnings will briefly be mentioned, while primarily discussing the relationship of unpaid work hours to female earnings. Following, theoretical framework and expectations are developed, trailed by description of the data and an explanation of the empirical methodology and findings. Finally, the article concludes with closing remarks and potential policy solutions.

**Literature Review**

The motherhood penalty impacts women from different angles. Most research on the motherhood penalty has focused on short-term wage penalties among women who are still raising relatively young children, typically when they are in their 20s and 30s(Kahn 2014). This can also continue on throughout their life, mothers may suffer a growing disadvantage over time if their lack of early investment in human capital and discontinuity in work experience keeps them out of high paying occupations and deny them opportunities for significant wage growth and occupational mobility later in life(Kahn 2014). In turn, this can affect the amount of education is received by the female. Higher levels of education allow people access to more specialized jobs that are often associated with high pay(Julian and Kominski 2011). In addition to higher earnings, people with higher levels of education are more likely to be employed full time, year-round, that is, they held a job for the entire year and worked in a full time capacity(Julian and Kominski 2011). These two variables will be discussed later in depth, but this clarifies the importance of them in this paper.

Beyond the age of child, education and other minor factors, the main focus point of this study is on the effects of increased house labor hours(through MealsW) has on a Female’s earnings and overall well-being. The literature examining the associations between unpaid labor hours and women wages has transformed recently. Some studies focus on the inequality of total hours as a burden on women, while others focus on the economic factors. As stated earlier in the paper, women are increasingly entering the workforce, and thus work-life balance is a struggle while rearing a child(Craig and Sawrikar 2009). The authors here also suggest that one of the main reasons for this struggling is that as women have entered the labor force, they have not seen a decrease in their amount of unpaid labor hours(Craig and Sawrikar 2009). Even when there is a man and a woman with equal work hours, the woman works twice as many unpaid hours as her spouse(Craig and Sawrikar 2009). One explanation for the imbalance in share of unpaid labor hours is the rearing of a child. However, the load is quite heavy when the child is younger, and lessens as the child ages and gains independence(Craig and Sawrikar 2009). We see that a chunk of women leave the labor force when a child is younger, yet return as the child ages(Craig and Sawrikar 2009). Some of these factors may be to model for the children, financial stress, boredom, a want to succeed, or other reasons. Results have shown that low income earners have the largest penalty for income classes relative to women from other economic groups(Budig and Hodges 2010). We can see the cyclical effects on education, as education of parents benefits their children’s own capital, however the mothers seem more beneficial as they tend to spend more time around their children(Becker 2010) This finding is also found where a study discovered that mothers are more likely than fathers to drop out of the labor force, cut back to part-time employment, take less demanding jobs, choose occupations that are more family-friendly, or pass up promotions, all of which affect their wage trajectories. (Raley, Bianchi, Wang 2010).An additional explanation to the gap can be explained by Discrimination, however due to this being hard to measure, it was not included in the regression. The effect of motherhood on wages might also be due to employer discrimination or to differences between mothers and other women in work effort(Staff and Jeylan 2011). This could be explained by family planning, or lack there-of of the low income earners. Many articles touch upon the impacts of how women’s unpaid labor affect them, but not a lot of studies touch upon how women’s unpaid labor hours affect men’s earnings. This article elaborates more on the previous work of analyzing unpaid work hours by analyzing the association between female wages and the amount of unpaid labor hours in two ways: showing nonlinearity and analyzing the variable not often covered in current literature, meals\_together.

**Data**

The data in this article was derived from the Panel Study of Income Dynamics (PSID), which is a longitudinal data set of a qualitative sample of individuals throughout the U.S. which consists of women, men, and children. The data is further broken down into family units and where they reside. The study was conducted at the University of Michigan. The survey originated in 1968 and has taken place ever since, and has been conducted biennially. The data can be found in the data mapping documents, using the 2011 set which was derived from 2010 sample.

The head of the family is interviewed each survey, and information is collected regarding the whole household in addition to each family member. Various personal factors are taken from the survey- such as if married, and other relevant family variables. The data does not include differentiation between head or a spouse, but that is separated by gender through coding. The variable that is unique to this survey is the MealW, and the interaction variable that focuses on women's unpaid labor in a subcategory of the amount of meals together as a family weekly.

The data has been cleaned so that values can be analyzed. For example, zeroes were usually coded as 5 in the data. This was fixed to better interpret the data, and allowed for the following variables to be used throughout the regression: *(Female, Married, hsewrk\_i, age2, age\_i, num\_child, MealW, religiosity\_high, years\_educ\_i, HouseworkW, meal\_together, age\_youngest\_child, youngchild, lnwage\_i)****.***

Detailed statistics regarding amount of housework hours for both males and females, the amount of meals together as a family per week, and other variables measured throughout this study started at the 2011 survey year. Also, this study used wage rates, education, housework hours, and other variables from the 2011 survey year.

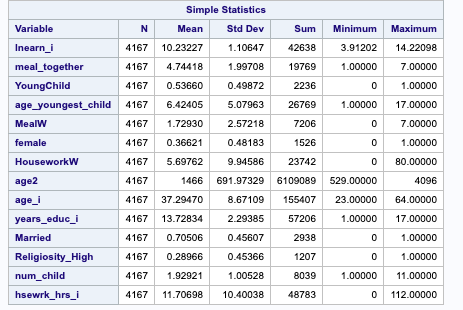
The age range used in this 2011 PSID data set is 14 to 120, which is the largest range used thus far. However, this study restricts the age group from 23 to 64. The reasoning behind this is to focus on those who are in the age range to be fully employed in the labor force. The restriction caused as well as the other observations that were dropped are listed below.

**Sample Selection**

The PSID data, 2011 survey year, surveyed men, women and children of all ages and marital statuses. The sample was restricted in order to most accurately capture the overall fit for the regression. A minimum age of 23 was selected, as well as a maximum age of 64. The reasoning behind this is that individuals under 23 are unlikely to be settled in an adequate job, and those over 64 are likely retired. After these restrictions were applied(as well as the restrictions/ dropped observations listed below), the sample was lowered to 4167, which consists of 1526 women and 2641 men.

There are certain variables in this model with missing information on key variables. The original dataset had 8,548 observations, but was restricted and the final regression has 4,167 entries. Earn\_i is missing 1,927 observations, but this is left in due to this being the dependent variable. Also, meal\_together is missing 2,247 observations, however this was left in as the main variable covered in the paper. Years\_educ\_i is missing 141 observations, which is not a lot. Also, years\_educ\_i is a great predictor of earnings. This would be an omitted variable bias. Married is only missing 1 variable also, and again this is a good fit for the model. Hsewrk\_hrs\_i is missing 65 observations. However, this is one of the main topics of the paper as well. Therefore, this was left in the model.

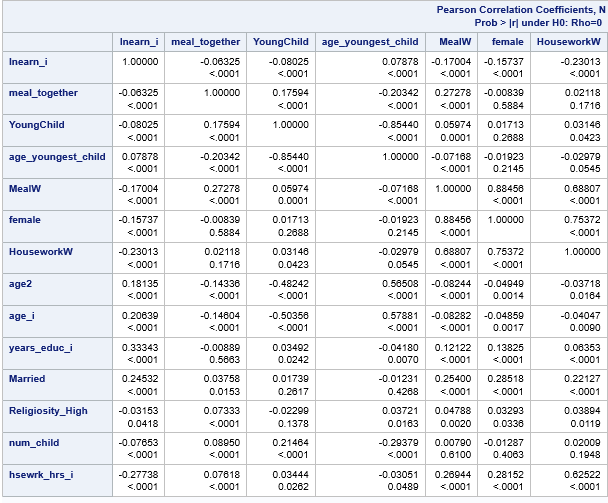
**Descriptive Statistics**

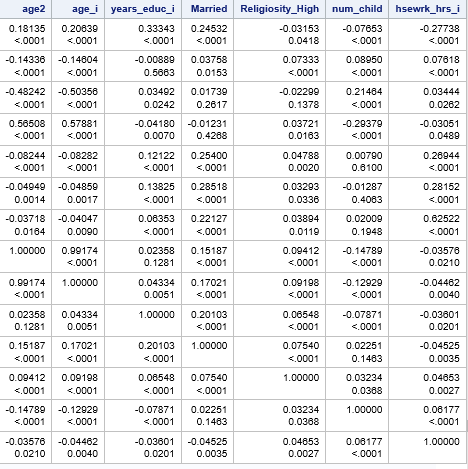
****

The descriptive statistics are listed above. The results above seem reasonable, however 112 hours of unpaid labor weekly seems excessive. It is also surprising that the highest man has a higher amount of housework hours than the highest woman. The number of children at 11 also seems pretty high. However, these cases should not mess up the overall data too much.

**Initial Regression**

In the Initial Regression phase, the variables were not properly assessed, and the regression was very confusing in terms of comprehending. There was no multicollinearity, which is a good sign. The chart that has VIFS is included for is below this section. Also, correlation between the variables was tested, and there is no high levels of correlation between the variables that would require us to look into it more. In the original regression, variables such as Drink\_How\_often\_i and smoke\_i were just not necessary and relevant to the research question at hand, and were dropped. Two variables were on the fence of being added, age\_youngest\_child and meals\_together. They were missing significant observations. However, these were included in the final regression to avoid omitted variable bias. The Initial Regression also had sample bias, and variables were improperly labeled. For example, The dependant variable had narrowed down to only select wifes, and the independent variables also selected wifes. In the Initial Regression, men and women regressions were separated, but now it is merged into one regression for simplicity. This was done by creating an individual variable that had the full sample. The regression now is transferred into individual variables on both sides of the equation, and certain slope dummies capture the message that was attempted to be displayed in the initial regression. Overall, the regression has been narrowed down to relevant variables to best capture the message for the research question, while removing the least amount of observations. *Also, the data below was found to be heteroscedastic. The p-values have been altered to reflect this variance.*

**

**



**Variable Construction**

*MealW* was constructed as a Slope Dummy variable. It grouped together *Female* as well as *Hsewrk\_hrs\_i.* This construction has not yet been done in modern papers(around the same topic) and is considered the main contribution this paper has to offer. Men were not included in this variable.

The *HouseworkW* variable is defined as if a female and the amount of associated earnings decrease based upon overall individual housework hours. Earlier studies have looked at housework hours on the individual level, as well as the split between each gender (Dush 2017; Avery 2010; Bianchi 2012). However, none of these studies used the slope dummy variable form, which turned out to be statistically significant. This a twist to the topic that has not been explored.

Also, some minor transformations happened to other variables. For example, *Age\_i* was squared(forming *Age2)* to capture the parabola shape of lifelong earnings. This was hypothesized that earnings increase up to a certain age, and decrease as the individual ages. Also, a dummy variable was formatted around *age\_youngest\_child* creating *youngchild.* The reason for this is that there is an unpaid labor difference in the housework associated with a child under 5, apart from a child above 5. When adding *Youngchild* into the regression, it turned *MealW* into a statistically significant variable. It is hypothesized that the mother has to perhaps cook more meals, and this is associated with a decrease in earnings.

**Methodology**

**Theoretical Model:**

Y=βo-β1X1+β2X2-β3X3-β4(X5\*X5)+β5X5-β6X6-β7(X12\*X1)-β8X8+β9X9-β10(X1\*X3)+β11(X11)-β12(X12)-B13(X13)+ƐI

Where:Y is The Log variable of labor income earnings which is the amount of labor income of the individual from 3.91202 to 14.22098

- X1 is Female, a dummy variable 1 if female, 0 if not.

X2 is married, a dummy variable, 1 if the individual is married, 0 if not.

-X3 is hsewrk\_hrs\_i, the amount of hours the ith individual spends doing housework hours per week, which ranges from 0 to 112.

-X5\*X5 is age2, the age of the individual, times the age of the individual. This is expressed in years. This is used to capture the parabola shape.

-X5 is age\_i, the years of age the individual has, which ranges from 23 to 64.

-X6 is the num\_child, which is the number of children the individual haswhich ranges from 1 to 11.

X13\*X1 is MealW. This is a slope dummy variable that captures the impact on a female and the amount of associated earnings decreased based upon amount of number of weekly meals prepared, ranging from zero times to seven times.

-X8 Is Religiosity\_High, which is coded as a 1 if the individual goes to church more than once a week, a 0 if not.

-X9 is the years\_educ\_i of the individual. Ranges from 0 to 16 years of school, and 17 is coded if at least some graduate school.

-X1\*X3 is HouseworkW in the female regression model, This is a slope dummy variable that captures the impact on a female and the amount of associated earnings decrease based upon overall individual housework hours.

-X11 is age\_youngest\_child, which is the age of the youngest child in the house, ranging from one to seventeen.

-X12 is youngchild, which is a dummy variable. 1 if the child is 5 or younger, 0 if not.

-X13 is meal\_together, which is the amount of times the family eats a main meal together weekly, ranging from 1 to seven.

-Ɛi=error term

**Theoretical Framework**

The model uses OLS with natural log regression analysis, using individual's earnings from the 2011 PSID dataset, comparing multiple variables that are listed above. The purpose of this functional form is to best fit the data representation curve. We also do this as we want the earnings in percentage points, for relevance. When in logarithmic form, the coefficients of the other variables become easier to comprehend and are easier to compare.

*Housework/unpaid labor hours*impacts individuals in various ways. It could cause the individual to not have as much time to dedicate to their profession. Therefore, it is predicted that an increase in *Housework/unpaid labor hours* will decrease the earnings of the individual. One study found that men enjoyed more leisure time, particularly on non work days, whereas their partners performed more nonmarket work. (Dush 2017). This could lead to women becoming burnt out at a quicker rate, relative to men. Secondly, women’s household participation can actually increase the earnings for men- as those with a dedicated housewife can focus more on their careers. Dush also theorized that exchange and bargaining perspectives suggest that because men have traditionally specialized in the market and have out-earned their wives, husbands have been able to bargain their way out of nonmarket work(Dush 2017). *HouseworkW* was constructed to to capture this effect. As a result, more hours of women’s contribution to housework could be associated with a higher wage rate for some men, and also linked to a decrease earning percentage for women

Looking more specifically, is the *age\_youngest\_child,* as well as the *num\_child* variables*. Num\_child* could be associated with a lower wage rate for women, as this is a time consuming activity for women. The more children a woman has, the more time is spent taking care of the child, and other unpaid labor hours. This could also have a negative impact on the women’s overall well-being as she has less time for herself. However, *age\_youngest\_child* is predicted to be associated with an increase the earnings, as the individual will have more time to dedicate to work. A study done on predicting work overload displayed the adverse effects on the health of overworked individuals (Avery 2010). The thought of having too much work and too little time is a psychological stressor, leading to burnout as well as work-life and work-family conflict. (Avery 2010). In 2009/10, women are estimated to do 1.6 times the amount of housework as men, on average(Bianchi 2012). These two findings put together in a context could explain the pay gap, or even possibly women’s withdrawl from the workforce. However, the main burden of childcare and associated housework occurs when children are very young and tapers off as they mature(Craig, Sawrikar 2009). Infants require the most extreme amounts of time and attention because of their age-related needs, but these interact with institutional settings that are also of importance. (Craig and Sawrikar 2009). This study will separate this variable using a dummy variable which captures a 1 if the child is 5 years or younger in age, a 0 if 6 to 17. The slope dummy variable, which is coded as *youngchild,* is predicted to decrease earnings for women, holding other variables constant.

*Years\_educ\_i* is a very important variable when considering earnings, and has been shown to lead to an increase in earnings. Through human capital theory in addition to signaling theory, education is associated with higher earnings. Another beneficial aspect of education for women is that educated women tend to earn much more, have fewer children, and appear more knowledgeable (Becker 2010). Therefore, it is predicted that education is associated to increase earnings for both sexes, but marginally decreasing as the years of schooling increases, holding the other variables constant.

As individuals continue to *age,* they gain crystallized knowledge and become more productive individuals. They continue to develop more human capital, and have less sunken costs. Thus, as the age variable increases, the earnings of the individual is predicted to increase, holding other variables constant. However, when individuals are at retirement age, they are not able to collect income, and live off of their retirement savings. This explains why the regression model uses the age2 variable, to capture the shape in order to fit the model most accurately.

*Highly Religious Individuals*have a different standard of values and economic participation than non religious individuals, especially those at a high participation level. High participation in religious activities is associated with decreased earnings, as it takes time away from time spent on work. This finding could also suggest that perhaps highly religious women are more passive, which is associated with lower earnings on the female side.

There is a plethora of studies on the gender gap, and negative economic associations with being a *Female.* One author boldly stated that in all western societies, women earn lower wages on average than men (Blau 2016). Therefore, in this model, women will be predicted to have less earnings than men. Possible explanations for this are the economic burden of housework hours, as well as rearing a child. This study was not able to measure for discrimination, but that is also a factor.

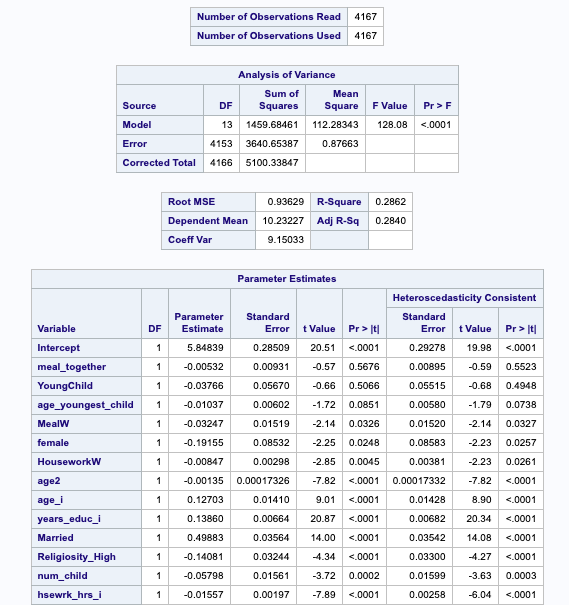
*Married* is a factor that will increase earnings. As marriage is a signal of a healthy relationship, this also will lead to shared housework duties. Having less time allocated to chores will allow the individual to focus more on income, thus yield higher labor earnings. One study found that married men may also have higher earnings than unmarried men because they work more hours(Mincy 2009). This could be explained by the unpaid labor hour difference mentioned above. Although women still benefit slightly, men are the ones who truly benefit from the marriage, shown from multiple empirical findings.

Finally, the main variable that is expected to have a negative impact on earnings is *meal\_together.* This is a form of unpaid labor expense, so as the amount of meals eaten together as a family weekly increase, it is predicted that earnings will decrease. Current literature does not mention this variable, but as listed above in other studies, higher unpaid labor is associated with a decrease in pay (Schneider 2011; Craig and Sawrikar 2009; Becker 2010; Raley, Bianchi, Wang 2010; Staff and Jeylan 2011). A slope dummy variable was constructed that includes *meal\_together* as well as *female.* This slope dummy is known as *MealW* and captures if the individual is a female, and the amount of times the family has a meal together per week. This will be associated with lower earnings for the female.

**Results and Final Regression**

When running the final regression, and all appeared good but the results turned out to be heteroscedastic. This was discovered by looking at the chi-squared probability to test for heteroscedasticity. The value is below that .05 threshold which indicates there is heteroscedasticity in the model. Below also include the white test that adjusts for this heteroscedasticity. This is reflected in the model. However, the values did not change a lot so the double log functional form was not needed to correct for this.



****

**28.40% percent of the variation in the regression in the natural log of the individual’s labor earnings can be explained by the independent variables, adjusting for degrees of freedom.**(Female, Married, hsewrk\_i, age2, age\_i, num\_child, MealW, religiosity\_high, years\_educ\_i, HouseworkW, meal\_together, age\_youngest\_child, youngchild)**.**

F-Test for Earnings: Ho: B1=B2=B3=B4=B5=B6=B7=B8=B9=B10=B11=B12=B13=0

Ha: one or more of the parameters isn’t equal to zero.

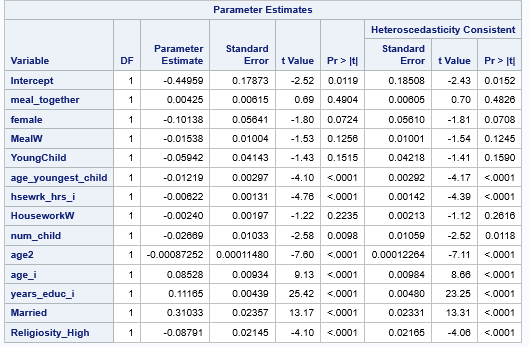
Where: B1 is coefficient of dummy variable (Female), B2 is coefficient of married, a dummy variable equal to 1 if married, 0 if not, B3 is the amount of housework hours the individual performs weekly(hsewrk\_hrs\_i), B4 is the coefficient of the age of the individual times the age of the individual(age2), B5 is the coefficient of the amount of years an individual has(age\_i), B6 is the coefficient of the number of children the individual has which ranges from 0 to 18 (num\_child), B7 is the coefficient of MealW, B8 is the coefficient of Religiosity High, B9 is the coefficient of years\_educ\_i, B10 is the coefficient of the interaction variable HouseworkW, B11 is the coefficient of age\_youngest\_child, B12 is the coefficient of youngchild, B13 is meal\_together.

**F\*=2.80**

F(4153)=128.08 and a P-value<.0001. We can reject the null hypothesis, as there is sufficient evidence to conclude that there is an overall significant relationship between the independent variables(Female, Married, hsewrk\_i, age2, age\_i, num\_child, MealW, religiosity\_high, years\_educ\_i, HouseworkW, meal\_together, age\_youngest\_child, youngchild) and the dependant variable natural log of total labor earnings.

* First, Female is in the expected direction.
  + Ho: B1=0, Ha: B1 is not equal to zero. Where: B1 is coefficient of female, a dummy variable equal to 1 if a female, 0 if not. We can reject the null hypothesis as This independant variable is statistically significant. We know this because the T(4153) is |-2.23|>**1.960(t\*)**, with a p-value of .0257. **If the individual is a female, the expected percentage of total labor earnings is expected to decrease by 19.15%, holding other variables constant.**
* Next, married is also in the expected direction.
  + Ho: B2=0, Ha: B2 is not equal to zero. Where: B2 is coefficient of married, a dummy variable equal to 1 if married, 0 if not. We can reject the null hypothesis as This independant variable is statistically significant. We know this because the T(4153) is |14.08|>**1.960(t\*)**, with a p-value of less than .0001. **If the individual is married, the expected percentage of total labor earnings is expected to increase 49.883%, holding other variables constant.**
* Next, housework\_hrs\_i is in the expected direction.
  + Ho: B3=0, Ha: B3 is not equal to zero. Where: B3 is the amount of housework hours the individual performs weekly(hsewrk\_hrs\_i). We can reject the null hypothesis as This independant variable is statistically significant. We know this because the T(4153) is |-6.04|>**1.960(t\*)**, with a p-value of less than .0001. **As the amount of housework hours a male performs increases by one hour, the expected percentage of total labor earnings is expected to decrease 1.56%, holding other variables constant. (Note that is further explained in B10.)**
* Next, age2 is in the expected direction.
  + Ho: B4=0, Ha: B4 is not equal to zero. Where: B4 is the coefficient of the age of the individual times the age of the individual(age2). We can reject the null hypothesis as This independant variable is statistically significant. We know this because the T(4153) is |-7.82|>**1.960(t\*)**, with a p-value of less than .0001. **This variable does not need to be interpreted.**
* Next, age\_i is in the expected direction.
  + Ho: B5=0, Ha: B5 is not equal to zero. Where: B5 is the coefficient of the amount of years an individual has(age\_i). We can reject the null hypothesis as This independant variable is statistically significant. We know this because the T(4153) is |8.90|>**1.960(t\*)**, with a p-value of less than .0001. **As the individual gains an additional year of life, the expected percentage of total labor earnings is expected to increase 12.7%, holding other variables constant.**
* Next, num\_child is in the expected direction.
  + Ho: B6=0, Ha: B6 is not equal to zero. Where: B6 is the coefficient of the number of individuals under 18 living in the household(num\_child). We can reject the null hypothesis as this independant variable is statistically significant. We know this because the T(4153) is |-3.63|>**1.960(t\*)**, with a p-value of .0003. **As the individual has an additional individual under 18 years of age living in their household, the expected percentage of total labor earnings is expected to decrease by 3.63%, holding other variables constant.**
* Next, *MealW*  is in the expected direction.
  + Ho: B7=0, Ha: B7 is not equal to zero. Where: B7 is the coefficient of a slope dummy variable that captures the impact on a female and the amount of associated earnings decreased based upon amount of times eating main meals together as a family weekly. We can reject the null hypothesis as this independent variable is statistically significant. We know this because the T(4153) is |-2.14|>**1.960(t\*)**, with a p-value of .0327. **If the individual is a female and as the amount of main meals eaten together as a family weekly increases by one meal weekly, her expected percentage of total labor earnings is expected to decrease by 3.25%, holding other variables constant.**
* Next, Religiosity\_High is in the expected direction.
  + Ho: B8=0, Ha: B8 is not equal to zero. Where: B8 is the coefficient of a dummy variable, which is coded as a 1 if the individual goes to church more than once a week, a 0 if not(Religiosity\_High). We can reject the null hypothesis as this independant variable is statistically significant. We know this because the T(4153) is |-4.27|>**1.960(t\*)**, with a p-value of less than .0001. **If the individual goes to church more than once a week, their expected percentage of total labor earnings is expected to decrease by 14.08%, holding other variables constant.**
* Next, Years\_educ\_i is in the expected direction.
  + Ho: B9=0, Ha: B9 is not equal to zero. Where: B9 is the coefficient of the years\_educ\_i of the individual. Ranges from 0 to 16 years of school, and 17 is coded if at least some graduate school(Years\_educ\_i). We can reject the null hypothesis as this independant variable is statistically significant. We know this because the T(4153) is |20.34|>**1.960(t\*)**, with a p-value of less than .0001. **As the individual gains an additional year of schooling, their expected percentage of total labor earnings is expected to increase by 13.86%, holding other variables constant.**
* Next, HouseworkW is in the expected direction.
  + Ho: B10=0, Ha: B10 is not equal to zero. Where: B10 is the coefficient of the HouseworkW, which is a slope dummy variable that captures the impact on a female and the amount of associated earnings decrease based upon overall individual housework hour. We can reject the null hypothesis as this independant variable is statistically significant. We know this because the T(4153) is |-2.23|>**1.960(t\*)**, with a p-value of .0261. **If the individual is a female and as the amount of housework hours increases by one hour weekly, her expected percentage of total labor earnings is expected to decrease by 2.4%, holding other variables constant.**
* Next, age\_youngest\_child is not in the expected direction.
  + Ho: B11=0, Ha: B11 is not equal to zero. Where: B11 is the coefficient age\_youngest\_child, which is the age of the youngest child in the house, ranging from one to seventeen. We cannot reject the null hypothesis as this independant variable is not statistically significant. We know this because the T(4153) is |-1.79|<**1.960(t\*)**, with a p-value of .0738.**We can not interpret this variable as it is not statistically significant.**
* Next, youngchild is in the expected direction.
  + Ho: B12=0, Ha: B12 is not equal to zero. Where: B12 is the coefficient of youngchild, which is a dummy variable. 1 if the child is 5 or younger, 0 if not. We can reject the null hypothesis as this independant variable is statistically significant. We know this because the T(4153) is |-.69|<**1.960(t\*)**, with a p-value of .4948. **We can not interpret this variable as it is not statistically significant.**
* Lastly, meal\_together is in the expected direction.
  + Ho: B13=0, Ha: B13 is not equal to zero. Where: B13 is the coefficient of meal\_together, which is the amount of times the family eats a main meal together weekly, ranging from 1 to seven. We cannot reject the null hypothesis as this independant variable is not statistically significant. We know this because the T(4153) is |-.59|<**1.960(t\*)**, with a p-value of .5523. **We can not interpret this variable as it is not statistically significant.**

**Sensitivity Analysis**

****

This study performed a robustness check to ensure that the results hold. Originally, this data set and study was going to use wage\_i instead of earn\_i. Wage\_i is defined as the hourly rate from one cent to 998.99 dollars and cents per hour. Only approximately .02% of the population earned above 999 dollars per hour. This variable accounts for hours worked, where *earn\_i i* does not. However, there was a decision not to due to the *female* variable not being statistically significant. The model is attached above, as *wage\_i* changed the signs on both of the two main variables, *MealsW* and *HouseworkW.*

The main conclusions drawn from the robustness check above is that the two variables of importance of this paper are no longer statistically significant. Therefore, there is a debate which model is a better fit. In the one with *lnwage\_i* as the dependant variable, *mealw* and *houseworkw* are no longer statistically significant. This is perhaps due to the fact that the dependant variable is accounting for hours worked. However, *Female* is a very important variable, and it is troubling that this variable isn’t statistically significant in the context of the model with *lnearn\_i* as the dependent variable.

In the model with *lnearn\_i* as the dependant variable however; *mealw, houseworkw,* and *female* are statistically significant. This model is not accounting for hours worked in the dependant variable, just overall labor earnings. Therefore, it would make sense for the above variables to be statistically significant.

**Conclusions**

This article examined the association between the amount of meals together as a family and females earning’s. Analysis of other variables were also provided. Focused on OLS, if an individual is a woman and as the amount of meals eaten together weekly as a family increases, is associated with lower earnings. The results were heteroscedastic, but this was accounted for. *Also, If the individual is a female and as the amount of housework hours increases by one hour weekly(decreasing her work assigned to job), her expected percentage of total labor earnings is expected to decrease.* These are the two main findings that this paper presents, although there are other statistically significant variables that have been shown which include; lower earnings associated with being a female, higher earnings associated with being married, higher earnings associated with higher years of education, higher earnings associated with higher age, lower earnings associated with high religiosity, lower earnings associated with an increased number of children, and males have less earnings associated with an increase in unpaid labor hours. However, further research that explores the sample that women that drop out of the labor force would be desirable.

**Contributions**

The relationships between the motherhood penalty and labor wages between women has been examined in the context of previous studies. These studies pick apart certain aspects such as number of children(Staff and Mortimer 2011). Another study includes men’s housework hours and time as well as number of children(Schneider 2011). In addition, to one study that focuses on time spent with children as the dependant variable(Raley, Bianchi, Wang 2012). Another finding measured wifes hours as cooking and cleaning as different independent variables(Killewald 2011).

However, this article has employed a slope dummy variable that captures amount of meals cooked weekly which is reflected on women’s earnings. This appears to be the first time this approach has been done to meals together, and the approach has yielded statistically significant results. None of the other studies have seemed to use a slope dummy variable either. The other variables used in the regression remain consistent with modern literature, so this shows that the slope dummy, MealW, is still statistically significant in the context of modern research.

The wage gap has been a topic of study of decades now, and getting to the root of the problem has not been accomplished, as the gap still exists. The findings in this article will hopefully allow individuals to reevaluate the work share done in their homes, and educate readers on the difference it reflects on pay. Societal norms have changed and will continue to change, and hopefully the pay gap will continue to lessen.

**Limitations**

This analysis was unable to account for unemployed individuals. It is hypothesized that as women reach a certain point of housework hours, they remove themselves from the labor force, in turn having an income of zero. This would be an interesting point for additional studies to build upon. In addition, this doesn’t account for if women reduce their work hours. Also, it would be interesting to see how this withdrawal would benefit men economically.

Also, the final regression in this study does not account for hours worked. This could explain why some of the results are statistically significant. While trying to control for this in the robustness check, three main variables transferred into not being statistically significant. Therefore, it would be worth exploring this in a further topic- however finding another way to account for labor hours on earnings.

**Policy Suggestions**

One topic that is heavily explored is universal child care. The would be considered as a pre-preschool aged tool that “levels the playing field.” One finding found that this leveled the playing field for children, in the long run(Tarjel 2015). This study also showed that certain countries, such as scandinavia have already implemented such a system. This is interesting, and an idea that was not heavily focused on throughout this research. The impact of a low earning mother/father continuing on through work while neglecting a child could lead to future educational issues as well as a cyclical-pattern. This policy suggested by Tarjel would help decrease the pay gap, and help out the low-earners(and mothers). This policy would free up more time for the mothers, and would prevent mothers from getting behind in human capital. Finally, this policy could possibly reduce the amount of women that withdraw from the labor force(for future research).

Another main area of focus is paid maternity leave. While this study was unable to control for women leaving the workforce, it is hypothesized that the impact of giving birth would cause a woman to leave the workforce for a certain amount of time. In combination of this paper’s empirical findings of women’s income decreasing based upon number of children, this is an area that should be focused upon. Even when women get a time of leave, they are expected to return soon. Currently, the only two states in America with paid maternity leave are California(2004) and New Jersey(2008) (Rossin 2011). Paid maternity leave has become a standard benefit in many countries throughout the world. (Aitzen, Garrett, Hewitt, Keogh, Hocking, Kavanagh 2015).With this dataset being structured in America, this could perhaps explain some of the reasoning behind the motherhood gap, and the statistically significance of *female.* This study is an analysis of seven separate studies from Australia, Sweden, Norway, USA, Canada, and Lebanon. The four studies that examined leave at an individual level showed evidence of maternal health benefits, whereas the three studies conducting policy-level comparisons reported either no association or evidence of a negative association. (Aitzen, Garrett, Hewitt, Keogh, Hocking, Kavanagh 2015). “France and Germany offer over 160 weeks of job-protected leave for mothers (20 weeks of which are paid in France, 42 weeks paid in Germany) whereas in the USA the federal government provides mothers with 12 weeks of job-protected leave, none of which is paid”(Aitzen, Garrett, Hewitt, Keogh, Hocking, Kavanagh 2015). This means that women in america are not given the same packages that other countries provide mothers with. Recent debates have focused on the ideal length of leave, and in some countries such as the USA and until recently, Australia, about whether or not mothers have a right to paid leave(Aitzen, Garrett, Hewitt, Keogh, Hocking, Kavanagh 2015). Further, this study found that the absence of paid leave or shorter paid leave can have detrimental effects on women’s rates of returning to work. Therefore, as this dataset uses data from the United States, it would be interesting to see the effects if this policy was put into place in America.

The last policy suggestion is a universal base income for mothers. Women have certain biological tools that leave them economically disadvantaged. This idea is similar to universal child care in an aspect, however, is a more direct solution to the motherhood penalty. There is not readily available peer reviewed articles for this topic, but it is rather intuitive. The subsidy would directly subsidize mothers, for having a biological tool that hurts them economically. A separate study could perhaps calculate a realistic amount for this subsidy, and would perhaps vary by region and cost of living in certain areas.

**Directions for Future Research**

A future topic that would be ideal for this data set/ topic would be an expansion from the limitations section. The greatest limitation of this paper was the inability to control for the amount of women that dropped from the labor force. Future research could possibly remove unemployed from the sample. This would entail a great deal of re-coding and organizing the data, but it could be done.

If this was done correctly, it could be a potentially better fit of this regression. This variable(unemployed women) isn’t the focus of this study, but could perhaps slightly alter the outcome. This would greatly contribute to the conversation, and would contribute to a section that nobody else has explored in this topic. This could be combined with possible policy recommendations in order to explore new ways to evaluate the gender gap.

Also, it would be interesting to see if the results change will controlling for amount of hours worked. This could perhaps be a better fit for the regression, and would better capture the coefficients of the variables HouseworkW, hse\_wrk\_i, as well as MealsW. The issues for the odd result in the robustness check warrants future research, but was not allowed in this study due to time constraints.

**References**

Aitken Z., Garrett C.C., Hewitt B., Keogh L., Hocking J.S., Kavanagh A.M.

(2015) The maternal health outcomes of paid maternity leave: A systematic review *Social Science and Medicine*, 130 , pp. 32-41.

**Avery, Derek. *Overworked in America?: How Work Hours, Immigrant Status ...* Journal Of Managerial Psychology, Feb. 2010, https://www.researchgate.net/publication/243459621\_Overworked\_in\_America\_How\_work\_hours\_immigrant\_status\_and\_interpersonal\_justice\_affect\_perceived\_work\_overload.**

Becker, Gary S., et al. “Explaining the Worldwide Boom in Higher Education of Women.” *SSRN Electronic Journal*, 2010, doi:10.2139/ssrn.1681606.

Bianchi, Suzanne M, et al. “Housework: Who Did, Does or Will Do It, and How Much Does It Matter?” *Social Forces; a Scientific Medium of Social Study and Interpretation*, U.S. National Library of Medicine, 1 Sept. 2012, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4242525/>.

Blau, Francine D., 2016. "[Gender, Inequality, and Wages](https://ideas.repec.org/b/oxp/obooks/9780198779971.html)," [OUP Catalogue](https://ideas.repec.org/s/oxp/obooks.html), Oxford University Press, number 9780198779971 edited by Gielen, Anne C. & Zimmermann, Klaus F..

Budig, Michelle J., and Melissa J. Hodges. “Differences in Disadvantage.” *American Sociological Review*, vol. 75, no. 5, 2010, pp. 705–728., doi:10.1177/0003122410381593.

Chiswick, Barry R., and Jidong Huang. “The Earnings of American Jewish Men: Human Capital, Denomination, and Religiosity.” *Wiley Online Library*, John Wiley & Sons, Ltd (10.1111), 4 Nov. 2008, <http://bs6vx4ge6d.scholar.serialssolutions.com/?sid=google&auinit=BR&aulast=Chiswick&atitle=The+earnings+of+American+Jewish+men:+Human+capital,+denomination,+and+religiosity&id=doi:10.1111/j.1468-5906.2008.00433.x&title=Journal+for+the+scientific+study+of+religion&volume=47&issue=4&date=2008&spage=694&issn=0021-8294>

Craig, Lyn, and Pooja Sawrikar. “Work and Family: How Does the (Gender) Balance Change as Children Grow?” *Gender, Work & Organization*, vol. 16, no. 6, 2009, pp. 684–709., doi:10.1111/j.1468-0432.2009.00481.x

Dush, Claire M. Kamp, et al. “What Are Men Doing While Women Perform Extra Unpaid Labor? Leisure and Specialization at the Transitions to Parenthood.” *SpringerLink*, Springer US, 3 Oct. 2017, https://link.springer.com/article/10.1007/s11199-017-0841-0.

Julian, Tiffany, and Robert Kominolski. “Education and Synthetic Work-Life Earnings Estimates.” *The United States Census Bureau*, 2 Aug. 2018, https://www.census.gov/library/publications/2011/acs/acs-14.html.

Kahn, Joan R., et al. “The Motherhood Penalty at Midlife: Long‐Term Effects of Children on Women's Careers.” *Wiley Online Library*, John Wiley & Sons, Ltd (10.1111), 13 Jan. 2014, https://onlinelibrary.wiley.com/doi/abs/10.1111/jomf.12086.

Killewald, Alexandra. “Opting Out and Buying Out: Wives Earnings and Housework Time.” *Journal of Marriage and Family*, vol. 73, no. 2, Oct. 2011, pp. 459–471., doi:10.1111/j.1741-3737.2010.00818.x.

Mincy, Ronald, et al. “Marriage: Cause or Mere Indicator of Future Earnings Growth?” *Journal of Policy Analysis and Management : [the Journal of the Association for Public Policy Analysis and Management]*, U.S. National Library of Medicine, 2009, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3043605/.

Pitt, Mark M, et al. “Human Capital Investment and the Gender Division of Labor in a Brawn-Based Economy.” *American Economic Review*, vol. 102, no. 7, 2012, pp. 3531–3560., doi:10.1257/aer.102.7.3531.

Raley, Sara, Bianchi, Suzanne, and Wang, Wendy, "When Do Fathers Care? Mothers’ Economic Contribution and Fathers’ Involvement in Child Care," American Journal of Sociology 117, no. 5 (March 2012): 1422-1459. <https://www.journals.uchicago.edu/doi/full/10.1086/663354>

Schneider, Daniel. “Market Earnings and Household Work: New Tests of Gender Performance Theory.” *Journal of Marriage and Family*, vol. 73, no. 4, 2011, pp. 845–860., doi:10.1111/j.1741-3737.2011.00851.x.

Rossin, Maya. “The Effects of Maternity Leave on Childrens Birth and Infant Health Outcomes in the United States.” *Journal of Health Economics*, vol. 30, no. 2, 2011, pp. 221–239., doi:10.1016/j.jhealeco.2011.01.005.

Staff, Jeremy, and Jeylan T. Mortimer. “Explaining the Motherhood Wage Penalty During the Early Occupational Career.” *SpringerLink*, Springer US, Demography Volume 49, Issue 1 pp 1-2126 Oct. 2011, <https://link.springer.com/article/10.1007/s13524-011-0068-6>.

Tarjel, Havnes. *Is Universal Child Care Leveling the Playing Field ...* Journal of Public Economics, July 2015, <http://ftp.iza.org/dp4978.pdf>. \*\*Policy

**Coding for Final Regression:**

PROC SQL;

CREATE TABLE WORK.query AS

SELECT ER27510A , ER27634A , ER27644A , ER27667A , ER28043A , S702A , S703A , S704A , S705A , S706A , S707A , S708A , S709A , S710A , S711A , S713A , S714A , S715A , S718A , S719A , S720 , S720A , S716A , S717A , ER33826A , ER33826B , ER33826C , ER33826D , ER33826E , ER33826F , ER33826G , ER33826H , ER33826I , ER33826J , ER33826K , ER33826L , ER33826M , ER33826N , ER33826O , ER33827A , ER33827B , ER33827C , ER33827D , ER33827E , ER33827F , ER33827G , ER33827H , ER33827I , ER33827J , ER33827K , ER33827L , ER33827M , ER33827N , ER33827O , ER33827P , ER33827Q , ER33827R , ER33827S , ER33827T , ER33827U , ER33827V , ER33828A , ER33828B , ER33828C , ER33828D , ER33828E , ER33828F , ER33828G , ER33828H , ER33828I , ER33828J , ER33828K , ER33828L , ER33833A , ER33833B , ER33833C , ER33833D , ER33833E , ER33833F , ER33833G , ER33833H , 'year'n , 'count'n , famid , personid , upid , in\_us\_68 , sample , gender\_male\_i , sex\_male\_hd , sex\_male\_i , rel\_to\_hd , seq\_num , head , indiv\_weight , respond\_i , respondent , interview\_num\_i , interview\_num\_f , age\_hd , age\_wf , no\_wife , age\_i , age\_sp , nowoman , 'move'n , year\_move\_in , marital\_pair , new\_hd , new\_wf , year\_new\_hd , year\_new\_wf , year\_new\_i , hd\_sp\_stat , split\_num , split\_id , famid\_other , release\_n , dob\_month\_i , dob\_year\_i , empl\_stat\_i , years\_educ\_hd , years\_educ\_wf , years\_educ\_i , years\_educ\_i1 , fam\_health , Hispanic\_hd , race\_hd1 , race\_hd2 , race\_hd3 , race\_hd4 , Hispanic\_wf , race\_wf1 , race\_wf2 , race\_wf3 , race\_wf4 , height\_inches\_wf , Hispanic , race\_i1 , race\_i2 , race\_i3 , race\_i4 , ethnic\_grp\_hd , ethnic\_grp\_wf , ethnic\_grp\_i , ethnic\_grp\_sp , ethnic\_grp\_hd1 , ethnic\_grp\_hd2 , ethnic\_grp\_wf1 , ethnic\_grp\_wf2 , ethnic\_grp\_i1 , ethnic\_grp\_i2 , fam\_size , num\_child , age\_youngest\_child , num\_nonfu\_hu , mar\_stat\_i , mar\_stat\_g , mar\_stat\_ch , cpl\_stat\_hd , educ\_hd , educ\_wf , educ\_i , educ\_sp , employ\_stat\_hd1 , employ\_stat\_hd2 , employ\_stat\_hd3 , employ\_stat\_wf1 , employ\_stat\_wf2 , employ\_stat\_wf3 , employ\_stat\_i1 , employ\_stat\_i2 , employ\_stat\_i3 , religion\_denom\_hd , relig\_hd , relig\_denom\_hd , religion\_denom\_wf , relig\_wf , relig\_denom\_wf , relig\_i , relig\_sp , relig\_denom\_i , relig\_denom\_sp , worship\_place\_hd , worship\_place\_wf , worship\_place\_i , relig\_svc\_num\_hd , relig\_svc\_num\_wf , relig\_svc\_time\_hd , relig\_svc\_time\_wf , relig\_svc\_num\_i , relig\_svc\_time\_i , relig\_svc\_per\_month\_i , relig\_svc\_per\_week\_i , relig\_donate , relig\_donate\_amt , 'state'n , geogr\_mobility , state\_grew\_hd , state\_other\_ever\_live , moved\_hd , moved\_reason\_hd , health\_hd , health\_wf , health\_i , health\_sp , hsewrk\_hrs\_hd , hsewrk\_hrs\_wf , hsewrk\_hrs\_i , hsewrk\_hrs\_sp , meal\_together , earn\_hd , earn\_wf , earn\_i , earn\_sp , wage\_hd , wage\_wf , wage\_i , wage\_sp , wage\_salary\_hd , wage\_salary\_wf , wage\_salary\_i , wage\_salary\_sp , fam\_inc , transfer\_inc\_hdwf , tax\_inc\_hdwf , tax\_itemize , wealth\_he\_no , wealth\_he\_yes , savings , savings\_amt , business\_farm\_own , business\_farm\_amt , debt\_cc , debt\_cc\_amt , re\_other\_own , re\_other\_own\_amt , stock\_own , stock\_own\_amt , veh\_amt , savings\_other , savings\_other\_amt , IRA\_own , IRA\_own\_amt , bonds , bonds\_amt , home\_ownership , home\_value , mortgage , mortgage\_type , mortgage\_amt , mortgage\_yr , mortgage\_2 , mortgage2\_type , mortgage2\_amt , debts\_other , debts\_other\_amt , IRA\_changed , IRA\_changed\_amt , pension\_cashed , pension\_cashed\_amt , home\_sold , home\_sold\_amt , re\_other\_bought , re\_other\_bought\_amt , re\_other\_sold , re\_other\_sold\_amt , re\_additions , re\_additions\_amt , business\_farm\_inv , business\_farm\_inv\_amt , business\_farm\_sold , business\_farm\_sold\_amt , business\_farm\_saving\_amt , stock\_bought , stock\_sold , stock\_more\_sold , stock\_amt , stock\_net\_amt , stock\_saving\_amt , stock\_inv\_amt , noIRA\_stock\_sold , noIRA\_stock\_amt , gift\_inher , gift\_inher\_yr , gift\_inher\_amt , gift\_inher2 , gift\_inher\_yr2 , gift\_inher\_amt2 , gift\_inher3 , gift\_inher\_yr3 , gift\_inher\_amt3 , pension\_job , pension\_yrs , pension\_yr\_joined , pension\_elig , pension\_yrs\_elig , pension\_vested , pension\_yrs\_vested , pension\_contr , pension\_contr\_req , pension\_contr\_req\_amt , pension\_contr\_req\_amt\_unit , pension\_contr\_vol , pension\_contr\_v\_amt , pension\_contr\_vol\_amt , pension\_contr\_empl , pension\_contr\_e\_amt , pension\_contr\_empl\_amt , pension\_now\_amt , fam\_comp , fth\_born\_US , fth\_educ\_foreign\_yrs , fth\_educ\_US , fth\_educ\_US\_yrs , fth\_educ\_yrs , mthr\_born\_US , mthr\_educ\_foreign\_yrs , mthr\_educ\_US , mthr\_educ\_US\_yrs , mthr\_educ\_yrs , brothers\_hd , sisters\_hd , siblings\_hd , grew\_up\_hd , parents\_income , grew\_up\_nat\_parents , step\_parent , urban\_rural , region , region\_grew\_up , worked\_weeks , drinks\_number , family\_needs , savings\_calc , w1 , w2 , w3 , w4 , w5 , w6 , w7 , w8 , s2 , s3 , s6 , years\_work\_hd , years\_work\_full\_time\_hd , years\_work\_wf , years\_work\_full\_time\_wf , years\_work\_i , years\_work\_sp , years\_work\_full\_time\_i , years\_work\_full\_time\_sp , health\_hd\_change\_2yrs\_ago , health\_hd\_change\_2yrs\_ago\_better , health\_hd\_change\_2yrs\_ago\_worse , health\_hd\_child , work\_limited\_hd , health\_wf\_change\_2yrs\_ago , health\_wf\_change\_2yrs\_ago\_better , health\_wf\_change\_2yrs\_ago\_worse , health\_wf\_child , work\_limited\_wf , health\_i\_change\_2yrs\_ago , health\_i\_change\_2yrs\_ago\_better , health\_i\_change\_2yrs\_ago\_worse , health\_i\_child , work\_limited\_i , fam\_health\_good , Smoke\_hd , drink\_hd , drink\_how\_often\_hd , drinks\_num\_hd , drinks\_45\_num\_days\_hd , Smoke\_wf , drink\_wf , drink\_how\_often\_wf , drinks\_num\_wf , drinks\_45\_num\_days\_wf , Smoke\_i , drink\_i , drink\_how\_often\_i , drinks\_num\_i , physical\_activity\_hvy\_num\_hd , physical\_activity\_hvy\_time\_hd , physical\_activity\_light\_num\_hd , physical\_activity\_light\_time\_hd , physical\_activity\_str\_num\_hd , physical\_activity\_str\_time\_hd , physical\_activity\_hvy\_num\_wf , physical\_activity\_hvy\_time\_wf , physical\_activity\_light\_num\_wf , physical\_activity\_light\_time\_wf , physical\_activity\_str\_num\_wf , physical\_activity\_str\_time\_wf , physical\_activity\_hvy\_num\_i , physical\_activity\_light\_num\_i , physical\_activity\_str\_num\_i , physical\_activity\_hvy\_time\_i , physical\_activity\_light\_time\_i , physical\_activity\_str\_time\_i , physical\_activity\_hvy\_week\_i , physical\_activity\_light\_week\_i , physical\_activity\_str\_week\_i , memory\_loss\_hd , learning\_disorder\_hd , cancer\_hd , memory\_loss\_wf , learning\_disorder\_wf , cancer\_wf , memory\_loss\_i , learning\_disorder\_i , cancer\_i , psych\_problem\_hd , psych\_problem\_age\_hd , psych\_problem\_type\_hd1 , psych\_problem\_type\_hd2 , psych\_problem\_type\_hd3 , psych\_problem\_wf , psych\_problem\_age\_wf , psych\_problem\_type\_wf1 , psych\_problem\_type\_wf2 , psych\_problem\_type\_wf3 , psych\_problem\_i , psych\_problem\_age\_i , psych\_problem\_type\_i1 , psych\_problem\_type\_i2 , psych\_problem\_type\_i3 , weight\_hd , height\_feet\_hd , height\_inches\_hd , height\_hd , weight\_wf , height\_feet\_wf , height\_wf , height\_i , weight\_i , BMI , census\_needs , food\_stamps\_2yrs\_ago , ER52020A , ER52037A , ER52045A , ER52055A , ER52063A , ER52399A , educ\_hs\_hd , employ\_stat\_sp1 , employ\_stat\_sp2 , employ\_stat\_sp3 , worship\_place\_sp , drink\_alchl\_hd , drink\_alchl\_wf , drink\_alchl\_i , drink\_alchl\_sp , IRA\_amt , inv\_cashed , inv\_cashed\_amt , re\_other , re\_other\_amt , weight\_lb\_hd , weight\_kg\_hd , weight\_kglb\_hd , height\_meters\_hd , height\_meters\_inches\_hd , weight\_lb\_wf , weight\_kg\_wf , weight\_kglb\_wf , height\_meters\_wf , height\_meters\_inches\_wf , relig\_i\_fmt , relig\_denom\_i\_fmt , relig\_denom\_i\_fmt1 , relig\_svc\_per\_month\_i\_fmt , relig\_svc\_per\_week\_i\_fmt , mar\_stat\_i\_fmt , mar\_stat\_ch\_fmt , Not\_Affiliated , Roman\_Catholic , Jewish , Mainline\_Protestant , Conservative\_Protestant , Mormon , Other\_Relig , test , relig\_svc\_i\_fmt , Religiosity\_High , Religiosity\_low , Religiosity\_vh , Religiosity\_h , Religiosity\_m\_l , Religiosity\_n , MP\_low , MP\_high , CP\_low , CP\_high , RC\_low , RC\_high , Jewish\_low , Jewish\_high , Mormon\_low , Mormon\_high , Other\_Relig\_low , Other\_Relig\_high , NA\_low , NA\_high , MP\_vh , MP\_h , MP\_m\_l , MP\_n , CP\_vh , CP\_h , CP\_m\_l , CP\_n , RC\_vh , RC\_h , RC\_m\_l , RC\_n , NA\_attend , Jewish\_vh , Jewish\_h , Jewish\_m\_l , Jewish\_n , Mormon\_vh , Mormon\_h , Mormon\_m\_l , Mormon\_n , Other\_Relig\_vh , Other\_Relig\_h , Other\_Relig\_m\_l , Other\_Relig\_n , NA\_vh , NA\_h , NA\_m\_l , NA\_n , region\_Northeast , region\_North\_Central , region\_South , region\_West , region\_Other , child\_under5 , child\_6\_11 , child\_12\_17 , child\_18\_over , Married , Single , Divorced\_Separated , Divorced , Separated , Widowed , health\_i\_fmt , health\_excellent , health\_very\_good , health\_good , health\_fair , health\_poor , under\_weight , normal\_weight , over\_weight , obese\_weight , race\_i\_fmt , White , Black , Other\_race , work\_employed , work\_unemployed , work\_nlf , edu\_lhs , edu\_hs , edu\_sclg , edu\_clg , edu\_mclg , age\_17\_under , age\_18\_24 , age\_25\_34 , age\_35\_44 , age\_45\_54 , age\_55\_64 , age\_65\_74 , age\_75\_over , age\_i\_sq , fam\_inc\_positive , fam\_inc\_log , fam\_inc\_positive\_sq , pov\_rate , pov\_rate\_0\_99 , pov\_rate\_100\_199 , pov\_rate\_200\_299 , pov\_rate\_300\_399 , pov\_rate\_400\_499 , pov\_rate\_500\_599 , pov\_rate\_600\_plus , pov\_rate\_groups , health\_e\_vg , health\_g\_f\_p , health\_e\_vg\_g , health\_f\_p , mental\_health , mental\_health1 , year2005 , year2011 , health\_worse , pension\_contrib\_time , pension\_contrib\_amt\_year , pension\_contrib\_amt , pension\_contrib\_employer\_time , pension\_contrib\_employer\_amt\_yr , pension\_contrib\_employer\_amt , FIPS , state1 , state\_fmt , State\_full , State\_abr , U\_rate\_2005 , U\_rate\_2011 , F6 , F7 , F8 , F9 , F10 , F11 , F12 , F13 , F14 , CivilianLF , CivilianLF\_2005 , CivilianLF\_2011 , Employed , Employed\_2005 , Employed\_2011 , Epop\_ratio , Epop\_\_ratio\_2005 , Epop\_\_ratio\_2011 , Population , Population\_2005 , Population\_2011 , U\_rate , Unemployed , Unemployed\_2005 , Unemployed\_2011 , work\_unemployed\_year2011 , work\_nlf\_year2011 , work\_employed\_year2011 , u\_rate\_year2011 FROM \_TEMP0.'all\_2011(2)'n;

RUN;

QUIT;

PROC DATASETS NOLIST NODETAILS;

CONTENTS DATA=WORK.query OUT=WORK.details;

RUN;

proc contents data=work.query;

run;

%web\_open\_table(work.query);

proc means data=work.query;

run;

PROC SQL;

CREATE TABLE WORK.query AS

SELECT ER27510A , ER27634A , ER27644A , ER27667A , ER28043A , S702A , S703A , S704A , S705A , S706A , S707A , S708A , S709A , S710A , S711A , S713A , S714A , S715A , S718A , S719A , S720 , S720A , S716A , S717A , ER33826A , ER33826B , ER33826C , ER33826D , ER33826E , ER33826F , ER33826G , ER33826H , ER33826I , ER33826J , ER33826K , ER33826L , ER33826M , ER33826N , ER33826O , ER33827A , ER33827B , ER33827C , ER33827D , ER33827E , ER33827F , ER33827G , ER33827H , ER33827I , ER33827J , ER33827K , ER33827L , ER33827M , ER33827N , ER33827O , ER33827P , ER33827Q , ER33827R , ER33827S , ER33827T , ER33827U , ER33827V , ER33828A , ER33828B , ER33828C , ER33828D , ER33828E , ER33828F , ER33828G , ER33828H , ER33828I , ER33828J , ER33828K , ER33828L , ER33833A , ER33833B , ER33833C , ER33833D , ER33833E , ER33833F , ER33833G , ER33833H , 'year'n , 'count'n , famid , personid , upid , in\_us\_68 , sample , gender\_male\_i , sex\_male\_hd , sex\_male\_i , rel\_to\_hd , seq\_num , head , indiv\_weight , respond\_i , respondent , interview\_num\_i , interview\_num\_f , age\_hd , age\_wf , no\_wife , age\_i , age\_sp , nowoman , 'move'n , year\_move\_in , marital\_pair , new\_hd , new\_wf , year\_new\_hd , year\_new\_wf , year\_new\_i , hd\_sp\_stat , split\_num , split\_id , famid\_other , release\_n , dob\_month\_i , dob\_year\_i , empl\_stat\_i , years\_educ\_hd , years\_educ\_wf , years\_educ\_i , years\_educ\_i1 , fam\_health , Hispanic\_hd , race\_hd1 , race\_hd2 , race\_hd3 , race\_hd4 , Hispanic\_wf , race\_wf1 , race\_wf2 , race\_wf3 , race\_wf4 , height\_inches\_wf , Hispanic , race\_i1 , race\_i2 , race\_i3 , race\_i4 , ethnic\_grp\_hd , ethnic\_grp\_wf , ethnic\_grp\_i , ethnic\_grp\_sp , ethnic\_grp\_hd1 , ethnic\_grp\_hd2 , ethnic\_grp\_wf1 , ethnic\_grp\_wf2 , ethnic\_grp\_i1 , ethnic\_grp\_i2 , fam\_size , num\_child , age\_youngest\_child , num\_nonfu\_hu , mar\_stat\_i , mar\_stat\_g , mar\_stat\_ch , cpl\_stat\_hd , educ\_hd , educ\_wf , educ\_i , educ\_sp , employ\_stat\_hd1 , employ\_stat\_hd2 , employ\_stat\_hd3 , employ\_stat\_wf1 , employ\_stat\_wf2 , employ\_stat\_wf3 , employ\_stat\_i1 , employ\_stat\_i2 , employ\_stat\_i3 , religion\_denom\_hd , relig\_hd , relig\_denom\_hd , religion\_denom\_wf , relig\_wf , relig\_denom\_wf , relig\_i , relig\_sp , relig\_denom\_i , relig\_denom\_sp , worship\_place\_hd , worship\_place\_wf , worship\_place\_i , relig\_svc\_num\_hd , relig\_svc\_num\_wf , relig\_svc\_time\_hd , relig\_svc\_time\_wf , relig\_svc\_num\_i , relig\_svc\_time\_i , relig\_svc\_per\_month\_i , relig\_svc\_per\_week\_i , relig\_donate , relig\_donate\_amt , 'state'n , geogr\_mobility , state\_grew\_hd , state\_other\_ever\_live , moved\_hd , moved\_reason\_hd , health\_hd , health\_wf , health\_i , health\_sp , hsewrk\_hrs\_hd , hsewrk\_hrs\_wf , hsewrk\_hrs\_i , hsewrk\_hrs\_sp , meal\_together , earn\_hd , earn\_wf , earn\_i , earn\_sp , wage\_hd , wage\_wf , wage\_i , wage\_sp , wage\_salary\_hd , wage\_salary\_wf , wage\_salary\_i , wage\_salary\_sp , fam\_inc , transfer\_inc\_hdwf , tax\_inc\_hdwf , tax\_itemize , wealth\_he\_no , wealth\_he\_yes , savings , savings\_amt , business\_farm\_own , business\_farm\_amt , debt\_cc , debt\_cc\_amt , re\_other\_own , re\_other\_own\_amt , stock\_own , stock\_own\_amt , veh\_amt , savings\_other , savings\_other\_amt , IRA\_own , IRA\_own\_amt , bonds , bonds\_amt , home\_ownership , home\_value , mortgage , mortgage\_type , mortgage\_amt , mortgage\_yr , mortgage\_2 , mortgage2\_type , mortgage2\_amt , debts\_other , debts\_other\_amt , IRA\_changed , IRA\_changed\_amt , pension\_cashed , pension\_cashed\_amt , home\_sold , home\_sold\_amt , re\_other\_bought , re\_other\_bought\_amt , re\_other\_sold , re\_other\_sold\_amt , re\_additions , re\_additions\_amt , business\_farm\_inv , business\_farm\_inv\_amt , business\_farm\_sold , business\_farm\_sold\_amt , business\_farm\_saving\_amt , stock\_bought , stock\_sold , stock\_more\_sold , stock\_amt , stock\_net\_amt , stock\_saving\_amt , stock\_inv\_amt , noIRA\_stock\_sold , noIRA\_stock\_amt , gift\_inher , gift\_inher\_yr , gift\_inher\_amt , gift\_inher2 , gift\_inher\_yr2 , gift\_inher\_amt2 , gift\_inher3 , gift\_inher\_yr3 , gift\_inher\_amt3 , pension\_job , pension\_yrs , pension\_yr\_joined , pension\_elig , pension\_yrs\_elig , pension\_vested , pension\_yrs\_vested , pension\_contr , pension\_contr\_req , pension\_contr\_req\_amt , pension\_contr\_req\_amt\_unit , pension\_contr\_vol , pension\_contr\_v\_amt , pension\_contr\_vol\_amt , pension\_contr\_empl , pension\_contr\_e\_amt , pension\_contr\_empl\_amt , pension\_now\_amt , fam\_comp , fth\_born\_US , fth\_educ\_foreign\_yrs , fth\_educ\_US , fth\_educ\_US\_yrs , fth\_educ\_yrs , mthr\_born\_US , mthr\_educ\_foreign\_yrs , mthr\_educ\_US , mthr\_educ\_US\_yrs , mthr\_educ\_yrs , brothers\_hd , sisters\_hd , siblings\_hd , grew\_up\_hd , parents\_income , grew\_up\_nat\_parents , step\_parent , urban\_rural , region , region\_grew\_up , worked\_weeks , drinks\_number , family\_needs , savings\_calc , w1 , w2 , w3 , w4 , w5 , w6 , w7 , w8 , s2 , s3 , s6 , years\_work\_hd , years\_work\_full\_time\_hd , years\_work\_wf , years\_work\_full\_time\_wf , years\_work\_i , years\_work\_sp , years\_work\_full\_time\_i , years\_work\_full\_time\_sp , health\_hd\_change\_2yrs\_ago , health\_hd\_change\_2yrs\_ago\_better , health\_hd\_change\_2yrs\_ago\_worse , health\_hd\_child , work\_limited\_hd , health\_wf\_change\_2yrs\_ago , health\_wf\_change\_2yrs\_ago\_better , health\_wf\_change\_2yrs\_ago\_worse , health\_wf\_child , work\_limited\_wf , health\_i\_change\_2yrs\_ago , health\_i\_change\_2yrs\_ago\_better , health\_i\_change\_2yrs\_ago\_worse , health\_i\_child , work\_limited\_i , fam\_health\_good , Smoke\_hd , drink\_hd , drink\_how\_often\_hd , drinks\_num\_hd , drinks\_45\_num\_days\_hd , Smoke\_wf , drink\_wf , drink\_how\_often\_wf , drinks\_num\_wf , drinks\_45\_num\_days\_wf , Smoke\_i , drink\_i , drink\_how\_often\_i , drinks\_num\_i , physical\_activity\_hvy\_num\_hd , physical\_activity\_hvy\_time\_hd , physical\_activity\_light\_num\_hd , physical\_activity\_light\_time\_hd , physical\_activity\_str\_num\_hd , physical\_activity\_str\_time\_hd , physical\_activity\_hvy\_num\_wf , physical\_activity\_hvy\_time\_wf , physical\_activity\_light\_num\_wf , physical\_activity\_light\_time\_wf , physical\_activity\_str\_num\_wf , physical\_activity\_str\_time\_wf , physical\_activity\_hvy\_num\_i , physical\_activity\_light\_num\_i , physical\_activity\_str\_num\_i , physical\_activity\_hvy\_time\_i , physical\_activity\_light\_time\_i , physical\_activity\_str\_time\_i , physical\_activity\_hvy\_week\_i , physical\_activity\_light\_week\_i , physical\_activity\_str\_week\_i , memory\_loss\_hd , learning\_disorder\_hd , cancer\_hd , memory\_loss\_wf , learning\_disorder\_wf , cancer\_wf , memory\_loss\_i , learning\_disorder\_i , cancer\_i , psych\_problem\_hd , psych\_problem\_age\_hd , psych\_problem\_type\_hd1 , psych\_problem\_type\_hd2 , psych\_problem\_type\_hd3 , psych\_problem\_wf , psych\_problem\_age\_wf , psych\_problem\_type\_wf1 , psych\_problem\_type\_wf2 , psych\_problem\_type\_wf3 , psych\_problem\_i , psych\_problem\_age\_i , psych\_problem\_type\_i1 , psych\_problem\_type\_i2 , psych\_problem\_type\_i3 , weight\_hd , height\_feet\_hd , height\_inches\_hd , height\_hd , weight\_wf , height\_feet\_wf , height\_wf , height\_i , weight\_i , BMI , census\_needs , food\_stamps\_2yrs\_ago , ER52020A , ER52037A , ER52045A , ER52055A , ER52063A , ER52399A , educ\_hs\_hd , employ\_stat\_sp1 , employ\_stat\_sp2 , employ\_stat\_sp3 , worship\_place\_sp , drink\_alchl\_hd , drink\_alchl\_wf , drink\_alchl\_i , drink\_alchl\_sp , IRA\_amt , inv\_cashed , inv\_cashed\_amt , re\_other , re\_other\_amt , weight\_lb\_hd , weight\_kg\_hd , weight\_kglb\_hd , height\_meters\_hd , height\_meters\_inches\_hd , weight\_lb\_wf , weight\_kg\_wf , weight\_kglb\_wf , height\_meters\_wf , height\_meters\_inches\_wf , relig\_i\_fmt , relig\_denom\_i\_fmt , relig\_denom\_i\_fmt1 , relig\_svc\_per\_month\_i\_fmt , relig\_svc\_per\_week\_i\_fmt , mar\_stat\_i\_fmt , mar\_stat\_ch\_fmt , Not\_Affiliated , Roman\_Catholic , Jewish , Mainline\_Protestant , Conservative\_Protestant , Mormon , Other\_Relig , test , relig\_svc\_i\_fmt , Religiosity\_High , Religiosity\_low , Religiosity\_vh , Religiosity\_h , Religiosity\_m\_l , Religiosity\_n , MP\_low , MP\_high , CP\_low , CP\_high , RC\_low , RC\_high , Jewish\_low , Jewish\_high , Mormon\_low , Mormon\_high , Other\_Relig\_low , Other\_Relig\_high , NA\_low , NA\_high , MP\_vh , MP\_h , MP\_m\_l , MP\_n , CP\_vh , CP\_h , CP\_m\_l , CP\_n , RC\_vh , RC\_h , RC\_m\_l , RC\_n , NA\_attend , Jewish\_vh , Jewish\_h , Jewish\_m\_l , Jewish\_n , Mormon\_vh , Mormon\_h , Mormon\_m\_l , Mormon\_n , Other\_Relig\_vh , Other\_Relig\_h , Other\_Relig\_m\_l , Other\_Relig\_n , NA\_vh , NA\_h , NA\_m\_l , NA\_n , region\_Northeast , region\_North\_Central , region\_South , region\_West , region\_Other , child\_under5 , child\_6\_11 , child\_12\_17 , child\_18\_over , Married , Single , Divorced\_Separated , Divorced , Separated , Widowed , health\_i\_fmt , health\_excellent , health\_very\_good , health\_good , health\_fair , health\_poor , under\_weight , normal\_weight , over\_weight , obese\_weight , race\_i\_fmt , White , Black , Other\_race , work\_employed , work\_unemployed , work\_nlf , edu\_lhs , edu\_hs , edu\_sclg , edu\_clg , edu\_mclg , age\_17\_under , age\_18\_24 , age\_25\_34 , age\_35\_44 , age\_45\_54 , age\_55\_64 , age\_65\_74 , age\_75\_over , age\_i\_sq , fam\_inc\_positive , fam\_inc\_log , fam\_inc\_positive\_sq , pov\_rate , pov\_rate\_0\_99 , pov\_rate\_100\_199 , pov\_rate\_200\_299 , pov\_rate\_300\_399 , pov\_rate\_400\_499 , pov\_rate\_500\_599 , pov\_rate\_600\_plus , pov\_rate\_groups , health\_e\_vg , health\_g\_f\_p , health\_e\_vg\_g , health\_f\_p , mental\_health , mental\_health1 , year2005 , year2011 , health\_worse , pension\_contrib\_time , pension\_contrib\_amt\_year , pension\_contrib\_amt , pension\_contrib\_employer\_time , pension\_contrib\_employer\_amt\_yr , pension\_contrib\_employer\_amt , FIPS , state1 , state\_fmt , State\_full , State\_abr , U\_rate\_2005 , U\_rate\_2011 , F6 , F7 , F8 , F9 , F10 , F11 , F12 , F13 , F14 , CivilianLF , CivilianLF\_2005 , CivilianLF\_2011 , Employed , Employed\_2005 , Employed\_2011 , Epop\_ratio , Epop\_\_ratio\_2005 , Epop\_\_ratio\_2011 , Population , Population\_2005 , Population\_2011 , U\_rate , Unemployed , Unemployed\_2005 , Unemployed\_2011 , work\_unemployed\_year2011 , work\_nlf\_year2011 , work\_employed\_year2011 , u\_rate\_year2011 FROM \_TEMP0.'all\_2011(2)'n;

RUN;

QUIT;

PROC DATASETS NOLIST NODETAILS;

CONTENTS DATA=WORK.query OUT=WORK.details;

RUN;

proc contents data=work.query;

run;

%web\_open\_table(work.query);

proc means data=work.query;

run;

data work.query; set work.query;

if age\_i=0 then delete;

if age\_i=1 then delete;

if age\_i=2 then delete;

if age\_i=3 then delete;

if age\_i=4 then delete;

if age\_i=5 then delete;

if age\_i=6 then delete;

if age\_i=7 then delete;

if age\_i=8 then delete;

if age\_i=9 then delete;

if age\_i=10 then delete;

if age\_i=11 then delete;

if age\_i=12 then delete;

if age\_i=13 then delete;

if age\_i=14 then delete;

if age\_i=15 then delete;

if age\_i=16 then delete;

if age\_i=17 then delete;

if age\_i=18 then delete;

if age\_i=19 then delete;

if age\_i=20 then delete;

if age\_i=21 then delete;

if age\_i=22 then delete;

if age\_i=65 then delete;

if age\_i=66 then delete;

if age\_i=67 then delete;

if age\_i=68 then delete;

if age\_i=69 then delete;

if age\_i=70 then delete;

if age\_i=71 then delete;

if age\_i=72 then delete;

if age\_i=73 then delete;

if age\_i=74 then delete;

if age\_i=75 then delete;

if age\_i=76 then delete;

if age\_i=77 then delete;

if age\_i=78 then delete;

if age\_i=79 then delete;

if age\_i=80 then delete;

if age\_i=81 then delete;

if age\_i=82 then delete;

if age\_i=83 then delete;

if age\_i=84 then delete;

if age\_i=85 then delete;

if age\_i=86 then delete;

if age\_i=87 then delete;

if age\_i=88 then delete;

if age\_i=89 then delete;

if age\_i=90 then delete;

if age\_i=91 then delete;

if age\_i=92 then delete;

if age\_i=93 then delete;

if age\_i=94 then delete;

if age\_i=95 then delete;

if age\_i=96 then delete;

if age\_i=97 then delete;

if age\_i=98 then delete;

if age\_i=99 then delete;

if age\_i=100 then delete;

if age\_i=101 then delete;

if age\_i=102 then delete;

if age\_i=103 then delete;

if age\_i=104 then delete;

if age\_i=105 then delete;

if age\_i=106 then delete;

if age\_i=107 then delete;

if age\_i=108 then delete;

if age\_i=109 then delete;

if age\_i=110 then delete;

if age\_i=111 then delete;

if age\_i=112 then delete;

if age\_i=113 then delete;

if age\_i=114 then delete;

if age\_i=115 then delete;

if age\_i=116 then delete;

if age\_i=117 then delete;

if age\_i=118 then delete;

if age\_i=119 then delete;

if age\_i=120 then delete;

run;

data work.query; set work.query;

if age\_youngest\_child=1 then YoungChild=1;

else if age\_youngest\_child=2 then YoungChild=1;

else if age\_youngest\_child=3 then YoungChild=1;

else if age\_youngest\_child=4 then YoungChild=1;

else if age\_youngest\_child=5 then YoungChild=1;

else if age\_youngest\_child=6 then YoungChild=0;

else if age\_youngest\_child=7 then YoungChild=0;

else if age\_youngest\_child=8 then YoungChild=0;

else if age\_youngest\_child=9 then YoungChild=0;

else if age\_youngest\_child=10 then YoungChild=0;

else if age\_youngest\_child=11 then YoungChild=0;

else if age\_youngest\_child=12 then YoungChild=0;

else if age\_youngest\_child=13 then YoungChild=0;

else if age\_youngest\_child=14 then YoungChild=0;

else if age\_youngest\_child=15 then YoungChild=0;

else if age\_youngest\_child=16 then YoungChild=0;

else if age\_youngest\_child=17 then YoungChild=0;

run;

data work.query; set work.query;

age2 = age\_i \* age\_i;

run;

data work.query; set work.query;

if meal\_together=8 then delete;

if meal\_together=9 then delete;

run;

data work.query; set work.query;

if hsewrk\_hrs\_i=998 then delete;

if hsewrk\_hrs\_i=999 then delete;

if hsrwrk\_hrs\_i=0 then delete;

run;

data work.query; set work.query;

drinks2 = drink\_how\_often\_i \* drink\_how\_often\_i;

run;

data work.query; set work.query;

if sex\_male\_i=0 then female=1;

else if sex\_male\_i=1 then female=0;

run;

data work.query; set work.query;

HouseworkW = female \* hsewrk\_hrs\_i;

run;

data work.query; set work.query;

MealW = female \* meal\_together;

run;

data work.query; set work.query;

HouseworkM = sex\_male\_i \* hsewrk\_hrs\_i;

run;

data work.query; set work.query;

MealM = sex\_male\_i \* meal\_together;

run;

data work.query; set work.query;

lnearn\_i = log(earn\_i);

run;

data work.query; set work.query;

age2 = age\_i \* age\_i;

run;

data work.query; set work.query;

if meal\_together=8 then delete;

if meal\_together=9 then delete;

run;

data work.query; set work.query;

if hsewrk\_hrs\_i=998 then delete;

if hsewrk\_hrs\_i=999 then delete;

if hsrwrk\_hrs\_i=0 then delete;

run;

data work.query; set work.query;

drinks2 = drink\_how\_often\_i \* drink\_how\_often\_i;

run;

data work.query; set work.query;

if meal\_together = . then delete;

run;

data work.query; set work.query;

if youngchild = . then delete;

run;

data work.query; set work.query;

if meal\_together=. then delete;

run;

proc corr data= work.query;

var lnearn\_i meal\_together youngchild age\_youngest\_child MealW female houseworkw age2 age\_i years\_educ\_i married religiosity\_high num\_child hsewrk\_hrs\_i;

run;

proc reg data=work.query;

model lnearn\_i= family\_needs food\_stamps\_2yrs\_ago meal\_together youngchild age\_youngest\_child MealW female houseworkw age2 age\_i years\_educ\_i married religiosity\_high num\_child hsewrk\_hrs\_i /vif;

run;

proc reg data=work.query;

model lnearn\_i= meal\_together hsewrk\_hrs\_i youngchild age\_youngest\_child MealW female houseworkw age2 age\_i years\_educ\_i married religiosity\_high num\_child /white;

run;

proc reg data= work.query;

model lnearn\_i= meal\_together youngchild age\_youngest\_child MealW female houseworkw age2 age\_i years\_educ\_i married religiosity\_high num\_child hsewrk\_hrs\_i /spec;

run;

proc reg data= work.query;

model lnearn\_i= meal\_together youngchild age\_youngest\_child MealW female houseworkw age2 age\_i years\_educ\_i married religiosity\_high num\_child hsewrk\_hrs\_i;

Run;

**Robustness Check Coding**

Everything is the same as the coding above, however note these changes:

**Delete these entries**

data work.query; set work.query;

lnearn\_i = log(earn\_i);

run;

proc corr data= work.query;

var lnearn\_i meal\_together youngchild age\_youngest\_child MealW female houseworkw age2 age\_i years\_educ\_i married religiosity\_high num\_child hsewrk\_hrs\_i;

run;

proc reg data=work.query;

model lnearn\_i= family\_needs food\_stamps\_2yrs\_ago meal\_together youngchild age\_youngest\_child MealW female houseworkw age2 age\_i years\_educ\_i married religiosity\_high num\_child hsewrk\_hrs\_i /vif;

run;

proc reg data=work.query;

model lnearn\_i= meal\_together hsewrk\_hrs\_i youngchild age\_youngest\_child MealW female houseworkw age2 age\_i years\_educ\_i married religiosity\_high num\_child /white;

run;

proc reg data= work.query;

model lnearn\_i= meal\_together youngchild age\_youngest\_child MealW female houseworkw age2 age\_i years\_educ\_i married religiosity\_high num\_child hsewrk\_hrs\_i /spec;

run;

proc reg data= work.query;

model lnearn\_i= meal\_together youngchild age\_youngest\_child MealW female houseworkw age2 age\_i years\_educ\_i married religiosity\_high num\_child hsewrk\_hrs\_i;

Run;

**Replace with**

data work.query; set work.query;

lnwage\_i = log(wage\_i);

run;

proc corr data= work.query;

var lnwage\_i meal\_together youngchild age\_youngest\_child MealW female houseworkw age2 age\_i years\_educ\_i married religiosity\_high num\_child hsewrk\_hrs\_i;

run;

proc reg data=work.query;

model lnwage\_i= family\_needs food\_stamps\_2yrs\_ago meal\_together youngchild age\_youngest\_child MealW female houseworkw age2 age\_i years\_educ\_i married religiosity\_high num\_child hsewrk\_hrs\_i /vif;

run;

proc reg data=work.query;

model lnwage\_i= meal\_together hsewrk\_hrs\_i youngchild age\_youngest\_child MealW female houseworkw age2 age\_i years\_educ\_i married religiosity\_high num\_child /white;

run;

proc reg data= work.query;

model lnwage\_i= meal\_together youngchild age\_youngest\_child MealW female houseworkw age2 age\_i years\_educ\_i married religiosity\_high num\_child hsewrk\_hrs\_i /spec;

run;

proc reg data= work.query;

model lnwage\_i= meal\_together youngchild age\_youngest\_child MealW female houseworkw age2 age\_i years\_educ\_i married religiosity\_high num\_child hsewrk\_hrs\_i;

Run;