

Lamberto Alberici

Professor Barreto

May 7th, 2025

Submitted in partial fulfilment of the requirements for BUSA 305: Regression with Microdata.

Do Mothers Face a Wage Penalty?

A Regression Analysis in 2025

Abstract

This paper investigates whether mothers in the U.S. faced a wage penalty in 2024 using regression analysis on microdata from the IPUMS CPS ASEC survey. The analysis models “LN_Inc wage” as a function of “Num_Children”, controlling for education, experience, hours worked, marital status, race, ethnicity, and occupation. While initial regressions indicate a negative correlation between number of children and income, the effect is not statistically significant when survey weights are applied. The findings suggest no conclusive evidence of a motherhood wage penalty, though limitations such as omitted variable bias and sampling concerns remain.

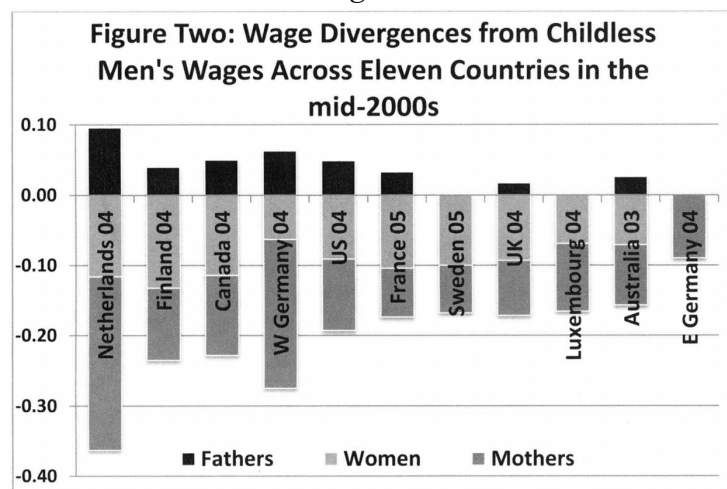
I. Introduction

The research question for this paper will be: “Do mothers suffer a wage penalty?” It comes as no surprise that gender discrimination is still present in today’s society, and I wanted to focus on how mothers get unfairly treated in the workplace compared to fathers. The previous studies that I found for this topic date to 10 to 15 years ago, making me wonder how the situation has changed in recent years. This paper will go over a literature review of past and related studies on the topic to get a basic understanding of background information. After it will jump into an empirical analysis, which will be split into two sections: the data gathering process and the result presentation and explanation. Lastly, it will provide a conclusion which will tie the research paper together.

II. Literature Review

The question that this paper will consider is: “Do Mothers Face a Wage Penalty?” The common conception is that when an individual becomes a parent, they must give attention to the child, whereas an individual who is solely based on work can direct their whole efforts to it. This statement serves as a basic explanation of why there is a difference in wages between parents and non-parents, but what Misra and Strader (2013) have found is that it is not as simple as that.

Figure 1



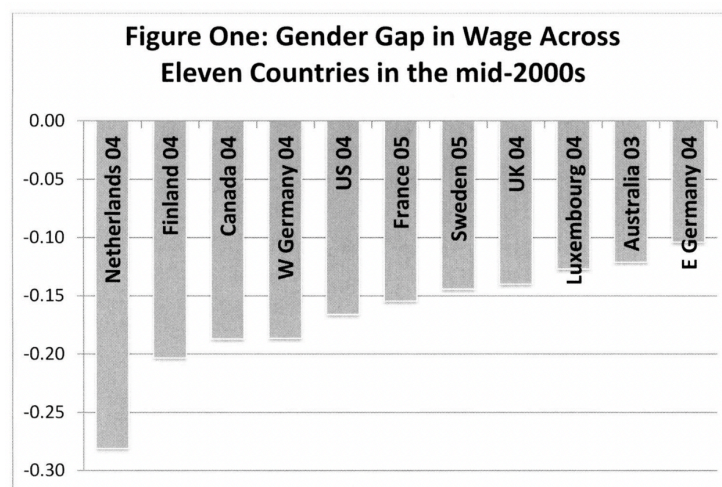
Source: Author's data analysis, Luxembourg Income Study Database (LIS).

Misra and Strader (2013, p. 32)

The study was conducted with data from the years 2004-2006, and it showed clear results that the difference in wages between parents and non-parents heavily depends on gender: as shown above in Figure 1, all of the countries taken for the study showed evidence of motherhood wage penalties, while most countries showed Fatherhood premiums. Misra and Strader's (2013) results portray a distinct difference from the claim that all parents will receive a wage penalty due to them having less time for work; the statement is only true when talking about mothers, as they are often seen and perceived as the primary caretaker of the child, but is that really the case? Misra and Strader (2013) also analysed the gender pay gap between 2004 and 2006 for the same

countries, as shown in Figure 2 below, where the parenthood penalties/premiums were recorded. The results show that the countries which have the most pay gap are also the ones which display the most motherhood wage penalty and the most fatherhood Premium, showing a strong correlation between gender pay gap and parenthood penalty/premium. This leads to the conclusion that the statement “Mothers experience a wage penalty due to them having to take care of their children” is heavily stigmatised.

Figure 2



Source: Author's data analysis, Luxembourg Income Study Database (LIS).

Misra and Strader (2013, p. 31)

In a different paper, Budig and Hodges (2010) have an interesting take on motherhood wage penalties based on their original wages. While Budig and Hodges (2010) don't necessarily answer the question “Do Mothers Face a Wage Penalty?” They assume that there is some type of penalty or premium, and they try to find a correlation between mothers and non-mothers across the earnings distribution. The article's conclusion is that “the wage penalty for motherhood is proportionally largest for the lowest-paid workers” (Budig et al., 2010), adding yet another layer to the motherhood penalties/premiums. This goes to show the complexity of the research

question of this paper and how different real-life aspects come into play when determining motherhood wage penalties.

One of the previous article's main limitations was that it only focuses on the study of White women's wages, whereas Glauber (2007) focuses on the motherhood wage penalty across White, African American and Hispanic women. The results from the paper found that "African American and Hispanic Mothers pay a smaller wage penalty than White mothers" (Galuber, 2007), which is another interesting point of interpretation on motherhood wage penalties. These articles have shown that there is a penalty/premium of parenthood and that there are a lot more niche analyses that can be done in order to really understand where the motherhood penalty/premium stems from. Galuber (2007), on top of taking into account different ethnicities/races, also takes into consideration the marital status of mothers and non-mothers. Interestingly, the articles showed that married mothers and unmarried mothers have a pay gap, but the motherhood wage penalty remains the same.

A lot of focus on related articles is around the motherhood wage penalty; on the other hand, Killwead (2013) focuses on the Fatherhood Premium. Fatherhood Premium is a strange concept. Why would having a child result in a wage premium for men while women receive a wage penalty? Killwead's (2013) results show that married fathers receive a wage bonus compared to single and childless fathers. If the stigmatised reason as to why mothers receive a wage penalty is that they lack commitment to their work, since they have to take care of their child, what is the commonly believed reason as to why men receive premiums? Possible explanations include that employers view fathers as more stable, committed, and responsible or that fathers do actually

increase their work hours, leading to wage increases, since they have to provide for their children, but that raises the question: why isn't that the case for women? Ultimately, it seems that gender norms in parenting lead to respective premiums and penalties for fathers and mothers.

Although a lot of other research and studies have been conducted on the difference in wage between parents and non-parents, there are still a lot of different variables that can be introduced to explain the results in more detail, an example could be that of comparing fathers with biological sons versus fathers who adopted their children. This paper will aim to determine the real effect on wages for mothers versus non-mothers.

III. Theoretical Section

For this paper's regression analysis, the dependent variable will always be "LN_Incwage," which translates to a semi-log functional form. The benefits of using a semi-log functional form over a standard linear form are the following: interpreting the coefficients as percentage changes, reducing heteroskedasticity, and capturing nonlinear relationships like diminishing returns. As for the independent variable, there are various ways to handle that, for instance, the variable could just be "Mother." Just regressing "Mother" on wages doesn't give the most detailed results, as the category "Mother" is quite broad; the number and age of the sons and daughters matter. A possible approach would be to use the age of the youngest child in the household as part of the independent variable, but that can attract a lot of selection and omitted variable bias. For the purpose of this paper, the independent variable "Num_Children" will be used to determine the impact each child has on wages.

In determining the control variables for the models of this paper, it is best to refer to the literature review and Becker's (1962) article about "Human Capital Theory". Becker's (1962) theoretical model highlights the importance of Education on Income, where higher levels of education are expected to give higher income due to greater productivity. Becker (1962) talks about other explanatory variables, but the main ones that will be included in this paper as control variables will be: experience, race & ethnicity, marital status, and occupation. Glauber's (2007) article agrees with Becker (1962) on the importance of race and ethnicity when assessing questions regarding income or wages, showing how important it is to keep race & ethnicity as a control variable. To conclude, all control variables are in place to prevent omitted variable bias as best as possible when answering the question "Do Mothers Face a Wage Penalty?"

IV. Empirical Analysis

Data

Source of Data

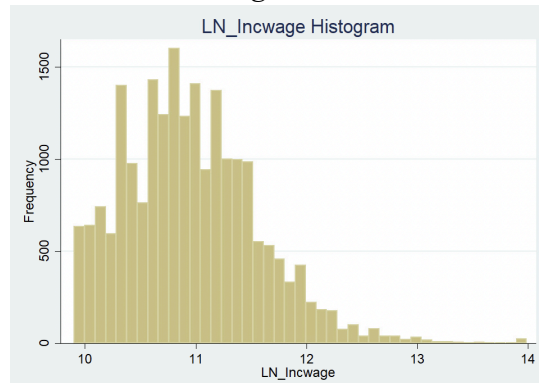
All the data for this paper were retrieved from IPUMS CPS. IPUMS CPS' official website describes itself as “an integrated set of data spanning more than 50 years (1962-forward) of the Current Population Survey (CPS). The CPS is a monthly U.S. household survey conducted jointly by the U.S. Census Bureau and the Bureau of Labor Statistics.” (“IPUMS CPS: FAQ,” 2025) Through IPUMS CPS, the paper acquires all the data from the CPS Annual Social and Economic (ASEC) Supplement 2024, which includes all respondents from the March Basic Monthly Survey.

Dependent Variable

LN_Inc wage

“LN_Inc wage” is the dependent variable, and as the name suggests, it is the natural Logarithm of Inc wage. It serves to linearise the relationship between Inc wage and the rest of the variables. All individuals in the data set with an Inc wage lower than 20,000\$ were removed to improve the model fit and to remove outliers. Figure 3 shows that LN_Inc wage has a skewed distribution; this is due to the extremely high earners, whose difference from the mean is higher than that of the low earners.

Figure 3

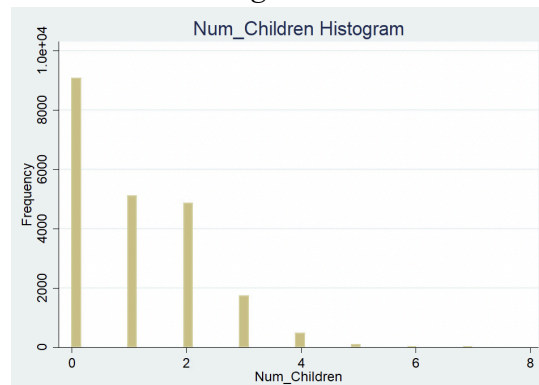


Independent Variable

Num_Children

“Num_Children” stands for Number of Children residing with each individual, where biological, adopted, and step children are included. “Num_Children” is the main variable that this paper will focus on, as it directly relates to the topic of the motherhood wage penalty. Similar to “LN_Incwage”, “Num_Children” has a skewed distribution as shown in Figure 4.

Figure 4



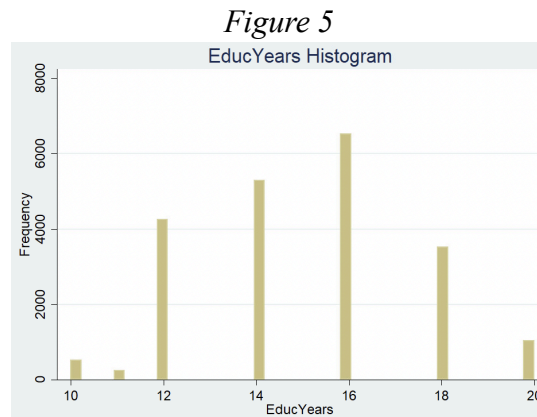
Control Variables

EducYears

“EducYears” is a recoded variable from the “Educ” variable from IPUMS CPS. “EducYears” refers to how many years of education an individual has had, ranging from 10 to 20 years.

“EducYears” definitely has an impact on “LN_Incwage”, since it is common knowledge that, on

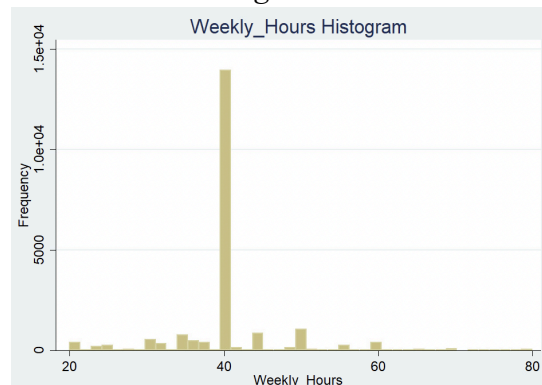
average, individuals who have more schooling will have a higher income, therefore, it must be kept constant in the regression. The histogram in Figure 5 shows that “EducYears” mainly fits the normally distributed model.



Weekly_Hours

“Weekly_Hours” represents the number of hours worked weekly by an Individual. All individuals with less than 20 or more than 100 “Weekly_Hours” were removed from the data set because they could be referring to retirees and students on one hand, and on the flip side, reports of above 100 “Weekly_Hours” are often unrealistic or reporting errors; to put it into perspective, 100 “Weekly_Hours” translates to just above 14 daily hours. Figure 6 shows that the vast majority of individuals in the data set work 40 hours per week, which is the standard “9 to 5” work schedule. “Weekly_Hours” has an impact on how much an individual earns, therefore, it needs to be kept constant in the regression

Figure 6

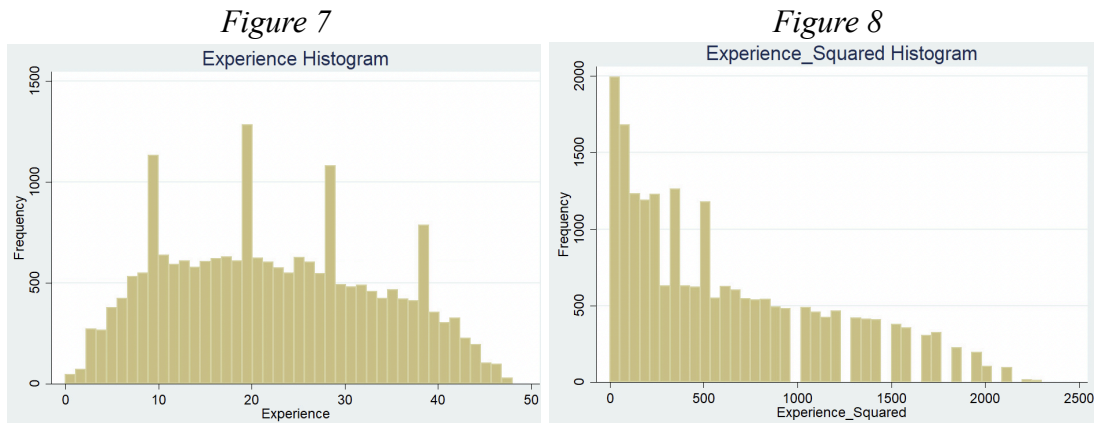


Experience and Experience_Squared

“Experience” is calculated by using the following formula: $Experience = Age - EducYears - 6$.

This equation takes into consideration how many years an individual has spent in the workforce by taking their age and subtracting the years they spent in school. Furthermore, 6 is subtracted, which is supposed to be the first 6 years of an individual’s life, during which they weren’t subject to the education system. Using this formula leads to rare individuals who have -1 “experience”; this, obviously, doesn’t make sense so the -1 was changed to 0. The individual with -1

“experience” could be child geniuses who started schooling earlier than most, resulting in an erroneous subtraction with the minus 6. “Experience_Squared” is in place to showcase how, over the course of time, the importance of Experience diminishes. Figure 7 shows that “Experience” is normally distributed, with the exception of the values 9, 19, 28, and 38 years, which have substantially more observations. Figure 8 shows that the histogram for “Experience_Squared” is skewed to the right due to the fact that squaring larger numbers increases them exponentially.



Marital Status

For “Marital Status”, there isn’t an exact variable; the use of dummy variables keeps constant the possible meanings of the IPUMS CPS variable “MARST.” “MARST” ranges from 1 to 7, with each number signifying a different marital status; with the use of STATA, “MARST” can be included in the regression with all the proper dummy variables by using “i.MARST” as a variable.

Race and Ethnicity

For “Race and Ethnicity”, the concept is more or less the same as marital status: there is a need for dummy variables. The variable “Race” showcases roughly 30 different races, which need to be broken down into dummy variables, while the variable “Hispan” has roughly 15 different Hispanic ethnicities. While the use of the variable “Race” to explain the effect of Race on “LN_Inc wage” is self-explanatory, using “Hispan” seems weird. The reason why only Hispanic ethnicities are included is that the IPUMS CPS doesn’t report any other ethnicities, meaning that “Hispan” is the only way to keep an ethnicity constant.

Occupation

The last control variable is “Occupation”. “Occupation” is another categorical variable with numerous possible numbers that have no numerical meaning; they only stand for an Occupation,

for example, Occ = 0010 stands for “Chief Executives.” “Occupation” has roughly 500 different values, meaning that it is important to keep it constant in the regression.

Additional Information and Limitations

Another variable that isn’t included in the regression is sex, and that’s solely because the data set was manipulated in a way that no men are present, so the focus can only be on mothers and non-mothers. A possible limitation to getting data from IPUMS CPS is that all the individuals who reply to the survey can lie if they wanted to, resulting in false data; adjusting the data set so that all individuals with less than \$20,000 “Incwage” and more than 100 “Weekly_Hours” is a way to try and filter those faulty observations out. Ultimately, the sample size of roughly 21,000 should be enough to remove the bias from those individuals. Another possible limitation to getting data from the CPS ASEC 2024 would be that it doesn’t have a way of showing change and trends over the years. Lastly, Figure 9 represents the summary statistics for most of the variables. Variables such as Marital Status and Occupation weren’t included as their numerical values have no real meaning.

Figure 9: Summary Statistics of Variables

	mean	sd	min	max
LN_Incwage	10.970	0.611	9.903	13.993
Num_Children	1.056	1.133	0.000	7.000
EducYears	15.027	2.415	10.000	20.000
Weekly_Hours	40.401	7.194	20.000	80.000
Experience	22.354	11.151	0.000	48.000
Experience~d	624.045	534.455	0.000	2304.000

Presentation and Interpretation of Results

Data Generating Process (DGP) and Error Term

This paper uses regression analysis to answer the question “Do mothers still suffer a wage penalty?” As stated in the data section, the dependent variable is “LN_Incwage”, while the independent variable is “Num_Children.” This essentially means that regressing “Num_Children” on “LN_Incwage” should be able to provide enough evidence to answer the question while keeping control variables constant. The DGP used for this paper takes data from IPUMS CPS, therefore, it takes a sample of the US population. The sampling used by IPUMS CPS is not simple random sampling, as that would imply that all observations would have the same weight; instead, IPUMS CPS uses a complicated weighted random sampling technique. The reasoning behind this is simple: the population distribution across different states in the US is not uniformly distributed; some US states, such as Wyoming, have populations of roughly 600,000 inhabitants, which is a far cry from the approximately 8 million inhabitants of New York City. To conclude, this means that an individual from Wyoming is going to have a higher weight than an individual from New York City because it implies that the individual from Wyoming is representative of more people.

After obtaining the raw sample data from IPUMS CPS, other parameters were implemented to provide a sample with the most relevant data; the changes that were applied are described in the “Data” section above. The table shown later in Figure 10 displays the three regression models run to determine whether mothers face a wage penalty. Model 1 is a simple OLS bivariate regression of “Number of Children” on “LN_Incwage,” whereas Model 2 is an OLS regression of “Number of Children” on “LN_Incwage” with all the other control variables, listed in the

“Data” section, kept constant. Model 3 is a more accurate representation of Model 2 as it applies the ASECWT, “a person-level weight that should be used in analyses of individual-level CPS supplement data” (IPUMS CPS: ASECWT, 2025), to Model 2. Furthermore, Model 2 and Model 3 hold “Race & Ethnicity,” “Marital Status,” and “Occupation” constant, even though the dummy variables are not shown in the table. Lastly, in all models, the dependent variable is “LN_Inc wage”, which implies that semi-log regression is used, meaning that the coefficients of the X variables represent the percentage change in “Inc wage” per unit change in the X variable.

For all these specific functional forms, the error term (ε) accounts for all unobserved variables that may influence “LN_Inc wage” but are not explicitly included in the analysis. Variables such as personal drive and public speaking skills would definitely have some effect on “LN_Inc wage”, but there is no real way to measure it. Furthermore, there are just too many variables to be included, so the analysis tried to include the major explanatory variables for “LN_Inc wage” as control variables. On top of taking into account all the omitted variables, the error term also serves to explain the DGP. The DGP is created from IPUMS CPS, therefore, it must take into account the imperfections when it comes to gathering data; it being a survey means that individuals may, voluntarily or not, report false information. Additionally, since the IPUMS CPS is a sample, it means that there is some sort of random sampling bias/error, which is taken into account by the error term in the functional form.

Regression Models Analysis

Figure 10

Regressions Number of Children on LN_Incwage			
	(1) Model 1	(2) Model 2	(3) Model 3
Num_Children	-0.0134*** (0.0037) [-0.0207,-0.0061]	-0.0063* (0.0032) [-0.0125,-0.0001]	-0.0025 (0.0037) [-0.0098,0.0047]
EducYears		0.0763*** (0.0021) [0.0722,0.0805]	0.0752*** (0.0025) [0.0703,0.0800]
Weekly_Hours		0.0138*** (0.0005) [0.0127,0.0148]	0.0140*** (0.0006) [0.0128,0.0153]
Experience		0.0176*** (0.0013) [0.015,0.0203]	0.0183*** (0.0016) [0.0152,0.0213]
Experience_Squared		-0.0003*** (<0.0000) [$<-0.0001, <0.0001$]	-0.0003*** (<0.0000) [$<-0.0001, <0.0001$]
Constant	10.9839*** (0.0056) [10.9729,10.995]	9.7394*** (0.0639) [9.6142,9.8646]	9.7351*** (0.0738) [9.5905,9.8797]
R-Squared	0.001	0.440	0.435
Adjusted R-Squared	0.001	0.426	0.421
F-statistic	13.075	445.55	322.77
Prob > F	<0.001	<0.001	<0.001
Root MSE	0.610	0.462	0.466

Robust Standard Errors in Parenthesis

95% Confidence Intervals in Square brackets

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

N=21,149

Model 3 is Model 2 with ASEC weights

The Model Summary Statistics for Model 4 cannot be directly compared to the ones in Model 1 and Model 2
553 "Marital Status", "Race & Ethnicity", and "Occupation" Dummies in Model 2 and Model 3

Coefficients of “Num_Children”

As shown above in Figure 10, the coefficient of “Num_Children” changes in each model. It starts at -0.0134 in Model 1, which means that as “Num_Children” increases, Incwage roughly decreases by 1.34%. The three asterisks next to the coefficient denote that when determining the statistical significance of the coefficient of “Num_Children,” the null hypothesis *coefficient of “Num_Children” = 0* can be safely rejected, meaning that on a bivariate regression, “Num_Children” is highly statistically significant. It is quite obvious that Model 1 suffers from omitted variable bias, so Model 2 and Model 3 try to fix that by holding control variables constant. Model 2 reports that when “Num_Children” increases, Incwage decreases by approximately 0.63% while keeping EducYears, WeeklyHours, Experience, Marital Status, Race & Ethnicity, and Occupation constant. The coefficient on Model 2 also rejects the null, showing that the coefficient is still statistically significant, although the same cannot be said for Model 3. Model 3 is the same regression as Model 2, but it uses the ASECWT to provide even more accurate results. The coefficient on “Num_Children” is -0.0025, which denotes a roughly 0.25% decrease in Incwage when “Num_Children” increases, while keeping EducYears, WeeklyHours, Experience, Marital Status, Race & Ethnicity, and Occupation constant. Although the coefficient explains that mothers suffer from a wage penalty, it fails to reject the null meaning that there is no statistical significance that can be drawn from it.

Standard Errors and Confidence Intervals

All three models shown in Figure 10 use Robust Standard Errors as opposed to normal Standard Errors. In a perfect DGP, normal SEs would perform better than Robust SEs, but the data retrieved from IPUMS CPS most definitely suffers from heteroskedasticity, resulting in Robust

SEs performing better. All three models portray the Estimated Robust SE of the coefficient of “Num_Children,” which is then used to calculate the 95% confidence intervals. In Model 1, the Robust SE is 0.37%, meaning that the confidence interval range is -2.07% to -0.61%; both the Robust SE and confidence interval are relatively high, but it still provides evidence that motherhood has a negative effect on “LN_Incwage.” Model 2’s Robust SE on the coefficient of “Num_Children” is about the same: 0.32%, providing a 95% confidence interval of -1.25% to -0.01%. This confidence interval is worrying as almost 0% is a possible estimator of the true value of the coefficient of “Num_Children.” Even more surprising are the Robust SE and confidence interval for Model 3. As explained before, Model 3 is supposed to be the most accurate model as it uses ASCWT to take into account IPUMS CPS’s complicated sampling method, yet the Robust SE and confidence interval on the coefficient of “Num_Children” are 0.37% and -0.098% to 0.47% respectively. This is clear evidence that with 95% confidence the true value of the coefficient of “Num_Children” lies within the confidence interval, which means that the true coefficient of “Num_Children” could very well be neutral or even positive, giving no real answer to the question “Do Mothers still suffer from a Wage Penalty?”

Coefficient of Determination

The coefficient of determination, also known as R^2 is also shown above in Figure 10 for all three models. R^2 ranges from 0 to 1, and it explains how well the independent variables in a regression model explain the variation in the dependent variable. The Adjusted R^2 refers to how the R^2 is adjusted for adding unnecessary explanatory variables. Model one has an Adjusted R^2 of 0.001, which is abysmal, but that is attributed to the fact that it’s a bivariate regression with an independent variable that isn’t the primary cause of “LN_Incwage.” For example, even if it were

a bivariate regression of “EducYears” on “LN_Incwage”, the R^2 would be somewhat higher, due to the fact that EducYears has a higher influence on “LN_Incwage” than “Num_Children.”

Model 2 and Model 3 provide similar results, even though they can’t be directly compared, with adjusted R^2 of 0.426 and 0.421 respectively. This means that, with the inclusion of more control variables, the x variables explain the y variable (“LN_Incwage”) better.

F Statistics

The whole-model F-statistic is used to determine if the x variables have any effect on the dependent y variable; all three models, shown in Figure 10, show that they are statistically significant. The null hypothesis for the F-statistic is that all X variables' coefficients are equal to 0, and Prob > F determines whether the null hypothesis is rejected or not. Since the p-value in all three models is >0.001, therefore, under 0.05, it means that the null hypothesis is safely rejected in all cases. To conclude, this means that all three models show that they are highly statistically significant.

Comparison to Previous Works

The results of this regression analysis report that there is no statistical significance in regressing “Num_Children” on “LN_IncWage.” This directly contrasts findings in the literature review, such as Misra and Strader’s (2013) claim that mothers do suffer from a wage penalty across different countries. The realistic answer is that the results of this paper possibly suffer from some type of bias, whether it may be omitted variable bias or selection bias in the sampling. Another possible reasoning jumps to mind when taking a look at the previously stated study from Budig and Hodges (2010), where the motherhood wage penalty was linked to an earnings distribution. Figure 11, shown below, shows a regression with two models: Model 4, where all the observations have an IncWage between \$20,000 to \$60,000, and Model 5, where all the observations have an IncWage over \$60,000. Both models are a replica of the complete Model 3, as shown above in Figure 10, where the ASECWT IPUMS CPS weights are used and the same control variables are kept. The results, albeit superficial, nicely align with Budig and Hodges’ (2010) findings. Model 4 shows that IncWage decreases by roughly 0.76% per “Num_Children” while keeping EducYears, WeeklyHours, Experience, Marital Status, Race & Ethnicity, and Occupation constant. Alternatively, Model 5 shows that IncWage increases by roughly 1.13% per “Num_Children” while keeping EducYears, WeeklyHours, Experience, Marital Status, Race & Ethnicity, and Occupation constant. These two coefficients show how higher-paid workers seem to receive a motherhood wage premium, while lower-paid workers seem to suffer a wage penalty. This analysis of the two income groups is very superficial and shouldn’t be used as a strong indicator, but it can serve as a temporary explanation as to why the results of this paper aren’t statistically significant.

Figure 11

Regression Results by Income Groups		
	(1) Model 4 \$20,000-\$60,000	(2) Model 5 >\$60,000
Num_Children	-0.0076** (0.0029) [-0.0133,-0.002]	0.0113* (0.0051) [0.0014,0.0213]
EducYears	0.0149*** (0.0019) [0.0113,0.0186]	0.0472*** (0.0037) [0.04,0.0544]
Weekly_Hours	0.0098*** (0.0005) [0.0089,0.0108]	0.0042*** (0.0008) [0.0027,0.0057]
Experience	0.0038** (0.0012) [0.0013,0.0062]	0.0195*** (0.0021) [0.0154,0.0235]
Experience_Squared	-0.0001* (<0.0001) [<-0.0001,<0.0001]	-0.0003*** (<0.0001) [-0.0004,-0.0002]
Constant	10.0997*** (0.0823) [9.938,10.261]	10.7580*** (0.0900) [10.582,10.934]
Observations	12206	8943
R-Squared	0.248	0.253
Adjusted R-Squared	0.219	0.216
F-statistic	99.97	65.78
Prob > F	<0.001	<0.001
Root MSE	0.274	0.390

Robust Standard Errors in Parenthesis

95% Confidence Intervals in Square brackets

*p<0.05, **p<0.01, ***p<0.001

553 "Marital Status", "Race & Ethnicity", and "Occupation" Dummies

V. Conclusion

This paper aimed to answer the research question: “Do mothers suffer a wage penalty?” After reading about previous studies from roughly 10-15 years ago, it seemed clear that mothers received a wage penalty while fathers received a wage premium, but what was unknown was the current situation. Taking data from the IPUMS CPS ASEC 2024, the paper regressed “Num_Children” on “LN_Inc wage” while keeping control variables constant; additionally, the ASECWT were applied to provide more accurate results. The results clearly showed that mothers’ Inc wage decreased by roughly 0.25% per “Num_Children,” while keeping EducYears, WeeklyHours, Experience, Marital Status, Race & Ethnicity, and Occupation constant. Although this shows a negative trend, the large Robust SEs, and confidence intervals make it so that there is no significant evidence to prove that mothers face a wage penalty.

Although this analysis tries to include the most important control variables, it's needless to say that it still heavily suffers from omitted variable bias, including variables such as motivation or drive that could've helped in keeping more of the unexplained constant would've been fantastic. Additionally, if it weren't for the time limitations, it would have been interesting to compare the motherhood wage penalty across different US states. Furthermore, it is correct to assume that the data gathered for this paper's analysis most likely suffers from a sampling selection bias; mothers, in all likelihood, still suffer a wage penalty, so it is important to keep that in mind when reading this paper. To conclude, this paper provides no evidence that mothers face a wage penalty or premium.

VI. Work Cited

- Budig M. J., Hodges M. J. (2010). "Differences in Disadvantage: Variation in the Motherhood Penalty across White Women's Earnings Distribution." *American Sociological Review*, 75(5) 705–728.
- Misra, J., & Strader, E. (2013). "Gender Pay Equity in Advanced Countries Revisited." *Politics & Society*, 41(1) 103–114.
- Glauber, R. (2007). "Marriage and the Motherhood Wage Penalty among African Americans, Hispanics, and Whites." *Journal of Marriage and Family*, 69(4) 951–961.
- Killewald, A. (2013). "A Reconsideration of the Fatherhood Premium: Marriage, Coresidence, Biology, and Fathers' Wages." *American Sociological Review*, 78(5) 96–116.
- Becker, G. 1962. *Investment in Human Capital: A Theoretical Analysis*. Chicago: The Journal of Political Economy.
- "Frequently Asked Questions (FAQ)." *IPUMS CPS: FAQ*, cps.ipums.org/cps-action/faq#. Accessed 29 Apr. 2025.
- IPUMS CPS: DESCR: Asecwt*, cps.ipums.org/cps-action/variables/ASECWT#description_section. Accessed 7 May 2025.