ISSN 2348-3156 (Print)
International Journal of Social Science and Humanities Research ISSN 2348-3164 (online)
Vol. 11, Issue 3, pp: (26-31), Month: July - September 2023, Available at: www.researchpublish.com

SON PREFERENCE AND WAGE DISCRIMINATION

Brian Towell

University of Tennessee at Martin

Martin, TN 38238

JEL Codes: J13, J16, J31

DOI: https://doi.org/10.5281/zenodo.8149353

Published Date: 15-July-2023

Abstract: **This paper investigates whether son preference influences the male-female wage gap in the United States. I compare fertility rates of families who have had only daughters to those who have had only sons and analyze the correlation between this difference and the gender wage gap. Using data from the Current Population Survey (CPS) and American Community Survey (ACS), I construct a wage regression that reduces the commonly cited wage gap of twenty-three cents by over half, and show that inclusion of son preference can further reduce the wage gap by approximately five percent.**

Keywords: **gender bias, wage discrimination, son preference, prejudice.**

1. INTRODUCTION

One of the most frequently cited statistics when dealing with gender issues is the claim that women, on average, earn 77 cents for each dollar earned by a man. This is based on the observation that the median female full-time worker earns 77% as much as the median male full-time worker in annual earnings. Of course, this does not account for differences in occupation, workplace characteristics, human capital accumulation, etc.

Much more meaningful estimates of the wage gap can be constructed by including these basic controls. Wood et al. (1993) estimate a wage gap of 13% by looking only at the wages of lawyers and controlling for hours worked to eliminate occupation self-sorting and differences in labor supply. Differences that occur in human capital accumulation are also suggested to play a large role. Weinberger (1998) looks at the wages of recent college graduates to reduce differences in experience, while including controls for grade point average, major, and college attended, and finds a gender wage gap ranging from 10% to 15%. While these and other explanations reduce the wage gap, they do not eliminate it. Thus, there is either some as of yet uncontrolled for factor that disproportionately affects the productivity of females, or gender bias still plays a significant role in wage differentials.

The typical approach to this problem of separating the male-female wage gap into explained and residual portions has serious drawbacks that diminish the usefulness of the produced results. One issue is that such regressions typically omit feedback effects; they ignore the effect that gender has on the regression's explanatory variables. A regression on wages including education, for instance, may understate the true male bias if discrimination causes females to get less education than they otherwise would. Such effects make meaningful interpretation of the female coefficient difficult, as the possibility of gender bias in the explanatory variables likely suggests that such estimates are not even robust as upper bounds. Additionally, endogeneity concerns are magnified when attempting to interpret the residual portion of the wage gap. Given the imprecision which permeates most wage regressions, it is unlikely that any such regression will be able to include all relevant variables. While this is of course a prominent concern in any regression, the interpretation of the female coefficient as a measure of discrimination is wholly dependent on the regression being properly specified.

ISSN 2348-3156 (Print)
International Journal of Social Science and Humanities Research ISSN 2348-3164 (online)
Vol. 11, Issue 3, pp: (26-31), Month: July - September 2023, Available at: www.researchpublish.com

This paper takes a different approach to estimating the portion of the wage gap that is due to discrimination. Rather than relying on the goal of a perfectly specified wage regression and interpreting the female coefficient as the level of wage discrimination, I utilize a measure of gender bias that is unlikely to have any correlation with productivity. The portion of the female coefficient that disappears when this measure is included can be interpreted as a lower bound on the portion of the wage gap that is due to discrimination. I measure the fertility patterns of families with only boys against those with only girls, and find that differences in this statistic can account for nearly five percent of the gender wage gap, even when controlling for differences in human capital, occupation, and state specific trends.

2. A NEW MEASURE OF GENDER BIAS

In this paper I utilize a measure of differential parental stopping behavior in the choice to have children. If parents are more likely to have another child when they have had only girls, this is indicative of son preference, and vice versa. Any wage differential tied to this measure would suggest a lower bound on the amount of the pay gap is based solely on discriminatory factors, as there is nothing to suggest that son preference should play a role in determining productivity.

The measure of son preference is constructed by first ranking the children in each household by age to observe the gender order. The fertility rate of parents who have had only boys is then subtracted from the fertility rate of parents who have had only girls. This measures the fertility decision of parents at each different composition of their household, until they have had a child of each gender. The observations are cut off after a child of each gender is observed so that when the fertility rates of all son families are subtracted from the fertility rates of all daughter families, any preference for having at least one child of each gender (as discussed in Angrist and Evans (1998)) is differenced out. Only households with between 1 and 4 children were included, and the son preference is computed at the state level for each year.

If there is assumed to be no correlation between the genders of a parent's current children and the gender of any of their future children, then son preference of this type will not affect the gender ratio. Consider a simple example in which there are *n* families, and all of them have children until they have a son, with each family having an equal chance of having a son or daughter. After each family has their first child, there will be on average $\frac{n}{2}$ boys and $\frac{n}{2}$ girls. The $\frac{n}{2}$ families having girls will have another child, which means another $\frac{n}{4}$ girls and $\frac{n}{4}$ boys will be born on average. This process will continue indefinitely, with an equal number of boys and girls being born in each iteration.

For the years 2005 - 2015, U.S. households were approximately 1.2 percentage points more likely to have another child if they had only girls as opposed to only boys, with a standard error of .04 percentage points. This likelihood varies significantly across geographic and demographic groups. Race, marital status, and education level all play a large role in determining the level of son preference. The different averages across these demographic groups are summarized in Table 1.

TABLE 1: Son Preference by Demographics

Race	***Mean***	***S.E.***	**Education**	***Mean***	***S.E.***	**Marital Status**	***Mean***	***S.E.***
Black	.0053	.0012	Neither HS	.0201	.0019	Married	.0114	.0005
Asian	.0371	.0019	One HS	.0182	.0014	Cohabiting	.008	.002
White	.0122	.0004	Both HS	.0096	.0005	Never Married	.0057	.001
						Divorced	.0158	.0011

For each demographic the proportion of families with only sons who have another child is subtracted from the proportion of families with only daughters who have another child. Cohabiting parents are those who report as being part of the same household but are not married. Educational attainment includes only households with two parents present, and is separated by the number of parents having graduated from high school. Being more inclined to have another child after having only daughters indicates son preference, as the family may want a son above all, and continue having children until this goal is reached, and vice versa. A positive result can be interpreted as a preference for sons, as it indicates that families are more willing to have additional children in order to have a son.

Families in which neither parent has graduated from high school are far more likely to exhibit this measure of son preference. Such families are 2 percentage points more likely to have another child after having only daughters, whereas families where both parents have graduated high school have a son preference measure that is about half as high, being .96 percentage points more likely to have another child after having only daughters. Sidanius et al. (1996) provide evidence supporting a

ISSN 2348-3156 (Print)
International Journal of Social Science and Humanities Research ISSN 2348-3164 (online)
Vol. 11, Issue 3, pp: (26-31), Month: July - September 2023, Available at: www.researchpublish.com

negative correlation between racism and education (though more recent work seems to suggest that higher cognitive ability may only make one more inclined or better able to hide any prejudices they have, see Wodtke (2013)). The differential in son preference by educational attainment supports the similar notion that education is negatively correlated with gender bias. Unlike other studies attempting to measure some form of prejudice, this measure avoids having to rely on polling data and self-reporting, thus the aforementioned Wodtke mechanism is not problematic.

Amongst racial groups, Asian families exhibit the largest preference for sons. Asian families who have had only girls are approximately 3.7 percentage points more likely to have another child than are those who have had only boys, more than three times the overall average. This is consistent with the findings of Abrevaya (2009), who finds that the gender ratio at birth is heavily skewed towards males in Asian families, suggesting the use of sex selective procedures which indicates a son preference. Unplanned pregnancies will of course draw the average son preference of a group towards zero, thus it is expected that demographics with high unplanned pregnancy rates will have the lowest son preference. The 2012 National Health Statistics Report shows that black women are most likely amongst racial groups to report a pregnancy as unplanned, and black families display the lowest son preference.

There are significant dissimilarities in son preference across households with differing marital statuses. Households with married parents display substantially more son preference than those with unmarried parents do. A 2012 Gallup poll asked the following question: *"Suppose you could only have one child. Would you prefer that it be a boy or a girl?"*. Amongst male respondents, 49% said they would prefer a son, as opposed to 22% who preferred a daughter. Female respondents preferred daughters to sons by a margin of 33% to 31%. This suggests that males are the primary drivers of son preference, and that either their influence on fertility decisions is limited if they are not married to their partner, or that men who do not have a strong son preference are less likely to get married.

In addition to demographic variation, there are also significant geographic disparities, as shown in Fig. 1. Substantial geographic clustering occurs with northern states generally exhibiting the highest level of son preference.

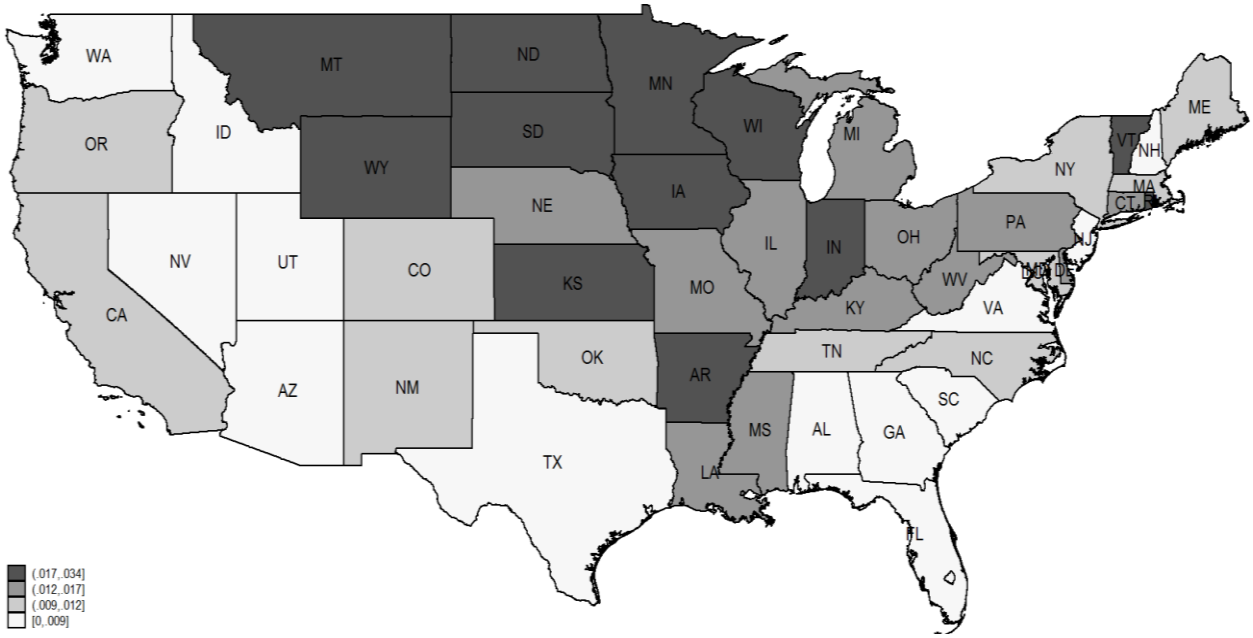

Fig. 1: Average Son Preference by State 2005-2015

Displayed is the difference between the percentage of families with only daughters who went on to have another child and the percentage of families with only sons who went on to have another child. Only households with 4 children or fewer are included.

One possible economic motivation for son preference within a region is the prevalence of agriculture. Farm owning families may feel that having a son to assist with farm work is economically beneficial. Table 2 investigates the relationship between the state's average son preference and the percent of land within the state that is farmable. Data on the percentage of farmland is from the United States Department of Agriculture Economic Research Service. The results show a coefficient on percentage of farmland of approximately .007. The standard deviation for percentage of farmland is approximately .25,

ISSN 2348-3156 (Print)
International Journal of Social Science and Humanities Research ISSN 2348-3164 (online)
Vol. 11, Issue 3, pp: (26-31), Month: July - September 2023, Available at: www.researchpublish.com

so an increase of one standard deviation in the percentage of farmland within a state would be expected to increase son preference by just under .2 percentage points. Thus, it appears that the presence of farmland and farming activities is playing a meaningful role in determining a state's level of son preference.

TABLE 2: Son Preference and Farmland

	1
%Farmland	0.007
	(0.004)*
R^2	0.41
N	51

* $p<0.1$; ** $p<0.05$; *** $p<0.01$

Standard error in parentheses. The dependent variable is the difference between the percentage of all daughter families within the state who had another child and the percentage of all son families who had another child. Controls included but not reported for the racial composition of the state.

The true preference for sons is likely to be understated by simple differences in fertility conditional on the gender of previous children. The results of Morgan et al. (1988) suggest that the probability of a marital disruption is 9 percent higher for married couples when having a daughter as compared to a son. Dahl and Moretti (2008) find that unmarried parents are less likely to get married if having a girl. These results provide a source of attenuation, as having a daughter reduces the chances of parents being in a marriage, which reduces the likelihood of their having further children.

3. METHODOLOGY

In order to identify the portion of the gender wage gap that is due to gender bias, son preference is included in a standard wage regression. As son preference should not play any role in determining productivity, any portion of the wage gap that disappears when son preference is included can be interpreted as a lower bound on the portion of the wage gap due to gender bias. The primary regression equation is:

$$w_{it} = \beta_1 female_{ist} + \beta_2 sonpreference_{st} + \beta_3 female_{ist} * sonpreference_{st} + \beta_4 X_{ist} + \theta_s + \gamma_t + \upsilon_{st} + \varepsilon_{it} \quad \textbf{(1)}$$

where i is an index for each individual, s is a state index, t is a time index, and w_{it} is the log of hourly wages. Son preference is calculated using data from the ACS, years 2005-2015. All other data is from the CPS Outgoing Rotation Group. X is a set of control variables including age, age squared, union status, education, number of children, marital status, race, month fixed effects, industry fixed effects, and occupation fixed effects. These results are reported in Table 3.

Any portion of the wage gap that can be explained by the son preference term is indicative of gender bias, as there is no reason to expect that differences across states and years in fertility choices should have any bearing on an individual's productivity. The inclusion of state and time fixed effects, as well as state specific time trends ensures that the son preference term is not simply reflecting a correlation between son preference and any other state level effects that may affect productivity.

One potential threat to this specification is reverse causality. Given the increasing portion of the population reaching old age, a larger percentage of parents may be counting on their children to assist them financially in their elder years. In that case son preference may be the result of wage discrimination as parents may have an incentive to have a son to increase the likelihood that they can be supported by one of their children. However, Silverstein and Gans (2006) indicate that daughters are more likely to provide for their aging parents.

The use of hourly earnings as the dependent variable is done to avoid fluctuations in earnings arising from differing supplies of labor, but leaves the regression vulnerable to the influence of part time workers, for whom it is often argued that work schedule flexibility is an important determinant of wage, as in Goldin (2014). As women are often suggested to possess less temporal flexibility when raising children, using hourly earnings may attach undue significance to gender, when the ability to adjust hours as needed is the true driver of wage differentials. In order to address this, the observations are limited to full-time workers, who generally have more linear pay schedules and more consistent hours.

ISSN 2348-3156 (Print)
International Journal of Social Science and Humanities Research ISSN 2348-3164 (online)
Vol. 11, Issue 3, pp: (26-31), Month: July - September 2023, Available at: www.researchpublish.com

4. RESULTS

Table 3 displays results from equation (1) in the final column. Preceding specifications reduce the number of controls included, and in each case the regression is run with and without the son preference terms. In the final specification with no son preference terms, the resultant female coefficient is -11.1%, which is similar to other wage regressions that control for occupation and human capital characteristics. For instance, Blau and Kahn (2007) estimate that if measured characteristics were held constant across men and women, women would earn 91% as much as men. When including the son preference terms in the regression the coefficient on female is reduced in absolute terms by .5 percentage points, suggesting that son preference accounts for approximately 4.5% of the gender wage gap. As this portion of the gender wage gap is ascribed to a characteristic that should have no relationship to productivity, this suggests a non-zero lower bound on the portion of the wage gap that is due to actual gender bias and cannot be explained away using appropriate control variables.

TABLE 3: Wage Regression

	1	2	3	4	5	6	7	8
female	-0.150	-0.149	-0.119	-0.114	-0.111	-0.106	-0.111	-0.106
	(0.009)***	(0.009)***	(0.003)***	(0.004)***	(0.003)**	(0.004)***	(0.003)***	(0.004)***
SonPref		0.899		0.550		0.157		0.154
		(0.478)*		(0.301)*		(0.131)		(0.122)
female*SonPref		-0.048		-0.417		-0.430		-0.429
		(0.251)		(0.175)**		(0.175)**		(0.176)**
Individual Controls	N	N	Y	Y	Y	Y	Y	Y
Year Fixed Effects	N	N	N	N	Y	Y	Y	Y
State Fixed Effects	N	N	N	N	Y	Y	Y	Y
State-Specific Trends	N	N	N	N	N	N	Y	Y
R^2	0.03	0.03	0.40	0.40	0.42	0.42	0.42	0.42

* $p<0.1$; ** $p<0.05$; *** $p<0.01$

There are 504,130 observations. The dependent variable is the log of hourly earnings. State clustered standard errors in parentheses. Individual control variables are age, age squared, education, race, month fixed effects, occupation fixed effects, industry fixed effects, union status, marital status, and number of children. Columns 7 and 8 include state specific linear and quadratic trends. Observations are full-time workers who report being paid hourly between the ages of 22 and 55.

5. SUMMARY

In this paper I find evidence that son preference can play a large role in explaining the existing male-female wage gap. Using data from the ACS and CPS, I show that incorporating son preference into regressions on the wage gap reduces the estimated wage differential by up to .5 percentage points, accounting for 4.5% of the gender wage gap. Improvements could be made if data were available on the financial costs of raising a boy versus a girl. If girls are more expensive to raise, this would provide yet another source of attenuation in the measure of son preference, whereas the converse could potentially eliminate much of the attenuation described in section 2. As previously referenced, Abrevaya (2009) finds evidence of a significant number of sex selective abortions and in vitro operations occurring amongst certain demographics, which also could skew true differences in fertility decisions that are taking place.

REFERENCES

[1] Abrevaya, Jason. "Are There Missing Girls in the United States? Evidence from Birth Data." *American Economic Journal: Applied Economics* 1.2 (2009): 1-34.

[2] Angrist, Joshua D., and William N. Evans. "Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size." *The American Economic Review* 88.3 (1998): 450-77.

[3] Blau, Francine D., and Lawrence M. Kahn. "The Gender Pay Gap: Have Women Gone as Far as They Can?" *Academy of Management Perspectives* 21.1 (2007): 7-23.

[4] Dahl, Gordon B and Enrico Moretti. "The Demand for Sons." *The Review of Economic Studies* 75.4 (2008): 1085-120.

[5] Goldin, Claudia. "A Grand Gender Convergence: Its Last Chapter." *American Economic Review* 104.4 (2014): 1091-119.

[6] Morgan, S. Philip, Diane Lye, and Gretchen Condron. "Sons, Daughters and the Risk of Marital Disruption." *American Journal of Sociology* 94.1 (1988): 110–29.

[7] Sidanius, Jim, Felicia Pratto, and Lawrence Bobo. "Racism, Conservatism, Affirmative Action, and Intellectual Sophistication: A Matter of Principled Conservatism or Group Dominance?" *Journal of Personality and Social Psychology* 70.3 (1996): 476-90.

[8] Silverstein, Merril, and Daphna Gans. "Intergenerational Support to Aging Parents." *Journal of Family Issues* 27.8 (2006): 1068-84.

[9] Weinberger, Catherine J. "Race and Gender Wage Gaps in the Market for Recent College Graduates." *Industrial Relations* 37.1 (1998): 67-84.

[10] Wodtke, Geoffrey. "Are Smart People Less Racist? Cognitive Ability, Anti-Black Prejudice, and the Principle-Policy Paradox." *Population Studies Center* Report 13-803 (2013).

[11] Wood, Robert G., Mary E. Corcoran, and Paul N. Courant. "Pay Differences among the Highly Paid: The Male-Female Earnings Gap in Lawyers' Salaries." *Journal of Labor Economics* 11.3 (1993): 417-41.