Empirical Analysis of a Policy reducing the Gender Pay Gap and Gender Employment Gap in the U.S.

&

Policy Recommendation for Germany

Machine Learning in Econometrics

Summer Semester 2023

Author:

Alexander Vogt
Beuthener Str. 33
81929, Munich

Matriculation number:
12071207

Supervisor:

Prof. Dr Daniel Wilhelm

Chair of Statistics and Econometrics
Department of Statistics
LMU Munich

Munich, 31.07.2023

1. Introduction

The Gender Pay Gap (GPG) has been intensively researched by labor economists over the last decades, due to its relevance for equal compensation and to contribute to a greater understanding the components and reasoning of its occurrence. Thereby, the GPG has declined in the U.S. over the last decades since 1985, which is attributed to a higher labor force participation and more education attainment among women (Blau & Khan, 2017). The same trend characterizes the European economies but despite the narrowing of the GPG, it is still significantly present in larger economies like Germany. Thereby, Germany had a GPG of 20% in 2014, which is 6 percentage points above the European average (Boll & Lagerman, 2019). This makes it important to intervene with an appropriate policy to target a reduction in the GPG in Germany.

In the U.S., a multi-layered policy was specifically implemented to reduce the GPG between men and women in the year 2007 by incentivizing firms to publish the salary of all workers and to encourage more negotiations about the salary. Additionally, the labor force participation of women and men should be increased by providing childcare to the workers, which is paid by the firms. Thus, the individuals were not directly targeted by the government, but firms were incentivized to improve working conditions in order to reduce gender inequality.

This report aims to answer several questions. Firstly, it is important to assess how effective this policy has been by estimating the impact on the GPG between the years 2005 and 2010. Moreover, an evaluation of how the effect differs between different parts of the population, e.g., urban or rural area, Northern or Southern part of the U.S. and being married or not being married. The last question to answer is which conclusions can be drawn from the policy in the U.S. that enables a greater understanding of the possible effectiveness of a similar policy in Germany.

The report proceeds as follows. In section 2, a short introduction to the underlying data is given as well as the data cleaning process. Section 3 provides insights into the decision-making process on the implemented econometric methodology. Section 4 presents the panel data results for the overall sample and the respective results for the different parts of the population. The second last section concludes on the policy implication for Germany. This empirical analysis ends with a short summary of the results.

2. Data

This empirical report uses a panel dataset with two periods for the analysis. It is important to stress out that there are not two consecutive periods available but only data of the years 2005 and 2010. Furthermore, the raw dataset consists of multiple relevant variables for the analysis of the GPG. The wage is given as a weekly wage and the employment is given as a dummy variable with one being employed. The treatment is only implemented in the year 2010 to women and men while the share of women benefiting from the policy is about 75 percent. Moreover, important variables are included, which are relevant to explain the magnitude of the wage of an individual: the IQ, the KWW[1], the years of education, the years of experience, the tenure in a firm, the age, the mother's and father's education in years, the number of siblings and the birth order.

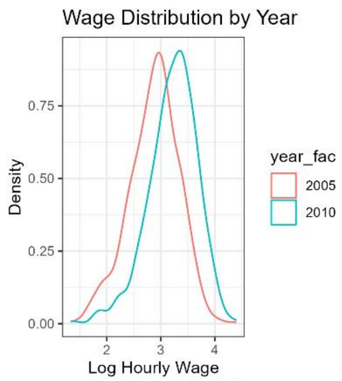
The data cleaning process is important for the later analysis to be reliable and not to be affected by outliers and missing values. The first step is to calculate the logarithmized hourly wage by dividing the weekly wage with the weekly working hours. Then, the outliers are removed using the interquartile range, which leads to dropping four individuals from the whole population set[2]. The next step is to replace the missing values in the mother's and father's education and birth order by the rounded mean values of the respective variables. This allows to preserve a large part of the sample size since the variable father's education has 448 missing values which is nearly half of the population size. Moreover, it is common practice to replace the missing values with the mean of the existing data. The final step is to create two datasets for the analysis of the GPG and GEG. The dataset for the GPG excludes all individuals which are unemployed and the dataset for the GEG is a full dataset after the data cleaning process.

Tbl. 1: Summary Statistics

Statistic	N	Mean	St. Dev.	Min	Max
IQ	909	101.402	15.020	50	145
KWW	909	35.723	7.640	12	56
Education	909	13.461	2.197	9	18
Experience	909	13.574	4.688	1	26
Tenure	909	9.245	5.308	0	23
Age	909	38.067	3.109	33	43
Siblings	909	2.953	2.315	0	14
Birthorder	909	2.950	2.650	1	10
Mother Educ.	909	10.623	2.730	0	18
Father Educ.	909	10.166	2.956	0	18
Log Hourly Wage	909	3.217	0.450	1.329	4.399

Tbl. 2: Shares

Statistic	Mean
Female	48.993
Urban	71.716
Black	12.464
South	33.557
Married	86.577



(Graph 1)

[1] KWW is a variable provided as the test score of occupational knowledge.
[2] The raw sample size consists of 1043 individuals.

In the following, a brief overview of the descriptive statistics on the GPG is given to allow for a better understanding of the results of this report. The panel data of GPG has a sample size of 933 in 2005 and 909 in 2010 after the data cleaning process. The log hourly wage has a mean of 3.22 in the year 2010 and has increased compared to the year 2005 as it is represented in Graph 1. In Table 1, a summary of the most important key values of the year 2010 are given. The average IQ lays at around 101 points resembling an average smart population and on average every individual has finished high school. Moreover, the sample consists of individuals of a small age range between 33 and 43, which is the most relevant age group since this benefit the most from wage increases and are more likely to have children compared to individuals above 50. Table 2 gives a short overview of the different parts of the population and allows to conclude that the data is balanced with respect to gender. Furthermore, 70% of the sample live in a city, 12% are black, 33% are located in the South of the U.S. and 86% are married. This information is relevant for assessing the heterogeneity of the impact of the policy on different parts of the population.

3. Econometric Methodology

In the following, the choices concerning the econometric methodology are explained and an argumentation for the use of Double Lasso and Double Machine Learning (DML) with random forest is given. Additionally, the strategy for the sensitivity analysis is provided.

The standard economical approach for determining the treatment effect of a policy in a panel data is the Difference-in-Difference (DiD) approach, in which the variation over time is exploited to gain insights into the policy effectiveness. However, the limited number of years make it impossible to gain reliable results for the treatment of women as the parallel assumption cannot be tested. Subsequently, the implementation of the DiD approach with an Ordinary Least Squared regression would lead to a biased estimator, due to unobserved explanatory variables. Therefore, another strategy has to be chosen to evaluate the effect of the policy on the GPG and GEG.

The strategy presented in this report is a DiD-like approach and implements Double Machine Learning to determine the causal effect of the policy in the U.S. Therefore, a variable measuring the policy effect on women needs to be created and is named policy variable. The first step is to divide the sample into a treatment group and a control group by creating a new variable called “TG”, based on whether an individual has received a treatment in the year 2010. Secondly, a time dummy variable called “Post” is needed, which takes the value one for the year 2010. Thirdly, the policy variable measuring the effectiveness of the policy on women earnings is an interaction term of the time-dummy variable, the dummy variable of the treatment group and the dummy variable for the gender, which takes the value one for being female.

The DML approach takes advantage of the Frisch-Waugh-Lovell (FWL) theorem, which allows to partial out the causal and unbiased coefficient β_1 of the policy on the GPG with Lasso and random forest. The standard regression approach is represented in (1), where the outcome variable log hourly wage Y is regressed on the policy variable D and all other explanatory variables W. The FWL-Theorem is applied as follows. Firstly, the policy variable D is regressed on the control variables W (see equation 2). Secondly, the outcome is regressed on all control variables as seen in equation (3). The third step is to regress the residual $\check{Y}$ from equation (3) on the residual $\check{D}$of equation (1) with the OLS, in order to obtain the causal estimator β_1 without any bias. In order to be able to allow a causal interpretation of the obtained β_1, a few assumptions need to be fulfilled, which are elaborated in the following.

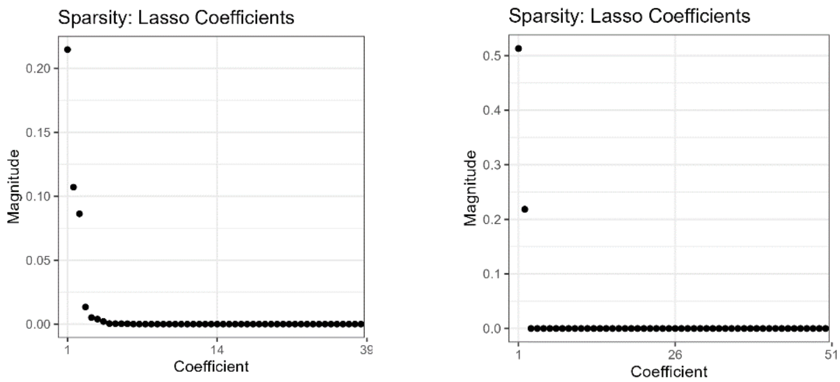
$$Y = D\beta_1 + W'\beta_2 + \varepsilon \qquad (1)$$

$$D = W'\hat{\gamma}_D + \check{D} \qquad (2)$$

$$Y = W'\hat{\gamma}_Y + \check{Y} \qquad (3)$$

First the assumptions imposed on Lasso and the general procedure are discussed, where the key assumption Adaptive Inference in High-Dimensions must hold. Therefore, the nuisance parameters $\hat{\gamma}_D$ and $\hat{\gamma}_Y$ need to be approximately sparse, which can be observed in Graphs 2 and 3. This assumption is fulfilled but the graphs itself need to be approached with cautiousness as it shows the sample approximate sparsity of the coefficient and not the true

Equation (3) Equation (2)



Graph 2: Sparsity Plots

sparsity of the population. Moreover, the explanatory variables W is assumed to fulfill the restricted isometry condition as it cannot be directly tested. Next to this assumption, it is important to determine the right penalty parameter λ in order to prevent over- and underfitting, which is achieved by k-fold-validation. In this empirical analysis, the penalty parameter that minimizes the mean squared error is used when determining the residuals from equation (2) and (3).

Secondly, the random forest method allows to handle sparse and dense data, which is the case as discussed in the previous paragraph. Moreover, it is important to use cross-fitting and small trees when estimating the residuals from equation (2) and (3) in order to not suffer from any regularization bias. The small-tree-condition is enforced by only allowing for a maximal depth of 5 and a minimal node size of 7 for each tree. The results are not changing for small variation in the maximal depth and minimal node size of a tree but for greater changes of the parameters the estimator suffers from overfitting and the coefficient doubles in size.

The final paragraph is about the general assumption on DML and the sensitivity analysis, in order to guarantee causal and robust estimators. The most important assumption is the conditional independence assumption, which assumes that the treatment conditional on the control variables is orthogonal to the outcome for male and female. This can be assumed as none of the receivers of the policy could expect the additional help from the government as the work-incentivizing-program is assigned on firm-level. Therefore, it can be concluded that the program is randomly assigned. Moreover, common support is assumed and rich covariates are used. The sensitivity analysis consists of two different control sets. The first control set is a simple control set used for the estimation process and a flexible control set with all possible interaction terms of the existing control variables. Further, DML with random forest is used as an additional sensitive analysis in the case of any non-linear relationship between variables.

4. Data Analysis

In the following, the results on how the GPG and GEG has developed over time are evaluated as well as the effectiveness of the policy. Additionally, an elaboration of the effectiveness with respect to different parts of the population is presented.

4.1 Effectiveness of the Policy

Firstly, a naïve evaluation is done to assess the GPG and GEG for the years 2005 and 2010 in order to reveal any changes over time. The first two columns in Table 3, GPG 2005 and GPG

2010, provide information about the gender pay gap respectively for the years 2005 and 2010. The causal coefficients in Table 3 and 4 are calculate by means of the DML with Lasso. Thereby, it can be observed that the GPG within the treatment group has switched from women earning less to them earning significantly more than men for the simple and the flexible control set. The concrete interpretation for the flexible model in GPG 2010 is as follows: women who benefited from the policy are earning on average 37.9% more than men with the coefficient being highly significant to the 5%-level. The same pattern can be observed in the GEG in 2010 with women in the treatment group being on average 30.4% more likely to be employed compared to men with the coefficient being highly significant to the 5%-level. This analysis does not allow for any conclusion on the effectiveness of the policy since a part of the increase in earnings can also be attributed to time effects, e.g., women have a higher average educational level in 2010 compared to 2005.

The second step is to analyze the effectiveness of the policy on women by interpreting the treatment effect introduced in the previous section. Thereby, the hourly wage has increased on average by 23.2% for women being exposed to the policy from the year 2005 to 2010, if the DML with Lasso is applied. Similar results can be found with the DML with random forest, where even an increase of 28% in hourly wages can be observed. Additionally, both coefficients are highly significant to the 5%-level. The analysis of the GEG suggests no significant increase in the probability to be employed as it is observed in Table 4 for Panel Data. Subsequently, the DML with random forest does not promise any reliable results as the coefficient of 32% is double the size of the coefficient in Table 4.

Overall, the conclusion can be drawn that the policy has a large positive effect on the hourly wage of women. Thus, this result indicates that the gender inequality with respect to wage has declined between the years 2005 and 2010. On the other hand, the policy does not seem to encourage more women to work as there is no significant coefficient given the data provided.

	GPG 2005	
	Log Hourly Wage	
	Simple	Flexible
	(1)	(2)
GPG	-0.157^{***}	-0.159^{***}
	(0.031)	(0.031)
Note:	$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$	

	GPG 2010	
	Log Hourly Wage	
	Simple	Flexible
	(1)	(2)
GPG	0.383^{***}	0.379^{***}
	(0.060)	(0.059)
Note:	$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$	

	Panel Data	
	Log Hourly Wage	
	Simple	Flexible
	(1)	(2)
Treatment Effect	0.232^{***}	0.230^{***}
	(0.061)	(0.058)
Note:	$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$	

Tbl. 3: Gender Pay Gap

	GEGap 2005	
	Employment	
	Simple	Flexible
	(1)	(2)
GEG	-0.104	-0.132*
	(0.064)	(0.071)
Note:	*p<0.1; **p<0.05; ***p<0.01	

	GEGap 2010	
	Employment	
	Simple	Flexible
	(1)	(2)
GEG	0.316**	0.304**
	(0.154)	(0.153)
Note:	*p<0.1; **p<0.05; ***p<0.01	

	Panel Data	
	Employment	
	Simple	Flexible
	(1)	(2)
Treatment Effect	0.164	0.168
	(0.143)	(0.142)
Note:	*p<0.1; **p<0.05; ***p<0.01	

Tbl. 4: Gender Employment Gap

4.2 Heterogeneity

The analysis of the heterogeneity with respect to different parts of the populations is conducted by excluding all other variables, e.g., if the treatment effect for the black population is assessed only black individuals are included in the dataset. The main interest lays on the heterogenous effects between the black and white population, urban and rural population, North and South population and the married and unmarried population. Thereby, the report concentrates on the treatment effects of the policy on hourly wage of women in Table 5 as there is no significant effect on the employment of women in any sub-population.

In the first sub-table, it is shown that the black female population significantly benefited from the treatment as their hourly wage has increased by 50%, while the white female population only earned 16.8% more. This result has to be handled with caution as the share of black in the sample size is only 25%, which could have led to an upwards bias of the coefficient. The second sub-table shows that there is no substantial difference between the female urban and female rural area population as both benefited equally. The third sub-table suggests that the hourly wage has increased by 40% for the female Southern population of the U.S. but there is no substantial increase in the North. This could be attributed to the fact that the Southern states are on average economically poorer compared to those in the North of the United States. The fourth sub-table evaluates a significant increase of the hourly wage for married women, while unmarried women do not experience any increase. The reason behind this

	Heterogeneity (1)	
	Log Hourly Wage	
	Black	White
	(1)	(2)
Treatment Effect	0.516***	0.168***
	(0.148)	(0.061)
	Heterogeneity (3)	
	Log Hourly Wage	
	South	North
	(1)	(2)
Treatment Effect	0.394***	0.107
	(0.107)	(0.066)
Note:	*p<0.1; **p<0.05; ***p<0.01	

	Heterogeneity (2)	
	Log Hourly Wage	
	Urban	Rural
	(1)	(2)
Treatment Effect	0.202***	0.207*
	(0.068)	(0.112)
	Heterogeneity (4)	
	Log Hourly Wage	
	Married	Not M.
	(1)	(2)
Treatment Effect	0.230***	0.185
	(0.062)	(0.136)
Note:	*p<0.1; **p<0.05; ***p<0.01	

Tbl. 5: Gender Pay Gap | Heterogeneity

observation could be a greater labor force participation of married women, due to the provided childcare since married women are more likely to have children.

5. Policy Implication

The results from the United States are very promising as the GPG has decreased. However, a one-to-one transfer of the policy to the German context does not imply an equal effectiveness. In order to be able to make a well-founded policy suggestion on what can be expected in Germany, it is important to compare the regional, ethnic and labor market differences of both countries with each other.

The heterogenous analysis of the GPG suggested that there was a significant decrease in the GPG for the Southern population while the North of the United States did not benefit to the same extent from the policy. This can be explained through a substantially higher GPG for Southern states compared to Northern states, e.g., Alabama (U.S. Census Bureau, 2023), due to a more conservative image of women represented by the fact that this state is governed by the Republicans since the 19^{th} century (Alabama Presidential Election Voting History - 270toWin, n.d.). Thus, a policy targeted on increasing labor force participation and introducing a greater visibility of the salaries has a greater effect in this part of the United States. These regional differences in the GPG are also found in Germany, when comparing West and East Germany with each other. Thereby, the GPG is substantially lower in the East of Germany, where women are earning only 6% less than men (Fuchs et al, 2019). One possible reason behind this division is similar as in the United States, because the West of Germany is considered more conservative in accordance with the South of the United States. Another reason is the fact that East Germany was the former socialist country GDR, where women had to go to work as they were seen as an important part of the labor force.

The benefits that black women get from the policy are remarkable and resemble the fact that the black population is economically poorer compared to the white population in the United States. Therefore, black women, who are earning significantly less, are benefiting to a grand extent from these rather light policy introductions. In Germany, this economic difference is observed among immigrants who have a substantially higher wage gap and, thus, it can be assumed that also a higher GPG exists among this population group (Ingwersen & Thomsen, 2021). Hence, it is of great importance to narrow this gap with a policy.

Despite finding the same patterns in regional and ethnic differences, the labor market is different with respect to various characteristics. The United States has a corporation culture marked by a hire-fire mentality with little protection for workers. The workers have to pay childcare by themselves if the firm does not provide it, they are not negotiating their salary because of the fear to not get hired and have a weaker social security system compared to Germany. Therefore, improvements like incentivizing firms to provide childcare and to encourage salary negotiations will have a greater impact in the United States compared to Germany, which has a strong social security system, e.g., a lot of counties are providing free childcare in Germany.

Overall, the predominant GPG of women earning 20% less compared to men in Germany makes a policy like the one in the United States desirable and even necessary (Boll & Lagerman, 2019). When looking at the empirical analysis it shows that the policy has a great potential to narrow the regional and ethnic differences as well as reducing the general GPG in Germany. On the other hand, the labor market difference suggest that the policy will be less effective in Germany. Thus, the recommendation of this report is to implement a similar policy in Germany, due to the promising results from the United States.

6. Conclusion

The topic of the gender pay gap and gender employment gap is widely discussed in the economic literature and of great importance for the emancipation of women and equal pay across genders. Thus, the empirical analysis of the policy targeted to reduce the GPG in the United States is a great opportunity to evaluate its effectiveness. The policy aimed to increase labor force participation and encouraged active negotiations of the salary in order to close the GPG and GEP. In order to evaluate the effectiveness of the policy a Difference-in-Difference-like approach is chosen. Thereby, the Double Machine Learning methods with Lasso and random forest are used to determine the causal effect of the policy on the hourly wage and employment of women. Hence, it was concluded that it had a great impact on reducing the GPG as the policy increased the hourly wage by 23-28% for women over time. In contrast, the policy did not have any significant effects on the employment of women. The successful results from the United States assessed in this report are a great motivation to introduce a similar policy in Germany. Thereby, it does not only promise to reduce the GPG across whole Germany but also to reduce regional and ethnic differences within the country.

7. References

Alabama Presidential Election Voting History - 270toWin. (n.d.). 270toWin.com. https://www.270towin.com/states/Alabama

Bach, P., Chernozhukov, V., Kurz, M., Spindler, M., & Sven, K. (2021). *DoubleML: Double Machine Learning in R* [Software]. https://github.com/DoubleML/doubleml-for-r/

Blau, F. D., & Kahn, L. M. (2017). The Gender Wage Gap: Extent, Trends, and Explanations. *Journal of Economic Literature*, *55*(3), 789–865. https://doi.org/10.1257/jel.20160995

Boll, C., & Lagemann, A. (2019). The gender pay gap in EU countries — new evidence based on EU-SES 2014 data. *Intereconomics*, *54*(2), 101–105. https://doi.org/10.1007/s10272-019-0802-7

Fuchs, M., Rossen, A., Weyh, A., & Wydra-Somaggio, G. (2019). Why do women earn more than men in some regions? Explaining regional differences in the gender pay gap in Germany. *RePEc: Research Papers in Economics*. https://www.econstor.eu/bitstream/10419/204859/1/1666019488.pdf

Ingwersen, K., & Thomsen, S. L. (2021). The immigrant-native wage gap in Germany revisited. *The Journal of Economic Inequality*, *19*(4), 825–854. https://doi.org/10.1007/s10888-021-09493-8

U.S. Census Bureau. (2023, June 12). *In Puerto Rico, No Gap in Median Earnings Between Men and Women*. Census.gov. https://www.census.gov/library/stories/2022/03/what-is-the-gender-wage-gap-in-your-state.html

8. Declaration of Autonomy

I confirm that this report is based on my own work. In preparing this report, I have not received any help from another human nor have I discussed any aspects of the empirical project with others.

Alexander Vogt

Munich, 31.07.2023

Place and date | Signature