Using econometrics to reduce gender discrimination: Evidence from a Difference-in-Discontinuity Design

Giannina Vaccaro*

December 15th, 2017

Abstract

Despite governments' enactment of multiple policies to reduce discrimination, it still persists in most countries. This paper estimates the effectiveness of a relatively inexpensive anti-discrimination policy. Introduced in Switzerland in 2006, this policy consists of a regression framework that firms can use to monitor their wage policies. Only firms with more than 50 employees are subject to the random controls, however sanctions may lead to exclusion from public procurement. By exploiting the cutting-off point of 50 and using periods before and after the reform, this article estimates the causal effect of this policy using a Difference-in-Discontinuity design. To analyze the potential manipulation employers may perform to comply with this regulation, and to explore the mechanisms behind the results, I implement an Oaxaca-Blinder decomposition of the raw gender wage gap and separately study the effects of this policy on multiple employment outcomes. Results show that this policy has reduced the gender wage gap. Also, neither evidence of employers' manipulation nor employment effects were found. However, particular groups of workers seem to be driving overall effects. In a nutshell, this paper suggests two critical lessons: (1) it shows a very clean test for the effectiveness of an anti-discrimination policy, and (2) it provides a good example of how to reduce the unexplained wage gap.

Keywords: Firm size, Wage differentials, Wage discrimination, and Wage inequality
JEL-Classification: J16, J21, J31, J71

*E-mail: gvaccaro@uci.edu. University of California, Irvine (UCI). Swiss National Centre of Competence in Research LIVES - Overcoming vulnerability: life course perspectives (NCCR LIVES). The current version of this paper has been updated and modified significantly with respect to previous versions. I thank deeply all who made comments and valuable suggestions. In particular, I thank Michele Pellizzari for his continual guidance and invaluable comments. Also, I thank David Neumark for his deep thoughts and useful feedback on the last version of this manuscript. I thank the participants at conferences and seminars in Geneva and Zurich (Switzerland), Izmir (Turkey), Mannheim and Nuremberg (Germany), Bergamo and Milano (Italy), SOLE (Seattle), and the 19th IZA European Summer School in Labour Economics (Ammersee), the Micro-Lunches and the Labor-Public Seminar at UCI and at the Vancouver School of Economics (VSE), and also to Damon Clark, Nicole Fortin, José Ramirez, Andrea Ichino, Rafael Lalive, Marco Francesconi, Ugo Trivellato, Enrico Rettore, Alan Manning, Abihishek Chakravarty, Wilbert van der Klaauw, Giovanni Mastrobuoni, Barbara Petrongolo, Silja Baller, Maurizio Bigotta, Florian Wendelspiess for thoughtful conversations and ideas. Special thanks to the Federal Office of Gender Equality (FOGE) for its insights and to Silvia Strub and Roman Graft for their advice and for sharing with me their domain expertise. My particular gratitude to Daniel Abler for all his comments and inspired discussions. I deeply thank the financial support of the Swiss National Science Foundation (SNSF) and the Swiss National Centre of Competence in Research LIVEs - Overcoming vulnerability: life course perspectives. An earlier version of this paper was distributed under the title "How to Reduce the Unexplained Gender Wage Gap?: Evidence from a Regression Discontinuity Design". All remaining errors are mine. This is a preliminary draft. Please do not cite or distribute without permission of the author.

1 Introduction

Numerous policy interventions have attempted to reduce or eliminate the gender wage gap.[1] However, women still earn substantially less than men and it has been very difficult to prevent discriminatory practices. A good share of the wage gap can normally be explained by differences in observables, but a non-negligible part of the gender wage difference remains unexplained, which is often viewed as signal of discrimination.[2] Pay-equity laws frequently mandate that otherwise similar workers will be paid the same. For example, the Federal Equal Pay Act (EPA) of 1963 prohibits sex-based wage discrimination between men and women who perform jobs that require "substantially equal skill, effort, and responsibility, and which are performed under similar working conditions".[3] People have tried different methods to prove discrimination. For instance, regression analysis has been used since the earlies 1970's (Bloom and Killingsworth 1982; Fogel 1986).[4]

Earlier research tries to estimate the effects of pay-equity laws, but evidence is mixed and not always compelling (Altonji and Blank 1999). For instance, Beller (1982) found that both the Title VII of the Civil Rights Act of 1964 and the federal contract compliance program decreased occupational segregation and the gender wage gap. Chay (1998) detected that the pay gap between black and white men in the South narrowed between 1.5% and 3.4% after the introduction of the Equal Employment Opportunity Act (EEOA) of 1972. More recently, Holzer and Neumark (2000) argued that Affirmative Action (AA) may alleviate discrimination and increase efficiency, but acknowledged that it was hard to establish its causal effect.[5] Furthermore, in the best scenario, when an effective wage policy is identified, it is likely to be very expensive and difficult to implement. Also, policies that only tackle wages may have negative employment effects. For this reason, there is need for studies of the impact of gender discrimination laws, their side effects, and the effectiveness of their particular enforcement mechanisms.

In 2006, without being anticipated by any firm or employer, Switzerland introduced a wage discrimination policy that aims to reduce the unexplained wage gap. This policy is inexpensive, easy to implement, and uses regression analysis. As typical in the literature, this regression includes different observable measures

1. Around the world, different legislations were implemented such as the Statement on Equal Pay of the International Human Rights law of 1951 of the International Labour Organization; the Equal Pay Act of 1963, the Title VII of the Civil Rights of 1946, and the Equal Employment Opportunity Act (EEOA) of 1972 for the US; the Equal Pay Act of 1970 for the UK; Art. 141 of the Treaty of Rome (1957) and its directly applicable equal pay directive of 1975 for the European Economic Community, the Canadian Human Rights Act of 1977; the Equal Treatment on Men and Women of 1979 for Austria, the Anti-discrimination legislation on the bases of sex of 1984 for Australia, among others.

2. Across OECD countries women earned on average 16.6% less than men in 2013, and this gap has only marginally declined from 18.6% in 2000. In the US, women working full-time earn just 80% of men's salaries. Switzerland, the country object of this study, is no exception. In 2012, women in Switzerland earned on average 23.2% (USD 1'800 approx.) less than their male colleagues in the private sector, down only 6.4 percentage points compared to 1996, and only 9 percentage points than in 1960. Blau and Kahn (2007) found that 41.1% of the gender pay difference was not explained neither by education, experience, occupation, industry, union status in 1998 in the US. In Switzerland, the unexplained gender wage gap has a similar magnitude than the one from the US and it has been stable over time. In 2014 the average wage difference between men and women in the private sector in Switzerland was on average 19.5% (USD 1,495). Only 60.9% of it was explained by objective factors, while 39.1% remained unexplained.

3. Equal Pay Act of 1963, Vol. 29 United States Code, section 2016(d).

4. Regression analysis have been used extensively in litigation cases, and already by 1980s seemed to have broad acceptance. Pioneer cases such as *United States v. US Steel* (1973) and *Wade v. Mississippi Cooperative Extension Service* (1974) already used regressions by expert witnesses and when appealing court's decisions.

5. Similar Pay Act policies to AA have been enacted in the US (i.e the *Lily Led better fair Pay Act of 2009*, but it has been criticized for attracting only well-off female employees who can afford a good lawyer to win their claim).

of productivity-related characteristics and a gender dummy. The latter represents unobservable factors that are associated with women's pay and it is attributed to gender wage discrimination (Bloom and Killingsworth 1982).[6] According to this Swiss policy, targeted firms should look at the magnitude and the statistical significance of their estimated gender coefficient to self-test if their wage policy is discriminatory or not. Random checks are performed by the government on a very small number of private firms among public tenders. Sanctions include revocation of licenses and/or cancellation of approved contracts, but they are rarely implemented. Furthermore, this Swiss wage policy is framed in the context of a very open and business-friendly Swiss culture which does not have the objective to punish non-compliant firms. It aims at facilitating self-surveillance of wage discrimination by allowing employers to determine whether their wage policy is discriminatory or not at almost no cost.

This paper has two contributions. First, it provides a very clean test of measuring the impact of a wage discrimination policy. The design and implementation of this Swiss anti-discrimination policy creates a unique opportunity to test the impact of such policy on gender wage discrimination. Specifically, I investigate if the unexplained gender wage gap decreases after the introduction of federal supervision using a Difference-in-Discontinuity Design (Diff-in-Disc), a combination of a Regression Discontinuity Design (RDD) and Differences-in-Differences (Diff-in-Diff) approach, and firm size as the running variable.[7] In other words, I study precisely the potential discontinuity of the unexplained wage gap at the threshold of 50, before and after the introduction of the policy, using exactly the same model specification employed by this Swiss policy. The Swiss Wage Structure Survey (SWSS) from 1998 to 2010, a biennial employer-employee data, and the Swiss Business Census (CS) are used to perform the analysis. Diff-in-Disc estimates show evidence that the unexplained gender wage gap reduces by 3.5 percentage points in firms with 50 or more employees after 2006. While analyzing the causal effect of this policy on the unexplained wage gaps is informative, we could be skeptic about the real effect of this policy on raw gender wage differences. Indeed, effects on raw wage gaps are smaller than for the unexplained wage gap, accounting for about a 1.5 percentage points decrease. These results can be interpreted as the success of this policy to reduce gender wage discrimination, but its limited effect to achieve gender pay equity. Moreover, where I find that the unexplained and explained wage gaps decreased, I typically do not find significant employment effects. Thus, this may be evidence that this pay regulation did help to reduce the unexplained gender wage gap without effects on employment.

Second, the results of this policy provide a good example of how to reduce the unexplained gender wage gap. Notwithstanding, these results cannot be generalized to all pay equity laws. Broader and less automated anti-discrimination laws may leave less scope for employers' discretion because other forms of discrimination are also regulated. Thus, the design of this Swiss wage policy should make us wonder

6. Measuring unexplained wage gaps using an imperfect model that relies on a statistical residual has its limitations (potential omitted-variables, reverse causality of observables, etc.), particularly when trying to assess the best way to measure wage discrimination. Nonetheless, results of studies using regression analysis to measure gender pay gaps are very instructive because they provide clear measures of unequal pay for equal work (Killingsworth 1993; Blau and Kahn 2007).

7. In a standard RDD setting, the condition for identification is guarantee when agents do not manipulate the running variable. This can be achieved by assuring the continuity of the conditional expectation of counter-factual outcomes in the running variable (McCrary 2008). However, in a setting where two (confounded) treatment effects exists, the cross-sectional RDD will provides a biased estimate of the ATE. Only by exploiting the time dimension (information pre-treatment, and post-treatment) running a DIFF-in-DIFF, the selection bias is removed (Grembi, Nannicini, and Troiano 2016). In our case, computing the Diff-in-Disc acts as placebo test to cross check the strength of the results.

about the potential manipulation that employers could perform to respect this law. Employers might be motivated to change their firm size to avoid being affected by the policy rule. Also, they may be tempted to use their discretion to adjust the values of any other employment variable to fulfill the regulation. For these reasons, I then analyze (a) how employers comply with this regulation, (b) if any potential side effects and other kinds of discrimination were generated after the implementation of this policy, and (c) if results were driven by particular groups of workers. In this regard, this study contributes to the literature not only by quantifying the effect of this policy, but also and foremost by highlighting the mechanisms behind the implementation of a simple policy which only targets wages.

To identify the main drivers of wage changes and determine to which extent these are manipulated, I study the distribution of firm size, and then decompose the wage gap using an Oaxaca-Blinder decomposition (OB) (Oaxaca 1973; Blinder 1973). The study documents two main results. First, graphical and statistical analysis reject the null hypothesis of continuity of the running variable before and after the introduction of the policy. These findings can be interpreted together as indication that employers neither change their firm size nor manipulate the composition of their workforce. Second, the OB decomposition shows that the most important factors behind the unexplained gender wage difference are due to returns to education and experience and not due to the temporal increase of other workers' characteristics. RDD and Diff-in-Disc decompositions on every group of workers show that results were driven by workers with particular characteristics (with upper secondary education, work independently, and have less than one year of tenure). As discussed more fully in the paper, these results may arise because after the introduction of this policy, employers did not adjust all female wages in their firms, but only those of recent hires.

2 Studies and Policies to reduce the Gender Wage Gap

Many studies have tried to estimate the effectiveness of wage policies, however the evidence is far from conclusive. Indeed, establishing the causal effect of these policies has been very difficult because of three main reasons. First, usually policies are not enacted in isolation and therefore it is difficult to disentangle the effect of these laws from other regulations. Second, anti-discriminatory regulations are framed in a very general context and it is not clear how exactly they can be implemented. Third, policies are sometimes designed in very particular contexts, which makes drawing general conclusions impossible.

Despite the great interest in gender discrimination around the world, most related economic literature studied the American labor market. The introduction of gender wage policies in the US date from almost 50 years ago. Bloom and Killingsworth (1982) for example, is one of the pioneer studies that investigated the use of regression analysis as analytic construct in courtrooms to argue against pay discrimination. They provided a historical overview about the use of this method and together with Killingsworth (1993) detail the relevance and main challenges of using regressions to fight gender discrimination. Since the late 90s, the literature has shifted direction and it no longer focuses on describing these policies, but mainly on analysing their effectiveness.

For example, Neumark and Stock (2006) used a triple difference strategy (DDD) and exploited the variation across time and states to investigate the effect of state sex and race discrimination laws passed prior to the federal anti-discrimination legislation on earnings and employment. Differently than here, they focused on long term effects and full-time workers. They found that only sex state discrimination laws, which not only targeted discrimination in terms of wages but also on employment, reduced relative employment of both black and white women.[8] These findings suggest that to reduce discrimination of wages and employment, laws must target explicitly both outcomes.

Hahn, Todd, and Klaauw (1999) is maybe the closest study to my research. In addition of their theoretical contributions, they used a Regression Discontinuity Design (RDD) to evaluate the effect of the US Equal Employment Opportunity Commission (EEOP) coverage on minority employment in small US firms. Particularly, they exploited the fact that by US law, only firms with at least 15 and 25 employees were covered by the Title VII of the Civil Rights Act of 1964 and by the 1976 amendment to the act, respectively. In contrast to my study, Hahn, Todd, and Klaauw (1999) only used cross-sectional variation by analyzing the effect of this policy year by year. They did not have data on individual firms but rather on individual workers which reported the size of the firms at which they were working. This resulted in two serious problems: first, only very few observations were subject of the study; and second, their results suffered from measurement error at the relevant threshold due to the fact that employees may have reported rounded and imprecise information of their firm size.

Studies that analyzed the effect of a policy based on firm size were also implemented outside the US. For the UK, Manning (1996) identified a large rise in the relative earnings of women due to introduction of the UK Equal Pay Act of 1970. However, he could not confirm the expected fall in relative employment due to monopsonic characteristics of the British labor market. Like the identification strategy employed in this study, Leonardi and Pica (2013) combined a RDD and Diff-in-Diff to determine the effect of an Employment Protection Legislation (EPL) Italian reform on wages. Similar to my study, they used employer-employee data and firm size as a running variable. The nature of their policy though is very different from the one studied here. It raised dismissal costs only for firms with 15 workers or fewer employees and left larger firms unaffected. Authors found that this policy had a small but significant negative effect on wages of firms affected by this policy. The effect was highly heterogeneous across firm size, which the authors suggested could be attributed to the relative bargaining power of workers versus firms.

In contrast to other studies, the design of the Swiss wage policy I study here, its relatively inexpensive implementation and its unprecedented design, make it an attractive and unique example to examine the impact of a pay equity law. The Diff-in-Disc design used in this paper is innovative and seems to be appropriate to identify the causal effect of this policy. It works well, and it provides interesting results, showing evidence of an easy and effective way to reduce gender wage discrimination. Furthermore, by using an OB decomposition, I can identify the drivers behind these results. The results provide useful insights into what might explain the evolution of wage differences, and inform how employers comply with this regulation.

8. Effects on earning was treated cautiously in that paper because selection problems generated by the sharp increase of female employment (Neumark and Stock 2006, p. 408).

3 Institutional Background: The Lohngleichheitsinstrument Bund *(Logib-CH)*

As in many other countries, gender equality laws in Switzerland have been framed in very general terms, and enforced at the federal level.[9] In April 2006, the Federal Office of Gender Equality (FOGE) launched a very detailed policy called *Lohngleichheitsinstrument Bund* (or *Logib*, due to its German name). *Logib* is a policy based on a wage regression as detailed in equation 1 in section 5, explicitly written in the legal recommendation, that aims to discourage wage discrimination among companies (See details in Section 5).[10] In order to facilitate companies to self-check whether their pay practices are discriminatory, the FOGE has developed an excel software that implements equation 1. The *Logib* excel tool is available in 4 languages (German, French, Italian and English) and it is provided along with a free helpline. It is an anonymous tool, since the data and information used for self-assessment is stored at a local level.[11] The goal of this study aims to identify the causal effects of *Logib* exactly as it is defined in the Swiss recommendation.

Similarly to the American Affirmative Action (AA) policies, the use of *Logib* is free of charge and voluntary. The FOGE only recommends its use to companies with more than 50 workers based on statistical reasons, and only a random selection of companies which have won public tender contracts are monitored.[12] Per year, approximately 30'000 companies win public tender contracts in Switzerland, and they represent about 10% of companies in the country. They provide goods and services at the federal and cantonal level for about CHF 80 billion per year.[13] In case a company is selected, the FOGE informs the company and requests information and data of all its workers to carry out the checks.

9. Previous to *Logib*, the principle of equal rights between men and women was introduced in Switzerland in 1981 in the Swiss Federal Constitution (Art. 8, al. 3 - RS 101). In 1994, the Federal Act on Public Procurement establishes that the contracting authority will only award public tender contracts to companies that guarantee equal pay between men and women.(FAPP 1995, Art 8c), on December 16, 1994 (172.056.1). In 1995, the Federal Act on Gender Equality was enacted to prohibit any type of discrimination between men and women in the labor market, particularly regarding to hiring, allocation of duties, working conditions, pay, basic and advanced training, promotion and dismissal (March 24th, 1995 (LEg -RS 151.1), LFE (1995)). This principle entitles equal pay for equal work.

10. *Logib* has been developed in 2004 by Strub (2005) on behalf of the FOGE based on a regression analysis to quantify wage discrimination and to promote gender equality within firms. The Strub's method follows the typical economic approach when studying the sources of the gender pay gap based on the Blinder-Oxaca decomposition. This method estimates wage regressions specifying the relationship between wages and productivity-related characteristics. Statistically, the gender pay gap can be decomposed into two components: an explained and unexplained part. The explained part is due to objective and measured characteristics such as observable factors like education, experience, tenure, hierarchical position, etc. The unexplained part refers to the portion of gender wage differences which is not accounted for (Blau and Kahn 2007), and it is usually attributed to labour market discrimination. Studies that preceded Strub (2005) in Switzerland are for example Flückiger and Ramirez (2000). The latter use a regression analysis to investigate the wage difference between men and women using the Swiss Labour Force Survey (SLFS) of 1994 and 1996. In their study, three components were identified as main drivers of wage differences: productivity characteristics given by human capital, price structure characteristics, and enhancement of those characteristics explained by variables such as education, experience and tenure, civil status, employment rate, hierarchical position, public or private employee, promotion system, and a discriminatory factor.

11. i.e. in each computer where the *Logib* tool has been installed.

12. Similarly, the American legislation mandates that AA must be taken by federal contractors and subcontractors "to recruit and advance qualified minorities, women, persons with disabilities, and covered veterans" (41 CFR Part 60-1 for Obligations of Contractors and Subcontractors, 41 CFR Part 60-2 for Affirmative Action Programs, and 41 CFR Part 60-741 for Affirmative Action and Nondiscrimination Obligations of Contractors and Subcontractors Regarding Individuals with Disabilities

13. The contracts of these companies are listed in the electronic platform of public procurement (www.simap.ch).

The FOGE verifies the provided information and runs the wage regression mentioned above using its own econometric software. In practice, this method consists in a very detailed wage gap regression (eq. 1 in section 5) controlling for observables and a female variable dummy (1 if women, 0 otherwise), which represents the coefficient of unexplained gender wage differences in the firm. The female dummy takes a negative value if the wage premium favors men and is positive otherwise. The FOGE considers an estimated female coefficient bigger than 5% and statistically significant at 95% of confidence level as an indication of gender wage discrimination. Firms therefore can verify if they are discriminating by looking at the magnitude and confidence level of this estimated coefficient resulting from the wage regression. For the purpose of this study, this coefficient will be called "Logib wage gap".

If a statistically Logib wage gap estimate is found (estimate ≥ 0.05 and p-value ≤ 0.05), the FOGE grants the monitored company a period of 6 to 12 months to correct its wage policy and to ensure wage equality. However, if after a second evaluation carried out by the FOGE similar Logib wage gap estimates are obtained, then legal sanctions are enforced (Art. 4, FOGE and Linder (2013)). They include exclusion from public procurement process, revocation of licenses and/or cancellation of approved contracts. Also, no new tender process will be settled until the firm reaches Logib wage gap estimates below 0.05.

Logib is similar to the Federal Contract Compliance Programs that monitor anti-discrimination and affirmative action provision of contractors and subcontractors in the US, but it is very weakly enforced. The process of requesting information from firms, verifying consistency of the provided data, estimating Logib wage gap coefficients and following up on the process is costly as it represents about 100 working hours of high skilled work.[14] For instance, assuming a conservative mean hourly wage of CHF 70, the cost of such an evaluation would be about CHF 7'000. For this reason, few cases have been analyzed since 2006. Table 1 shows the number of monitored firms between 2006 and 2014. Only 43 firms have been evaluated and no firm has been sanctioned.

Table 1: Monitoring Statistics

Checks 2006-2014	**Number of cases**
Total ongoing checks	15
Total completed checks	28
First round:	
No discrimination	9
Discrimination $< 5\%$	16
Discrimination $> 5\%$	3
Second round:	
Sanctioned after correction	0

Source: FOGE.

14. Information provided by the FOGE.

4 Data and Descriptive Statistics

4.1 Data

This paper uses the Swiss Wage Structure Survey (SWSS), a biannual survey among firms that is administered by the Swiss Federal Statistical Office (FSO). The SWSS is one of the largest official surveys in Switzerland that collects not only information about size and geographic location of a firm, but also about socio-demographic characteristics of its workers. The SWSS is based on a written questionnaire sent to companies, and it is conducted every two years in October. It provides representative data for all economic branches except agriculture, thus depicting the structure of salaries in Switzerland on a regular basis. Eight waves between 1996 to 2010 are used in the analysis.[15] The structure of the data used in the analysis is a repeated cross-section. Firm identifiers are not publicly available in this survey and therefore, it is impossible to use a panel structure based on the firm ID. To analyze the data before and after the introduction of *Logib*, I have pooled all cross-sectional information using firm size as a reference for testing the potential discontinuities of gender wage gap.

The SWSS is designed based on random selection of firms following different criteria for each firm class. The stratification sample is made in two levels: firms and workers. All firms are used to construct the drawing and response rate at the company level.[16] Since 2000, companies are divided in 3 classes: less than 20 workers (small), from 20 to 49 (medium), and 50 workers or more (large). The sample has been constructed using an average drawing rate of 20% for small, 58.3% for medium, and 87% for large companies. Large firms are required to report information for at least of 33% of their employees, medium size businesses 50%, and small businesses provide them all. Different sampling rates that change at the cut-off point will potentially increase the confidence intervals of the estimated coefficients at the threshold and, consequently, it will be less likely to find significant RDD estimates.

Survey details are presented in Table 2. This table reports the number of workers and firms for each year available in the SWSS. To provide a good idea of the information collected in this survey, the sum of observations for the period before the introduction of *Logib* (1996-2004) and after (2006-2010) is reported. In total, the data set contains approximate information for nine million workers which belong to 223,962 firms: 155,249 are small (69.32%), 37,968 are medium (16.95%), and about 41,188 (18.39%) are large companies. The number of workers in small firms represent 11.44%, those in medium firms 9.77%, and workers in large firms 78.79% of the total workforce. For the period before *Logib* (1996 - 2004), information of 4,283,073 workers which belong to 95,988 firms were reported in the data. When considering the period from 2006 to 2010, the SWSS provides information for 4,861,435 workers which belong to 127,974 firms.

Although in practice firms report wage information for most of their workers regardless of their company size group; other variables such as education, experience and other worker characteristics might be missing. For this reason, the sample of analysis was restricted to firms that provide all worker information (Table A1 in Appendix).

15. The wave of 1994 was not included in the study because it uses a different industry classification than the more recent ones starting from 1996. Data after 2010 was not available at the time of the study.

16. This is constructed based on the Business and Enterprise Register (BER).

The first part of the investigation consists in an evaluation of the wage policy of each firm of the SWSS using exactly the *Logib* specification. In other words, I carry out a within-firm analysis for each year that exploits information about wages and socio-demographic characteristics of its workers. Only individuals of working age (between 16 and 65 years) are included. The study employs standardized monthly wages excluding night and Sunday work, as well as allowances. In a second stage, the study uses company information. The *total sample* used in this analysis includes all firms that employ at least 5 men and 5 women. To remove outliers, the 5% tails of the distribution of dependent variables (Logib wage gap and Raw wage gap estimates) have been excluded from the RDD. Various analyses on different sub-samples are performed. The data set is further restricted to a *local sample* which includes only companies with less than 250 workers.[17] Also Diff-in-Disc estimations use *local samples* of firms with 100 workers or less. Table A2 in the Appendix details the information used in this study after implementing the different constraints. A local RDD analysis at the threshold is performed.

Since sampling rates in the SWSS change by firm size, I use the Swiss Business Census (BC) for the years between 1991 and 2008. This allows me to examine the potential manipulation of firm size when implementing the RDD. The BC is one of the oldest data sources in Switzerland that collects information about the universe of all firms in this country. It includes information of workplaces and persons employed in all companies from the industrial, trade and service sectors in Switzerland. Its objective is to provide information about the performance of economic, social and geographic across sectors of all the economy. The BC is collected 3 times per decade, and it has been made available at the end of September each of those years. In this article, I use company information from years 1998, 2001, 2005, 2008. In 2011 the BC is replaced by the STATENT (from its French name, Statistique Structurelle des Entreprises), which is collected exclusively on-line. The STATENT has not been included in the analysis because it has been collected using a different criteria than the BC.

4.2 Descriptive Statistics: evolution of wages and covariates

In Switzerland, the gross wage difference between men and women working in the private sector has been stable since 1998, and in 2010, women earned on average 23% less than men.[18] Table 3 reports gross average hourly wages measured in CHF, and the percentage of wage difference per year. After 2006, wage differences are smaller than before, particularly for firms with 50 workers or more. Table 4 reports the mean average wage of men and women, as well as the percentage wage difference for each type of company separately for the period before and after the introduction of *Logib*. Small and medium companies have higher percentage wage difference than large companies. The percentage wage difference has fallen more in large than in medium and small companies. The unexplained wage gap has reduced slowly, from 41.1% in 1998 to 37.6% in 2010, but it remains still remarkable.[19]

17. Firms with 250 workers or fewer employees was used as a natural threshold to analyse firms with similar characteristics. Small and medium-sized enterprises (SMEs) in Switzerland have less than 250 full-time employees.

18. In the US, the wage gap against women in the labor market has narrowed since 1970s and particularly in the early 2000s, but it is still significant. In 1963 women earned about 59% of men wages, and in 2012 they earned about 80.9% (Brunner and Rowen 2007).

19. Data provided by the FSO in 2015 available in http://www.bfs.admin.ch/bfs/portal/en

Table 2: Survey Details by company size

By type of company	**Before *Logib*** (1996-2004)		**After *Logib*** (2006-2010)		**All period** (1996-2010)	
Number of workers	4,283,073		4,861,435		9,144,508	
Number of firms	95,988		127,974		223,962	
	N^0	%	N^0	%	N^0	%
Small (≤ 20)						
Workers	483,367	11.29%	562,574	11.57%	1,045,941	11.44%
Firms	73,168	1.71%	82,081	64.14%	155,249	69.32%
Medium ($20-49$)						
Workers	385,079	8.99%	508,699	10.46%	893,778	9.77%
Firms	16,606	0.39%	21,362	16.69%	37,968	16.95%
Large (≥ 50)						
Workers	3,414,627	79.72%	3,790,162	77.96%	7,204,789	78.79%
Firms	19,352	11.29%	17,989	14.06%	41,188	18.39%

Notes: N^0 refers to number of observations, and % to the percentage. It considers all information of the SWSS without any restriction.

Table 3: Wage evolution

Hourly wage[1]	**Years** 1998	2000	2006	2008	2010
	Total				
Men	37.95	41.36	45.01	47.09	48.64
Women	27.75	29.43	34.47	36.32	37.33
% wage difference[2]	26.9%	28.8%	23.4%	22.9%	23.2%
	Firms with ≥ 50 workers				
Men	37.91	41.32	45.09	47.14	48.73
Women	27.77	29.39	34.77	36.61	37.51
% wage difference[2]	26.7%	28.9%	22.9%	22.3%	23.0%
	Firms with < 50 workers				
Men	39.24	39.02	44.43	46.75	47.72
Women	27.87	27.30	32.59	34.02	35.63
% wage difference[2]	30.0%	28.9%	26.7%	27.2%	25.3%

Notes:
[1] Hourly wage refers to the average hourly gross wage measured in CHF.
[2] % wage difference refers to the percentage wage difference between men (w_m) and women (w_f), measured as $(w_m - w_f)/w_m$.
[3] Source: SWSS.

Table 4: Descriptive statistics of wages by company size

Descriptive statistics	**Before *Logib*** (1996-2004)	**After *Logib*** (2006-2010)
	Small Firm (≤ 20)	
Men	43.62	46.98
Women	30.36	33.67
% wage difference[2]	30.40%	28.33%
	Medium Firm ($20-49$)	
Men	41.67	45.96
Women	30.78	34.10
% wage difference[2]	26.12%	25.81%
	Large Firm (≥ 50)	
Men	40.56	46.62
Women	30.22	36.12
% wage difference[2]	25.50%	22.52%

Notes:
[1] Hourly wage refers to the average hourly gross wage measured in CHF.
[2] % wage difference refers to the percentage wage difference between men (w_m) and women (w_f), measured as $(w_m - w_f)/w_m$.
[3] Source: SWSS.

On the other hand, employment rate is significantly different between male and female employment in many European countries. Switzerland has a high proportion of part-time workers (35.6%) (FOGE and Linder 2013).[20] Since 2010, the majority of women employed in the labor market in Switzerland work part-time (61%), while most men work full-time (85%).[21] Firms with less than 250 workers represent 99.9% of the total number of companies in Switzerland and account for 66.6% of the total employment.[22] Among them, very small companies (up to 9 workers) and medium firms (20-49) account respectively for 87% and 10.6% of all companies in Switzerland. Accordingly, they gather 25% and 22% of the workers. Large firms (> 50) represent only a small proportion of companies in Switzerland (2.4%), but they account for more than half of the total employment.

20. According to the European Labour Force Survey (2009), Netherlands is another remarkable country with high proportion of part-time workers (48.3%), 76% of women working part-time, and 24.9% of men in similar positions (Sandor, Eszter 2011).

21. These percentages are based on the male(female) working population, where all persons who work at least one hour per week are considered to be employed. Source FSO - Employment Statistics.

22. This considers all the companies of the Swiss BC of 2008.

5 Methodology: Mincerian Analysis, RDD and Diff-in-Disc Design

5.1 Gender discrimination through wage regression analysis

As a first step of the analysis, I estimate a wage regression for each company and each year from 1996 to 2010 using the same specifications as in *Logib*, detailed in equation 1 (Strub 2005).

$$\ln(w_{ij}) = \alpha_j + \beta_{\mathrm{j}}\mathrm{fem}_{\mathrm{ij}} + \theta_j X_{ij} + \mu_{ij} \qquad \forall j \tag{1}$$

Equation 1 is an Ordinary Least Squared (OLS) regression that uses an extended Mincerian specification.[23] The index i refers to information of each worker, and j to each firm, $ln(w_{ij})$ refers to the logarithm of hourly wage, *fem* to a dummy variable: 1 if the worker is a woman, and 0 if is a man, X_{ij} refers to the specific vector of control variables as detailed in *Logib* (education, experience, tenure, hierarchical position, and job difficulty), and μ_{ij} to the error term that is assumed to be normally distributed with mean zero and constant variance. Hierarchical position is a categorical variable that depends on the level of professional responsibility of a worker and attributes the following values: (1) senior, (2) middle-management, (3) junior workers, (4) low management, and (5) no management functions. The level of difficulty of the job takes the values of: (1) the most difficult, (2) independent work, (3) professional knowledge, and (4) simple and repetitive tasks. The gender estimated coefficient ($\widehat{\beta_j}$) from this regression refers to the Logib wage gap estimate of each firm in each year, and $\widehat{\theta_j}$ indicates the vector of estimated coefficients of the control variables.[24]

After computation of the Logib wage gap estimates ($\widehat{\beta_j}$) and their respective standard errors $se(\widehat{\beta_j})$ for each firm (j) and each year, I create a new data set merging these results with other important firm characteristics, such as size, industry, and a binary variable indicating whether the company belongs to the private or public sector. Then, I apply RDD to this combined dataset. The peculiarity of the RDD setting employed here consists in using an estimated parameter as the dependent variable. Since the variance of ($\widehat{\beta_j}$), and therefore its statistical significance, depends inversely on the number of observations in each firm, one might want to correct for the uncertainty in the *Logib* estimates by giving higher weight to those estimates that have less uncertainty. For this reason, I weight the RDD estimations by the inverse of the standard error of each Logib wage gap coefficient $se(\widehat{\beta_j})$. As a result, similar than any

23. Where the exact policy recommendation as detailed in the documentation is $\ln(wage) = \beta_0 + \beta_1\, educ + \beta_2\, exp + \beta_3\, exp^2 + \beta_4\, ten + \beta_5\, hie + \beta_6\, dif + \beta_7\, fem + \mu$ where: *educ* refers to education, *exp* to experience, *ten* to tenure, *hie* to hierarchical position, *dif* to level of difficulty of the post, and *fem* to a female dummy.

24. I am aware that a residual regression such *Logib* has its limitations and one would like to control for all potential factors that can make female and male workers more alike and better attribute unexplained wage gaps to discrimination. As pointed out by Blau and Kahn (2007), any study based on a statistical residual will open the question of whether all necessary explanatory variables were included in the regression. It is probable that even the more extensive regressions with almost all observable workers' characteristics still provide inaccurate estimates of wage discrimination. However, residual regressions are a good starting point to understand the sources of wage inequality. Other factors such as civil status, and the number and age of children can also affect the residual gap. The literature suggests a negative impact of children on female wages, especially when women are young. Also, children matter for labor participation and experience accumulation. Related to that, the availability of childcare facilities as well as extended-family support positively influence female labor participation. Finally, Ryder (2014) suggests that paternity leave can have positive effects for gender equality at work as well as at home, but much more research needs to be done. However, all these factors do not represent a problem for this study since this paper does not aim to identify all explanatory variables that can explain gender wage differences, but rather estimate the causal effect of this policy exactly as designed here.

other weighted least squared estimation, our RDD estimates will be consistent. However, their variance will be larger and suffer from heteroskedasticity; implying that it will be less likely to find statistically significant RDD estimates

5.2 Regression Discontinuity Design (RDD)

To infer how strongly the Swiss wage policy affected firms, we need to identify the effect of this law on firms with at least 50 workers. I use a RDD to test the effect of the federal recommendation at the aggregate level by studying the potential discontinuity of these Logib wage gap estimates at the relevant threshold (firm size = 50). At this cut-off, differences in gender wage estimates can be expected to originate from treatment differences between companies with less than 50 workers and bigger ones.

The treatment here is twofold: first, the FOGE recommendation to test firms wage policy using *Logib*, and second, the monitoring rule. Although the FOGE encourages all companies not to discriminate, for statistical reasons, the implementation of *Logib* is only recommended to firms with at least 50 workers. Of those companies, only a random selection of private firms who have won public tender contracts in Switzerland will be monitored. Since it is not possible to disentangle the effect of the *Logib* recommendation from the effect of monitoring, all firms with 50 workers or more are affected by this regulation. Also, even if only public tenders are randomly checked, potential contractors might want to comply with this regulation.[25] This is the reason why a *sharp* RDD is used here.[26]

The assignment rule of the recommendation to check a firm wage policy via *Logib* can be described as:

$$D_j = \begin{cases} 1 & \text{if} \quad S_j \geq 50 \\ 0 & \text{if} \quad S_j < 50 \end{cases} \tag{2}$$

Due to the selection threshold, one would generally expect that firms with less than 50 workers ($S_j < 50$) have higher Logib wage gap estimated coefficients $\widehat{\beta_j}$ than larger firms after the introduction of *Logib*. The arguments that lead to this hypothesis are as follows. First, firms with 50 workers or more will face higher marginal costs than smaller firms to respect this federal recommendation. To minimize costs and to avoid being punished, firms can adjust their wage structure, which will translate into lower wage

25. One can reasonably argue that all companies categorized in the treatment group are indeed treated because the use of the *Logib* tool was recommended to all firms above the threshold and the presence of non-compliers cannot be accounted for in the analysis. It has been very difficult though to test this hypothesis without precise information of public tender firms. The Federal Office for Building and Logistic does provide information concerning neither size nor characteristics of companies that work under public procurement. Moreover, data on public procurement is reported at the federal, canton and municipal level.

26. Indeed, one may hypothesize that companies that do not qualify for governmental controls will not implement the *Logib* recommendation despite being in the treatment group. If someone would be interested to identify specifically the effect of *Logib* only on public tenders, then the causal effect of being assigned to the treatment would be identified by the Average Treatment Effect on the Treated (ATT) as $ATT = ITT/Pr[(\widehat{\beta_j}) = 0]$, where ITT refers to the Average Intent-to-Treat effect. Data to obtain the exact proportion of firms with $Pr(\widehat{\beta_j}|D_j = 1) = 0$ is not available. Information of public tenders is only provided by contracts and not by firms (see www.simap.ch) and there is not such available data by firm or size identifier. Therefore, providing accurate ATT estimates has not been possible. However, to have an idea of this effect, we can provide a rough estimate. Let's assume that the ITT effect is 5 percentage points, and we know that approximately 10% of companies in Switzerland work under public procurement; then the *ATT* estimate would be about 0.5 percentage points. Then, due to the lack of information on public tenders, estimations presented here are likely to be lower bounded.

gap coefficients. Whether firms accommodate the composition of labor force is discussed in section 6.3. Second, firms may want to be seen as examples of having complied with this regulation to obtain future tender contracts and to build a good reputation by showing their commitment to gender equity.

The empirically estimated RDD equation can be written as:

$$\widehat{\beta_j} = f(S_j) + \rho\,[\mathrm{S_j} \geq 50] + \gamma[S_j \geq 50]f(S_j) + \Gamma\, Z_j + \eta_j \tag{3}$$

where $\widehat{\beta_j}$ represents an element of the vector of Logib wage estimates, $f(S_j)$ is a non-linear function of firm size, ρ the coefficient of interest, $[S_j \geq 50]$ refers to the treatment assignment, Γ represents the coefficient vector of control variables (year, industry, sector dummies, and their interaction with $[S_j \geq 50]$), and η_j the error term. As mentioned before, Eq. 3 is weighted by the inverse of the standard error of the $\widehat{\beta_j}$ to account for the statistical significance of the gender wage gap estimate. The identification of the treatment effect will be strongly valid especially for firms around the cut-off point, i.e firms located marginally below and above the threshold of 50 workers. This means that the results will have strong internal validity.[27] However, if any discontinuity is found at the relevant threshold, one could worry about systematically different characteristics between these two groups of firms. Although it is not very clear why these characteristics should vary exactly between firms with 49 and 51 workers in Switzerland, where no other firm-size based regulation was enacted; we could reasonably argue that Human Resources departments operate differently in smaller companies than bigger ones. Below, I address this concern.

5.3 Differences-in-Discontinuities (Diff-in-Disc)

To ensure we capture exactly the effect of *Logib* which only affects firms with at least 50 workers after 2006, as a third step of the analysis, I implement the Diff-in-Disc estimator developed by Grembi, Nannicini, and Troiano (2016). This will allow us to further test the effect of *Logib* on wage gap estimates before and after its introduction. If companies with less than 50 workers are structurally different than bigger companies, differences should persist after the introduction of *Logib*. Otherwise, if no discontinuity is found at the threshold of 50 before the introduction of *Logib* but only after 2006, we can be confident that the causal effect of this policy has been identified.

In this set-up, the Diff-in-Disc estimator takes the difference between the discontinuities before the introduction of *Logib* (1994-2004) and after the introduction of *Logib* (2006-2010), by fitting a linear regression function to the observations distributed within the distance (measured in company size) on either side of the threshold (firm size $\in [S_{50} - h, S_{50} + h]$) before and after the introduction of *Logib* Grembi, Nannicini, and Troiano (2016). The estimated model is now as follows:

$$\widehat{\beta_j} = \delta_0 + \delta_1 f(S_j) + [S_j \geq 50](\gamma_0 + \gamma_1 f(S_j)) + T_t\{\alpha_0 + \alpha_1 f(S_j) + [S_j \geq 50](\beta_0 + \beta_1 f(S_j))\} + \xi_{it} \tag{4}$$

where $[S_j \geq 50]$ refers to the dummy for firms bigger than 50 workers, T_t to the post-treatment indicator

27. To be sure that the treatment effect ρ at the threshold is still the coefficient on $[S_j \geq 50]$, I normalize the cutting-off point at size 50 to 0. Making the cut-off at 0 assures to identify the difference in observed mean outcomes marginally above and below the threshold. Having a cut-off point different from 0 would bias the intercept and therefore $\hat{\rho}$.

(after 2006). Similarly as before, S_j refers to the normalized firm size, centered to the mean, and regressions are weighed by the inverse of the standard errors of the Logib wage gap estimates. The Diff-in-Disc estimator will be captured by β_0 and identifies the treatment effect $D_{it} = [S_j \geq 50] * T_t$. Different bandwidths are used to test the results.

To allow for flexible functional forms to fit the relationship between $\widehat{\beta_{jt}}$ and $f(S_j)$ below and above 50 workers, before and after 2006, I run the following model for the local and total sample:

$$\widehat{\beta_j} = \sum_{k=0}^{p}(\delta_k f(S_j)^k) + [S_j \geq 50]\sum_{k=0}^{p}(\gamma_k f(S_j)^k) + T_t[\sum_{k=0}^{p}(\alpha_k f(S_j)^k) + [S_j \geq 50]\sum_{k=0}^{p}(\beta_k f(S_j)^k)] + \xi_{it} \quad (5)$$

Although a parametric approach can provide more precise sample average estimates when using a large data set like the SWSS, it entails the risk of generating biased estimates in the neighborhood of the boundary due to an inaccurate model specification (Lee and Lemieux 2010). To avoid potential poor finite sample properties of standard Wald estimates and boundary bias of traditional kernel estimators, local linear non-parametric regressions are estimated. When the running variable is continuous, they are typically based on weighted local linear at both sides of the cut-off. Implementation of non-parametric estimations is, however, not possible here since the running variable of the design is discrete.

5.4 Hints from an Oaxaca-Blinder (OB) decomposition

We are primarily interested in identifying the causal effect of *Logib* on unexplained wage gaps, but we may also want to know if it indeed reduces the raw wage differences between men and women, and how it affects other labor outcomes such as employment. Furthermore, in the program evaluation literature, the causal analysis of a policy studies mainly the impact on the target outcome of interest, but it does not provide explanations about the mechanisms underlying the relationship between factors and outcomes. However, it is often of scientific interest to investigate the side effects of a policy and why such impacts occur (Keele, Tingley, and Yamamoto 1961).

To identify the direct and indirect effects of this policy, I measure and decompose the Raw wage gap before and after 2006 using an OB decomposition. A detailed OB method decomposes the mean raw wage difference between men and women, and attributes this difference to inequalities in the different covariates for which I control. This is a standard tool that applied economists typically use to understand to what extent gender wage gaps can be attributed to returns to skills or unexplained factors (Fortin, Lemieux, and Firpo 2011). Estimated parameters will inform to which extent raw wage differences are due to inequalities in education, experience, hierarchy, etc. For example, finding a significant change after 2006 of job difficulty will shed light on this variable, pointing out two potential interpretations. On the one hand, one can infer that the increase in the contribution of job difficulty was because employers adjusted wages of women who perform very difficult jobs. On the other hand, this change can inform as to potential adjustments in the composition of the workforce, leading to increasing employment of women who perform difficult tasks. Ultimately, estimates from the OB decomposition will tell us which part of the gender wage gap can be associated with differences in characteristics (explained gap), or discrimination (unexplained gap). OB decompositions usually assume that the outcome variable Y linearly depends on related covariates ($X_i = [X_{i1}, ..., X_{iK}]$], and the error term v. Following Fortin, Lemieux, and Firpo (2011), by disaggregating the gender wage gap, we can identify which variables are more important to explain differences in slope coefficients ($\widehat{\beta}_{gk}(k = 1, ..., K)$) and intercepts ($\widehat{\beta}_{g0}$).

This can be summarized as follows:

$$Y_{gi} = \beta_{g0} + \sum_{k=1}^{K} X_{ik}\beta_{gk} + v_{gi}, g = M, W \tag{6}$$

where $E(v_{gi}|X_i) = 0$. Then, the overall difference in average outcomes $\widehat{\Delta}^{\mu}_{O} = \overline{Y}_W - \overline{Y}_M$ between Men (M) and Women (W) can be written as

$$\widehat{\Delta}^{\mu}_{O} = \underbrace{(\widehat{\beta_{W0}} - \widehat{\beta_{M0}}) + \sum_{k=1}^{K} \overline{X}_{Wk}(\widehat{\beta_{Wk}} - \widehat{\beta_{Mk}})}_{\Delta^{\mu}_{S}\text{(Unexplained)}} + \underbrace{\sum_{k=1}^{K}(\overline{X}_{Wk} - \overline{X}_{Mk})\widehat{\beta_{Mk}}}_{\Delta^{\mu}_{X}\text{(Explained)}} \tag{7}$$

When detailed decompositions use dummy variables, they can sometimes be problematic because coefficient effects attributed to these dummy variables are not invariant of the choice groups (Jones 1983). Given the specification of the *Logib* policy, this critique is relevant for this analysis. To overcome this problem, I follow Yun (2005) and use normalized regressions that allow us to identify the intercept and slope coefficients of each dummy variable.[28]

6 Identification

Indeed, the main concern is the potential manipulation firms may perform to change their size in order to avoid being affected by the federal recommendation. To assert that the treatment assignment based on firm size is as-good-as random, I am first interested to show that the distribution of firms in Switzerland is continuous (i.e the distribution of firms with less than 50 workers and the one of firms with 50 workers or more did not change) before and after the introduction of *Logib*. However, if there is no manipulation of firm size, we may also wonder about the possible manipulation of other covariates' distribution. Therefore, I investigate whether the distribution of other covariates significantly changed for firms with at least 50 workers after 2006. Finally, one could wonder if employers assign particular values to other covariates such as hierarchical position and job difficulty. To be sure employment manipulation of these categories does not take place, I examine whether a discontinuity exists in the proportion of female employment across groups of workers classified by the job difficulty and hierarchical position.[29]

28. See Jann (2008) for implementation details in Stata.

29. Additionally, one may wonder about the role of military service on work composition. In Switzerland, military service is compulsory for all Swiss men aged between 18 and 34. Participation of women is voluntary. Compulsory military training lasts 260 days. The minimum initial military training lasts 21 weeks (124 or 145 days), and is followed by training periods of 19 days each (7 for those opting for the shorter initial training and 6 for those choosing the latter.) Swiss militia calls its recruits at the age of 20 and training can not be postponed to finish university (Mannitz and Hass 2009). Most of Swiss men take this initial training before starting a contractual employment. In case Swiss men choose to enroll in civil defense, they have to serve until the age of 40. Men who can attest their physical or mental inability to serve the military must pay

6.1 Firm size manipulation

In contrast to the French legislation or similar laws, neither the Swiss Federal Labour Act (RS82) nor other Swiss regulation establish legal obligations that differentiate between firms with less than 50 workers and firms with 50 workers or more.[31] Obligations for companies in Switzerland vary according to their legal structure (Sole proprietorship, Limited liability Companies (GmbH), Public Limited Companies (AG), Limited Partnership or General Partnership). Furthermore, taxes which are levied on federal, cantonal and municipal level do not depend on company size, but vary across cantons and municipalities. This provides initial evidence for the absence of employers' incentives to control firm size at the cut-off point of 50.

The SWSS considers different drawing and response rates for each company group ($1-19, 20-49, 50+$ employees) resulting in reduced representation of large companies, particularly for companies with 20 and 50 employees.[32] Since these structural characteristics of the SWSS made analysis of the distribution of firms particularly at the threshold impossible, I use the BC data. Employing the BC, a survey of all second and third sector business and companies in Switzerland, has two advantages: First, it collects information of all companies in Switzerland. And second, it avoids potential sample design problems of the SWSS.[33] After analyzing graphically and statistically the continuity of the running variable (firm size) using the McCrary (2008) test, results fail to reject the null hypothesis of continuity of firm density function at 50. This allows to conclude that firms are continuously distributed at this threshold. Consequently, if any discontinuity is found at the size of 50, we could be confident that it is due to the introduction of *Logib*. Graphical distribution of firm size and statistical results from McCrary (2008) tests are presented in Figure 1.[34] Figures show clearly a continuous distribution of firm size and results fail to reject the null hypothesis of continuity of the running variable.

To ensure that results are not driven by specific industries in the economy (i.e. industries with higher proportion of public tenders), I study the distribution of firms at the threshold of 50 within different industries (Figure A1 and A2, the Appendix). Evidence shows that firm size is also continuously distributed within industries. Due to data limitations, it is not possible to compare industries that intensively used public tenders versus other industries in the economy.

---

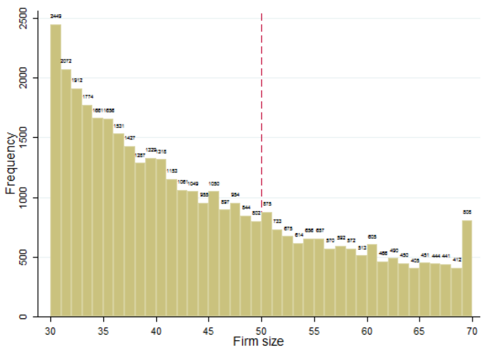
an additional tax of 3% when they are aged between 20 and 30. Similar to maternity leave insurance, the Swiss Federal Law establishes the duration and amount of compensation for loss of earnings due to participation in military and civilian services. The beneficiaries are persons serving in the Swiss army or the Swiss Red Cross, serving in civilian service or civil defense.[30] However, since companies are not allowed to dismiss an employee when workers are serving in the military, civil defense, or civilian service; we rule out this scenario. Furthermore, military service laws did not change after the introduction of *Logib* and therefore, it is very unlikely they have a confounding effect on wage gaps and employment dismissal. I acknowledge Kevin Milligan for pointing out the potential effect of military service on relative male/female employment dismissal.

31. For example, Garicano, LeLarge, and Reenen (2013) use French employment protection legislation restrict firms with 50 or more employees to identify equilibrium and welfare effects. According to the Federal Audit Oversight Authority (FAOA), articles 727 of the 1^{st} and 2^{nd} chapter and following articles of the Code of Obligations, which states the obligations of auditors, stipulates that firms must have 250 full-time equivalent workers in average per year to be subject of audits from 1st of January 2012.

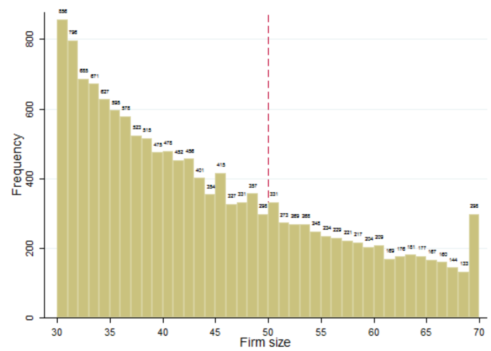
32. Regardless their size group, provided most information about wages, response rate design matters when reporting information about worker and firm characteristics.

33. As mentioned before, firms are sampled differently below and above 50 workers in the SWSS.

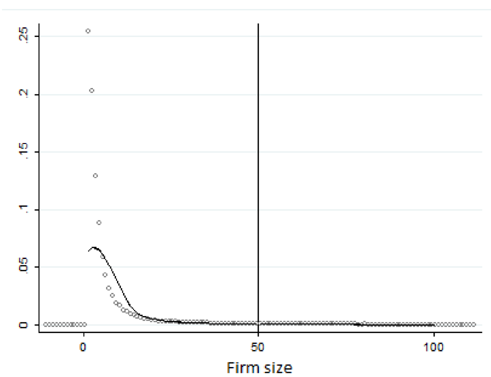
34. A similar analysis has been conducted using the SWSS. Details of firm size histograms and McCrary (2008) tests using the SWSS are available on request. In this case, an indication of non-sorting can be attributed to the presence of similar McCrary (2008) estimates for periods before and after the introduction of *Logib*.



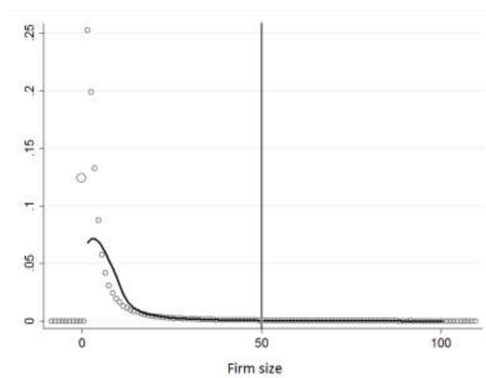
(a) Firm distribution (Before *Logib*)



(b) Firm distribution (After *Logib*)



(c) McCrary Test (Before *Logib*)



(d) McCrary Test (After *Logib*)

Figure 1: Firm distribution by company size
Source: Business Census (BC). Upper panels (a) and (b) show clearly a continuous distribution of firm size. Results of the McCrary (2008) test, presented in the bottom panels (c) and (d), fail to reject the null hypothesis of continuity of the running variable.

Table 5: Estimations from McCrary Test

McCrary estimates	**All years 1998-2008**	**Before Logib (1998-2005)**	**After Logib (2008)**
$\hat{\theta}$	0.059	0.021	0.047
se $\hat{\theta}$	0.034	0.033	0.056
No obs.	1,729,484	1,277,833	451,651

Notes: A bin size of 1 is imposed for the computations. Source: Business Census (BC).

6.2 Selection on Observables: RDD on covariates

To be sure the treatment effect is completely randomized and that we are identifying the causal effect of the policy, I am interested to show that no discontinuities in observable variables are observed prior the assignment. To test whether there is manipulation or sorting of those variables, I perform a RDD analysis on other covariates (X_j) from regression policy 1. Table 6 reports the RDD and Diff-in-Disc estimated coefficients. Columns (1) and (2) report RDD estimates of the parameter of interest $[S_j \geq 50]$, and column (3), the Diff-in-Disc estimates of $[S_j \geq 50] * [T_t \geq 2006]$. Each row of the table refers to a particular covariate estimate of the policy recommendation which is used as dependent variable.[35] Neither RDD after 2006 nor Diff-in-Disc estimates of any covariate (differently than the gender estimate) are statistically significant, which provides evidence of continuity of these observables from which we can infer that the treatment is randomized.

Table 6: Diff-in-Disc estimates of Baseline Covariates (θ_j)

Outcome or Parameter	RDD $\leq$ 2006	RDD $\geq$ 2006	DIFF-in-DISC
θ_j	(1)	(2)	(3)
Education	-0.028	-0.010	-0.012
Experience	0.006	0.001	-0.002
Experience squared	0.143	0.021	-0.022
Tenure	-0.001	0.001	-0.000
Senior	0.326***	-0.074	0.106
Middle management	0.057	-0.011	-0.014
Junior worker	0.119***	-0.012	0.007
Responsible of execution	0.101***	0.007	-0.004
The more exigent	0.192*	0.146	-0.058
Independent & very qualified	-0.055	-0.028	-0.034
Simple & repetitive work	0.070	0.065	-0.027

Note: RDD regressions consider similar controls as performed for female or *Logib* estimates. Reported DIFF-in-DISC estimated regressions refer to specifications that include multiplicative dummies. Reported values refer to estimate from local samples in the range of company size between $[10, 90]$. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Variables such as employment rate, proportion of female workers, or proportion of skilled workers in the firm cannot be studied as baseline covariates because they can be determined endogenously. They are instead studied as potential outcomes and used to explain how firms adjust the composition of their labor

35. These covariate refer to all other control variables different than the gender dummy detailed in the policy recommendation.

force when facing a recommendation such as *Logib*.

6.3 Potential manipulation of employment variables

Although the policy documentation clearly defines work categories, the documentation does not provide details on how to classify workers into these categories. By hierarchical position, workers can be classified in senior, middle management, junior workers, low management (responsible for work execution) or without management position. Using job difficulty, workers can be sorted into groups of workers who perform the most difficult tasks, work independently, use professional knowledge or perform repetitive tasks. In other words, firms have the discretionary power to adjust the value of some categorical variables such as hierarchical position and job difficulty. Manipulation of these variables could be used to comply with the policy without improving gender pay equity.

If employers were motivated by this incentive, one could expect changes in the composition of the workforce of firms with 50 workers or more. For example, women in senior and managerial positions will be recoded as 'responsible for execution'; and women doing difficult tasks as those doing 'routine or repetitive work'. As consequence, one might predict: (a) an increase in the proportion of female workers in low skilled categories, with respect to the total female labor force, and (b) a decrease in the proportion of female workers in high skilled categories with respect to the total female labor force. To test this behavior, I study the employment rate of workers categorized exactly in those work categories. Specifically, I analyze the proportion of women employed in these particular work categories divided by the total female labor force in each firm. As a result, Diff-in-Disc estimates of the proportion of female workers in these occupational categories is not statistically significant. This can be interpreted as evidence that employers do not chose to manipulate those variables to be in compliance with the *Logib* regulation (See Tables A3 and A4 in the Appendix).

7 Results

This section informs on the effect of *Logib* on gender wage gaps, whether wage gap effects resulted in a decrease of employment, and if overall results were driven by particular groups of workers. This section starts by providing descriptive evidence about the direction and magnitude of the Logib wage gap estimate for firms below and above 50 workers to inform about who (men or women) mostly received an unexplained wage gap. Second, for each outcome I provide visual information on the RDD estimates by graphing the evolution of mean estimates of each dependent variable across firm size before and after the introduction of the policy. While the size of the discontinuity at the threshold of 50 refers to the corresponding RDD estimate for each period, the resulting Disc-in-Disc regression estimates allow to derive the difference between the size of discontinuities at the threshold of 50 before and after the introduction of *Logib* . Third, I interpret the regression estimates, which allows to make statistical inference on the estimated effect of *Logib*. Regression specifications presented here use multiple polynomial specifications and control for industry type, and for the firm belonging to the private or public sector. Finally, I turn to the

Oaxaca-Blinder (OB) decomposition for the same periods of analysis, which enables a sharper focus on the side effects of *Logib*. Results are computed on the total and local samples.

7.1 Logib wage gap estimates

OLS estimations

In this section, I first analyze the distribution of Logib wage gap estimates for treated (firm size ≥ 50) and control group (firm size ≤ 50) firms.[36] Results show that estimates are highly concentrated between -0.5 and 0.5, suggesting unexplained wage penalties for men and women. However, estimates are more disperse in companies with 50 workers or more, than in small companies (Figure A3 and A4). Even though the distribution of the estimates before and after the introduction of *Logib* is not very different, this finding suggests differences in gender estimates between small firms and companies with at least 50 workers. Table A5 in the Appendix reports the number of firms by the sign of their Logib wage gap estimate. The upper panel of this table shows gender wage estimates for the total sample, while the lower panel shows gender wage estimates for firms that exceed the 5% tolerance level for all the years of the analysis.[37] In most of the cases, gender estimates are negative regardless of firm size and year, indicating an unexplained wage gap mostly against women.

RDD estimations

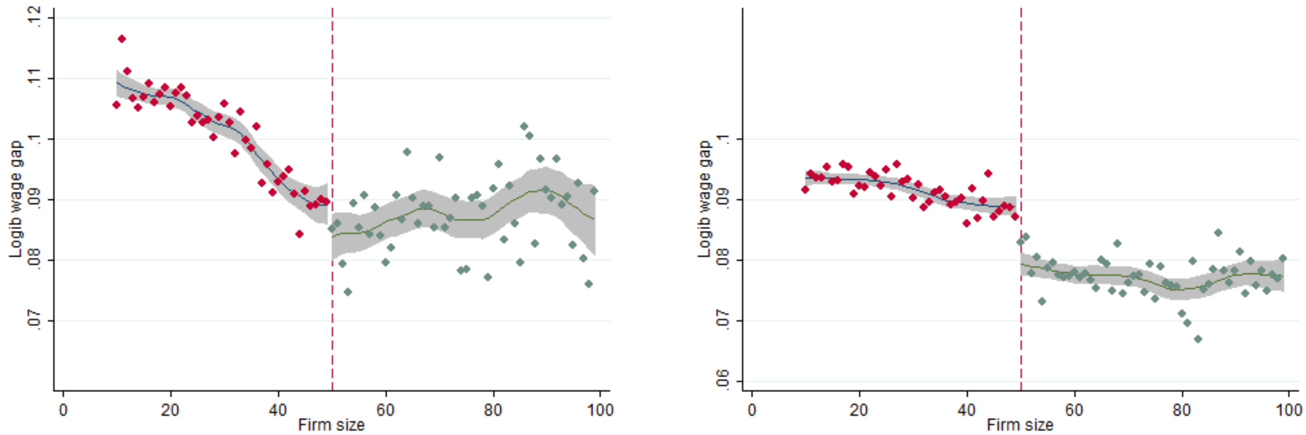
I start by analyzing the impact of *Logib* on the Logib wage gap, which is the main outcome of interest for which *Logib* was designed. Figure 2 graphs the mean Logib wage gap estimates by firm size, with solid lines referring to polynomial fits of third degree and shadowed areas to the 95% confidence interval of the fit. These figures show smaller mean Logib wage gap estimates for firms with 50 workers or more.[38] However, only after the introduction of the *Logib* tool, significant discontinuity in Logib wage gap estimates at company size of 50 employees is observed. Accordingly, firms with 50 workers or more display smaller unexplained gender wage gap estimates than smaller firms, which indicates the effectiveness of the policy. Logib wage gap estimates for these firms are decreasing after the introduction of *Logib*, which is consistent with the idea that firms strive to comply with the regulation as they increase in size in order to qualify for public tenders. Similar results are obtained using the Local Sample (Figure A5 in the Appendix). Subsequently, these results are tested statistically using RDD regressions.[39] Baseline regressions include as explanatory variables, firm size, its interactions with year dummies, and a dummy variable indicating if the firm has at least 50 workers. Table 7 reports RDD estimates for the period before the introduction of *Logib* (from 1996 to 2004), and Table 8 for the period after its introduction (from 2006 to 2010). For the period before the introduction of *Logib*, Table 7 reports positive and non significant point estimates

36. Strictly speaking, the first dependent variable of interest is the result of merging in one dataset all estimated coefficients of Logib wage gap estimates ($\widehat{\beta_j}$) from equation 1 in section 5.

37. Disaggregated results for the periods before and after the introduction of *Logib* are presented in Table A6 in the Appendix.

38. Details of ATE and RDD graphical computations at the threshold before and after 2006 are shown in Table A8.

39. Since the regulation tackles wage discrimination in general (independently to be against men or women), the RDD analysis is performed using absolute values of Logib wage gap estimates. The use of absolute values will allow to maximize number of firms used in the analysis. As noted before, to account for the significance of firms' Logib estimates and against whom firms discriminate, the RDD is performed on all firms weighted by their standard errors.



(a) Before the introduction of Wage Control (1996, 1998, 2000, 2002, and 2004)

(b) After the introduction of Wage Control (2006, 2008, and 2010)

Figure 2: The effect of firm size on *Logib* wage gap
Notes: Graphs are made using the *total sample.* Scatter points show the average of the mean gender coefficient by firm size. Solid lines refer to a third degree polynomial fit, and the shadowed areas indicate the 95% confidence interval of the fit. RD regressions control for industry and private/public sector.

in all specifications. While for the period after the introduction of *Logib*, Table 8 shows negative and significant coefficients in all specifications. These results show evidence that *Logib* affected firms with at least 50 workers after its introduction, but there is essentially no sign of the effect of this policy before 2006. Using RDD estimates, the magnitude of the effect of *Logib* after 2006, in the most conservative specifications range between -0.051 and -0.073, reveals a decrease in the unexplained wage gap for firms tackled by this regulation. To verify if the estimated effect is caused by the introduction of the *Logib* recommendation, I implement a "placebo test" using a Diff-in-Disc design. Diff-in-Disc estimates the relationship between the unexplained wage gap and firm size for the period before and after the launch of *Logib*. The difference between RDD estimates before and after 2006 should be about the size of Diff-in-Disc estimates. Finding significant Diff-in-Disc estimates indicates that the Logib wage gap reduced by about 3.5 percentage points after the introduction of *Logib*.

Finally, an important concern arises from weighting the estimations by the inverse of the standard error. As noted before, weighted OLS regressions increase the confidence interval of the estimated coefficients. This generated an inference problem because larger standard errors will reduce the likelihood of finding significant results. However, having found significant Diff-in-Disc estimates, this represents a minor problem.

Robustness checks

I have found evidence that *Logib* was associated with a statistically significant decrease of unexplained wage gaps. This shows that some of the changes in Logib wage gap estimates were large enough to be detected by my estimates. However, these findings could be affected by particular attributes (polynomial specifications, exclusion of some relevant control variables) of the regression specifications, anticipation effects, etc. To cross-check the stability of the results, I carry out the following robustness checks.

Table 7: Effect of *Logib* on Logib wage gap estimates: Period 1996 - 2004

	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
Total Sample										
firm size > 50	0.0108	0.0108	0.0123	0.0135	0.0130	0.0226	0.0228	0.0250	0.0258	0.0242
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Industry dummies	yes	yes	yes	yes	yes	no	no	no	no	no
Sectoral dummy	yes	yes	yes	yes	yes	no	no	no	no	no
N					18230					
$N < 50$					11413					
$N \geq 50$					6817					
Local Sample ($N < 250$)										
firm size > 50	0.0198	0.0034	0.0118	0.0141	0.0208	0.0154	-0.0156	0.0026	0.0128	0.0175
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Industry dummies	yes	yes	yes	yes	yes	no	no	no	no	no
Sectoral dummy	yes	yes	yes	yes	yes	no	no	no	no	no
N					15311					
$N < 50$					11413					
$N \geq 50$					3898					
Robust check: exclude $47 < N < 53$										
Total Sample										
firm size > 50	0.0090	0.0089	0.0104	0.0116	0.0111	0.0232	0.0235	0.0257	0.0264	0.0249
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Industry dummies	yes	yes	yes	yes	yes	no	no	no	no	no
Sectoral dummy	yes	yes	yes	yes	yes	no	no	no	no	no
N					17834					
$N < 50$					11123					
$N \geq 50$					6711					
Local Sample ($N < 250$)										
firm size > 50	0.0179	-0.0030	0.087	0.0113	0.0252	0.0163	-0.0227	0.0020	0.0262	0.0523
	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.02)	(0.03)	(0.03)	(0.04)	(0.06)
Industry dummies	yes	yes	yes	yes	yes	no	no	no	no	no
Sectoral dummy	yes	yes	yes	yes	yes	no	no	no	no	no
N					14915					
$N < 50$					11123					
$N \geq 50$					3792					

Notes:

[1] Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

[2] Columns (1), (2), (3), (4), and (5) refer to regressions of gender discrimination that contains the first, second, third, fourth and fifth polynomial degree of size respectively.

Table 8: Effect of *Logib* on Logib wage gap estimates: Period 2006-2010

	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
Total Sample										
firm size > 50	-0.0693***	-0.0683***	-0.0677***	-0.0668***	-0.0656***	-0.0670***	-0.0641***	-0.0627***	-0.0610***	-0.0591***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Industry dummies	yes	yes	yes	yes	yes	no	no	no	no	no
Sectoral dummy	yes	yes	yes	yes	yes	no	no	no	no	no
N					27783					
$N < 50$					16901					
$N \geq 50$					10882					
Local Sample ($N < 250$)										
firm size > 50	-0.0630***	-0.0511**	-0.0731***	-0.0316	-0.0692***	-0.0592***	0.0254***	-0.0475***	0.0332***	-0.0333*
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)
Industry dummies	yes	yes	yes	yes	yes	no	no	no	no	no
Sectoral dummy	yes	yes	yes	yes	yes	no	no	no	no	no
N					24066					
$N < 50$					16901					
$N \geq 50$					7165					
Robust check: exclude $47 < N < 53$										
Total Sample										
firm size > 50	-0.0781***	-0.0772***	-0.0766***	-0.0757***	-0.0745***	-0.0714***	-0.0685***	-0.0671***	-0.0654***	-0.0634***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Industry dummies	yes	yes	yes	yes	yes	no	no	no	no	no
Sectoral dummy	yes	yes	yes	yes	yes	no	no	no	no	no
N					27049					
$N < 50$					16478					
$N \geq 50$					10571					
Local Sample ($N < 250$)										
firm size > 50	-0.0768***	-0.0629***	-0.1012***	-0.0267	-0.1323***	-0.0636***	0.0368***	-0.0788***	0.0681***	-0.1084**
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.00)	(0.01)	(0.01)	(0.01)	(0.04)
Industry dummies	yes	yes	yes	yes	yes	no	no	no	no	no
Sectoral dummy	yes	yes	yes	yes	yes	no	no	no	no	no
N					23332					
$N < 50$					16478					
$N \geq 50$					6854					

Notes:

[1] Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

[2] Columns (1), (2), (3), (4), and (5) refer to regressions of gender discrimination that contains the first, second, third, fourth and fifth polynomial degree of size respectively.

First, I test the discontinuity of gender discrimination by firm size by employing a parametric approach using five different polynomial specification models. Resulting point estimates remain stable.[40] Second, I test the results on Logib wage gap estimates by interacting the dummy of interest with other firm characteristics: industry (using 2 digits of disaggregation of the General Classification of Economic Activities, NOGA) and a dummy for private or public sector (1 if private, 0 otherwise). I obtained similar estimates for the treatment effect, independent of the functional form of firm size ($f(S_j)$). A third robustness check is performed by restricting the analysis to a local sample that includes only small firms and companies with up to 250 workers in RDD regressions, and up to 100 workers in Diff-in-Disc estimations. Considering firms of smaller size results in asymmetric observations at the threshold, and very small sample of firms. Estimating regressions for firms of this size also allows to compare more homogeneous firms that represent approximately 99% of the companies in Switzerland. [41] Fourth, since firm size is key for the identification of the effects of *Logib*, one might wonder about potential changes in firm size particularly at the relevant threshold between the collection period and the time at which data become available. For this reason, I then run similar regressions excluding firms that have more than 47 and less than 53 workers, to account for firms whose size may have changed slightly between January and July (collection period) and October (time of data availability). Fifth, one could worry about potential anticipation effects. The FOGE implemented a preliminary version of *Logib* on five companies between 2001 and 2003. After the implementation of this pilot program, *Logib* was modified and launched in its final version in 2006.[42] This shows that essentially no one could have anticipated the introduction of *Logib*. However, one might wonder about the immediate response of employers to use *Logib* to test their wage policies, and if the effect remains when excluding year 2006 from the analysis. Effects of *Logib* after 2008 remain significant. As a cross-check, I have shifted one period forward the introduction of *Logib*, and use 2008 as a cut-off in the time line to compute the Diff-in-Disc estimates. Here, the period before *Logib* lays between 1996 to 2006, and the period after *Logib*, between 2008 and 2010. As a result, the size of discontinuities in Logib wage gap estimates in the period after the introduction of *Logib* is bigger than when including 2006 as post-reform period. These results suggest that *Logib* has been significant in reducing the unexplained wage gap, and even more so in the years after its implementation, when companies could be informed about the availability of this tool. (Figure A6 in the Appendix).

7.2 Raw wage gap

I investigate the net effect of *Logib* on wage gaps in general by analyzing the impact of the policy on the raw wage gap, defined as the log difference of wages of men and women. Like the analysis of Logib wage gaps, the analysis of Raw wage gaps only includes companies with at least 5 male and 5 female employees. However, in this analysis no information of worker characteristics is used in RDD computations.[43] From

40. As before, as explanatory variables, I include the treatment assignment $[S_j \geq 50]$, firm size, and its interactions with year dummies (Angrist and Pischke 2008).

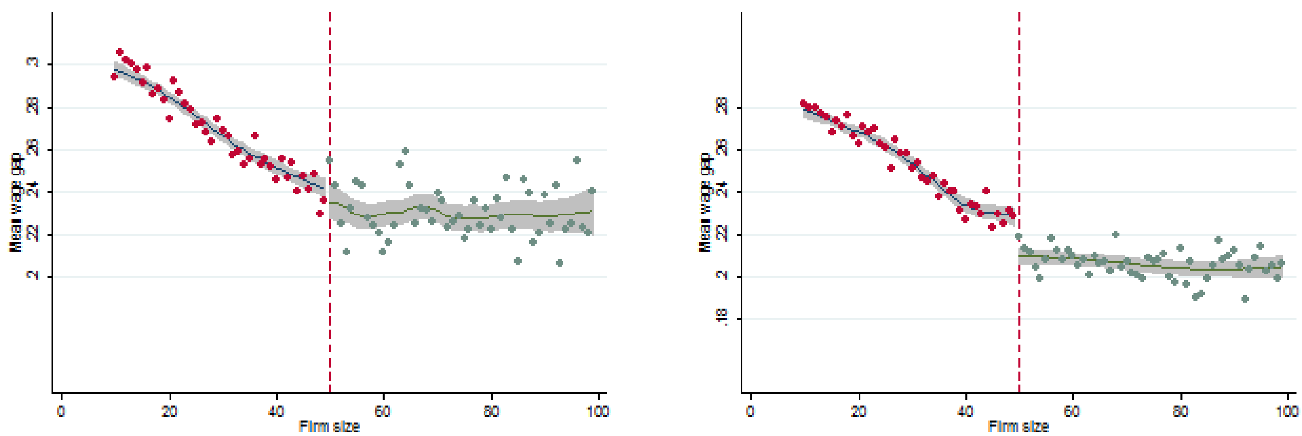
41. As stated by Winter-Ebmer and Zweimüller (1999) large firms might structurally differs from small firms having for example dedicated Human Resources departments that might pay more attention to gender equality.

42. Also, after having interviewed Silvia Strub, who originally designed this policy and followed in detail the introduction of the policy in 2006, confirmed that companies did not foreseen the implementation of *Logib*.

43. I use absolute values of raw wage differences between men and women. I thank Lorenzo Cappellari for suggesting using absolute values of Raw Wage Gaps. Earlier versions of this article used simple raw wage gaps. ATE effects by using

now on, I will refer to Raw wage gap when indeed use absolute raw wage differences ($|ln(w_M/w_F)|$) in the computations.

If the introduction of the *Logib* recommendation reduced unexplained gender wage gaps of firms with at least 50 workers by paying fairer wages to women, it can be expected that Raw wage gap ratios decrease as well. Graphical results show a small discontinuity in Raw wage gap after the introduction of *Logib* but not before (Figure 3). Overall, the magnitude and sign of RDD estimates indicates a slight reduction of gender wage differences. RDD estimations of Raw wage gaps for different periods do not provide statistical inference on the differences of the effect before and after the reform. By pooling the data together from the periods before and after the reform, Diff-in-Disc estimates instead, provide statistical evidence of a slight reduction of Raw gender wage differences. At the threshold, Diff-in-Disc estimates show a decrease in Raw wages gaps of about 1.4 percentage points. These findings show that the impact of *Logib* on Raw wage gap was not as strong as it was for Logib wage gap estimates.



(a) Before introduction of Wage Control (1996, 1998, 2000, 2002, and 2004)

(b) After introduction of Wage Control (2006, 2008, and 2010)

Figure 3: The effect of firm size on Raw wage gap

Notes:

Absolute Raw Wage Gap is measured as $|ln(\frac{W_M}{W_F})|$. Absolute Raw Wage differences are based on standardized monthly salaries. Scatter points show the average of the mean gender coefficient by firm size. Solid lines refer to a third degree polynomial fit, and shadowed area indicates the 95% confidence interval of the fit. RD regressions control for industry and private/public sector. Graphs are made using the same inclusion criteria (number of women bigger than 5, number of men bigger than 5, firm size bigger than 10) as before. Here 5% tails of the distribution of raw wage gaps are included.

These results are in line with the findings of Manning (1996), Chay (1998), Hahn, Todd, and Klaauw (1999) and Carrington, McCue, and Pierce (2000), who confirmed the positive effect of anti-discriminatory laws on reducing raw wage differences. Neumark and Stock (2006), in most of their specifications, found no statistically significant evidence for the effect of equal pay laws on earning effects. Under a particular specification however, they obtained a positive effect of discrimination laws on women's relative earnings of about 0.26% per year. Chay (1998) found that due to the EEOA of 1972, black-white earnings gap narrowed on average 0.11-0.18 log points more than previous years before the introduction of EEOA. My

simple and absolute values of Raw Wage Gaps are however very similar in magnitudes and direction. Computations using simple Raw Wage Gaps are available on request.

results can be explained by the main objective of *Logib*. Indeed, *Logib* was not created to reduce gender wage differences in general, but only to reduce the unexplained part of gender wage differences after taking into account education, experience, tenure, hierarchical position, and job difficulty. The small effects obtained for Raw wage gap estimates can be explained as a consequence of reducing the unexplained wage gap. Employers may only do the minimum required to comply with the *Logib* recommendation, but do not aim to reduce the Raw wage gap. Also, firms which are not public tenders may have no incentives to self-check and verify if their wage policy is discriminatory or not. Other explanations could be attributed to the change in firms behavior or to other ways of discrimination to only comply with *Logib*. I explore this argument in section 7.3.

Summary Results

Table 9 summarizes the findings of the main analysis and reports the range of the estimates in brackets. The first two rows report estimates for the Logib wage gap, and the other two rows report estimates for the Raw wage gap estimates. Table 10 summarizes Diff-in-Disc estimates for multiple specifications of Logib and Raw Wage Gaps. For the most conservative Diff-in-Disc estimator of the Logib wage gap, I find significant evidence of a decrease of 3.5 percentage points. The magnitude of this estimate is exactly to the one accounted by graphs and corresponds to the difference between RDD before and after 2006 at the relevant threshold. RDD estimates after 2006 are all significant. For Raw wage gaps, the evidence is more mixed. On the one hand, the sign of RDD estimates indicates a reduction of Raw wage gaps after 2006, but they are not statistically significant. Judging from the magnitude of Diff-in-Disc effects, one could infer that *Logib* was effective at reducing the unexplained wage gap but its effects were marginal on Raw wage gaps. In addition to the *Logib* controls, Raw wage differences can also be explained by occupational differences which are not accounted for in the analysis.

Table 9: Summary Wage Results: *Logib* and Raw Wage Gaps

	Method	Before 2006 RDD (1)	After 2006 RDD (2)	DIFF-in-DISC (pooled data) (3)
Logib Wage Gap	Graphs & Regressions	0.001 [0.012,0.021]	-0.045 $[-0.051^{**}, -0.073^{**}]$	-0.046 -0.035^{**}
Raw Wage Gap	Graphs & Regressions	0.006 [0.092,0.106]	-0.008 $[-0.0001, -0.0207]$	-0.014 -0.017^{*}

[1] Regressions include year, industry and sectoral dummies.
[2] Prefered specifications for RDD and DIFF-in-DISC regressions.
[3] * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

7.3 Mechanisms behind the effect of *Logib*

Results from the Oaxaca-Blinder Decomposition

While RDD and Diff-in-Disc can identify the effect of *Logib* on wage gaps, they do not provide evidence on how wage gaps were adjusted. To allow for the difference in wage structures between men and women, I estimate separately wage regressions for both genders. Using the coefficients of each estimation, I decompose the difference in mean log wages for men and women using a OB decomposition, and run separate regressions for the treated and untreated firms, before and after the introduction of *Logib*.[44]

Table A12 in the Appendix reports the results from a detailed OB estimation. Results show that returns to education explain the biggest part of differences in endowments for all firms before and after 2006, and the difference in its contribution is similar for both groups of firms. The change in returns to experience is larger for firms with 50 workers or more. However, differences in unexplained returns to hierarchical positions and to job difficulty seem to be remarkably different between firms in the treated and control group. It is also worth pointing out that the difference in returns to no-management position seems to be one of the drivers of change in the unexplained gender wage difference in companies with 50 workers or more, but not in smaller firms. This can be viewed as a sign of employers adjusting gender wage differences by modifying the pay structure of workers start their careers and are relatively new in the firm. This finding is consistent with the possibility that employers may not decide to adjust all workers wage differences, which may be costly and difficult to implement, but instead to pay higher salaries to women who were recently hired.

To complement the OB analysis and understand the drivers of the results, I looked at the effect for different sample of workers. I apply similar RDD and Diff-in-Disc to study next the effect of *Logib* on all group of workers separately. Figures 4 presents graphically the RDD estimations, for the periods before and after 2006, on Logib wage gap and Raw wage gap for groups of workers with no more than one year of tenure. Similar results are found for workers with upper secondary education (Figure A8). They show a discontinuity on Logib wage gap and Raw wage gap at the size of 50 after 2006, which can suggest a positive effect of *Logib* on reducing the wage gap of these group of workers. In Tables A10 and A11 in the Appendix, I report only summarized results of the Diff-in-Disc estimations, which are the main effects of interest. Estimates show a significant decrease of wage gaps of workers with upper secondary education, who perform repetitive tasks, and who have less than one year of tenure.

These RDD and Diff-in-Disc results, are consistent with the findings of the OB decomposition, and they provide evidence of groups of workers responsible for overall effects. It is possible to infer then, that firms do not seem to have raised all female salaries, but instead only adjust wages of newly hired workers in the firm (with less than 1 year of tenure), who do not hold a managerial position, who have on average upper-secondary (VET) education, and who on average perform repetitive tasks.

7.4 Employment Effects

The *Logib* setup can bring incentives to employers of large firms to adjust their behavior in order to reduce unexplained wage gaps among their employees and comply with the regulation, but at the same time, fall into other forms of discrimination.[45] While the government tries to keep female wages on par with male

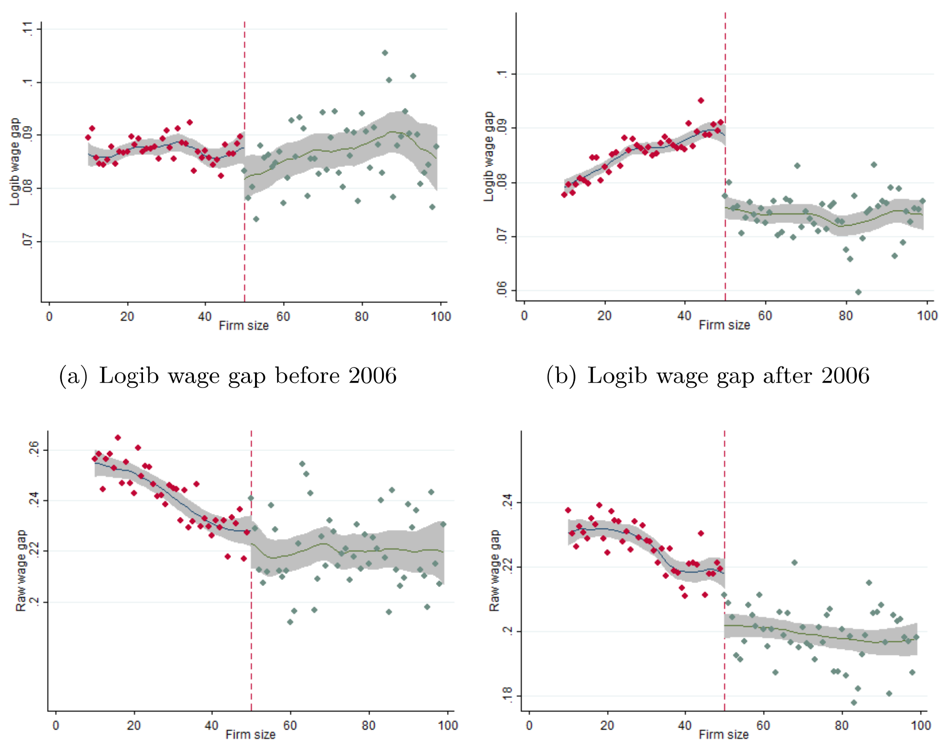
44. This method is also known as the "averaging approach" and based on the idea that if estimates depend on the specification of reference groups, then the true effect can be provided by the average of these estimates.

45. Anti-discrimination wage policies have been criticized for a number of reasons. Some people question the real improvement of labor market outcomes of women and minorities, how socially efficiently are affirmative actions, and if discrimination really exists or if it is the result of labor market competition.

Table 10: Diff-in-Disc Estimates: Logib and Raw Wage Gaps

	Dependent variable			
	***Logib* Gap**		**Raw Wage Gap**[2]	
Total Sample	(1)	(2)	(3)	(4)
DIFF-in-DISC	-0.026**	-0.024**	0.007*	0.008*
	(0.009)	(0.009)	(0.003)	(0.003)
Obs.	46013	46013	47453	47453
Local sample[1]	(1)	(2)	(3)	(4)
DIFF-in-DISC	-0.020*	-0.035**	-0.010	-0.017*
	(0.01)	(0.01)	(0.006)	(0.008)
Obs.	33366	33366	40760	40760

Notes:
[1] Companies with only less than 100 workers.
[2] Computations were performed using simple Raw Wage Gap.
Regressions include additive year, industry and sectoral dummies.
Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
Column (1) and (3) refer to linear specifications, while columns (2) and (4) refer to specifications of third polynomial degree.



(a) Logib wage gap before 2006 (b) Logib wage gap after 2006

(c) Raw wage gap before 2006 (d) Raw wage gap after 2006

Figure 4: The effect of firm size Logib wage gap and Raw wage gap of workers with 1 year of tenure
Notes: Notes Figures (a) and (b) from the upper panel refer to Logib wage gap , while figures (c) and (d) in the bottom panel refer to Raw wage gap . Graphs are made using the local sample. Scatter points show the average of the mean gender coefficient by firm size. Solid lines refer to the linear fit, and shadow area the 95% confidence interval of the fit.

wages, adverse effects on female employment can be created.[46] Regulating female wages could increase the cost of labor and, therefore, the cost of hiring women. As a consequence, employers can decide to reduce female employment in the firm. This argument is mainly theoretical, and empirical evidence is mixed regarding the side effects of wage discrimination laws on employment.[47]

In this section, I study whether the *Logib* recommendation, which only tackles wages, decreased female employment. To test this effect, I analyze first the effect of *Logib* on total employment, and second on employment of particular work categories. To measure the effect of *Logib* on total employment, I use three outcome variables: (i) Total employment, (ii) Female employment relative to total employment in the firm, and (iii) Relative female/male employment ratio. Similarly, to measure the effect of *Logib* on particular work categories, I use two outcome variables: (i) Female employment in a particular work category relative to total employment in the firm, and (ii) Relative female/male employment ratio in particular work category (similar to Zabalza and Tzannatos (1985)). RDD results on total employment are plotted in Figure A7. Graphs show no discontinuities at the size of 50 in none of the employment ratios, neither before nor after 2006; which can be seen as an indication of *Logib* having no effect on this outcome. Diff-in-Disc estimates, presented in Table A13 in the Appendix, provide statistical evidence of this finding. These findings may be explained by the design of the policy and its very weak enforcement. They are in-line with evidence supported by policies designed as AA laws. They usually involve voluntary actions to promote employment of discriminated groups instead of punishing employers who decide not to higher a similarly qualified person because of her sex. Results can be explained, therefore, because employers may not have real incentives to displace these workers.

Similar to wages gaps, *Logib* maybe have side effects on employment of particular groups of workers. One might expect, for example, a decrease of employment of less qualified women or a reduction of relative female employment in groups categorized by hierarchical position or job difficulty. Diff-in-Disc estimations on employment outcomes of particular groups of workers show no significant evidence of decrease of a relative female employment. Tables A14, A3, and A4 in the Appendix show Diff-in-Disc estimates for workers categorized by level of education, hierarchical position and job difficulty, respectively.

Overall, empirical evidence of the *Logib* effects on employment go in-line with Manning (1996), who found that female employment did not fall after the introduction of the Equal Pay Act in the UK. Neumark and Stock (2006), however, found that gender equal pay laws decreased by 2% to 6% the relative employment of women. Others who studied the effect of specific anti-discrimination laws on employment found significant effects. For instance, Chay (1998) found that black employment grew 0.5-1.1 log points per year between 1973 and 1979, while Hahn, Todd, and Klaauw (1999) found that the effect on growth of minority employment was between 3% and 11% depending on the estimation model. The absence of adverse effects of *Logib* on employment may be explained by the very weak enforcement of the policy. Also, the Swiss Federal Act on Gender Equality (enacted in 1995), which prohibits any type of discrimination including employment discrimination, might provide no real incentives to employers to

46. Indeed, concerns about gender discrimination laws were pointed out early in the literature, wage policies can have strong negative effects on the targeted population (Posner 1989).

47. Some empirical evidence assessed the effect of anti-discrimination laws on wages and employment. For example, Neumark and Stock (2006) found robust evidence that state sex discrimination laws, that targeted discrimination in pay only, reduced relative employment of both black and white women.

discriminate against women when hiring new workers.

8 Conclusion and Discussion

This paper has two main contributions: First, it overcomes one of the limitations of the literature, which has been mixed and inconclusive regarding the effects of anti-discrimination policies (Beller 1982; Manning 1996; Chay 1998; Altonji and Blank 1999; Hahn, Todd, and Klaauw 1999). Since *Logib*, the Swiss wage anti-discrimination policy studied here, was enacted in isolation to other policies, I have been able to isolate its effect from any other policy on gender wage gaps and employment. Studying the effect of *Logib* is also appealing because of its interesting design. *Logib* is based on a residual model, which punishes discrimination practices based on the estimated unexplained wage gap. Residual models were usually used to argue against discrimination, but they have never been employed to be the design of a policy themselves.

Second, by exploiting the design of the policy, and using a combination of RDD and Diff-in-Diff (Diff-in-Disc) designs, this paper provides a very clean test to measure the effectiveness of an anti-discrimination policy. Results of my preferred specification show that after the introduction of this Swiss policy, unexplained wage gaps of firms with 50 workers or more decreased, on average, by 3.5 percentage points. While *Logib* seems to do a good job of reducing unexplained wage gaps, its effects are rather marginal on raw gender wage differences (about 1.5 percentage points). Furthermore, Oaxaca-Blinder (OB) results seem to indicate that employers mainly adjust the wage gap of workers with upper-secondary (VET) education, who perform routine work, and who are relatively new in the firm (with less than 1 year of tenure). Although wage policies are sometimes criticized in the literature because they can lead to other forms of discrimination; *Logib* does not seem to affect employment. One of the main advantages of using RDD in this study is the strong internal validity of its results. Results usually are strongly valid for firms around the relevant threshold (firm size = 50). Nevertheless, RDD findings cannot be easily generalized to wider set firms with structurally different characteristics than firms with about 50 workers.

Different threats may compromise the confidence in the causal relationship of *Logib* on the studied outcomes. First, causality might be questioned if there were other confounding variables that influence gender wage gap estimates for firms around the threshold. I showed in this paper, first descriptively and then using the McCrary (2008) test, that the distribution of firms is continuous around the threshold of 50. Also, as confirmed by the FOGE, in Switzerland no other policy was implemented based on company size with this cut-off point. Second, after implementing a Diff-in-Disc design, cross-checking the results, and exploring if companies with at least 50 workers were able to anticipate the introduction of *Logib*, results suggest that companies did not anticipate the launch of this recommendation. Furthermore, Diff-in-Disc results on other covariates provide evidence of no discontinuities on observables. All this evidence suggests that employers may not have manipulated the size of their firm or the structure of their labor force to comply with this regulation. It would be interesting to additionally explore the effect of *Logib* on the p-value of the estimated parameter, and use this statistic as a dependent variable. Since *Logib* does not only address the size of the estimated unexplained wage gap, but also on its statistical significance, this would be interesting because one could test if instead of reducing the wage gap, employers change the statistical significance of the parameter of interest. Future work can implement a Probit model and similar RDD and Diff-in-Disc estimations on the p-value of the Logib wage gap estimate as a final test. Future research will explore the evolution of firm profits by which companies adjust to federal wage policies, and contribute to the improvement of existing models of gender wage discrimination.

Finally, identifying the effect of *Logib* helps to clarify the current debate about this recommendation in Switzerland.[48] Multiple articles have examined the evolution of wage gap and wage discrimination in this country, but studies of the empirical effects of federal policies are missing. This paper constitutes a first attempt to understand the effects of this control instrument on wage (in)equality. Also, the results of this paper might be relevant for the implementation of similar tools at the national and multinational levels. Based on the Swiss *Logib* tool, similar instruments of wage inequality have been developed in various European countries. In 2009, Germany developed its own tool, *logib-d*, and in 2011 an excel version was adapted for companies in Luxembourg under the name *Logib-lux*. A broader project that enables companies to voluntarily analyze their pay structures to detect a potential gender pay gap and its causes is under development on the EU level. This web-based tool, "equal gender pay analysis for a competitive Europe *(equal pacE)*" has been developed recently for the United Kingdom, and it will soon be available for other European countries like Finland, France, Poland, and Portugal.

48. The public debate questions the enforcement mechanisms, the variables taken into account in the analysis, and the methodology that is used. Sources of current public debate are in Postulat Noser Ruedi (2014).

References

Altonji, Joseph G., and Rebecca M. Blank. 1999. "Race and gender in the labor market," edited by O. Ashenfelter and D. Card, 3:3143–3259. Handbook of Labor Economics. Elsevier. https://ideas.repec.org/h/eee/labchp/3-48.html.

Angrist, Joshua D., and Jorn-Steffen Pischke. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion.* 1st ed. Princeton University Press, December. ISBN: 0-691-12035-8.

Beller, Andrea H. 1982. "Occupational Segregation by Sex: Determinants and Changes." *The Journal of Human Resources* 17 (3): 371–392. ISSN: 0022166X. JSTOR: 145586.

Blau, Francine D., and Lawrence M. Kahn. 2007. "The Gender Pay Gap: Have Women Gone as Far as They Can?" *Academy of Management Perspectives* 21 (1): 7–23. ISSN: 15589080. JSTOR: 4166284.

Blinder, Alan S. 1973. "Wage Discrimination: Reduced Form and Structural Estimates." *The Journal of Human Resources* 8 (4): 436–455.

Bloom, David E., and Mark R. Killingsworth. 1982. "Pay Discrimination Research and Litigation: The Use of Regression." *Industrial Relations: A Journal of Economy and Society,* no. 3: 318–339. ISSN: 1468-232X. doi:10.1111/j.1468-232X.1982.tb00240.x. http://dx.doi.org/10.1111/j.1468-232X.1982.tb00240.x.

Brunner, B., and Beth Rowen. 2007. *The Equal Pay Act: A History of Pay Inequity in the U.S.* http://www.infoplease.com/spot/equalpayact1.html.

Carrington, William J., Kristin McCue, and Brooks Pierce. 2000. "Using Establishment Size to Measure the Impact of Title VII and Affirmative Action." *The Journal of Human Resources* 35 (3): 503–523. ISSN: 0022166X. JSTOR: 146390.

Chay, Kenneth Y. 1998. "The Impact of Federal Civil Rights Policy on Black Economic Progress: Evidence from the Equal Employment Opportunity Act of 1972." *Industrial and Labor Relations Review* 51 (4): 608–632. ISSN: 00197939. JSTOR: 2525011.

FAPP. 1995. "Federal Act on Public Procurement (172.056.1)."

Flückiger, Yves, and José Ramirez. 2000. "Analyse comparative des salaires entre les hommes et les femmes sur la base de la LSE 1994 et 1996," University of Geneva. Observatoire Universitaire de l'Emploi.

FOGE and Erika Linder. 2013. "Vers l'égalité des salaires! Faits et tendances," no. 301.919.f.

Fogel, Walter. 1986. "Class Pay Discrimination and Multiple Regression Proofs." *Nebraska Law Review* 65:289–329.

Fortin, Nicole, Thomas Lemieux, and Sergio Firpo. 2011. "Decomposition Methods in Economics," edited by O. Ashenfelter and D. Card, 4:1–102. Handbook of Labor Economics. Elsevier. https://ideas.repec.org/h/eee/labchp/4-01.html.

Garicano, Luis, Claire LeLarge, and John Van Reenen. 2013. “Firm Size Distortions and the Productivity Distribution: Evidence from France.” National Bureau of Economic Research, Working Paper Series, no. 18841 (February). doi:10.3386/w18841.

Graf, Monique. 2002. “Enquête suisse sur la structure des salaires 2000. Plan d’échantillonnage, pondération et méthode d’estimation pour le secteur privé.”

———. 2004. “Enquête suisse sur la structure des salaires 2002. Plan d’échantillonnage, pondération et méthode d’estimation pour le secteur privé.”

Grembi, Veronica, Tommaso Nannicini, and Ugo Troiano. 2016. “Do Fiscal Rules Matter?” *American Economic Journal: Applied Economics.* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1852523.

Hahn, Jinyong, Petra Todd, and Wilbert Van der Klaauw. 1999. “Evaluating the Effect of an Antidiscrimination Law Using a Regression-Discontinuity Design,” no. 7131 (May). http://ideas.repec.org/p/nbr/nberwo/7131.html.

Holzer, Harry, and David Neumark. 2000. “Assessing Affirmative Action.” *Journal of Economic Literature* XXXVIII:483–568.

Jann, Ben. 2008. “OAXACA: Stata module to compute the Blinder-Oaxaca decomposition.” *Stata Journal* 8 (4): 453–479.

Jones, F. L. 1983. “On Decomposing the Wage Gap: A Critical Comment on Blinder’s Method.” *Journal of Human Resources* 18 (1): 126–130. http://EconPapers.repec.org/RePEc:uwp:jhriss:v:18:y:1983:i:1:p:126-130.

Keele, Luke, Dustin Tingley, and Tappei Yamamoto. 1961. “Identifying mechanisms behind policy interventions via causal mediation analysis,” edited by Burt S Barnow, 34:937–963. Methodos for Policy Analysis, August.

Killingsworth, Mark R. 1993. “Analyzing Employment Discrimination: From the Seminar Room to the Courtroom.” *American Economic Review Papers and proceedings* 83:67–72.

Lee, David S., and Thomas Lemieux. 2010. “Regression Discontinuity Designs in Economics.” *Journal of Economic Literature* 48 (2): 281–355.

Leonardi, Marco, and Giovanni Pica. 2013. “Who Pays for it? The Heterogeneous Wage Effects of Employment Protection Legislation.” *Economic Journal* 123 (12): 1236–1278.

LFE. 1995. “Loi fédérale sur l’égalité entre femmes et hommes (151.1).”

Manning, Alan. 1996. “The Equal Pay Act as an Experiment to Test Theories of the Labour Market.” *Economica,* New Series, 63 (250): p191–212. ISSN: 00130427. JSTOR: 2554759.

Mannitz, Sabine, and Rabea Hass. 2009. “The Swiss Concep of Soldering in Practice: The Swiss case.” *Prif HSFK. Peace Research Institute Franckfurt* XLIII (III): 2–12.

McCrary, Justin. 2008. “Manipulation of the running variable in the regression discontinuity design: A density test.” *Journal of Econometrics* 142, no. 2 (February): 698–714.

Neumark, David, and Wendy A. Stock. 2006. “The Labor Market Effects of Sex and Race Discrimination Laws.” *Economic Inquiry* 44, no. 3 (July): 385–419.

Oaxaca, Ronald. 1973. “Male - Female Wage Differentials in Urban Labor Markets.” *International Economic Review* 14 (3): 693–709.

Posner, Richard. 1989. “An Economic Analysis of Sex Discrimination Laws.” *University of Chicago Law Review* (Vol. 56, No. 4.): 1311–1335.

Postulat Noser Ruedi. 2014. “Egalité salariale. Améliorer la pertinence des statistiques. Bulletin Officiel - Le procès-verbaux du Conseil national et du Conseil des Etats.”

Ryder, Guy. 2014. *Maternity and paternity at work: Law and practice across the world.* May.

Sandor, Eszter. 2011. “Part-time work in Europe.” *European Company Survey 2009. European Foundation for the Improvement of Living and Working Conditions.*

Strub, Silvia. 2005. “Approche méthodologique relative au contrôle d’égalité salariale entre femmes et hommes dans les marchés publiques de la Confédération - Instructions pour procéder au contrôle standardisé,” Bureau détudes de politique du travail et de politique sociale (BASS). Bureau fédérale de l’égalité entre femmes et hommes.

Winter-Ebmer, Rudolf, and Josef Zweimüller. 1999. “Firm-Size Wage Differentials in Switzerland: Evidence from Job-Changers.” Papers and Proceedings of the One Hundred Eleventh Annual Meeting of the American Economic Association, *The American Economic Review* Vol. 89, No. 2 (May): 89–93.

Yun, Myeong-Su. 2005. “A Simple Solution to the Identification Problem in Detailed Wage Decompositions.” *Economic Inquiry* 43 (4): 766–772.

Zabalza, A., and Z. Tzannatos. 1985. “The Effect of Britain’s Anti-Discriminatory Legislation on Relative Pay and Employment.” *The Economic Journal* 95 (379): 679–699.