Using econometrics to reduce gender discrimination: Evidence from a Regression Discontinuity Design

Giannina Vaccaro*

22nd February 2016

This is a preliminary draft.
Please do not cite or distribute without permission of the author.

Abstract

Despite numerous policy interventions, women still earn substantially less than men in most countries around the world. The income difference, however, can only partially be explained by observables. This paper estimates the causal effect of a relatively inexpensive anti-discrimination policy introduced in Switzerland in 2006 consisting of a simple regression framework that firms can use to monitor their wage policies. Random checks are implemented by the government and sanctions may lead to exclusion from public procurement. Only firms with more than 50 employees are subject to the random controls and our data span periods before and after the introduction of the policy. Using a combination of regression discontinuity and difference-in-differences methodologies, I found that this policy has significantly reduced the unexplained gender wage gap.

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

*E-mail: Giannina.Vaccaro@unige.ch. Geneva School of Economics and Management (GSEM), University of Geneva. Swiss National Centre of Competence in Research LIVES - Overcoming vulnerability: life course perspectives (NCCR LIVES). I would like to thanks Michele Pellizzari for his continual guidance and invaluable comments. Thanks to the participants at conferences and seminars in Geneva and Zurich (Switzerland), Ismir (Turkey), Mannheim and Nuremberg (Germany), Bergamo and Milano (Italy), and also to 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 and Florian Wendelspiess for thoughtful conversations and ideas. Special thanks to the Federal Office of Gender Equality (FOGE) for his insights and to Roman Graft for his advice and for sharing with me his 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 Centre of Competence in Research LIVEs - Overcoming vulnerability: life course perspectives, granted by the Swiss National Science Foundation. 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.

1 Introduction

Gender related differences in earnings are a persistent stylized fact of virtually any labour market around the world. 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.[1] Switzerland, the country object of this study, is no exception. In 2012, women in Switzerland earned on average 23.2% (CHF 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.[2]

A good share of the wage gap can normally be explained by differences in observables, although selection into labour market participation often exacerbates it (Olivetti and Petrongolo, 2008). However, a non-negligible part of the gender earning difference remains unexplained, which is often viewed as a signal of discrimination. In the US in 1998, about 41.1% of the gender pay difference was explained by neither education, experience, occupation, industry, union status, nor race; emphasizing that even after controlling for whatever measured variables, there is a considerable pay difference between men and women that remains unexplained (Blau and Kahn, 2007). In Switzerland, the unexplained gender wage gap is very important and it has been stable over time. According to the Swiss Federal Statistical Office (FSO) in 2012, the average wage difference between men and women was about CHF 1,770. Only 59.1% (CHF 1,046 aprox.) of it was explained by objective factors, while 40.9% (CHF 723.93 approx.) remained unexplained. Numerous policy interventions have attempted to reduce or eliminate the gender wage gap by preventing discriminatory practices.

Most studies of government policies use state and regional data to determine their impact on labour market outcomes, and it has been very difficult to establish the causal effect of a policy on reducing unexplained wage gaps (Altonji and Blank, 1999, p. 3245- 3246). Also, most related economic literature studied the American labour market. For instance, Chay (1998) examines the effects of the Equal Employment Opportunity Act (EEOA) of 1972 on earnings and employment of individual workers.[3] He finds that after the introduction of this law, the earning gap between black and white men in the South narrowed between 1.5% and 3.4%, depending on the year of birth. Beller (1982) finds that both the Title VII of the Civil Rights Act of 1964 and the federal contract compliance program decreased occupational segregation and gender wage gap.

Hahn et al. (1999) is maybe the closest study to my research. In addition of their theoretical contributions, they used RDD to evaluate the effect of the US Equal Employment Opportunity Comission (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 Act of 1964 and by the 1976 amendment to the act, respectively. In contrast to my study, Hahn et al. (1999) only used cross-sectional variation by analysing 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 on two serious problems: first, only very few observations were subject of the study; and second, their results suffered of measurement error

[1] Percentages refer to measure used by the OECD, the unadjusted gender wage gap, defined as the difference between median wages of men and women relative to the median wages of men.

[2] Based on BFEG (2013) and on the estimation of Flückiger and Ramirez (2000).

[3] Similar Pay Act policies have been enacted recently in the US, i.e. the Lily Led better Fair Pay Act of 2009, but it has been criticized for attracting asserted only well-off female employees who can afford a good lawyer to win their claim.

at the relevant threshold due to the fact that employees may have reported rounded and imprecise information of their firm size.

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 the 1970. However, he could not confirm the expected fall in relative employment due to monopsonic characteristics of the British labour market. More recently, some research evaluated the impact of affirmative action on gender wage discrimination. The evidence is far from conclusive. Holzer and Neumark (2000) review the effects of affirmative action, understood as the policy with the goal to improve the conditions of minorities and women in the labour market, educational institutions or business procurement. They infer that affirmative action may alleviate discrimination and increase efficiency, but they recognize that evidence regarding the effects of affirmative action is mixed, and significant labour market discrimination against minorities and women persists. Indeed, there is still a need of studies on the impact of gender discrimination laws and the effectiveness of their particular enforcement mechanisms.

This paper estimates the causal effect of an anti-discrimination tool introduced in Switzerland in 2006 and finds a significant reduction in the unexplained gender wage gap. The characteristics of this tool and its relatively inexpensive implementation make it an attractive and unique example. It shows an easy and effective way to reduce gender wage discrimination, from which much can be learned about how labour markets work.

The policy consists in a simple regression framework that firms can use to monitor their wage policies. Random checks are performed by the government on a very small number of private firms among public tenders. For instance, from 2006 to 2014, only information of 43 firms has been checked and no firm has even been sanctioned. Sanctions include revocation of licenses and/or cancellation of approved contracts, but they are rarely implemented. The design of this policy is based on the requirements demanded to firms to fulfil the Swiss Wage Structure Survey (SWSS). The 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. Instead 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. Furthermore, the introduction of this policy was the result of a profound academic reflection and its launch on 16 of may 2006 was not anticipated by any firm or employer in Switzerland.[4]

Specifically, the regression model used in this policy controls for observables and includes a gender dummy to which gender discrimination is attributed. After running this regression, targeted firms should look at the magnitude and the statistical significance of their estimated gender coefficient to self-check if their wage policy is discriminatory or not.[5] I test the causal effect of this policy by using two different outcome variables. First, I run exactly this same model specification and use the estimated gender coefficient and standard error of each firm to construct the first dependent variable. Then, since the system is prone companies with discretion to adjust their behaviour, I used raw gender wage gaps, measured as log wage differences between men and women, as a second dependent variable to test the real effect of this policy on gender discrimination. I particularly wonder, for example, how

[4]The first legal claim under which this method was used to prove wage discrimination in Switzerland was LEG (2003) and it was proposed by Flückiger and Ramirez (2000).

[5]Measuring unexplained wage gaps using an imperfect model that relies on a statistical residual has its limitations, particularly when trying to assess the best way to measure wage discrimination. Nonetheless, results of statistical studies of gender pay gap are very instructive (Blau and Kahn, 2007).

firms can change their hiring decisions to comply with this regulation.

The design of this Swiss anti-discrimination policy creates the ideal opportunity to test the impact of such policy on gender wage discrimination. Using data from the SWSS from 1998 to 2010, a biennial Swiss survey of firms and workers characteristics, I implement a combination of a regression discontinuity design (RDD) and differences-in-differences (DIFF-in-DIFF) approach.[6] A number of robustness checks are carried out to test the validity of the results.

I check precisely the potential discontinuity of the estimated vector of gender firm dummies at the threshold of 50, before and after the introduction of the policy. In my preferred specification, I find that, once the policy is in place, women are paid 12.1% less than men in firms with fewer than 50 employees and that this gap declines to 7.6% in larger firms. Before the introduction of the policy in 2006, there was no detectable change in the wage gap at the threshold of 50 employees. Effects are smaller when looking at the raw wage gaps. I found that firms with more than 50 workers reduce the raw gender wage difference by 1% after the implementation of the policy. As before, no change in raw wage gaps is found for smaller firms.

This paper contributes to the literature in a number of ways. First of all, I show that a low-cost, weakly-enforced policy on gender wage discrimination is effective and easy to implement. Multiple other wage policies have been evaluated in the literature, but their results are not conclusive. Current efforts on affirmative action try to redress the disadvantages associated with discrimination, but such policies have been criticized for causing reverse discrimination and being economically inefficient. Second, I help to clarify the effects of this Swiss policy and to enlighten the current debate on the impact of similar recommendations at the European level. Similar policies have been already implemented in Germany (logib-d), Luxembourg (logib-lux) and recently there have been efforts towards developing a similar tool at the European level (equal-pacE.eu). Third, looking at how firms adjust when imposing an additional cost that guarantees gender wage equality can provide useful insights to understand the sources of wage inequality. Identifying the effects of such policies can help us to determine other side effects on employment. Finally, this type of settings can help us to understand why discrimination is present in the labour markets, which is prominent for policy designs.

[6] 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 counterfactual 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 et al., 2015). In our case, computing the DIFF-in-DISC acts as placebo test to cross check the strength of the results.

2 Regulation on Gender Equality

The problem of inequality and wage discrimination affects all countries, all economic sectors and all types of workers. Switzerland is not an exception.

To guarantee gender equality, the Swiss regulation is designed and enforced from the federal level. Maybe the most important milestone in the history of Switzerland with regard to gender equality was the enactment of the Federal Act on Gender Equality, which prohibits any type of discrimination between men and women in the labour 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)).[7]

2.1 The Lohngleichheitsinstrument Bund - Schweiz (Logib-CH)

In April 2006, the FOGE launched the *Lohngleichheitsinstrument Bund* (or *Logib*, due to its German name). *Logib* is a policy based on a wage regression as detailed in equation 1, explicitly written in the legal recommendation that aims to discourage wage discrimination among companies. It allows companies to verify their wage policy by estimating a gender coefficient ($\hat{\beta}_7$) of a wage regression (Mincer, 1974).

$$\ln(wage) = \beta_0 + \beta_1\, educ + \beta_2\, exp + \beta_3\, exp^2 + \beta_4\, ten + \beta_5\, hie + \beta_6\, dif + \beta_7\, fem + \mu \tag{1}$$

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 (1 if woman, 0 if man). Variables such as hierarchical position and level of difficulty are precisely defined in *Logib* and in the data used in this study, so employers are constraint when coding these variables. 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 post takes the values of: (1) the most difficult, (2) independent work, (3) professional knowledge, and (4) simple and repetitive tasks. However, employers have discretion to assign particular values of these variables to the job of their employees. Future work will explore the potential discontinuity in these variables to investigate other ways of gender discrimination such as in terms of employment.

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 use of *Logib* is free of charge and voluntary, but the FOGE only recommends its use to companies with more than 50 workers based on statistical reasons, and only this subset of companies are monitored. The *Logib* excel tool is available in 4 languages (German, French, Italian and English) and it is provided along with a free helpline.

[7]Indeed, several efforts have been made to achieve gender equality in Switzerland since almost four decades ago. From the legal point of view, 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). This principle entitles equal pay for equal work. With the objective to promote gender equality in all areas of life by eliminating all forms of direct and indirect discrimination, the Federal Council established the Federal Office of Gender Equality (FOGE) in 1988. One of the main goals of the FOGE is to guarantee equal opportunities as well as equal pay at work. Later, different actions to promote gender equality have been introduced. The Federal Act on Public Procurement, on December 16, 1994 (172.056.1), establishes that the contracting authority will only award a contract to tenders in Switzerland who guarantee equal treatment of men and women with respect to salary (FAPP, 1995, Art 8c).

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. It is an anonymous tool, since the data and information used for self-assessment remains in the hands of the user.[8] The Strub's method follows the typical economic approach when studying the sources of the gender pay gap. 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.[9] In practice, this method consists of a very detailed wage gap regression (eq. 1) 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 favours men and is positive otherwise. For the purpose of this study, this coefficient will be called *Logib wage gap*. 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 coefficient resulting from the wage regression.

Firms with more than 50 workers are strongly recommended to use *Logib* to check for wage discrimination; however, enforcement mechanisms are very weak. The FOGE only monitors a random selection of private companies with at least 50 employees which have won public tender contracts. 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.[10] 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. The FOGE verifies the provided information and runs the wage regression mentioned above using its own econometric software. If a statistically significant (95%) Logib wage gap estimate above 0.05 is found, the FOGE grants the monitored company a period of 6 to 12 months to correct its wage policy and to ensure wage equality. If after a second evaluation carried out by the FOGE, similar Logib wage gap estimates are obtained, legal sanctions are enforced (Art. 4, BFEG (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 5%.

The process of requesting information from firms, verifying consistency of the provided data, estimating Logib wage gap coefficients and following up the process is costly as it represents about 100 working hours of high skilled work (private communication with FOGE). For instance, assuming a conservative mean hourly wage of CHF 70, then the cost of such an evaluation would be about CHF 7'000. For this reason, only few cases have been analysed since 2006. Table 1 shows the number of

[8] Storage and data processing take place at the local level i.e. in each computer where the *Logib* tool has been installed.

[9] Strub (2005) developed her model based on Flückiger and Ramirez (2000). They use a regression analysis to investigate the wage difference between men and women using the Swiss Labour Force Survey 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.

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

monitored firms between 2006 and 2014. Only 43 firms have been evaluated, and no firm has even 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.

3 Data and Descriptive Statistics

3.1 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.[11] Firm identifiers are not publicly available in this survey, therefore, it is impossible to use a panel structure based on the firm ID. To analyse 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. The total universe of firms is used to construct the drawing and response rate at the company level.[12] 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. Having different sampling rates that change particularly at the relevant cut-off point will not be a serious issue, but they will create an inference problem. They 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.

[11] 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.

[12] This is constructed based on the Business and Enterprise Register (BER).

Survey details are presented in Table 2. This table reports number of workers and number of 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 approximately information for nine million workers which belong to 223,962 firms is reported. Of which 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 workers. For the period before *Logib* (1996 - 2004), information of 4,283,073 workers which belong to 95,988 firms are reported in the data. When considering the period from 2006 to 2010, the SWSS provides information for almost 4,861,435 workers which belong to 127,974 firms.

Table 2: Survey Details by company size

By type of company *Total firms and workers*	**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	
Small (≤ 20)						
Number of workers	483,367	11.29%	562,574	11.57%	1,045,941	11.44%
Number of companies	73,168	1.71%	82,081	64.14%	155,249	69.32%
Medium ($20-49$)						
Number of workers	385,079	8.99%	508,699	10.46%	893,778	9.77%
Number of companies	16,606	0.39%	21,362	16.69%	37,968	16.95%
Large (≥ 50)						
Number of workers	3,414,627	79.72%	3,790,162	77.96%	7,204,789	78.79%
Number of companies	19,352	11.29%	17,989	14.06%	41,188	18.39%

Note: Number of observations of the SWSS without any restriction.

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

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 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. Table A2 details the information used in this study after implementing the different constraints. A local RDD analysis at the threshold is performed. In this case, the percentage of missing observations for companies at the threshold [49, 51] is lower than the one of the total sample of firms; however, missing information on workers characteristics does not represent a problem for the internal validity of the results.

Since sampling rates in the SWSS change by firm size, to examine the potential manipulation of firm size when implementing the RDD, I use the Swiss Business Census (BC) for the years between 1991 and 2008. The BC is one of the oldest data sources in Switzerland that collects information about the universe of all firms in Switzerland. 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 sectors of all the economy. The BC has been 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's information of 1998, 2001, 2005, 2008. In 2011 the BC has been replaced by the STATENT (from its French name, Statistique Structurelle des Entreprises), which is collected exclusively on-line. This data has not been included in the analysis because it has been collected using a different criteria than the SWSS.

3.2 Descriptive Statistics: evolution of wages and covariates

In the US, the wage gap against women in the labour market has narrowed since 1970s and particularly in the early 2000s, but it still significant.[13] In Switzerland, the gross wage difference between men and women working in the private sector has been stable since 1998. In 2010 in Switzerland, women earned on average 23% less than men. Employment rate is one of the crucial differences between male and female employment. Switzerland like other developed countries has a high proportion of part-time workers (36.5%) (BFEG, 2013).[14] Since 2010, the majority of women employed in the labour market in Switzerland work part-time (61%), while most men work full-time (85%).[15] After the introduction of *Logib*, we observe differences in average gender wage differences between workers from small and medium companies. On average, gender wage differences within companies with more than 50 workers seem to have clearly decreased in 2006 and 2008 (Table 3). However, after controlling for many objective factors, an important unexplained wage difference after implementing *Logib* is observed. Unexplained wage gap has reduced slowly, from 41.1% in 1998 to 37.6% in 2010, but it remains still remarkable.[16]

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.

[13] In 1963 women earned about 59% of men wages, and in 2012 they earned about 80.9% (Brunner and Rowen, 2007).

[14] 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 (ECS, 2011).

[15] 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.

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

Table 3: Wage evolution

Hourly wage	1998	2000	2006	2008	2010
	Years				
	Total				
Men	37.95	41.36	45.01	47.09	48.64
Women	27.75	29.43	34.47	36.32	37.33
% diff	**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
% diff	**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
% diff	**30.0%**	**28.9%**	**26.7%**	**27.2%**	**25.3%**

Source: SWSS. Hourly wage refers to the average hourly gross wage measured in CHF. *% diff* refers to the percentage wage difference between men (w_m) and women (w_f).

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	30.40%	28.33%
Medium Firm ($20 - 49$)		
Men	41.67	45.96
Women	30.78	34.10
% wage difference	26.12%	25.81%
Large Firm (≥ 50)		
Men	40.56	46.62
Women	30.22	36.12
% wage difference	25.50%	22.52%

Note: Source: SWSS. Wages are measured in CHF as hourly gross wage without allowances, Sunday, or night work. Percentage wage differences refer to the percentage difference between men (w_m) and women (w_f), measured as $(w_m - w_f)/w_m$.

4 Methodology: LOGIB and DIFF-in-DISC

The methodology employed in this paper involves two steps. First, I evaluate the wage policy of each firm in each year accordingly to *Logib*. After estimating these wage regressions, Logib wage gap estimates for each firm and each year are obtained. Then, I test the effect of the federal recommendation at the aggregate level by studying the discontinuity of these Logib wage gap estimates at the threshold (firm size = 50) and implementing a Difference-in-Discontinuity (DIFF-in-DISC) design to further test the effect of *Logib* on gender wage gap before and after its introduction (Grembi et al., 2015).

Hahn et al. (1999) found several advantages of using RDD when studying the effect of Title VII of the Civil Rights Act of 1964, a federal anti-discrimination law on firm employment of minority workers that covers only firms with 15 or more employees. They use RDD to test the potential discontinuity of wage differences between firms with 15 employees and smaller ones.

4.1 Gender discrimination through wage regression analysis

As first step of the analysis, I estimate a wage regression for each company (j) and each year from 1996 to 2010 using the same specifications as in *Logib* and described by Strub (2005). For explanatory purposes, I summarize equation 1 in equation 2.

$$ln(w_{ij}) = \alpha_j + \beta_{\mathrm{j}}\mathrm{fem}_{\mathrm{ij}} + \theta_j X_{ij} + \mu_{ij} \qquad \forall j \tag{2}$$

Equation 2 is a simple OLS regression that uses an extended Mincerian specification. 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 person is female, and 0 if male, X_{ij} refers to the specific vector of control variables detailed in *Logib* (education, experience, tenure, hierarchical position, and level of difficulty of the post), and μ_{ij} to the error term. 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.

As mentioned in section 2, it is at the discretion of employers to assign a particular category of hierarchical position and level of difficulty to the post of each of their employees. Analysing the evolution of these variables can bring some useful insights about the effects of an anti-discriminatory wage policy such as *Logib* on other forms of discrimination like employment. This analysis will be carried out in future research.

Then, I extract the Logib wage gap estimates $(\widehat{\beta_j})$ and its respective standard errors $se(\widehat{\beta_j})$ for each firm (j) and each year. Using this information, I create a new data set by merging with other important firm characteristics, such as size, industry, and a binary variable indicating whether the company belongs to the private or public sector. Using this resulting data, I then study the potential discontinuity of Logib wage gap estimates at the firm size of 50 using an Regression Discontinuity Design (RDD). Results are presented in section 6.

4.2 Regression Discontinuity Design (RDD)

Baseline

The objective here is to identify the causal effect of *Logib* on gender wage differences. The design is based on a comparison of gender wage gap estimates of firms with less than 50 workers and those with 50 workers or more, respectively. At the threshold of 50, differences in gender wage estimates can be expected due to treatment differences among both groups.

The treatment 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 least 50 workers. Of those companies, only a random selection will be monitored. However, it is not possible to disentangle the effect of *Logib* recommendation from the effect of monitoring. For this reason, the causal effect of the treatment will be identified by estimating the Average Treatment Effect (ATE).

Moreover, the FOGE checks randomly only private companies in Switzerland who have won public tender contracts, not all companies above the threshold are subject to federal supervision. Nevertheless, one can reasonably argue that all companies categorized in the treatment group are indeed treated because the presence of non-compliers cannot be accounted for in the analysis, and the use of the *Logib* tool was recommended to all firms above the threshold. Therefore, a *sharp* RDD is used here.[17]

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

$$D_j = \begin{cases} 1 & \text{if} \quad S_j \geq 50 \\ 0 & \text{if} \quad S_j < 50 \end{cases} \tag{3}$$

Due to the selection threshold, one would generally expect that small firms ($<$ 50 workers) have higher Logib wage gap estimated coefficients $\widehat{\beta_j}$ than larger firms after the introduction of *Logib*. The argument that leads to this hypothesis is as follows. Firms with 50 workers or more will face higher marginal costs than smaller firms to comply with this federal recommendation. To minimize costs, firms affected by this recommendation can adjust the composition of their inputs by changing the structure of their labour force (hiring more qualified women, less qualified men, etc.) or adjusting wages. In this article, I will not focus on how firms adjust or accommodate the composition of labour force, but only on the causal effect of this federal recommendation. Then, in any scenario, one would expect that the minimization of marginal costs translates into lower wage gap coefficients.

The hypothesis of having lower Logib wage gap estimates for firms with 50 workers or more can be tested by fitting a regression to the relationship between Logib wage gap (or Raw wage gap) estimated

[17] Indeed, one may hypothesise 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. Since information of public tenders is only provided by contracts and not by firms (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%, 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.

coefficients and firm size. The empirically estimated RD equation can be written as:

$$\widehat{\beta_j} = f(S_j) + \rho\,[\mathrm{S_j} \geq 50] + \Gamma\, Z_j + \eta_j \tag{4}$$

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 (industry and sector dummies), and η_j the error term. Regression 4 is weighted by the inverse of the standard error of the estimated β_j to account for its statistical significance.

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.[18]

5 Identification

Indeed, our main concern is the potential manipulation firms could perform to change their size to avoid being affected by the federal recommendation. To assert that the treatment assignment based on firm size is as-good-as random, we are interested to show that the distribution of firms in Switzerland is continuous or that the distribution of firms with less than 50 workers and the one of firms with 50 workers or more do not show a systematically different behaviour before and after the introduction of *Logib*. Also, if no precise manipulation or sorting of firm size is confirmed and the change in the evolution of gender wage difference is identified by *Logib*, we should expect no jump in other baseline covariates.

5.1 Firm size and Covariates

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.[19] Among them, very small companies (up to 9 workers) and medium firms (20-49) account for 87% and 10.6% of the total companies in Switzerland. Respectively, 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.

In contrast to legislation in France or other countries, neither the Swiss Federal Labour Act (RS82) nor other Swiss regulation establish different legal obligations between firms with less than 50 workers and firms with 50 workers or more.[20] Obligations for companies in Switzerland vary according to their legal structure (sole, simple, general and limited partnership vs. limited liability companies or

[18] 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 $\hat{\rho}$.

[19] This considers all the companies of the Swiss Business Census of 2008.

[20] For example, Garicano et al. (2013) use French employment protection legislation restrict firms with 50 or more employees to identify equilibrium and welfare effects. According to the 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. Before 2012, this regulation demanded to have 50 full-time equivalent workers in average per year, but this does not matter for the hiring decisions to the firms.

corporation limited by shares). Furthermore, taxes which are levied on federal, cantonal and municipal level do not depend on company size, but vary across cantons and at the municipality level. This provides initial evidence for the absence of incentives of firms to control their size at the cut-off point of 50.

The SWSS considers different drawing and response rates for each company group ($1-19, 20-49, 50+$) resulting in reduced representation of large companies (Drops in the firm size histogram at firm size of 20 and 50 show this, Figure A5 in the Appendix).[21] Since these structural characteristics of the SWSS made impossible to analyse the distribution of firms particularly at the threshold, 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.[22] After analysing graphically and statistically the continuity of the running variable using the McCrary (2008) test, results fail to reject the null hypothesis of continuity of firm size density function at 50, and therefore allow us to conclude that firms are continuously distributed at the threshold. Then, if any discontinuity is found at the size of 50, we can 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 and Table 5.[23]

Table 5: Estimations from McCrary Test

	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).

Variables such as employment rate, proportion of female workers, or proportion of skilled workers in the firm cannot be studied as baseline covariates to explore the sensitivity of the results because they can be determined endogenously. They could be potential outcomes that explain how firms adjust the composition of its labour force when facing a recommendation like *Logib*. Instead, I analyse the distribution of firms at the threshold of 50 within the main industries (Figure A8 in Appendix).[24]

6 Results

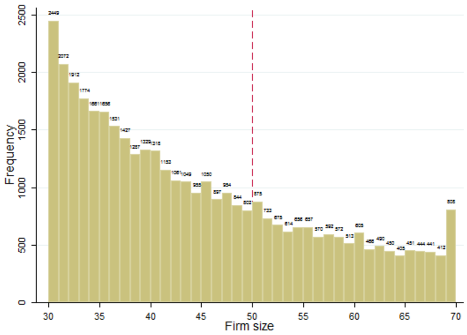
The peculiarity of the RDD setting employed here consists on the use of an estimated parameter as a dependent variable. Strictly speaking, the dependent variable of the RDD setting is the result of

[21] Although we mention early that firms, regardless their size group, provide most information about wages, response rate design matters when reporting information about worker and firm characteristics.

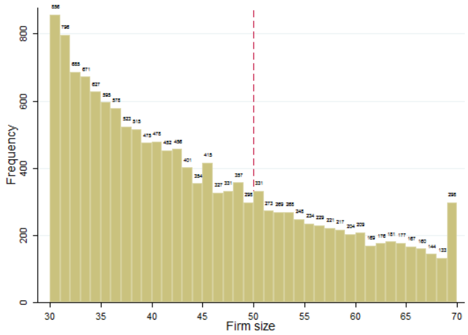
[22] As mentioned before, firms are sampled differently below and above 50 workers in the SWSS.

[23] 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*.

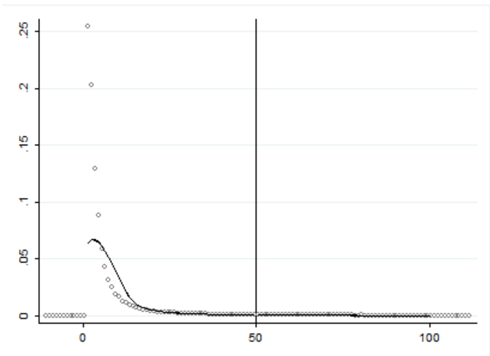
[24] I thank Abihishek Chakravarty for suggesting this cross-check to me.



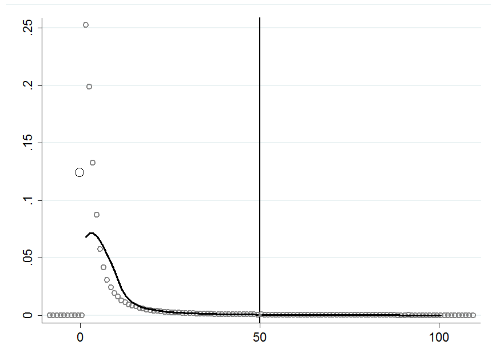
(a) Firm distribution (Before *Logib*)



(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)

merging in one data set all estimated coefficients of β_7 from equation 1.

In the following paragraphs, I comment the main trends of the *Logib* wage estimates and the RDD results.

6.1 Wage gap using *Logib* estimations

In this section, I analyse the evolution of the unexplained gender wage gap in each firm captured by the estimated coefficients of the variable fem from equation 1 in section 4. Results show highly concentrated Logib gender wage gap estimates between $]-0.5, 0.5[$ with mean around 0. Throughout the period of analysis, the distribution of those estimates is more disperse in companies with 50 workers or more than in small companies (Figure A1). Even though the distribution of the estimates before and after the introduction of *Logib* is not very different, this finding suggests different behaviour of gender estimates between small firms and companies with at least 50 workers.

In many of the cases, gender estimates are negative regardless of firm size and year, indicating wage discrimination against women. However, also positive Logib wage gap estimates were found, meaning discrimination against men. Table 6 reports the number of firms by the sign of their Logib wage gap estimate. For the purpose of the RD analysis, Logib wage gap estimates are used in absolute values.

The upper panel of Table 6 shows gender wage estimates for the total sample. The lower panel of this table shows gender wage estimates for firms that exceed the 5% tolerance level for all the years of the analysis. Disaggregated results for the periods before and after the introduction of *Logib* are presented in Table A3, Appendix. Regardless of the tolerance level and to maximize number of observations, RDD is performed on all firms (upper panel) weighted by their standard errors.

Table 6: Number of firms and *Logib Wage Gap* coefficients (β_1) for all the years (1996-2010)

	Against Women ($\beta_1 < 0$)		**Against men** ($\beta_1 > 0$)	
Total Firms[1]	**Cases**	**%**	**Cases**	**%**
Firms < 50 workers	21,114	58.35	8,480	76.28
Firms > 50 workers	15,074	41.65	2,637	23.72
Total number of firms	36,188	100.00	11,117	100.00
Considering Tolerance Level[2]	**Cases**	**%**	**Cases**	**%**
Firms < 50 workers	3,679	42.49	-	
Firms > 50 workers	4,980	57.51	-	
Total number of firms	8,659	100.00	-	

[1] Consider all coefficients across years for which firm size and the Logib wage gap estimated parameter are not missing.
[2] Refers to the subset of firms for which the Logib wage gap estimates is significantly larger than 5%, and which therefore exceed the tolerance level.

6.2 RDD estimates

The setting of *Logib* can bring incentives to employers of large firms to adjust their behaviour in order to reduce their unexplained wage gap and comply with the regulation, but at the same time, fall into other forms of discrimination. An attempt to investigate the net effect of *Logib* on wage discrimination in general, is by looking at the potential discontinuities on raw wage gaps.

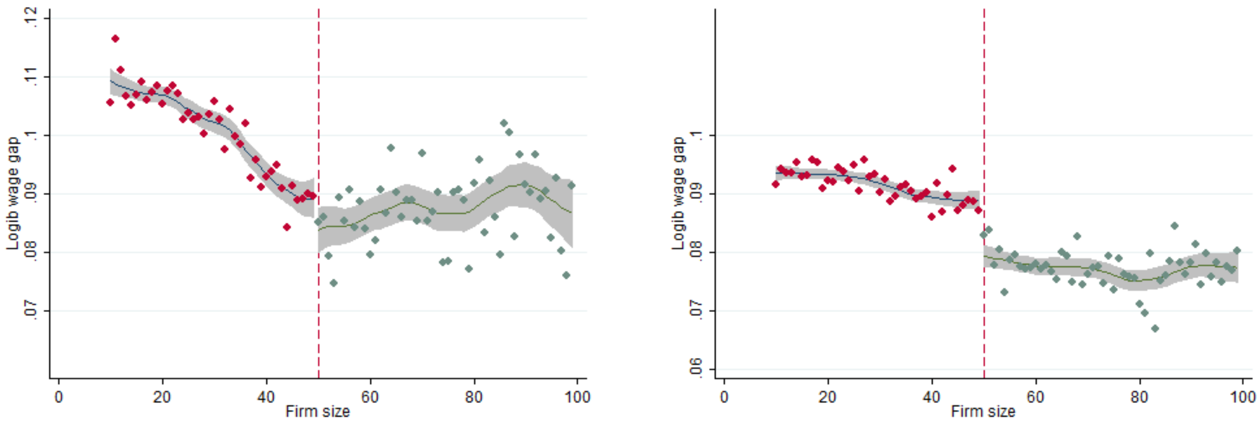
Therefore, the impact of the introduction of the *Logib* recommendation on wage discrimination is tested here using RDD employing two different dependent variables: Logib wage gap and Raw wage gap estimates. Graphical evidence and regression results show a clear discontinuity at the threshold. Firms with 50 workers or more display smaller gender wage gap estimates (Logib and Raw wage gap) than smaller firms, as would be expected if the policy was effective. Results are stable under all different specifications. Also, we can observe a decreasing trend of wage gaps, before the introduction of *Logib*, which is consistent with the evolution of raw wages presented in Table 3.

To evaluate if the *Logib* recommendation had an impact on wage discrimination in Switzerland, I test the results on the Total and Local sample; and then apply DIFF-in-DISC, as described in section 4, by splitting the sample in two: before and after the introduction of *Logib*.

RDD on Logib wage gap estimates

Here, baseline regressions employ Logib wage gap coefficients as dependent variable. Firm size, its interactions with year dummies, and a dummy variable indicating if the firm has at least 50 workers are included as explanatory variables. ATE and DIFF-in-DISC estimates for Logib Wage Gap are reported in Table 7. Negative and significant dummy estimates are found after the introduction of *Logib*. Sign and magnitude of point estimates remain stable using different robustness checks described in section 4.2. After 2006, ATE estimate of Logib Wage Gap at the threshold is statistically significant and equal to -4.5, suggesting a significant decrease in the unexplained wage gap for firms tackled by *Logib*.

Graphical evidence using Logib wage gap estimates as dependent variable before and after the introduction of *Logib* is presented in Figure 2. This figure exhibits smaller mean Logib wage gap estimates for firms with 50 workers or more. However, only after the introduction of the *Logib* tool, significant discontinuity in Logib wage gap estimates at company size of 50 employees is observed. This indicates that the discontinuity in the Logib wage gap estimates after the reform, is actually due to the introduction of *Logib* and not caused by other characteristics of firms with around 50 workers. Similar results are obtained using the the Local Sample (Figure A3 in the Appendix).



(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.

Table 7: ATE Results using Logib Wage Gap estimates

	Pre *Logib* (1996-2004)		**Post *Logib* (2006-2010)**	
	All firms	**At** $[48, 52]$	**All firms**	**At** $[48, 52]$
$E[\hat{\beta}_0 \mid D = 0]$	0.100	0.084	0.116	0.121
$E[\hat{\beta}_1 \mid D = 1]$	0.124	0.085	0.084	0.076
ATE	0.023	0.001	-0.031	-0.045

	All firms	**At** $[48, 52]$
DIFF-in-DISC	-0.054	-0.046

[1] Computations use Logib wage gap estimates as dependent variable.
[2] The Average Treatment Effect (ATE) estimates defined as $ATE = E[Y_1|D_1] - E[Y_0|D_0]$. Strictly, ATE estimates would be achieved if the probability of receiving the treatment at the threshold was available.
[3] Estimations are computed using fifth polynomial degree specifications without controlling for industry nor private/public sector and based on the total sample of firms.

These results are tested statistically using RD regressions. Table 8 and 9 report parametric estimates of the effect of *Logib* on gender wage discrimination for companies with 50 workers or more. Specifically, table 8 reports the effect for the period before the introduction of *Logib* (from 1996 to 2004), and table 9 for the period after its introduction (from 2006 to 2010).

These tables are divided in four panels. The first two report results of the estimations for the total and local sample respectively, without any other restrictions than the ones described in section 3. The third and fourth panel report the results after implementing a robust check. In the latter panels, samples exclude firms with more than 47 and less than 53 workers. Columns (1), (2), (3), (4), and (5) refer to regressions specifications that contain, respectively, a first, second, third, fourth and fifth polynomial degree of firm size. For the period before the introduction of *Logib*, Table 8 reports positive and non significant point estimates in all specifications. While for the period after the introduction of *Logib*, Table 9 shows negative and significant coefficients in all specifications. Since *Logib* only applies to companies with 50 workers or more after 2006, these results confirm the significant impact of *Logib* to reduce gender wage discrimination.

RDD on Raw wage gap ratios

Since employers may assign values of hierarchical position and level of difficulty of the post to their workers at their discretion, we should wonder about the real effects of *Logib* on raw wage gap. I implement a RDD, similarly to the one for Logib wage gap estimates, but using now $ln(w_M/w_F)$ as dependent variable: I test the effect of firm size, particularly beyond the treatment threshold of 50 employees. As before, to verify the results, I implement a DIFF-in-DISC design, that analyses the evolution of raw wage gaps across firm size for the period before and after the launch of *Logib*. Here computations are based on a more flexible selection criteria than before. Now a higher number of

Table 8: Effect of *Logib* recommendation on Logib wage gap estimates (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									

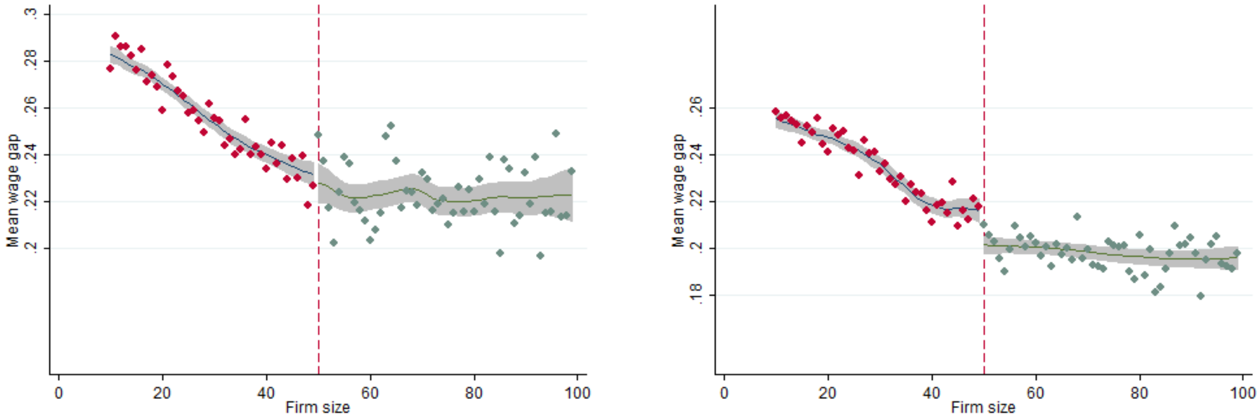
[1] Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

[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 9: Effect of *Logib* recommendation on Logib wage gap estimates (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									

[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.



(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:

Raw wage gap is measured as $ln(\frac{W_M}{W_F})$. Log wage differences are based on standardised monthly salaries. Graphs are made using the same inclusion criteria (number of women>5, number of men>5, firm size>10) as before. Here 5% tails of the distribution of raw wage gaps are included. 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.

observations is obtained since no information on workers characteristics is used for RDD computations. However, the restriction of having 5 men and 5 women in the firm is kept. A robust test was carried out using the same sample as for the computation of Logib wage gaps.

Graphical results for the pooled data show a small discontinuity in Raw wage differences after the introduction of *Logib*. A discontinuity of Raw wage gap estimates at the relevant threshold is confirmed after the reform but not before (Figure 3). ATE estimates of Raw wage gaps are negative only when using the local sample, but its magnitude is smaller compared to ATE estimates of Logib wage gap. When looking at the signs of ATE coefficients of Raw Wage Gap after the introduction of *Logib*, it takes the value of -1.9%, which mirror a decrease in the wage difference between men and women (Table 10).

Table 10: ATE Results using Raw wage gap estimates

	Pre *Logib* (1996-2004)		Post *Logib* (2006-2010)	
	All firms	**At** $[48, 52]$	**All firms**	**At** $[48, 52]$
$E[Y_0 \mid D = 0]$	0.124	0.218	0.035	0.221
$E[Y_1 \mid D = 1]$	0.178	0.218	0.048	0.202
ATE	0.054	0.000	0.013	-0.019

	All firms	**At** $[48, 52]$
DIFF-in-DISC	-0.041	-0.019

[1] Computations use Raw Wage Gap estimates as dependent variable. Y refers to the Raw Wage Gap.
[2] The Average Treatment Effect (ATE) estimates defined as $ATE = E[Y_1|D_1] - E[Y_0|D_0]$. Strictly, ATE estimates would be achieved if the probability of receiving the treatment at the threshold was available.
[3] Estimations are computed using fifth polynomial degree specifications including controls for industry nor private/public sector and based on the total sample of firms.

Moreover, this small effect seems to be non-significant under some parametric and non-parametric specifications. Signs of regression coefficients in parametric regressions confirm the graphical results, however significance was not achieved (Table A5 in the Appendix.). Statistically, these results do not allow us to reject the null hypothesis that the effect of *Logib* on reducing Raw wage gap is different from 0. However, when performing the analysis per year, statistical significance is achieved. Non-parametric specifications for 2010 confirm statistically the null effect of *Logib* on Raw wage gaps. Also, under non-parametric specifications, ATE estimates of Raw wage gap are not statistically significant different from 0 (Table A4 in Appendix).

In a nutshell, using the local sample and parametric estimations, we observe that after the introduction of *Logib* , the magnitude of *Logib* wage gap estimated coefficients decreased by 4.5% at the threshold (Table 7). When looking at the effect of *Logib* on Raw wage gaps (Table 10), the effect at the threshold was of smaller magnitude (a reduction of 1.9%), than Logib wage gap estimates. Using non-parametric specifications by year, negative and significant ATE results were found in the case of Logib wage gap (about 1% for 2010). But adding up similar effects per year, results seem to be consistent with the ones obtained under parametric specifications.

Robustness checks

To test the discontinuity of gender discrimination on firm size, following Angrist and Pischke (2008), I employ a parametric approach using five different polynomial specification models. As before, the dependent variables refer to *Logib* and Raw Wage Gap. As explanatory variables, I include the treatment assignment $[S_j \geq 50]$, firm size, and its interactions with year dummies.

When using *Logib* Wage Gap, the stability of the results is examined using different tests. First, I interact 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). A second robustness check is performed by restricting the analysis to a local sample that includes only small firms and companies up to 250 workers. I expect to obtain similar estimates for the treatment effect, independent of the specification of $f(S_j)$. By including only these firms, we will be able to compare more homogeneous firms that represent approximately 99% of the companies in Switzerland.[25] Third, I run similar RD regressions excluding firms that have more than 47 and less than 53 workers, to account for firms whose size may have changed slightly between the collection period (between January and July) and the time of data collection (October).

Between 2001 and 2003, the FOGE implemented a preliminary version of *Logib* on five companies. After the implementation of this pilot programme, *Logib* was modified and lunched in its final version in 2006. Therefore, the influence of any anticipation effects is ruled out. However, one might wonder about the immediate response of employers to use *Logib* to test their wage policies. 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. After performing a similar RDD as described before, similar discontinuities in Logib wage gap estimates in the period after the introduction of *Logib* are observed (Figure A4 in the Appendix).

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 neighbourhood 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, I follow Hahn et al. (2001) and Porter (2003) and estimate local linear non-parametric regressions based on weighted local linear or polynomial regressions at both sides of the cut-off. Triangular kernel functions are used to weight observations at both sides of the threshold. Because local polynomial estimations do not allow the inclusion of year dummies and pooled estimations without year effects may be biased, separated estimations per year are carried out using the total sample and a local sample for firms with at least 45 and no more than 55 workers. Robust standard errors are obtained after bootstrapping using 999 repetitions.[26] Non-parametric results for years 2006, 2008 and 2010 evidence negative and significant ATE estimates (Table A4 in the Appendix).

To verify if the estimated effect is caused by the introduction of the *Logib* recommendation, I implement a "placebo test" using a Difference-in-Discontinuity (DIFF-in-DISC) design by studying the relationship between wage gap and firm size for the period before and after the launch of *Logib*. DIFF-in-DISC design applies the DIFF-in-DIFF approach of the standard literature of Program Evaluation to RDD (Grembi et al., 2015).

Finally, an important concern arises from the fact that the SWSS has different sampling rates particularly at the threshold of firm size equals to 50. They increase the confidence interval of the estimated coefficients generating an inference problem. Larger standard errors will make less likely to find significant results. However, since results here are statistically significant, this represents a minor problem in this study.

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

[26] Although non-parametric techniques may reduce bias when observations closely approach the discontinuity, they reduce precision (Imbens and Lemieux, 2008).

7 Discussion: The effect of *Logib* on wage discrimination

Logib is a simple policy, easy to implement and very weakly-enforced that, as it has been shown in this article, has a significant effect on reducing the unexplained gender wage gap. Here, I discuss the key methodological steps, the magnitude of the results and different factors that may affect the validity of the findings.

Logib recommendation addresses all private firms with 50 workers or more. However, only those firms that have at least 50 workers and won public tender contracts are monitored. Average Treatment Effects (ATE), informative and useful for policy recommendations, are computed in order to identify the effect of *Logib* on this group. As demonstrated by Hahn et al. (1999), RDD Wald estimates are unbiased. In this study the precision of Logib wage gap coefficients and ATE wage gap estimates is affected by the SWSS response rates since they depend on firm size (section 3): as response rates decrease with firm size, ATE standard errors increase. However, this does not represent a problem here because the results are statistically significant and the interest of the research is to focus on the percentage of the treated firms conditional on a given firm size.

If the introduction of *Logib* recommendation indeed reduced unexplained gender wage gaps of firms with at least 50 workers by paying fair wages to women, it can be expected that Raw wage gap ratios decreased as well. This trend was observed in figure 3. These results are in line with the findings of Manning (1996), Chay (1998), Hahn et al. (1999) and Carrington et al. (2000), who confirmed the positive effect of anti-discriminatory laws. The impact of *Logib* on raw wage gaps is not as strong as it was for Logib wage gap estimates though. These results can be explained due to 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 level of difficulty of the post. If employers only do the minimum required to comply with the *Logib* recommendation but not to guarantee gender wage equality, we can also expect no effects on Raw Wage gap ratios. Also, one might hypothesize that firms can alter the composition of their labour force.[27] Furthermore, the small effects obtained for Raw wage gap estimates can be explained because firms who do not work in public procurement could have lower incentives to self-check and verify if their wage policy is discriminatory or not. Other explanation could be attributed to the change in firm behaviour to only comply with this *Logib* , but do discriminate in other ways. These results can be interpreted as the positive impact of *Logib* in reducing unexplained gender wage gap across firm size, and a very small non-statistically significant effect on gender wage gap in general. Judging from the magnitude of ITT effects, we can conclude that the introduction of wage policy recommendations and the implementations of tools like *Logib* are a good start to reduce gender wage discrimination, but they are small to achieve non discriminatory wage policies.

The small magnitude (and almost non significant effects) of Raw Wage gap effects agree with findings in the literature. Neumark and Stock (2006), in most of their specifications, found no statistical significance evidence of equal pay laws on earning effects. Under a particular specification they obtained however, 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. Most

[27] This hypothesis will be explored in further research.

of the related literature found statistically significant positive effects of anti-discriminatory laws on employment of minorities. However, the effects of anti-discriminatory policies on employment are not very clear. For instance, Chay (1998) found that black employment grew 0.5-1.1 log points per year between 1973 and 1979. Hahn et al. (1999) found that the effect on growth of minority employment was between 3% and 11% depending on the estimation model. For Neumark and Stock (2006) equal pay laws decreased by 2% to 6% the relative employment of women. While Manning (1996) found that female relative employment did not fall after the introduction of the Equal Pay Act in the UK. In future research, I will look at the effects of *Logib* on other relevant labour outcomes such as employment, level of education, hierarchical position, etc.

One of the main advantages of our study by employing RDD consists in the strong internal validity of its results. However, different threats may compromise the confidence in the causal relationship of *Logib* on wage discrimination. First, causality might be questioned if there were other confounding variables that influence gender wage gap estimates for firms around the threshold. As confirmed by the FOGE and the Federal Audit Oversight Authority (FAOA), in Switzerland there is no other policy targeted to companies with at least 50 workers. Second, by implementing a DIFF-in-DISC design, cross checking the results and exploring if companies with at least 50 workers had been able to anticipate the introduction of *Logib*, I exclude history or anticipation as a threat of internal validity. Results from the DIFF-in-DISC design show no discontinuity on *Logib* wage gap estimates before the introduction of *Logib*, which indicates that companies did not anticipate the launch of this recommendation. Third, looking at figures 2 and 3 we can observe that after the introduction of *Logib*, firm wage gaps decreased for all firms, regardless of their size. This could be interpreted as a maturation process in which all companies progressively reduce gender discrimination following a normal developmental process over time. Since all companies follow the same trend, there is not threat to internal validity.

Other concerns might be related to the design of *Logib* itself. A residual regression such *Logib* has its limitations. As pointed out by Blau and Kahn (2007), any study based on a statistical residual will open the questions 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, number and age of children can also affect the residual gap. The evidence suggests negative impact of children in female wages, especially when they are young. Also, children matter for labour participation and experience accumulation. Related to that, the availability of caring-children facilities as well as extended-family support influence positively female labour 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, they do not represent a problem for my study since this paper does not aim to identify all explanatory variables that can explain gender wage differences, instead its purpose is to identify the causal effect of *Logib* on reducing the unexplained gender wage gap.

8 Conclusions

Since 1990s, convergence of female wages relatively to men has slowed considerably in the US after a rapid confluence since 1970. In Europe, gender pay gap also remains important (17.4% in 2009) and has not decreased much over the last years. The enactment of Title VII of the Civil Rights Act of 1964 and the Equal Pay Act of 1973 in the US, or the Equal Pay Principle of 1957 in Europe were some examples of regulations implemented to reduce gender wage discrimination.
The literature finds that government regulations reduce wage inequality (Beller, 1982; Manning, 1996; Chay, 1998; Altonji and Blank, 1999; Hahn et al., 1999); however, the magnitude of the impact of those laws and their effectiveness depend in particular on their enforcement mechanisms (Holzer and Neumark, 2000). Residual models explain gender wage differences by observable and unexplained factors. The latter is usually associated to gender wage discrimination. Despite the limitations of these models, they are very informative and commonly used in the literature (Blau and Kahn, 2007).

This paper overcomes the limitations of policies analyses in the literature by studying the impact of an inexpensive and very weakly-enforced Swiss recommendation: *Logib* uses a residual model and targets companies with at least 50 workers. The main contribution of this paper is to show the effectiveness of *Logib* in reducing unexplained wage gaps within firms using a combination of RDD and DIFF-in-DIFF techniques. The ATE estimates are robust under different model specifications and multiple tests. Two RDD outcome variables are used to test the effect of this policy: first, estimated coefficients for unexplained wage gap of the firm exactly as described in the *Logib* regulation (called in this paper Logib wage gap); and second, Raw wage gaps measured as the differences of log wages between men and women. Results of my preferred specification show that after the introduction of this Swiss policy, unexplained wage gap of firms with 50 workers or more decreased by 4.5% points in average. The magnitude of *Logib*'s effect on raw wage gaps is small, but consistent with the literature.

There are reasons to wonder about the causal identification of *Logib* to reduce gender wage discrimination. In particular the potential manipulation employers could perform to adjust the size of their firms in order to not be affected by this recommendation, and the presence of other confounding factors that could overestimate the effect. 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. Second, no other Swiss policy establishes a cut-off point at the firm size of 50. Third, the distribution of firms at the size of 50 in different sectors is always continuous.

In addition, identifying the effect of *Logib* helps to clarify the current debate of this recommendation in Switzerland.[28] Multiple articles have examined the evolution of wage gap and wage discrimination in this country, but studies on 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 other national and at European level. Based on the Swiss *Logib* tool, similar instruments of wage inequality have been developed in some European countries. In 2008, Germany developed its own tool *logib-d*, in 2009 and 2011 an excel version was adapted to companies in Luxembourg under the name *Logib-lux*. A broader project that enables companies to voluntarily analyse their pay structures to detect a potential gender pay gap and its causes at the European level is under development. This web tool-based "equal

[28]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 SP (2014).

gender pay analysis for a competitive Europe *(equal pacE)*" has been recently developed for the United Kingdom and it will be soon available for other European countries like Finland, France, Poland and Portugal.

Finally, exploiting settings like *Logib* can help us to understand better the sources of wage inequality and why wage discrimination is still present. Further research can explore the mechanisms by which companies adjust to federal wage policies. For instance, studying the change of gender and skill composition of companies' work force after the introduction of *Logib* will be an avenue for future research. Together with an examination of the evolution of firm profits, this research can contribute to the improvement of existing models of gender wage discrimination.

References

Altonji, Joseph G. and Rebecca M. Blank, "Race and gender in the labor market," in O. Ashenfelter and D. Card, eds., *O. Ashenfelter and D. Card, eds.*, Vol. 3 of *Handbook of Labor Economics*, Elsevier, 1999, chapter 48, pp. 3143–3259.

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

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

BFEG, "Vers l'égalité des salaires! Faits et tendances," Technical Report, Bureau fédéral de l'égalité entre femmes et hommes (BFEG). Office fédéral de la statistique (OFS). June 2013.

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

Brunner, B. and Beth Rowen, "The Equal Pay Act: A History of Pay Inequity in the U.S.," 2007.

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

Chay, Kenneth Y., "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*, 1998, *51* (4), pp. 608–632.

ECS, "Part-time work in Europe," Technical Report, European Company Survey 2009. European Foundation for the Improvement of Living and Working Conditions 2011.

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

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

Garicano, Luis, Claire LeLarge, and John Van Reenen, "Firm Size Distortions and the Productivity Distribution: Evidence from France," Working Paper 18841, National Bureau of Economic Research February 2013.

Graf, Monique, "Enqute suisse sur la structure des salaires 2000. Plan d'échantillonnage, pondération et méthode d'estimation pour le secteur privé," Technical Report, Office fédéral de la statistique 2002.

_ , "Enqute suisse sur la structure des salaires 2002. Plan d'échantillonnage, pondération et méthode d'estimation pour le secteur privé," Technical Report, Office fédéral de la statistique 2004.

Grembi, Veronica, Tommaso Nannicini, and Ugo Troiano, "Do Fiscal Rules Matter?," *American Economic Journal: Applied Economics*, (forthcoming) 2015.

Hahn, Jinyong, Petra Todd, and Wilbert Van der Klaauw, "Evaluating the Effect of an Antidiscrimination Law Using a Regression-Discontinuity Design," NBER Working Papers 7131, National Bureau of Economic Research, Inc May 1999.

_ , _ , and Wilbert Van der Klaauw, "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design," *Econometrica*, 2001, *69* (1), 201–209.

Holzer, Harry and Da Neumark, "Assessing Affirmative Action," *Journal of Economic Literature*, 2000, *XXXVIII*, 483–568.

Imbens, Guido W. and Thomas Lemieux, "Regression discontinuity designs: A guide to practice," *Journal of Econometrics*, 2008, *142* (2), 615 – 635.

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

LEG, "VD 22.12.2003 Discrimination Salariale," 2003.

LFE, "Loi fdrale sur l'galit entre femmes et hommes (151.1)," 1995.

Manning, Alan, "The Equal Pay Act as an Experiment to Test Theories of the Labour Market," *Economica*, 1996, *63* (250), pp. 191–212.

McCrary, Justin, "Manipulation of the running variable in the regression discontinuity design: A density test," *Journal of Econometrics*, February 2008, *142* (2), 698–714.

Mincer, Jacob A., *Schooling, Experience, and Earnings*, Columbia University Press, 1974.

Neumark, David and Wendy A. Stock, "The Labor Market Effects of Sex and Race Discrimination Laws," *Economic Inquiry*, July 2006, *44* (3), 385–419.

Olivetti, Claudia and Barbara Petrongolo, "Unequal Pay or Unequal Employment? A Cross-Country Analysis of Gender Gaps," *Journal of Labor Economics*, 2008, *26* (4), pp. 621–654.

Porter, Jack, "Estimation in the Regression Discontinuity Model," Technical Report, Harvard University. Department of Economics 2003.

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

SP, "Postulat Noser Ruedi. Egalit salariale. Amliorer la pertinence des statistiques. Bulletin Officiel - Le procs-verbaux du Conseil national et du Conseil des Etats," 2014.

Strub, Silvia, "Approache méthodologique relative au contrôle entre femme et hommes dans les marchés publiques de la Confédétion," Technical Report, Bureau détudes de politique du travail et de politique social BASS. Bureau fédérale de l'égalitéentre femmes et hommes 2005.

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

9 Appendix

Variables Description

- **Standardised monthly wage**: Continuous variable that represents sandardised monthly salary (4 1/3 weeks, 40 hours). It does not take into account wage earned in extra hours usually from night and Sunday work. Log standardised wages are used in the computations. As suggested in the documentation of the SWSS, standard weights for workers paid per hour and for workers paid per month are taken into account.
- **Gender**: Categorical variable: (1) men, (2) women.
- **Education**: Categorical variable that can take 9 values, representing the highest level of education achieved: (1) Tertiary academic, (2) Tertiary vocational; (3) High secondary vocational, (4) Teaching certificate, (5) High secondary academic, (6) Secondary vocational (apprenticeship), (7) Secondary vocational (Not apprenticeship), (8) Compulsory education and (9) Others.
- **Experience**: Is built by decreasing to age, years of education and additionally 6 years.
- **Tenure**: Refers to the number of years spent in the current firm. A value of 0 is attributed if the worker has only months of experience in the current firm.
- **Hierarchical position**: Refers to the hierarchical position of the post. (1) Senior; (2) Middle management, (3) Junior workers, (4) Low management, and (5) no management functions.
- **Level of difficulty of the post**: Categorical variable that can take 4 values: (1) The most difficult, (2) Independent work, (3) Professional knowledge, (4) Simple and repetitive work.
- **Firm size**: Refer to the number of employees the firm reported to have.
- **Industrial sector**: Categorical variable based in NOGA (2 codes).
- **Private**: dummy variable that is equal to 1 if the individual is working in the public sector and 0 otherwise.

9.1 A1: Sample details

Table A1: Percentage of reported wages by firm size (%)

	25% p-tile	**50% p-tile**	**75% p-tile**
size < 20	100	100	100
$20 \leq$ **size** < 50	96.43	100	100
size > 50	94.24	99.79	100

[1] Percentages refer to the proportion of reported wage information over firm size for each firm size group of pooled data.

Table A2: Descriptive statistics of data used

Year	1996	1998	2000	2002	2004	2006	2008	2010
Survey details								
Average reported information of employees[1]	73%	83%	73.3%	85%	87%	85%	83%	85%
Total number of workers	560,665	509,854	634,054	1,215,689	1,362,811	1,481,282	1,608,087	1,772,066
Total firms	8,058	7,035	16,669	39,080	40,146	42,018	40,908	45,048
Firms < 50	3,509	5,085	14,232	33,000	33,948	35,476	17,528	18,317
Firms ≥ 50	4,549	1,950	2,437	6,080	6,198	6,542	5,366	8,610
Total sample[2]								
Total number of workers	384,204	307,137	356,670	794,473	103,7047	1,209,188	1,426,745	1,630,494
Total firms (after run *Logib*)	4,306	3,668	8,246	21,142	23,580	25,056	25,116	28,686
Firms ≥ 50	349	544	1,408	4,201	5,495	5,835	6,047	5,723
Firms < 50	3,957	3,124	6,838	16,941	18,085	19,221	19,069	22,963
***RD* analysis**								
Total sample[3]								
Total firms (after run *Logib*)	1,039	1,121	2,181	6,270	8,205	8,902	9,406	10,189
Firms ≥ 50	349	544	1,408	4,201	5,495	5,835	6,047	5,723
Firms < 50	690	577	773	2,069	2,710	3,067	3,359	4,466
Local sample[4]								
Total firms (after run *Logib*)	600	747	1,699	5,471	7,354	7,976	8,376	8,401
Firms ≥ 50	349	544	1,408	4,201	5,495	5,835	6,047	5,723
Firms < 50	251	203	291	1,270	1,859	2,141	2,329	2676
At the threshold[5]	-5.73%	7.45%	4.55%	4.69%	11.56%	12.17%	11.64%	18.40%

[1] Refers to average percentage of workers for whom employers report information. According to the Federal Statistical Office (FSO), the net response rate of 1996 and 2000 corresponds to the average response rate by firm size group. See (Graf, 2002) and (Graf, 2004) respectively.
[2] Refers to the number of observations with no missing information for *Logib* wage computations, with at least 10 workers, including among them at least 5 men and 5 women.
[3] Refers to the number of observations with complete information for the *Logib* wage regressions, but also with complete information about firm size, industry, region, and private or public sector.
[4] The local sample includes only small and medium-sized enterprises. i.e. Firms with 250 employees or less. Only enterprises that have more than 10 workers, including among them at least 5 men and 5 women.
[5] Reported additional observations at the threshold compared to average observations in the total sample.

A2: ***Logib*** estimate distribution

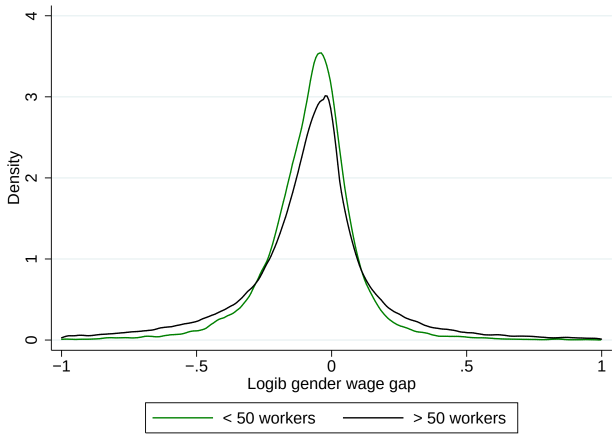
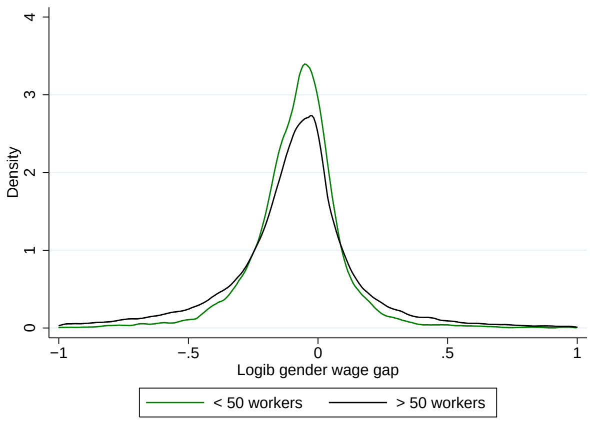
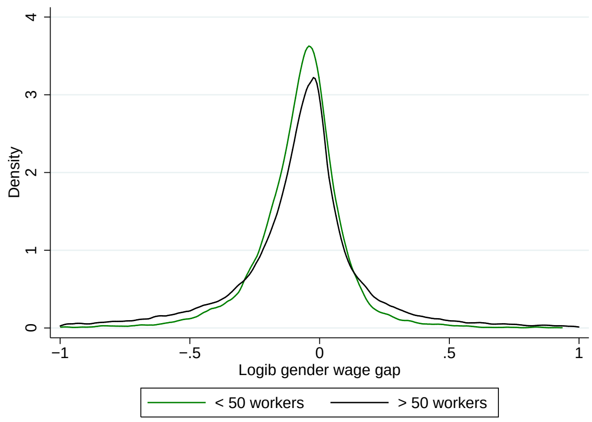


Figure A1: Distribution of unexplained *Logib* wage gap estimates by firm size
Notes: Plot based on total sample gender coefficients estimates restricted from -1 to 1. Red line refers to the mean gender estimate.



(a) Before the introduction of *Logib*



(b) After the introduction of *Logib*

Figure A2: Distribution of unexplained *Logib* wage gap estimates by firm size before and after the introduction of *Logib*
Notes: Plot based on total sample gender coefficients estimates restricted from -1 to 1. Red line refers to the mean gender estimate.

A3: Preliminary results

Table A3: Number of firms and Logib Wage Gap estimates ($\hat{\beta}_1$) before and after the introduction of *Logib*

	Against Women ($\beta_1 < 0$)		**Against men** ($\beta_1 > 0$)	
Before *Logib* (1996-2004)	**Cases**	**%**	**Cases**	**%**
Total Firms[1]				
Firms < 50 workers	8,683	59.22	3,313	79.78
Firms > 50 workers	5,979	40.78	840	30.23
Total number of firms	14,662	100	4,153	100
Considering Tolerance Level[2]				
Firms < 50 workers	1,645	40.53	-	
Firms > 50 workers	2,414	59.47	-	
Total number of firms	4,059	100	-	
After *Logib* (2006-2010)	**Cases**	**%**	**Cases**	**%**
Total Firms[1]				
Firms < 50 workers	12,431	57.75	5,167	74.20
Firms > 50 workers	9,095	42.25	1,797	25.37
Total number of firms	21,526	100	6,964	100
Considering Tolerance Level[2]				
Firms < 50 workers	2,034	44.22	-	
Firms > 50 workers	2,566	55.78	-	
Total number of firms	4,600	100	-	

[1] Obtained after pooling all coefficients across years for which size of the firm and the estimated parameter is not missing.

[2] Refers to the subset of firms for which the gender discrimination coefficient is significantly bigger than 5%, and therefore are above the tolerance level.

Only 1 firm with gender equality ($\hat{\beta} = 0$) was found before the introduction of *Logib*, while 7 firms with gender equality were found after the introduction of *Logib*.

Table A4: Non Parametric Estimations

	Dependent Variable: Logib Wage Gap	
	2010	
	Total Sample	**Local Sample** $45 < N < 55$
ATE coefficient	-0.021	-0.010
p-value$^{(}1)$	0.000	0.019
No rep	999	999

	Dependent Variable: Raw Wage Gap $ln(w_m)/ln(w_f)$	
	2010	
	Total Sample	**Local Sample** $45 < N < 55$
ATE coefficient	0.000	0.001
p-value$^{(}1)$	0.000	0.010
No rep	999	999

[1] P-value calculated using two-tail critical values and 1 degree of freedom. They are computed as (n-k), where is the number of observations, and k the number of dependent variables in the regression, and it is constructed from bootstrapped t-statistics.
[2] Estimations are made using nonparametric specifications per year using triangular kernel, and 1 as bandwidh.

Table A5: Effect of *Logib* recommendation on Raw Wage Gaps

	After *Logib* (2006-2010)				
	(1)	(2)	(3)	(4)	(5)
Total Sample					
firm size > 50	-0.0219	-0.0195	-0.0172	-0.0169	-0.0177
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
N			28566		
$N < 50$			17674		
$N \geq 50$			10892		
Local Sample ($N < 250$)					
firm size > 50	-0.0189	-0.0207	-0.0193	-0.0001	0.0107
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
N			24821		
$N < 50$			17674		
$N \geq 50$			7147		

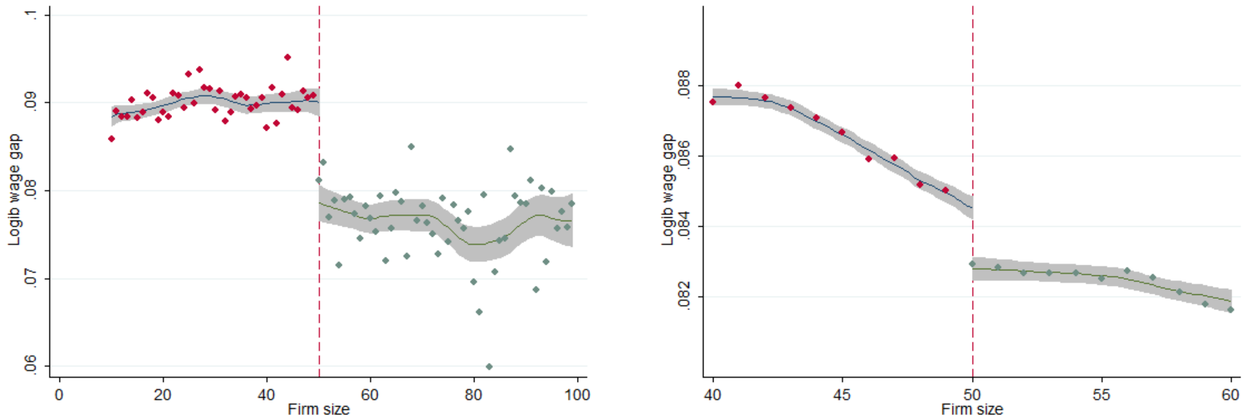
	Before *Logib* (1996-2004)				
	(1)	(2)	(3)	(4)	(5)
Total Sample					
firm size > 50	0.0071	0.0075	0.0082	0.0085	0.0085
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
N			18887		
$N < 50$			12065		
$N \geq 50$			6822		
Local Sample ($N < 250$)					
firm size > 50	0.1064	0.0967	0.0923	0.0931	0.1045
	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)
ITT at [48,52]			0.011		
N			15965		
$N < 50$			12065		
$N \geq 50$			3900		

[1] Raw wage gap is measured as $ln(\frac{W_M}{W_F})$. Log wage differences are based on standardised monthly salaries. Graphs are made using the sames exclusions (number of women>5, number of men>5, firm size>10) than before. 5% tails of the distribution of Raw wage gap are included.

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

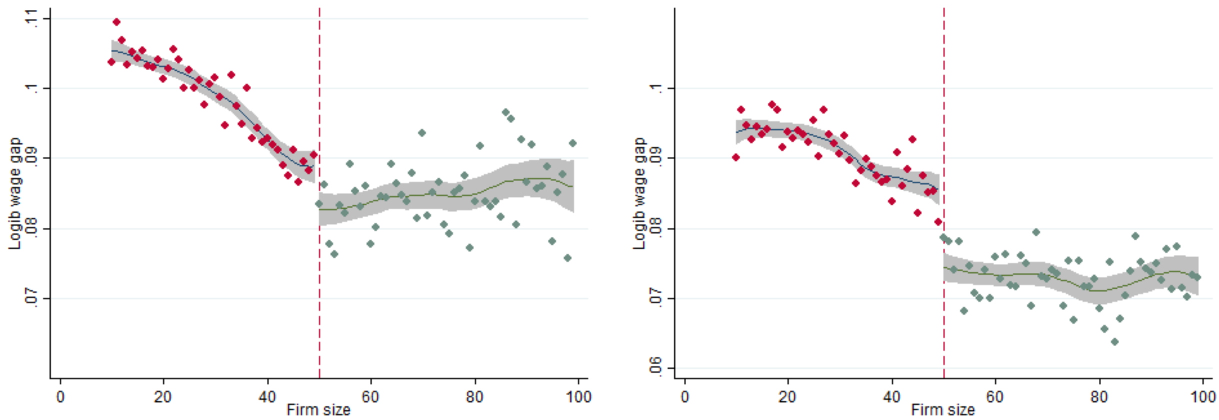
[3] Columns (1), (2), (3), (4), and (5) refer to regressions of first, second, third, fourth and fifth polynomial degree of size respectively.

A4: Stability of RDD



(a) Total Sample (b) Local Sample

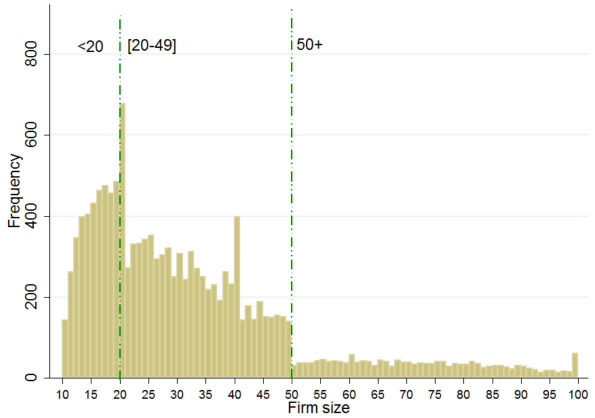
Figure A3: The effect of firm size on Logib wage gap after for the period between 2006-2010
Notes: Graphs are made using the *total sample.* Scatter points show the average of the mean gender coefficient by firm size. Solid lines refer to a third degree polynomial fit, and shadow area the 95% confidence interval of the fit.
Reported graphs are shown in its simplest specification without considering controls variables. However, when considering big firms (with more than 250 workers), the industrial sector become more important, therefore RDD specifications for the *Total sample* include an additive dummy for the industrial sector.



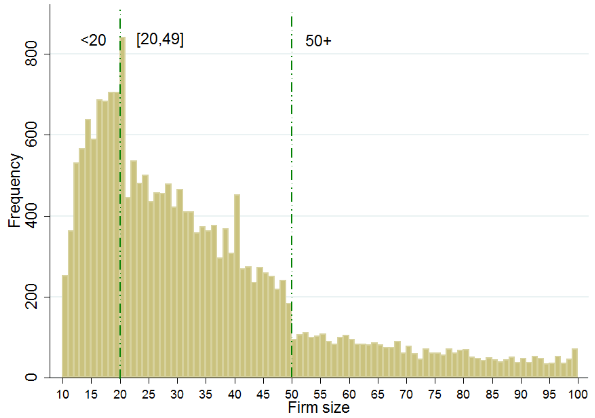
(a) Before 2008 (b) After 2008

Figure A4: The effect of firm size on gender discrimination before and after the implementation of *Logib*
Notes: Logib wage gap coefficient is obtained as the gender variable of the Mincerian (*Logib*) Regression. Graphs are made using the *total sample*. Scatter points show the average of wage gap by firm size. Log wage differences are based on standardised monthly salaries. Solid lines refer to a third degree polynomial fit, and shadow area the 95% confidence interval of the fit.

A4: McCrary test



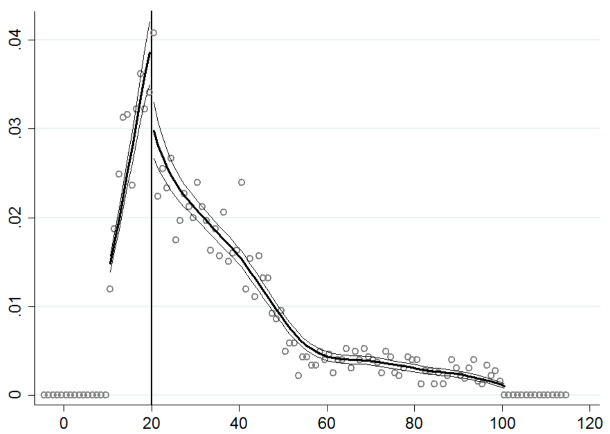
(a) Before the introduction of the Wage Control (1996-2004)



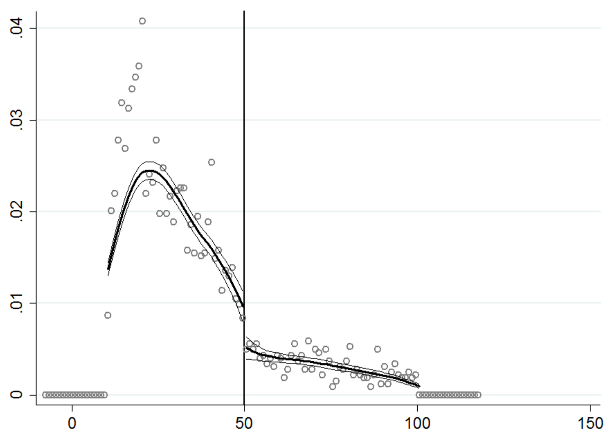
(b) After introduction of the Wage Control (2006-2010)

Figure A5: Firm distribution by company size

Notes: Based on *total sample* and same exclusion assumption used in the regression analysis (Section 3). Doted lines refer to the threshold between firm class categories. Source: SWSS.



(a) Breakpoint at 20



(b) Breakpoint at 50

Figure A6: Graphical representation of McCrary Test with two breakpoints

Notes: Test consider firms up to 100 workers and binsize of 1.
Source: SWSS.

Table A6: Estimations from McCrary Test per year

	Year							
McCrary estimates	**1996**	**1998**	**2000**	**2002**	**2004**	**2006**	**2008**	**2010**
$\hat{\theta}$	-2.50	-0.89	-1.60	-3.85	-0.78	-0.68	-0.92	-0.40
se $\hat{\theta}$	1.08	0.55	0.52	0.65	0.19	0.18	0.17	0.16
No obs.	3 946	3 379	7 295	18 989	21 304	22 584	22 568	25 896

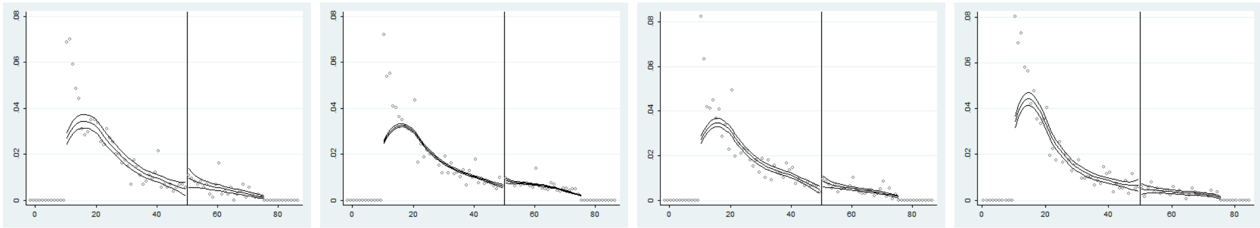
Notes:
(1) Computations are not pondered neither by drawing nor by response rates because unavailability of this information.
(2) A binsize of 1 is imposed for the computations, and test the discontinuity at 50 as breakpoint.
(3) Source: SWSS

Table A7: Estimations from McCrary Test

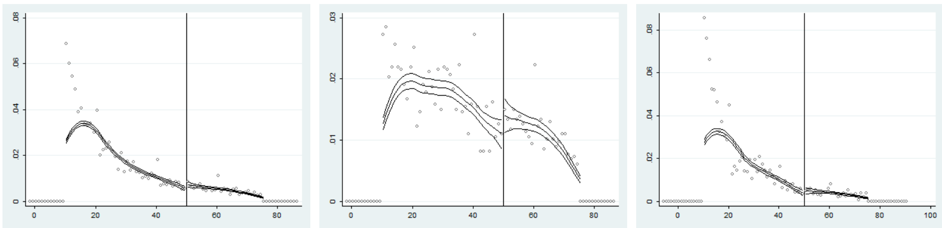
	Breakpoints	
All period (1996-2010)	**20**	**50**
$\hat{\theta}$	-0.22	-0.64
se $\hat{\theta}$	0.07	0.15
No obs.	125961	

	Breakpoints	
Post reform (2006-2010)	**20**	**50**
$\hat{\theta}$	-0.21	-0.44
se $\hat{\theta}$	0.08	0.17
No obs.	71048	

Notes:
(1) Computations are not pondered nor by drawing nor response rates because unavailability of this information.
(2) A bin-size of 1 is imposed for the computations.
(3) Source: SWSS

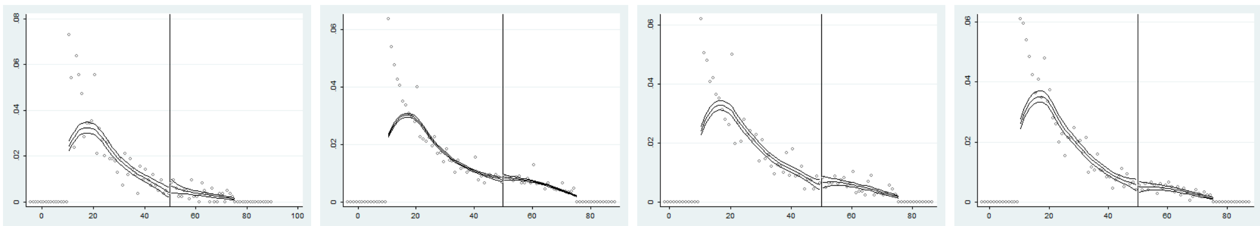


(a) Manufacture and extractive industries (b) Retail trade, transport, hotels and restaurants (c) Information et communication (d) Financial and insurance activities

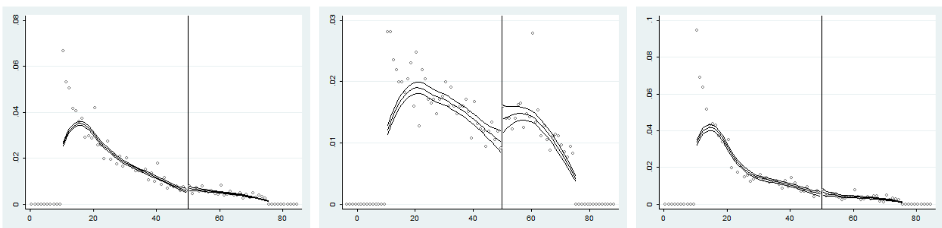


(e) Professional, scientific and technical activities (f) Public administration, defence, education, human health and social work (g) Other activities and services

Figure A7: Firm distribution per sector before the introduction of *Logib* (1996-2004)
Notes: Graphs are outputs after having run the McCrary (2008) per industry. For ilustration purpose, NOGA Industries have been aggregated in 11 categories according the categoriyation used by Eurostat. Only industries where McCrary (2008) test was achieved are shown here. A binsize of 1 is imposed for the computations, and test the discontinuity at 50 as breakpoint. Source: SWSS.



(a) Manufacture and extractive industries (b) Retail trade, transport, hotels and restaurants (c) Information et communication (d) Financial and insurance activities



(e) Professional, scientific and technical activities (f) Public administration, defence, education, human health and social work (g) Other activities and services

Figure A8: Firm distribution per sector after the introduction of *Logib* (2006-2010)
Notes: Graphs are outputs after having run the McCrary (2008) per industry. For ilustration purpose, NOGA Industries have been aggregated in 11 categories according the categoriyation used by Eurostat. Only industries where McCrary (2008) test was achieved are shown here. A binsize of 1 is imposed for the computations, and test the discontinuity at 50 as breakpoint. Source: SWSS.