Fair-ON-Pay

Analysis for equal pay
between women and men

Log-ON: description of methodology

Version 1.2
December, 2024

Owner:
The Fair-ON-Pay Association

Table of contents

1 Introduction

This document describes in detail the methodological principles of the Log-ON analysis model.

The Log-ON instrument is based on the Confederation's standard analysis tool for the analysis of equal pay between women and men (Logib Module 1 for short). Logib Module 1 was developed in Switzerland by the Federal Office for Gender Equality (EBG) in the early 2000s with the support of private specialized institutions (Strub, 2004). Logib Module 1 has been validated in various evaluations (INFRAS, 2013, Felfe, Trageser & Iten, 2015). The theoretical basis of Logib Module 1 is the human capital theory (Becker, 1993). With the help of a statistical procedure (multiple regression analysis), it is examined to what extent objective and non-discriminatory personal and role-related characteristics determine salary, and whether or to what extent gender has an influence on salaries (cf. chapter 2). For statistical reasons, the instrument is particularly suitable for larger companies.

The OLS regression analysis on which Logib Module 1 (and thus also Log-ON) is based was approved by the Swiss Federal Supreme Court to clarify the question of the existence of salary discrimination[1].

The present method description serves the transparency and comprehensibility of Log-ON, but builds directly on the method description of Logib.

2 Log-ON

2.1 Overview

Log-ON was methodically designed to be applied from 50 valid records.

It consists of the following four components:	
A	the dependent variable: standardized gross salary based on a salary specification;
B	several independent variables: Factors justifying salary differentials between women and men (education, seniority, potential work experience, region, competency level and professional position) and the variable gender;
C	a statistical analysis procedure (semi-logarithmic OLS regression analysis);
D	a given tolerance threshold “x” for the gender factor, which must not be statistically significantly exceeded[2] to uphold equal pay.

[1] For Logib module 1 cf. BGE 130 III 145. Regarding scientific and legal conformity of Logib cf. declaration of conformity of the FOGE: https://www.ebg.admin.ch/ebg/de/home/dienstleistungen/logib-triage/logib-modul-1/dokumentation-logib.html

[2] This tolerance threshold of ±X% must not be confused with the significance level α = 5% of the hypothesis test, which is 5% in this approach

Log-ON is used to explain workers' salaries based on personal qualification characteristics (education, seniority and potential work experience), role-related factors (company skill level, occupational position, region) and gender. All other things being equal, this method can be used to determine which part of the salary difference is exclusively due to the gender of employees. In other words, this method makes it possible to determine which salary differences exist between women and men in a company with comparable personal and role-related characteristics. Log-ON is limited to the six variables mentioned above. At the same time, a tolerance threshold for the resulting salary inequality is introduced (cf. chapter 2.2.6), which compensates for the part that could possibly be explained by other objective, company-specific factors, but that are not reflected in the to some extent generic statistical model. The relevant tolerance threshold depends on the be achieved Fair-ON-Pay certification level.

2.2 Method

Log-ON is based on the *multiple linear regression* approach[3] , a statistical model used to study the relationship between a dependent variable (e.g. salary log) and independent variables (e.g. potential work experience, education, occupational status). The dependent variable is also called the explained variable and the independent variables are called explanatory variables.

2.2.1 Multiple linear regression model

The general form of the multiple linear regression model[4] with *n* observations and *p* independent variables is as follows:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \ldots + \beta_p x_{ip} + \epsilon_i$$

where i = 1,2, ..., n.

y_i	dependent variable for the i^{th} scalar value (e.g. logarithm for the salary of person i)
$x_{i1}, \ldots, x_{ip}$	independent variables for the i^{th} single value (e.g. age, potential work experience, education, … of person i)
β_0	the constant or output coordinate
$\beta_1, \ldots, \beta_p$	coefficients to be estimated. The value of a coefficient β_j indicates the marginal influence of an increase of one unit of the variable x_{ij} on the dependent variable (ceteris paribus).
ϵ_i	random error term for the i-th mean-free scalar value with constant variance

[3] The basics of this approach are described in this chapter for a more detailed treatment of regression analysis; see J. M. Wooldridge (2006)

[4] See decision BGE 130 III 145 of the Federal Supreme Court for Logib

The regression coefficients $\beta_0, \ldots, \beta_p$ are estimated by the least squares (OLS) method as $\beta_0, \ldots, \beta_p$. That is, the coefficient estimates are based on minimizing the sum of the residual squares[5]:

$$\sum_{i=1}^{n} \hat{e}_i^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \text{ , where } \hat{y}_i = \hat{B}_0 + \hat{B}_1 x_{i1} + \ldots + \hat{B}_p x_{ip}$$

2.2.2 Data basis

Log-ON is applied at the enterprise level, so the following data is needed for the reference month of the individual employees of the enterprise:

2.2.2.1 Gender

The legal gender of the employee, female or male (male = 1, female = 2).

2.2.2.2 Age

The biological age of the employee as a figure (in years).

2.2.2.3 Education

The highest completed degree of the employee; a code between 1 and 8 that is also used to calculate the years of education for each employee:

EQF Level / Description		Common education type / Examples	Log-ON code	Years of education
8 & 7	Academic level	**Master** or equivalent (also incl. PHD, EMBA)	1	17
6		**Bachelor** or equivalent	2	15
5	Post upper secondary level	**Higher national diploma** or equivalent	3	14
4	Upper secondary level	**Higher national certificate, upper secondary diploma** or equivalent	4	15
			5	13
3	Secondary level	Completed **vocational education** leading to a nationally accepted certificate of proficiency (i.e., apprenticeship)	6	12
2	Primary level	Completed school AND/OR exclusively "in-house", generally unrecognized training	7	11
		Compulsory school (**secondary school**) <u>without</u> completed vocational training	8	7

2.2.2.4 Region

The region in which the employee works; a code between 1 and 8 that allows to differentiate up to 8 regions (region 1 = 1; region 2 = 2, etc.).

[5] The error term indicates the deviation between the dependent variable and the population regression, while the residual stands for the deviation between the dependent variable and the sample regression. Therefore, an error term is not observable, while a residual is observable and quantifiable.

2.2.2.5 *Potential employment experience*

The biological age of the employee minus the "Years of education" minus 6 (years = generally start of schooling).

2.2.2.6 *Years of Service / Seniority at the company*

Years between Entry date and Reference date (cut-off date) of the employee.

2.2.2.7 *Competency level (skill level)*

The competency level is a classification by employee according to the requirements of their role. In theory, the classification can be made, based on the role profile and role specification (tasks) as well as the competency profile (required skills and abilities). However, practically, it is rather decided based on information such as: job evaluations, grading, requirement level, rank, etc.). The level of competency is recorded in Log-ON according to the following codification:

Log-ON Code	ISCO-08[6]	Simplified description / guideline (may be more or less suitable depending on company orientation)	Typical required level of education
1	Skill level IV	Key activities with most complex problem-solving/decision-making tasks; position with extremely demanding and difficult tasks	Academic or post-secondary level
2			
3	Skill level III	Activities with very complex technical or practical tasks; position that requires very high independence and skilled work	Academic level or upper secondary level
4			
5	Skill level II	Activities with tasks that require professional or specialist knowledge; position that requires professional or specialist knowledge	Vocational training or higher
6			
7	Skill level I	Activities with simple manual or routine tasks; Position with mainly simple and/or repetitive activities	Basic vocational training
8			None

The competency level should refer to the role and not the individual employee. A uniform and systematic system should therefore be used as a basis for the development of the coding.

2.2.2.8 *Professional Position*

Captures how much responsibility a role entails. Up to 8 levels of professional position are distinguished:

Log-ON Code	Simplified Description / Guideline
1a	Executive (highest responsibility) / "Top" Management
1b*	
2a	"Upper" / Senior Management
2b*	

[6] The content of the competence levels was originally based on the international occupational classification ISCO-08 with 4 levels. ISCO-08 is currently under revision (as of December 2024). In any case, the application at company level must always take place in a relative context in order to best reflect the professional reality.

3a	"Mid" Management / Non-management with high responsibility
3b*	
4	"Lower" Management / Non-management with increased responsibility
5	Employee without any management role / increased responsibility

* Unlike Logib Module 1, Log-ON offers the possibility to further differentiate the professional position with 3 additional levels (total of 8 instead of 5). This can be particularly interesting for complex or large companies, for which only five levels are too little differentiated or do not sufficiently reflect the different levels or responsibility.

The professional position is primarily derived from the management hierarchy in the organization, as it is usually depicted in the organizational chart. The focus should not only be on "Management", but on the underlying "concept": to map the organization coherently on the levels according to different degrees of (professional) responsibilities / decision-making autonomies. Therefore, there may also be employees without management responsibility who have other specific responsibilities (e.g. technical responsibility). Such employees can be classified accordingly higher than professional position 8 (employees without increased responsibility / without a management role).

The professional position should not refer to the individual employee. A uniform and systematic system should therefore be used as a basis for the development of the coding.

2.2.2.9 Workload

- Employees on monthly salaries: degree of employment, a (decimal) number corresponding to the contractual degree of employment in percentage points.
- Hourly-paid employees: hours paid in the reference month or average over the reference year, a (decimal) number that must correspond to the number of hours paid by the enterprise in the reference month or the average over the reference year (excluding holiday or public holiday compensation).

2.2.2.10 Gross salary[7]

The salary earned by the employee in the reference month, consisting of:

- **Basic salary** (incl. a monthly share of any 13th, 14th or x-th monthly salary if applicable)
- **Difficulty / hardship allowances**
- **Special payments**

Please see the Fair-ON-Pay Norm salary definition for details on any of the above elements.

2.2.2.11 Usual weekly working hours

The usual working time in hours per week at full time equivalent (employment rate of 100%).

2.2.3 Standardization of the salary

For all employees, the sum of the monthly salary components is standardized to correspond to a full-time position with the company's modal standard weekly working hours (i.e. the most frequent weekly working hours in the organization). The standardized salary is calculated:

[7] For a detailed salary specification in Switzerland, see PricewaterhouseCoopers AG (2020).

Employees on monthly salaries:

$$Standardized\ salary \equiv gross\ salary * \frac{100}{Employment\ level} * \frac{modal\ weekly\ working\ time\ customary\ in\ the\ company}{employee's\ usual\ weekly\ working\ hours}$$

Hourly-paid workers:

$$Standardized\ salary \equiv gross\ salary * \frac{52 * \ modal\ weekly\ working\ time\ customary\ in\ the\ company}{12 * \ paid\ hours}$$

2.2.4 Specification of the Log-ON

The Log-ON model is based on the formula[8] below, using the 5 variables (objective explanatory factors) that are also used in the Swiss federal government's standard analysis tool (Logib, Module 1), the possible additional variable / explanatory factor region, and the factor of gender:

$$\ln(Sal_i) = \beta_0 + \beta_{Train}Train_i + \beta_{Exp}Exp_i + \beta_{Exp^2}Exp^2{}_i + \beta_{Tnr}Tnr_i + \beta_{Skl}Skl_i + \beta_{Pos}Pos_i + \beta_{Reg}Reg_i + \beta_{Gender}Gender_i + \epsilon_i$$

Whereby

$\ln(Sal_i)$	Logarithmical gross salary of person i standardized to full-time with the modal weekly working time of the enterprise (see chapter 2.2.3).
$Train_i$	Number of years of education of person i, converted using the education code according to explanations in chapter 2.2.2.
Exp_i	Potential work experience for person i, calculated as $Max\left[Age_{i)} - Train_i - 6{,}0\right]$. This variable is also used in the formula squared, as the influence of work experience on salaries is usually non-linear according to economic theory.
Tnr_i	Number of years of service (tenure) of person i.
Skl_i	Competency / Skill level for person i. This variable is included in the regression model as a categorical variable with up to 8 levels (i.e. 7 dummy variables, see chapter 2.2.2 for a description of the levels).
Pos_i	Professional Position of person i. This variable is included in the regression model as a categorical variable with up to 8 levels (i.e. 7 dummy variables, see chapter 2.2.2 for a description of the levels).
Reg_i	Region in which person i works. This variable is included in the regression model as a categorical variable with up to 8 levels (i.e. 7 dummy variables).
$Gender_i$	Dummy variable for the gender of person i, 1 for women and 0 for men.

The model uses the logarithmical form of the dependent variable, i.e. in this case $\ln(Sal_i)$. It is common practice to logarithmize the salary as the dependent variable, since salaries follow a log-normal distribution from an empirical point of view. The logarithm of the salary therefore

[8] For hourly-paid workers, the gross salary is given without holiday and public holiday pay. In addition, the categorical dummies are each replaced by a single variable to simplify the formula. Categorical variables such as Skl_i or Pos_i are used as multiple dummy variables in the model. For a company that has e.g. 4 competency / skill levels in its data, $\beta_{Skl}Skl_i$ effectively reads as $\beta_{Skl_1}Skl_{1i} + \beta_{Skl_2}Skl_{2i} + \beta_{Skl_3}Skl_{3i}$ in the formula above, where Skl_{1i}, Skl_{2i}, Skl_{3i} are dummy variables (0/1 variables) representing the requirement level of person i.

has a normal distribution[9] . This log transformation also reduces the influence of outliers (extremely high or low salaries), which makes the estimates more reliable and allows a better interpretation of the results. If the logarithm of the dependent variable is used, the regression formula is called semi-logarithmic or log-level. In a semi-logarithmic regression, the interpretation of the estimates β_1 , ... , β_p changes slightly: in this form, β_j can be interpreted as an approximation of the percentage salary increase when the variable x_{ij} is raised by one unit. Around $\beta_{Train} = 0.02$ the model indicates that one more year of education will raise the standardized gross salary by $100 \times (e^{\beta_{Train}} - 1) = 2.02\%$, which corresponds to the approximation $100 \times \beta_{Train} = 2\%$. The approximation error occurs because the approximation becomes less accurate as the logarithm changes.

2.2.5 Estimation of the gender-specific effect

To estimate the gender impact on salaries, the Kennedy (1981) estimator is used, which is shown to be a consistent and nearly non-skewed estimator (assuming normal errors). Its formula is as follows:

$$\hat{\tau}_K \equiv \exp\left(\hat{\beta}_{Gender} - \frac{\hat{\mathbb{V}}[\hat{\beta}_{Gender}]}{2}\right) - 1,$$

where exp(·) corresponds to the exponential function and $\mathbb{V}[\beta_{Gender}]$ to the square of the estimated standard error of the estimated discrimination coefficient.

2.2.6 Significance of the gender-specific effect

When a parameter such as the discrimination coefficient is estimated, this always includes some uncertainty, as the effective value of the parameter β_{Gender} remains unknown. Nevertheless, a hypothesis can be made about the value of β_{Gender} and a statistical inference can be used to test this hypothesis. In a hypothesis, a result is considered statistically significant if it is extremely unlikely to have occurred by chance. Therefore, the estimate of the discrimination coefficient should be interpreted together with its statistical significance.

Basic assumption: the null hypothesis applies

$$H_0: \beta_{Gender} = 0$$

which corresponds to the hypothesis that gender has no effect on salaries once all other explanatory variables are controlled. If this hypothesis were correct, it would imply that gender has no effect on salaries once the other variables are taken into account. If the hypothesis were false, there would have to be a gender-specific effect.

One cannot know with certainty whether H_0 is true or false, but a data-based rule can be established by which the hypothesis is rejected or not. Consider the following test statistic:

$$t_{\beta_{Gender}} = \frac{\hat{\beta}_{Gender} - \beta_{Gender}}{se(\hat{\beta}_{Gender})},$$

[9] The log-normal distribution is used in particular when the dependent variable takes on only positive values and when the frequency distribution on the left-hand side is larger than on the right-hand side, i.e. when there is a right-skewed distribution. This is typically the case with the salary distribution when there are many low salaries on the left side and few very high salaries on the right side of the scale.

where $se(\beta_{Gender})$ is the standard error of the estimated discrimination coefficient. Furthermore, we assume a significance level, given by α, that represents the probability of rejecting the null hypothesis if it is indeed true. The choice of α = 5% is extremely common, and this value is also used with Log-ON for the Fair-ON-Pay certification.

Given this significance level α, a critical t-value t_{crit} can be calculated so that H_0 can be rejected in favor of the alternative hypothesis H_A as soon as $t_{\beta Gender} > t_{crit}$.

Log-ON assesses the statistical significance of the gender effect in two steps:

1. the null hypothesis of the discrimination coefficient of the order **of less than or equal to** the defined tolerance threshold of ±2.5% is tested (significance level α = 5%)
2. the null hypothesis of the discrimination coefficient of the order **of less than or equal to** the defined tolerance threshold of ±5% is tested (significance level α = 5%)

In each of both assessments above, the following hypotheses is formulated:

- H_0 : β_{Gender} ≤ tolerance threshold, the gender effect is less than or equal to the tolerance threshold
- H_A : β_{Gender} > tolerance threshold, the gender-specific effect is greater than the tolerance threshold

In both null hypothesizes above, it is stated that the gender effect does not exceed the defined tolerance threshold. Consequently, the alternative hypothesis implies that the gender effect is above the threshold. If the null hypothesis is rejected, the statistical method used has sufficiently proven that there is a large gender effect for work of equal value.

2.2.6.1 Applicable tolerance threshold

For Fair-ON-Pay there are two gradations of the certificate. Accordingly, there is a defined tolerance threshold per certificate, which is defined in the Fair-ON-Pay norm.

However, the determination of the tolerance threshold is based on the fact that the statistical model "Logib", which is used as a methodical basis for Log-ON, has been applying a threshold of +/-5.0% in legal proceedings in Switzerland for more than 15 years.

3 Open-Source code for “Logib module 1”

Log-ON is based on the Confederation's standard analysis tool for the analysis of equal pay between women and men (Logib Module 1 for short).

The regression model basics and the statistical calculations of Log-ON are identical to the online application of Logib Module 1, which is implemented with the programming language R. A package providing the open source version of the Logib Module 1 code is freely available on CRAN[10] . Instructions for using the R package can be found in its README[11].

[10] https://cran.r-project.org/web/packages/logib

[11] https://cran.r-project.org/web/packages/logib/readme/README.html

4 Bibliography of the underlying Logib method description

- Becker, Gary S. (1993). Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education, 3rd edition, Chicago: University of Chicago Press.
- Chicha, Marie-Thérèse (2016). Strengthening gender justice. Ensuring equal pay. An ILO guide to gender-neutral job evaluation. Berlin: International Labour Organization.
- Federal Office for Gender Equality FOGE (2021). Standard analysis model of the Confederation - Method description for Logib module 1 and 2, version 2021.1
- Felfe, Christina, Trageser, Judith & Iten, Rolf (2015). Study on the statistical analyses of the Confederation concerning equal pay for women and men Final report. Commissioned by the Federal Office for Gender Equality FOGE. St. Gallen and Zurich: Swiss Institute for Empirical Research in Economics University of St. Gallen, Infras.
- Hirschi, Andreas & Ghetta, Anja (2020). Development Report Logib Modul 2 Analyse der Lohngleichheit zwischen Frau und Mann für kleinere Unternehmen insbesondere mit weniger als 50 Mitarbeitern. Commissioned by the Federal Office for Gender Equality FOGE. Bern: Department of Work and Organisational Psychology University of Bern.
- Humphrey, S. E., Nahrgang, J. D., & Morgeson, F. P. (2007). Integrating motivational, social, and contextual work design features: A meta-analytic summary and theoretical extension of the work design literature. Journal of Applied Psychology, 92, 1332-1356.
- INFRAS (2011). Evaluation of Controls in Procurement. Final report. Prepared on behalf of the Federal Office for Gender Equality FOGE. Zurich, Infras.
- International Labour Office (2008). Promoting equity: Gender-neutral job evaluation for equal pay: A step-by-step guide. Geneva: International Labour Office.
- Kass, Robert E. & Raftery, Adrian E. (1995). Bayes Factors. Journal of the American Statistical Association, 90(430), 773-795.
- Katz, Christian & Baitsch, Christof (1996). Equal pay for practice. Two instruments for gender-independent job evaluation. Eidgenössisches Büro für die Gleichstellung von Frau und Mann EBG (ed.). Zurich: vdf.
- Kendall, Maurice G. (1938) A new measure of rank correlation, Biometrika, 30, pp. 81-93.
- Kennedy, Peter E. (1981). Estimation with Correctly Interpreted Dummy Variables in Semilogarithmic Equations, American Economic Review, 71(4), p. 801.
- Krell, Gertraude & Winter, Regula (2011). Anforderungsabhängige Entgeltdifferenzierung: Orientierungshilfen auf dem Weg zu einer. discriminierungsfreien Arbeitsbewertung. In: G. Krell, R. Ortlieb & B. Sieben (Eds.). Equal opportunities through personnel policy. 6th completely revised and expanded edition. Munich: Gabler. S. 343-360.
- PricewaterhouseCoopers AG (2020). Technical description of the salary specification in the Confederation's standard analysis model. Legally compliant salary specification. Commissioned by the Federal Office for Gender Equality FOGE. Zurich: PWC AG.
- Schär Moser, Marianne (2019). Analytical job evaluation: a reliable, gender-neutral basis for setting functional salaries. Federal Office for Gender Equality (ed.). Bern: EBG.
- Strub, Silvia (2004). Review of compliance with equal pay between women and men in federal procurement. Report on the pilot phase for the implementation of Art. 8 para. 1 let. C of the Federal Law on Public Procurement. Commissioned by the Federal Office for Gender Equality (FOGE) and the Federal Procurement Commission (FPC). Bern: BASS Office.
- Wooldridge, Jeffrey M. (2006). Introductory Econometrics: A Modern Approach. Mason: South-Western.