

Summary of the basics of the tolerance threshold in equal pay analyses using the Logib standard analysis tool

Working paper FAO FOGE (information as at June 2022), version 1

Commissioned by the
Federal Office for Gender Equality (FOGE)

Patrik Gajta, Lena Liechti, Silvia Strub and Markus Schärner
Bern, 9 January 2023

Contents

Executive Summary	III
Summary	V
1 Introduction and description of issue	1
1.1 Background	1
1.2 Questions	2
1.3 Methodological approach	3
1.3.1 Analysis of the literature and documentation	3
1.3.2 Questionnaire for legal practitioners	3
1.3.3 Collection and analysis of data from reviews carried out by the authorities	4
1.3.4 Questionnaire for authorities and qualitative interviews	4
1.3.5 Discussion of the results in a meeting of equal pay specialists	5
1.4 Structure of the report	5
2 Equal pay analyses: methodological foundations and context	5
2.1 Methods for measuring wage equality	5
2.2 The standard analysis model in Logib module 1	6
2.3 Development of the standard analysis model and definition of the tolerance threshold	8
2.4 Developments in the standard analysis tool	11
2.5 Developments in control practices	13
2.6 Developments of the context	14
3 The tolerance threshold in the literature and other domains	16
3.1 Evaluations of the tolerance threshold	17
3.1.1 Evaluation of the pilot review phase (2006-2011)	17
3.1.2 Evaluation of federal government equal pay analyses (2015)	19
3.1.3 Assessing the impact of regulation – amendment to the GEA (2015)	21
3.1.4 Assessments from the legal perspective	21
3.1.5 Other analyses and publications on Logib module 1	22
3.2 Empirical results concerning the tolerance threshold	23
3.2.1 Reviews in the public procurement system	23
3.2.2 Swiss Earnings Structure Survey	24
3.2.3 Analyses in the context of a private certification	26
3.2.4 Monte-Carlo simulations	26
3.3 Tolerance thresholds and limit values in other countries and fields	27
3.3.1 Equal pay analyses in neighbouring countries	27
3.3.2 Other tolerance thresholds and limit values	29
4 The tolerance threshold as applied in state reviews	34
4.1 Results of equal pay reviews	34
4.1.1 Descriptive data and statistics	34
4.1.2 Results of reviews and unexplained gender-specific differences	37
4.1.3 Adjustments to the tolerance threshold	43

4.2	Experiences of supervisory authorities	50
4.3	Assessments from a legal perspective	54
5	Summary and recommendations	55
5.1	Methodological aspects	55
5.2	Empirical aspects	57
5.3	Legal aspects	58
5.4	Political aspects	58
5.5	General conclusion	59
6	Bibliography	60

Executive Summary

Background

The Federal Office for Gender Equality (FOGE) commissioned the Centre for Labour and Social Policy Studies BASS to draw up a working paper collating the basic information on the tolerance threshold in equal pay analyses and also the experiences accrued during reviews of equal pay compliance in the public procurement and/or subsidies system at the municipal, cantonal and federal level. This report is intended to serve as a basis for deliberations carried out by FOGE equal pay specialists and external experts with regard to a possible adjustment to the tolerance threshold.

Results and conclusions

■ **Methodological aspects:** the tolerance threshold was set during a multi-stage procedure within the context of reviews of equal pay compliance in the Swiss Confederation's public procurement system. Since then, the analysis model has been externally assessed on several occasions at the national (Felfe et al, 2015; Pärli & Oberhauser, 2019) and international level (ILO, Equal Pay International Coalition [EPIC]) and deemed scientifically sound and legally compliant.

The analysis tool and control practices have been updated on numerous occasions; according to experts this has resulted in a notable improvement in the quality and completeness of the review data. With these gains in accuracy, the majority of experts and supervisory authorities feel that the tolerance threshold of +/-5% only allows the most serious cases of wage inequality within companies to be detected. Indeed, the tool was designed to analyse internal wage equality at the level of the whole company but does not provide any indicators concerning wage discrimination at the individual or group level. We know today that it is almost impossible for companies to be falsely identified by Logib as practising wage discrimination. Even if the tolerance threshold were to be lowered, the impact on this would be negligible.¹ In contrast, due to the methodology used, a great many companies are being "exonerated" even with large unexplained gender wage gaps. With the current tolerance threshold, the sensitisation

effect is being significantly reduced and has adverse consequences: a high degree of wage discrimination is being tolerated.

In the light of the various changes made to the tool and gains in knowledge, it seems appropriate for the FOGE's equal pay experts and external experts to consider adjusting the tolerance threshold. Other methodological aspects are currently being discussed and will also need to be taken into account, such as the impact of the omitted variable bias, the definition of and number of tiers in the "skill level" variable, as well as the question of robust standard errors. According to several studies, the tolerance threshold must be set based primarily on expert assessments, and it is not possible to precisely quantify the adjustment to the threshold based on changes made to the method or the tool. Therefore, the following factors should also be considered.

■ **Empirical aspects:** nearly all the studies presenting empirical results regarding the tolerance threshold show relatively similar proportions of entities displaying a gender effect²: almost half of entities display gender effects which differ significantly from zero, whether this relates to all Swiss entities, those reviewed within the context of the public procurement and/or subsidies system, or entities seeking to obtain equal pay certification. However, the proportion of entities significantly exceeding the +/-5% tolerance threshold varies considerably between the different sources: this figure stands at 25% and 19% in the studies based on the ESS, but at only 8% in the context of reviews and 4% in the context of certification analyses. Furthermore, the proportion of entities exceeding the +/-5% tolerance threshold has decreased substantially over time, falling from 8% on average to 3.3% for reviews conducted since 2019. Among these reviews, 60.7% do not show any gender effect.

In terms of empirical data, it thus seems that the FOGE specialists now have enough information to be able to judge the impact of a potential adjustment to the tolerance threshold on the results of equal pay reviews and analyses at different levels. Furthermore, this information should soon be supplemented by the results of tests relating to the impact of the omitted variable bias. The data in their current state seem to indicate that the FOGE should consider moving towards a lowering of the tolerance threshold.

¹ A BSS study shows that with a tolerance threshold of 5% the probability of false accusations with Logib module 1 is 0.2%. If the tolerance threshold were to be lowered to 3%,

the probability would still be vanishingly small, at less than 0.8%.

² An exception is the study of Trageser et al. (2011), which focused on only a smaller sample of 14 reviews.

■ **Legal aspects:** from a legal perspective, no gender discrimination whatsoever is tolerated by the law, no matter how small. The +/-5% tolerance threshold thus has no basis in law.

In the opinion of the legal equal pay experts consulted within the context of this assignment, the FOGE is the competent body to define the appropriate tolerance threshold for equal pay analyses using the standard analysis model. Furthermore, the legal experts share the opinion that not only the various technical and methodological developments, but also those relating to sensitisation and social and political context require the FOGE to look into possibly adjusting the tolerance threshold. What is more, they feel that lowering the tolerance threshold to +/-2.5% would be appropriate. In addition to this, current discussions at the European level are also heading in the same direction.

Based on the information obtained, it does not seem necessary to have any additional bases for a legal assessment. Defining a threshold is first and foremost a political task, and the legal experts have confirmed the competence of the FOGE to deal with this matter and to define the appropriate threshold level.

■ **Political aspects:** when the tolerance threshold was introduced within the context of the standardised review procedure, it had two objectives: to provide some degree of legal certainty by avoiding “false accusations” and to encourage political acceptance by giving companies time to adjust and to build awareness of the topic. In the meantime, almost 300 reviews have been conducted at different levels and a great many self-tests performed, in particular following the revision of the Gender Equality Act, which in July 2020 introduced an analysis obligation for all companies with 100 or more employees. Furthermore, many equal pay labels and certifications have sprung up in recent years, with growing success. The argument for a period of adjustment and the need to build awareness of the topic in order to justify the existence of a tolerance threshold thus seems hard to accept today.

In addition, the Federal Council and the experts consulted at the time felt that this threshold was set in a rather generous manner. Today, the supervisory authorities are of the same opinion. Many of them, who have been conducting reviews for a number of years, feel that certain infractions are not being detected with this tolerance threshold, as there is a significant gap between the low number of reviewed entities exceeding the threshold set and the number of

those showing a statistically significant gender effect.

Furthermore, in recent years there have been efforts at the political level to move towards greater wage equality between women and men at both the Swiss and international level. In addition, discussions are currently under way in the European Parliament to improve pay transparency. In particular, a tolerance threshold of +/-2.5% deviation in gross pay between women and men holding positions with comparable requirements is being discussed. In view of the various developments, deliberations regarding a possible adjustment to the tolerance threshold also seem necessary from a political point of view and are pertinent to the current context.

■ **General conclusion:** in conclusion, all the aspects considered within the context of this working paper justify giving some thought to adjusting the tolerance threshold. Moreover, the 2.5% level seems appropriate from the perspective of the various developments in the tool and the context, and would be in line with discussions at European level. However, this is primarily a normative decision which must be taken by the competent authorities.

Taking into account this working paper as well as the other work recently carried out on the subject, we believe that the FOGE now has sufficiently robust information bases to be able to launch the discussion process on a potential adjustment to the tolerance threshold.

Moreover, other work is currently being carried out in parallel and will need to be taken into account in this process. In particular, we can cite the work on the effects of the omitted variable bias, the drafting of a report on the strategy for strengthening the Charter for equal pay in the public sector, deliberations regarding the definition of and number of tiers in the “skill level” variable, and the upcoming integration of Logib module 1 into the Swiss ELM electronic salary reporting system through ERP systems. In addition, the FOGE had to prepare a report by the end of 2022 on the status of European legislation and possible effects on Switzerland.

Summary

Background

The Swiss Confederation provides a free standard analysis tool for verifying equal pay between women and men. While this tool was being developed between 2001 and 2005, a tolerance threshold of $\pm 5\%$ was introduced for the gender coefficient which must be exceeded in a statistically significant manner for the entity to be deemed non-compliant with the provisions of the law.

Since then, the tool has undergone many changes, as have control practices and the context surrounding equal pay analyses. Further, a number of studies have focused on the appropriateness of the model, the variables taken into account and the tolerance threshold, and have provided empirical data on this based on different data sources. Finally, numerous reviews at the federal, cantonal and municipal levels since 2006 as well as the Logib helpline, among other things, have allowed a vast body of experience to be acquired on both the topic and the analysis tool. With these changes and substantial gains in experience comes the question of whether the $\pm 5\%$ tolerance threshold is still appropriate to the standard analysis tool and its instructions in their current form.

The FOGE thus commissioned the Centre for Labour and Social Policy Studies BASS to draw up a working paper to help answer this question and aimed at collating the basic information on the tolerance threshold and also the experiences gained with this when conducting reviews of equal pay compliance in the public procurement and/or subsidies system at the municipal, cantonal and federal levels.

Methodological approach

This working paper is based on the following empirical methods and information sources: an analysis of the literature and documentation, collection and analysis of data from reviews conducted by the 7 federal, cantonal and municipal authorities having already carried out reviews, a questionnaire and qualitative interviews with these authorities, and a questionnaire sent to legal specialists in the field of equal pay.

Equal pay analyses: methodological foundations and context

Logib module 1 can be used with 50 or more valid data sets. It comprises the following four components:

■ **A dependent variable:** standardised gross wage based on a wage specification;

■ **Multiple independent variables:** factors to justify wage differences between women and men (education, seniority, potential work experience, skill level and professional position) as well as the “gender” variable;

■ **A statistical analysis method:** semi-logarithmic OLS regression analysis;

■ **A tolerance threshold of $\pm 5\%$** for the “gender” factor that must be exceeded by a statistically significant amount. On top of the tolerance threshold, the standard analysis model provides for **statistical significance tests** in order to take into account the statistical uncertainty inherent in the regression analysis. Thus, for the tolerance threshold to be exceeded, it must be exceeded in a statistically significant manner, that is with a degree of certainty of at least 95%.

The $\pm 5\%$ tolerance threshold was defined during a **multi-stage procedure** in the context of state reviews carried out within the Swiss Confederation’s public procurement system. There is no tolerance threshold provided for under the law. This $\pm 5\%$ threshold was initially set to offset the simplicity of the model, which took into account only three **human capital factors**, namely education, seniority and potential work experience. The intention was for this threshold to serve as a basis for identifying cases where discrimination was potentially suspected. However, it should be clarified that at the time this analysis with a $\pm 5\%$ tolerance threshold was only the first step in the three-step procedure envisaged. The plan was to abandon it later on in favour of analyses incorporating **job-related variables** and in-depth analyses. Strub (2004) thus argued that a $\pm 5\%$ tolerance threshold was justified when the “simple” model was being used, but not in the context of a “complete” model. Nevertheless, this threshold has continued to be used in the context of the **standardised review**, as well as in the **“expanded regression”**, which also includes the “skill level” and “professional position” variables, and the results of which are decisive for determining whether or not an entity is compliant within the context of a review. This is the regression retained since then in the standard analysis model.

The standard analysis model has used the same variables ever since the reviews were introduced, but these variables have been refined and verified repeatedly in the meantime. The definitions of the variables have thus been adjusted to developments in national and international understanding in the field of equal pay. Several assessments

Summary

(Felfe et al, 2015; Pärli & Oberhauser, 2019) dealing with this question demonstrated that the variables taken into account had significant **explanatory potential** and were **non-discriminatory**, while adding new variables was **not advised** as the majority of these are discriminatory or would require too great an effort on the part of companies to implement in an objective and neutral way. The standard analysis model was deemed to be **scientific** and **legally compliant**.

However, the tool has undergone many changes, as have control practices and the context surrounding equal pay analyses.

Changes to the tool, control practices and context

Table 1 below gives an overview of the main changes in these three aspects since 2006.

Table 1: Overview of changes

Aspect	Main changes	Impact
Analysis tool	- Integration of job titles and detailed instructions for coding of job-related variables - Pay specification - Modernisation of the tool and validation aids - Expert opinions guaranteeing the tool's scientific rigour and legal compliance - Kennedy Estimator - Integration into ERP systems from 2023	- Improved accuracy and standardisation in usage of tool, error limitation - Improvement in data quality - Improved accessibility and understanding of the tool
Control practices	- Directive regulating legal aspects - Documents for data collection and validation (questionnaires, checklist, internal validation (IVP)) - Collaboration with specialists and ongoing development of practices - Standardised report - Modernisation of the tool, validation aids, sensitivity analyses, integration into ERP systems from 2023	- Standardisation of process - Better quality and completeness of data - Practices up to date with current knowledge
Context	- Revision of the Gender Equality Act: analysis mandatory for companies with 100 or more employees, new PPA (Federal Act on Public Procurement) - Charter for equal pay in the public sector - Introduction of reviews in municipal areas/cantons - Equality Strategy 2030	- Greater sensitisation to the topic - Strengthening of the fight against inequalities

Source: BASS presentation

The tool has notably undergone many improvements over the years in order to achieve greater **accuracy** and **standardisation**. Likewise, control practices have also evolved in this direction, in particular with **continuous sharing of knowledge** and **good practices** between equal pay specialists. These developments are helping to decrease the room for manoeuvre for correct usage of the standard analysis model and thus increase the probability of obtaining results which concur with the company's reality.

The developments in the legal, social and political context have contributed to greater **awareness** of the topic of wage equality, to a growing interest on the part of companies to prove their compliance with this provision, as well as to a strengthened **political commitment** to combat wage discrimination.

Evaluations of the tolerance threshold

On top of these numerous developments, equal pay analyses have also formed the subject matter of many studies (in particular Trageser et al., 2011; Felfe et al., 2015; Rüegge et al., 2018) with the aim of evaluating the model used, the variables taken into account or the appropriateness of the tolerance threshold.

Broadly, both the authors of the different studies and the consulted specialists agreed that the tolerance threshold had been set in a **prudent** and **generous** way in order to avoid **unjustified sanctions**, with no real theoretical, empirical or legal foundation. Nevertheless, it had been deemed adequate at the time for various reasons. On the one hand, it fostered **acceptance of the process by companies** and gave them time to build awareness of the topic. On the other, it provided a way to offset the **lack of empirical experience** regarding the appropriateness of the model in practice and the fact that certain other potential non-discriminatory company-specific variables are not taken into account in the analysis. According to these various evaluations, however, this threshold would need to be lowered if the model were to be optimised, expanded or to include in-depth analyses. Nevertheless, various studies have demonstrated that it was **not appropriate** to expand the standard analysis model as the majority of other variables proposed had the potential to be discriminatory or implied too great an effort for companies to implement. Still, some studies provided scope to identify **potential ways to improve** the standard analysis model, in particular with regard to the specification of the variables included, more precise evaluation of functions or the minimum number of data sets required for use of Logib module 1.

Summary

Some of these recommendations have already been implemented as part of the tool's development, while others are currently being considered.

Furthermore, the studies have highlighted the need to **regularly examine** the appropriateness of the tolerance threshold in the light of new experiences within the context of equal pay reviews or analyses. In the absence of a legal or empirical foundation, the threshold should primarily be set based on **assessments by experts** in the field. Nonetheless, this decision remains to a very large extent a **political** one. For this reason, it is important to recall that the **law does not provide for a tolerance threshold** and prohibits any discrimination on the grounds of gender. Likewise, from an econometric perspective, it is not possible to determine the **appropriate value** for the tolerance threshold, insofar as any regression analysis already includes a confidence interval aimed specifically at limiting the risks of reaching a false conclusion. Finally, other countries using similar tools for analysing wage equality do not apply any tolerance threshold at all.

Empirical results concerning the tolerance threshold

The various studies which have applied the tolerance threshold to data from the ESS (Graf and Garibian, 2014; Felder and Wunsch, 2021; Chávez-Juárez and Graf, 2021) have reached similar conclusions: **between 20 and 25% of entities significantly exceed the tolerance threshold if it is set at +/-5%**, with substantial differences noted according to sector (private/public), line of business, size of entity and proportion of women within the entity. These last two elements are also identified as key drivers influencing the power of the tests performed by the Logib tool.

The report by Comp-On (2021), based on analyses conducted in the context of a certification, presents markedly different results, finding that **just 4% of entities significantly exceed the tolerance threshold**. This could be explained, in particular, by the differences in how certain variables are implemented (wage, skill level), by the heightened awareness of the topic in companies which are keen to obtain this type of certification and are thus not representative of the economy as a whole, as well as by the data validation process carried out by an expert which avoids certain errors in the data capture process. However, their analyses show that 25% of companies display **borderline results**, i.e. results which stand at 5% or more, but which do not exceed the threshold by a significant amount.

The studies by Graf and Garibian (2014) and Chávez-Juárez and Graf (2021) show that lowering the tolerance threshold would lead to a **linear evolution** in the proportion of entities significantly exceeding the threshold. This figure stands at about 50% when no threshold is applied. According to Kaiser (2022), lowering the tolerance threshold **would strengthen the probability** of detecting gender effects, while the probability of obtaining false positive results would increase only very slightly (from 0.2% to a maximum of 1% probability of error for a tolerance threshold of +/-3%). Finally, Chávez-Juárez and Graf (2021) find that, from an econometric perspective, it is not justifiable to have a tolerance threshold on top of a confidence interval, as these two elements have the same objective: to avoid **concluding erroneously that an effect exists** when in reality it does not. They would thus recommend **removing** the tolerance threshold while integrating certain additional statistical tests to verify the robustness of the model and the results, in particular with regard to the standard errors. This process could be implemented gradually and via a series of intermediate tolerance thresholds.

Tolerance thresholds and limit values in other countries and fields

In Switzerland's neighbour countries, the majority of equal pay analyses are based on **unrefined comparisons of pay** between women and men in equivalent employment categories. The lessons to be drawn from them are thus relatively limited. However, Germany provides Logib-D for equal pay analyses; this tool is based on Logib module 1 with variables which in some instances are implemented differently. Logib-D does **not use a tolerance threshold** and only operates a single significance test to determine if the gender-specific wage gap is significantly different from zero.

Furthermore, discussions are under way at the European Parliament level to compel companies to put in place measures from a certain threshold difference in pay between women and men performing work of equal value, and in cases where the wage difference cannot be justified by objective and gender-neutral factors. The Parliament is looking to set this threshold at **2.5%** for all companies with 50 or more employees.

To sum up, an examination of further reviews carried out by the public authorities shows that due to **technical, scientific and social developments** as well as **experiences** with a specific review process, it may be necessary to adjust threshold values, thresholds for penalties, tolerance ranges or the review procedure in order to

Summary

be able to continue to fulfil the execution mandate. Having said that, direct conclusions from a methodological perspective seem limited, as no area could be identified which is also based on regression analyses (cf. also Infras, 2011).

Results of equal pay reviews

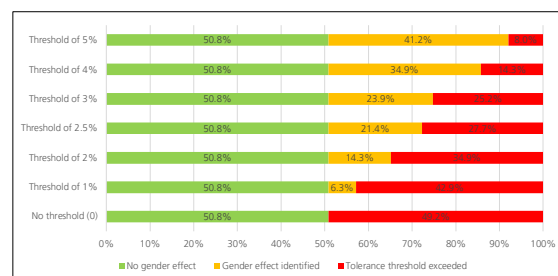
As part of this assignment, we collected and analysed the data from 260 reviews carried out at the federal, cantonal and municipal levels.

We find that **more than 50% of the entities reviewed** do not display **any gender effect** and would not have significantly exceeded the tolerance threshold even if a 0% threshold were applied. These results are relatively similar to those of the studies based on the ESS. The proportion of entities **exceeding the tolerance threshold** when it is set to $\pm 5\%$ stands at around **8%** and is markedly lower than that of the studies mentioned above. This can be explained by the fact that the reviewed entities are not representative of the economy as a whole and that the bidders for public procurement contracts commit to complying with wage equality and undergoing potential reviews by the authorities. Another explanatory factor is that certain variables operate differently between the ESS and analyses using the Logib module 1 tool. In addition, there is a data validation process included in the scope of the reviews which is not carried out in the ESS. It is also interesting to note that among the 40% of entities displaying a gender effect which does not exceed the $\pm 5\%$ tolerance threshold in a statistically significant manner, more than **two out of every three entities** have a discrimination coefficient of 5% or more and can thus be considered as **borderline cases**.

Furthermore, significant deviations are found in the average unexplained wage difference, notably in relation to the line of business, the proportion of women, the R^2 and the year of the review.

Adjusting the tolerance threshold downwards would have a considerable impact on the proportion of entities exceeding it. **Figure 1** shows the **impact** of such an adjustment on the review results.

Figure 1: Impact of adjusting the tolerance threshold on review results



Source: Data from reviews carried out by the authorities with information on standard error (n=238), BASS calculations

If a 0% threshold had been applied, this would have implied that 49.2% of the reviewed entities would have exceeded this threshold. This percentage decreases in an almost linear fashion in relation to the tolerance threshold: 42.9% with a $\pm 1\%$ threshold, 27.7% with a $\pm 2.5\%$ threshold, 14.3% with a $\pm 4\%$ threshold, and finally 8% with a $\pm 5\%$ threshold.

These adjustments would have had a differentiated impact on the results of the reviews in relation to various factors. In particular, they would have had a greater impact on the results of entities with at least 1,000 employees, those with 60% women or fewer, and reviews carried out between 2006 and 2010. Conversely, the results for small entities, for those employing more than 60% women, and for reviews carried out since 2020 would have been impacted considerably less by a lowering of the tolerance threshold.

Experiences of supervisory authorities

The experiences of supervisory authorities were also collected and analysed as part of this assignment.

The supervisory authorities consider that numerous **developments in the tool** and in **control practices** have helped to markedly enhance the **quality** and **completeness of the data**, or at least the speed at which these criteria could be achieved. The processes are now more standardised and there is a better guarantee that the results will be robust.

The authorities also agree on the important role that the existence of a tolerance threshold has played in **acceptance** of the review process by the reviewed entities.

They feel that **awareness** of the topic has developed greatly since the reviews were introduced and that it is more and more important for companies to show that they are upholding wage equality. According to one authority, lowering

Summary

the tolerance threshold would nevertheless still be viewed poorly by entities that could potentially be reviewed.

Furthermore, the majority of the supervisory authorities consider the $\pm 5\%$ tolerance threshold to be **too generous**. Apart from the developments mentioned which allow for more accurate results to be obtained, many authorities lament the fact that virtually all entities reviewed “pass” the review while in some cases displaying gender effects higher than $\pm 5\%$, which suggests that some infractions are not being detected. Some authorities feel that the standard analysis model with this $\pm 5\%$ tolerance threshold only allows them to identify the **most serious cases** of wage discrimination, as an enormous number of individual instances within the entity would be required to significantly exceed the tolerance threshold. This is considered a problem as there is no tolerance threshold envisaged under the law.

The majority of supervisory authorities agree on the **need to lower** the tolerance threshold; however, they are unable to give a precise idea of what an appropriate value for it might be. Among others, a tolerance threshold of **$\pm 2.5\%$** has been mentioned, as it is easy to communicate (threshold reduced by half) and corresponds in part to discussions at the European level. Other authorities would even argue for a lower threshold, set at 1 or 2%, if not for removing it altogether.

If the majority of the authorities argue in favour of lowering the tolerance threshold in order to be better able to detect infractions, they remain relatively sceptical with regard to acceptance of this move at the cantonal and municipal levels, whether in certain political circles or by companies.

Assessments from the legal perspective

All 3 of the legal experts consulted feel that the FOGE is competent to review and if necessary adjust the currently valid tolerance threshold for Logib module 1. Likewise, they agreed with the assertions that an adjustment to the currently valid 5% tolerance threshold would be appropriate (if not essential), for the following reasons: technical progress, various clarifications, numerous guidance documents, training courses and information materials, new validation processes for scientific quality and legal compliance, changes to

legal and political framework conditions, international developments and greater awareness among employers.

Summary and recommendations

■ **Methodological aspects:** the tolerance threshold was set during a multi-stage procedure within the context of reviews of equal pay compliance in the Swiss Confederation’s public procurement system. Since then, the analysis model has been externally assessed on several occasions at the national (Felfe et al, 2015; Pärli & Oberhauser, 2019) and international level (ILO, EPIC) and deemed scientifically sound and legally compliant.

The analysis tool and control practices have been updated on numerous occasions; according to experts this has resulted in a notable improvement in the quality and completeness of the review data. With these gains in accuracy, the majority of experts and supervisory authorities feel that the tolerance threshold of $\pm 5\%$ only allows the most serious cases of wage inequality within companies to be detected. Indeed, the tool was designed to analyse internal wage equality at the level of the whole company but does not provide any indicators concerning wage discrimination at the individual or group level. We know today that it is almost impossible for companies to be falsely identified by Logib as practising wage discrimination. Even if the tolerance threshold were to be lowered, the impact on this would be negligible³. In contrast, due to the methodology used, a great many companies are being “exonerated” even with large unexplained gender wage gaps. With the current tolerance threshold, the sensitisation effect is being significantly reduced and has adverse consequences: a high degree of wage discrimination is being tolerated.

In the light of the various changes made to the tool and gains in knowledge, it seems appropriate for the FOGE’s equal pay experts and external experts to consider adjusting the tolerance threshold. Other methodological aspects are currently being discussed and will need to be taken into account, such as the impact of the omitted variable bias, the definition of and number of tiers in the “skill level” variable, as well as the question of robust standard errors.

According to several studies, the tolerance threshold must be set based primarily on expert assessments, and it is not possible to precisely

³ A BSS study shows that with a tolerance threshold of 5% the probability of false accusations with Logib module 1 is 0.2%. If the tolerance threshold were to be lowered to 3%,

the probability would still be vanishingly small, at less than 0.8%.

Summary

quantify the adjustment to the threshold based on changes made to the method or the tool. Therefore, the following factors should also be considered.

■ **Empirical aspects:** nearly all the studies presenting empirical results regarding the tolerance threshold show relatively similar proportions of entities displaying a gender effect⁴: almost half of entities display gender effects which differ significantly from zero, whether this relates to all Swiss entities, those reviewed within the context of the public procurement and/or subsidies system, or entities seeking to obtain equal pay certification. However, the proportion of entities significantly exceeding the $\pm 5\%$ tolerance threshold varies considerably between the different sources: this figure stands at 25% and 19% in the studies based on the ESS, but at only 8% in the context of reviews and 4% in the context of certification analyses. Furthermore, the proportion of entities exceeding the $\pm 5\%$ tolerance threshold has decreased substantially over time, falling from 8% on average to 3.3% for reviews conducted since 2019. Among these reviews, 60.7% do not show any gender effect.

In terms of empirical data, it thus seems that the FOGE specialists now have enough information to be able to judge the impact of a potential adjustment to the tolerance threshold on the results of equal pay reviews and analyses at different levels. Furthermore, this information should soon be supplemented by the results of tests relating to the impact of the omitted variable bias. The data in their current state seem to indicate that the FOGE should consider moving towards a lowering of the tolerance threshold.

■ **Legal aspects:** from a legal perspective, no gender discrimination whatsoever is tolerated by the law, no matter how small. The $\pm 5\%$ tolerance threshold thus has no basis in law.

In the opinion of the legal equal pay experts consulted within the context of this assignment, the FOGE is the competent body to define the appropriate tolerance threshold for equal pay analyses using the standard analysis model. Furthermore, the legal experts share the opinion that not only the various technical and methodological developments, but also those relating to sensitisation and social and political context require the FOGE to look into possibly adjusting the tolerance threshold. What is more, they feel that lowering the tolerance threshold to $\pm 2.5\%$ would be appropriate. In addition to this, current discussions

at the European level are also heading in the same direction.

Based on the information obtained, it does not seem necessary to have any additional bases for a legal assessment. Defining a threshold is first and foremost a political task, and the legal experts have confirmed the competence of the FOGE to deal with this matter and to define the appropriate threshold level.

■ **Political aspects:** when the tolerance threshold was introduced within the context of the standardised review procedure, it had two objectives: to provide some degree of legal certainty by avoiding “false accusations” and to encourage political acceptance by giving companies time to adjust and to build awareness of the topic. In the meantime, almost 300 reviews have been conducted at different levels and a great many self-tests performed, in particular following the revision of the Gender Equality Act, which in July 2020 introduced an analysis obligation for all companies with 100 or more employees. Furthermore, many equal pay labels and certifications have sprung up in recent years, with growing success. The argument for a period of adjustment and the need to build awareness of the topic in order to justify the existence of a tolerance threshold thus seems hard to accept today.

In addition, the Federal Council and the experts consulted at the time felt that this threshold was set in a rather generous manner. Today, the supervisory authorities are of the same opinion. Many of them, who have been conducting reviews for a number of years, feel that certain infractions are not being detected with this tolerance threshold, as there is a significant gap between the low number of reviewed entities exceeding the threshold set and the number of those showing a statistically significant gender effect.

Furthermore, in recent years there have been efforts at the political level to move towards greater wage equality between women and men at both the Swiss and international level. In addition, discussions are currently under way in the European Parliament to improve pay transparency. In particular, a tolerance threshold of $\pm 2.5\%$ deviation in gross pay between women and men holding positions with comparable requirements is being discussed.

In view of the various developments, deliberations regarding a possible adjustment to the tol-

⁴ An exception is the study of Trageser et al. (2011), which focused on only a smaller sample of 14 reviews.

Summary

erance threshold also seem necessary from a political point of view and are pertinent to the current context.

■ **General conclusion:** in conclusion, all the aspects considered within the context of this working paper justify giving some thought to adjusting the tolerance threshold. Moreover, the 2.5% level seems appropriate from the perspective of the various developments in the tool and the context, and would be in line with discussions at European level. However, this is primarily a normative decision which must be taken by the competent authorities.

In order to provide the best conditions for making this decision, this working paper aims to provide a documentary base that is as wide-ranging as possible, bringing together an analysis of the literature, empirical data, experiences and assessments by experts, supervisory authorities and legal equal pay specialists. Taking into account this working paper as well as the other work recently

carried out on the subject, in particular by Chávez-Juárez and Graf (2021) and Kaiser (2022), we believe that the FOGE now has a sufficiently robust base of information to be able to launch the discussion process on a potential adjustment to the tolerance threshold.

Moreover, other work is currently being carried out in parallel and will need to be taken into account in this process. In particular, we can cite the work on the effects of the omitted variable bias, the drafting of a report on the strategy for strengthening the Charter for equal pay in the public sector, deliberations regarding the definition of and number of tiers in the “skill level” variable, and the upcoming integration of Logib module 1 into the Swiss ELM electronic salary reporting system through ERP systems. In addition, the FOGE had to prepare a report by the end of 2022 on the status of European legislation and possible effects on Switzerland.

1 Introduction and description of issue

1.1 Background

In the context of the tendering procedure for public procurement contracts, it should be ensured that certain conditions are met by the providers, in particular in order to preserve social gains and industrial peace, as well as avoiding unfair competition. The Federal Act on Public Procurement (PPA) specifies that one of the required conditions for winning public procurement contracts is compliance on equal pay between men and women (Art. 12 para. 1 PPA). Equal pay is enshrined in the Federal Constitution, which notably stipulates that discrimination is prohibited (Art. 8 para. 2 Cst.) and that “men and women have the right to equal pay for work of equal value.” (Art. 8 para. 3 Cst.). The Gender Equality Act (GEA) specifies that the prohibition of all discrimination on the basis of sex applies to all employment relationships, including pay (Art. 3 para. 2 GEA).

In order to verify equal pay compliance in the Confederation’s public procurement system, the awarder can, as per Art. 4 of the Ordinance on Public Procurement (PPO), entrust equal pay reviews in particular to the Federal Office for Gender Equality (FOGE). Besides the FOGE, other cantonal and municipal authorities conduct regular equal pay reviews during award procedures for public procurement contracts or subsidies within their canton or municipality. To allow these reviews to be carried out, the Swiss Confederation provides a free analysis tool for verifying wage equality between women and men, comprising two modules. Module 1, which can be used with 50 or more valid data sets, is based on a semi-logarithmic regression analysis which assesses the impact on pay of different objective and non-discriminatory attributes such as years of education, years of service, potential work experience, skill level and professional position, as well as the gender variable to determine the gender-specific wage difference.

In the course of developing the standard analysis tool (Strub, 2004; 2005), a tolerance threshold of $\pm 5\%$ was introduced for the gender coefficient which must be exceeded in a statistically significant manner. This tolerance threshold was introduced following a pilot phase conducted with 5 companies. However, the $\pm 5\%$ value of the threshold has no theoretical, empirical or legal foundation, according to Felfe et al. (2015), and can thus be considered as a political threshold. Consequently, it was recommended in the Federal Council’s report in response to the postulate from Noser (14.3388) to re-examine the level of the tolerance threshold based on review experiences.

According to the different studies carried out on the tolerance threshold in recent years, and in particular those by Trageser et al. (2011) and Felfe et al. (2015), the $\pm 5\%$ tolerance threshold was considered appropriate in relation to the accuracy of the analysis model and the variables included within it. However, these two studies recommended that authorities re-evaluate the appropriateness of the tolerance threshold with the aid of new experiences gained in the context of reviews and consider a potential adjustment to this threshold in line with these new experiences and the potential optimisations added to the analysis tool.

In recent years, Logib’s module 1 has been further developed and optimised, in particular by incorporating job titles, developing detailed instructions for how to code jobs, providing a detailed specification for pay⁵ and its components, introducing a Kennedy Estimator into the regression, and modernising the tool, which is now available online and includes validation aids. Furthermore, several reports (Felfe et al., 2015; Pärli & Oberhausser, 2019) have assessed the appropriateness of adding new variables to the model and

⁵ PricewaterhouseCoopers (2019). Technical description of the pay specification used in the Confederation’s standard analysis model, FOGE

1 Introduction and description of issue

reached the following conclusion: the existing model is scientific and legally compliant and the variables included have a high explanatory potential and are non-discriminatory. The majority of the other proposed variables have the potential to be discriminatory or entail a disproportionate effort on the part of companies to calculate in an objective and non-discriminatory way and thus may not be included in such a model.

Furthermore, the numerous reviews and self-tests carried out and the operation of the Logib helpline (around 2,700 enquiries received) have allowed a substantial body of experience to be acquired which can no longer be compared to that of 2005 after the pilot phase with 5 companies. Thus, with these changes and substantial gains in experience comes the question of whether the $\pm 5\%$ tolerance threshold is still appropriate to the standard analysis tool and its instructions in their current form.

In response to the inquiry on “Wage inequality between women and men. A question of method” (21.4190) from Eva Herzog⁶, the Federal Council has furthermore affirmed that “by 2023 the FOGE will investigate, in collaboration with specialists, the need to adjust this threshold” on the basis of the reviews carried out in the context of the public procurement system and with the beneficiaries of subsidies awarded by the Confederation.

This working paper is intended to form part of this procedure and aims to collate the basic information on the tolerance threshold as well as experiences of it during reviews of equal pay compliance in the public procurement and/or subsidies system at the municipal, cantonal and federal levels.

1.2 Questions

One of the particular objectives of this assignment is to provide answers to the following questions, which can be divided into three thematic areas:

Current knowledge on the tolerance threshold and comparison with other fields

What is currently known about the tolerance threshold within the context of equal pay reviews or analyses? Are there, within other review procedures, thresholds which could be compared to the tolerance threshold for equal pay reviews? Are tolerance thresholds applied elsewhere to fulfil the enforcement mandate? How are limit values set in other fields? How are they applied from a legal and methodological perspective, and how are their values objectively justified? Is it possible to draw conclusions from these other fields for the tolerance threshold?

Results arising from completed reviews

What is the composition of the companies reviewed in terms of size, linguistic region, economic sector? What are the results of the reviews? In how many companies was no gender effect observed? In how many companies is there a gender effect exceeding the tolerance threshold? In how many companies is this gender effect below the tolerance threshold? What are the results concerning the significance level of the gender-specific difference? What would the results of reviews look like if the tolerance threshold changed?

Evaluation of the tolerance threshold

What effect does the tolerance threshold have during reviews in the field of public procurement contracts and/or subsidies? What questions arise from this? How have the modernisation of Logib and the clarification of the instructions for data collection impacted the quality of the data and the results of reviews?

⁶ Retrievable online: <https://www.parlament.ch/fr/ratsbetrieb/suche-curia-vista/geschaeft?AffairId=20214190>

1 Introduction and description of issue

Should this have an impact on the tolerance threshold value? Are there recommendations in place for adjusting the tolerance threshold? Is other information necessary to allow the FOGE and specialists to investigate whether the tolerance threshold should be adjusted by 2023?

In order to provide answers to these questions, we developed a multi-stage methodological approach. This is presented in detail in the next section.

1.3 Methodological approach

In order to prepare this working paper, we used the following empirical methods and sources of information:

- Analysis of the literature and documentation
- Questionnaire for legal practitioners
- Collection and analysis of data from reviews carried out by the authorities
- Questionnaire for authorities and qualitative interviews
- Discussion of the results in a meeting of equal pay specialists

1.3.1 Analysis of the literature and documentation

As part of the analysis of the literature and documentation, we first took into account the various studies by Strub defining the analysis method and the tolerance threshold, in particular the report on the pilot review phase (2004) and the description of the method (2005). We also analysed the various studies conducted since then with the aim of evaluating the analysis model or the review process as well as the tolerance threshold from the methodological and legal perspective, in particular those by Trageser et al. (2011), Felfe et al. (2015), Stern et al. (2015), Rügge et al. (2018), Binggeli et al. (2018), Pärli & Oberhauser (2019) and Felder & Wunsch (2021).

We also took into account the development of the legal foundations relating to equal pay and public procurement contracts as well as that of the documentation for the Logib tool and the control practices. To this end, we analysed the studies regarding the pay specification (PwC, 2019).

Finally, we also studied and summarised the various reports dealing with the results of analyses carried out using Logib. These relate to different aspects: the results of empirical analyses conducted based on data from the Swiss Earnings Structure Survey with the Confederation's standard analysis model (Graf & Garibian, 2014; Felder & Wunsch, 2021; Chávez-Juárez & Graf, 2021), the results of equal pay analyses conducted within the context of a certification (Comp-On, 2020), as well as the results of simulations aiming to measure the power of the tests used by the Logib tool (Kaiser, 2022).

We also analysed the existing literature and documentation in other countries and fields with experience of tolerance thresholds or limit values in order to determine whether potential conclusions can be drawn for the tolerance threshold in equal pay analyses in the Swiss context. Exploratory interviews with representatives of the respective fields were also conducted in order to round off the observations.

1.3.2 Questionnaire for legal practitioners

In order to optimally integrate the legal perspective into this assignment, it was decided to address a questionnaire to three legal equal pay specialists with substantial experience in this field. This questionnaire comprised five questions dealing with the following aspects:

- The FOGE's competence to set and evaluate the tolerance threshold for equal pay analyses;
- The developments which would justify a potential review of the tolerance threshold;

1 Introduction and description of issue

- Adjustment of the value of the tolerance threshold;
- The information needed for a potential adjustment of the tolerance threshold;
- The conclusions which can be drawn from limit values and thresholds present in other fields.

The legal experts consulted responded concisely to the questions; the responses were not fully developed legal appraisals.

1.3.3 Collection and analysis of data from reviews carried out by the authorities

In order to provide the most complete background data possible, it was decided to collect anonymised results from reviews carried out by the authorities at the municipal, cantonal and federal levels. An Excel table to be completed by the authorities was sent out by e-mail along with an instruction document for the data entry process. The Excel table included the following information:

- Year and reference month of review
- Economic sector and size of company
- Number of employees, men and women in the analysis
- Results: discrimination coefficient, significance in relation to 0% and +/-5%, R^2 and standard error of the discrimination coefficient

The authorities contacted are as follows:

- Swiss Confederation: Federal Office for Gender Equality (FOGE)
- Canton of Basel-Stadt: Abteilung Gleichstellung von Frauen und Männern
- Canton of Bern: Fachstelle für die Gleichstellung von Frauen und Männern
- Canton of Geneva: Bureau de promotion de l'égalité et de prévention des violences (BPEV) und Office cantonal de l'inspection et des relations du travail (OCIRT)
- Canton of Vaud: Bureau de l'égalité entre les femmes et les hommes (BEFH)
- City of Bern: Fachstelle für Gleichstellung von Frau und Mann der Stadt Bern
- City of Zurich: Fachstelle für Gleichstellung der Stadt Zürich

Based on the data collected, analyses were conducted with a view, in particular, to determining the composition of the reviewed companies in terms of size, economic sector and proportion of women, and to evaluating the results based on the gender-specific effects detected. Simulations were also carried out in order to determine what the review results would be if the tolerance threshold changed.

1.3.4 Questionnaire for authorities and qualitative interviews

When the Excel table was sent out to the authorities, a questionnaire containing five questions was sent at the same time. The idea here was to obtain written feedback regarding their experiences with the review process and the tolerance threshold as well as their appraisals of these. Questions on the quality of the data and the developments in the tool and its instructions as well as on current discussions being held by the authorities were also asked in order to benefit from their input on these topics.

Based on the responses obtained, four interviews were conducted with the authorities, as the three others felt that the written responses provided were sufficient to express their point of view on the topic. The aim of the interviews was primarily to further explore the responses obtained in writing.

2 Equal pay analyses: methodological foundations and context

It should be highlighted here that the authorities' experiences are relatively different, as some of them have been conducting reviews since 2006 (at the federal level) or 2011 (at the cantonal level), while others have only followed a few reviews over the years or have only introduced a review procedure within their municipality/canton recently.

1.3.5 Discussion of the results in a meeting of equal pay specialists

Finally, the provisional results for this report were discussed during a meeting of equal pay experts from the FOGE and independent external specialist. Based on these discussions, the results in this working paper were revised and supplemented in order to incorporate the experts' assessments.

1.4 Structure of the report

The rest of this report is structured as follows. Section 2 provides a summary of the methodological bases and the context surrounding equal pay analyses. After a brief recap of the method and the standard analysis model, as well as the definition of the tolerance threshold, in sections 2.1, 2.2 and 2.3, we describe the main developments in the standard analysis tool (2.4), the control practices (2.5) and the context (2.6).

Section 3 reviews the existing documentation and literature on the tolerance threshold in equal pay analyses but also in other fields. We thus analyse the various evaluations of the tolerance threshold carried out in the literature (3.1) as well as the empirical results concerning the tolerance threshold presented by different studies (3.2). Next, the documentation on other thresholds or limit values currently in effect in other fields at the national or international level is reviewed in Section 3.3.

Section 4 focuses on the tolerance threshold in practice. The first part (4.1) provides a detailed analysis of the empirical results of equal pay reviews conducted at the federal, cantonal and municipal levels. The second (4.2) focuses on the experiences and assessments of the authorities with regard to control practices and the tolerance threshold. Finally, the assessments of legal equal pay experts on the subject are presented (4.3).

Section 5 provides a summary of the preceding sections and presents the recommendations formulated on the basis of the various analyses conducted within the context of this assignment.

2 Equal pay analyses: methodological foundations and context

This section seeks to provide an overview of the methodological bases and the context surrounding equal pay analyses and, consequently, the tolerance threshold. First, it briefly presents the standard analysis model and how the tolerance threshold was defined. In a second part, we present the various developments in the standard analysis tool but also in the context and control practices, which are relevant for any considerations relating to the tolerance threshold.

2.1 Methods for measuring wage equality

The equal pay principle involves three distinct levels, namely the individual, company and national levels. For each of these levels, the legal bases, competent authorities, groups to be compared, measurement methods and legal consequences for non-compliance differ greatly (Marti Whitebread, 2016). In the context of reviews within the public procurement system and, more broadly, equal pay analyses carried out using the Logib tool, we are dealing with the **company level**, which means that the method used must measure internal wage equality within a specific company. This working paper thus focuses on this level.

2 Equal pay analyses: methodological foundations and context

There are different options for measuring equal pay compliance within a company. The first of these is to conduct an analytical job assessment based on occupational science. The second is the statistical economic procedure known as a **regression analysis**. The scientific and legal conformity of these two approaches has been documented and they are accepted by the Swiss Federal Tribunal to respond to questions regarding wage discrimination.⁷

We will be focusing here on the regression analysis as it is the procedure on which the standard analysis model in Logib module 1 and thus the tolerance threshold is based. According to Strub (2004) and Bauer et al. (2001), this approach is used to evaluate instances of discrimination in the economy in general and may be applied to companies under certain conditions, notably a large enough company size. Furthermore, this approach is also used regularly by the courts in Switzerland and various other countries such as the USA to deal with the question of wage discrimination. The regression analysis calculates pay equations in order to ascertain the impact on pay of a variety of **explanatory factors** such as human capital factors, job-related characteristics and gender in order to reveal the gender-specific wage difference. This part of the wage difference, which can only be attributed to the “gender” factor after the effects of the other factors have been removed, is considered discriminatory. The following section describes the standard analysis model on which Logib module 1 is based in more detail.

2.2 The standard analysis model in Logib module 1

The **standard analysis model** is made up of the following four components:

- A dependent variable: standardised gross wage based on a wage specification;
- Multiple independent variables: factors to justify wage differences between men and women (education, seniority, potential work experience, skill or qualification level and professional position) as well as the gender variable;
- A statistical analysis method: semi-logarithmic OLS regression analysis;
- A tolerance threshold of +/-5% for the “gender” factor which must be exceeded by a statistically significant amount.

The standard analysis model explains employee wages by means of their **personal characteristics** (education, seniority, potential work experience), **job-related factors** (skill level, professional position) and **gender**. It thus provides a way to calculate, all other things being equal, the **gender-specific wage difference** between women and men. In order to achieve this, the standard analysis model uses the **multiple linear regression** approach, which is a statistical procedure used to study the relationship between a dependent variable (here the logarithm of the wage) and a number of independent variables. The regression coefficients are calculated using the **ordinary least squares** (OLS) method.⁸ The standard analysis model has been recognised as being scientific and legally compliant (see Sections 2.4 and 3.1).

The standard analysis model applies to the **company as a whole**, in other words all employees with a standard employment contract. In contrast, it is not applicable at the level of individuals or groups. The result obtained by the standard analysis model thus gives no indicators in terms of discrimination against individuals or groups.

The standard analysis model is implemented in concrete form in the **Logib module 1** analysis tool, which can be used with a minimum of 50 valid data sets. An Excel version of Logib module 1 has been freely

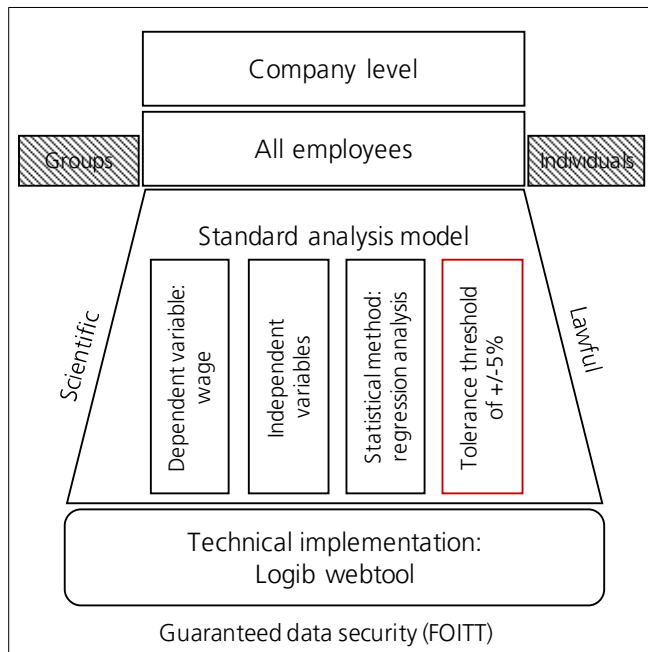
⁷ See the FOG declaration of conformity: <https://www.ebg.admin.ch/ebg/en/home/services/logib-triage/logib-modul-1/dokumentation-logib.html>

⁸ For details of the method and regression equations used, see the FOG's description of its methodological approach: <https://www.ebg.admin.ch/ebg/en/home/services/logib-triage/logib-modul-1/standardanalysemodell-bund.html>

2 Equal pay analyses: methodological foundations and context

available since 2004, and it has been available as a web-based tool since July 2020. The web-based application is managed by the Federal Office of Information Technology, Systems and Telecommunication (FOITT) and complies with all federal government security requirements. **Figure 2** below depicts the scope of application of the standard analysis model.

Figure 2: Scope of application of the standard analysis model



Source: BASS presentation

It should be clarified here that the discrimination coefficient, which measures the gender-specific wage difference, is estimated and thus lies within a certain area of statistical uncertainty. The standard analysis model thus runs two **statistical significance tests**: first to determine if the estimated value is significantly different to zero and if there is therefore a statistically significant gender effect; then, if that is the case, a second test is run to determine if the estimated value is significantly greater than the tolerance threshold, in this case $\pm 5\%$. The model uses a 95% confidence level, which is standard in this type of approach.⁹ Thus, for a company to exceed the $\pm 5\%$ tolerance threshold, it must exceed it in a statistically significant manner with a 95% confidence level.

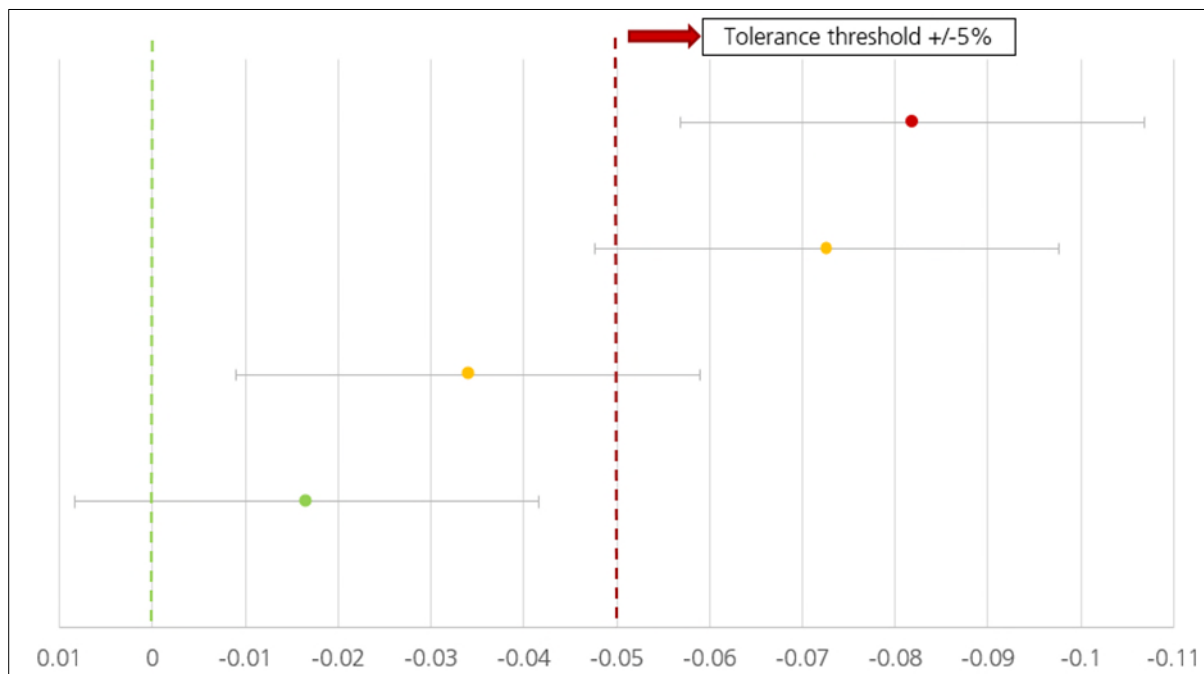
Figure 3 below illustrates the significance tests and the different possible scenarios with regard to the results of equal pay analyses. The four dots represent the discrimination coefficients calculated by the analysis model. The grey horizontal bars on either side of these dots represent the confidence intervals. For the green dot, we can see that the zero value is situated within the confidence interval, signifying that the discrimination coefficient in question is not significantly different from zero and there is thus no confirmed gender effect at the company level. For the first orange dot from the bottom, the discrimination coefficient differs significantly from zero as the zero value does not fall within the confidence interval, thus demonstrating the existence of a confirmed gender effect. However, this is below the tolerance threshold of $\pm 5\%$. The second orange dot shows a discrimination coefficient of -0.072 , which represents a gender-specific wage difference of 7.2% to the disadvantage of women. This coefficient differs significantly

⁹ The $\pm 5\%$ tolerance threshold for unexplained wage differences should not be confused with the significance level (or α level) which is also set at 5% and which determines if the result of the analysis is statistically significant. Up until now, adapting the α level has never been on the agenda.

2 Equal pay analyses: methodological foundations and context

from zero but nevertheless does not exceed the tolerance threshold as the 0.05 value is within the confidence interval. Finally, the red dot represents a discrimination coefficient of -0.082. This coefficient differs significantly from zero and is significantly greater than the tolerance threshold of $\pm 5\%$. There is thus a significant confirmed gender effect in this case.

Figure 3: Discrimination coefficients and results of significance tests



Source: Dummy data, BASS presentation

The following section describes in more detail the development of the standard analysis model and the definition of the tolerance threshold derived from it.

2.3 Development of the standard analysis model and definition of the tolerance threshold

The analysis tool for verifying compliance with wage equality between women and men was developed in the early 2000s by the Center for Labor and Social Policy Studies BASS on behalf of the FOGE and the Federal Procurement Commission (FPC) as part of the implementation of the Federal Act on Public Procurement (PPA; SR 172.056.1). This tool is based on a **statistical economic analysis procedure** which measures the impact of different factors on wages including, on the one hand, the share of wage differences attributable to **objective characteristics** relating to personal qualifications (education, seniority, potential experience) or to the job held (skill level, professional position) and, on the other, the share of the wage difference which is **gender-specific** and thus discriminatory. This procedure is based on the OLS regression method, which has been approved by the Federal Tribunal for dealing with matters relating to wage discrimination.¹⁰ The analysis tool was tested with 5 companies¹¹ in the context of a pilot phase carried out between 2001 and 2003. The report by Strub (2004) gives a detailed description of the procedure applied and the results of the pilot phase.

¹⁰ See, in particular, Decision ATF 130 III 145.

¹¹ Of the 5 companies, one was medium-sized (between 50 and 249 employees) and the 4 others were large companies. On average, these 5 companies employed 374 men and 109 women.

2 Equal pay analyses: methodological foundations and context

During the development phase, the process for analysing wage equality took place in several steps. In the first step, the regression analysis was run with a view to determining the effect on wages of **human capital factors** (education, seniority and potential work experience) on the one hand, and gender on the other. This initial analysis did not include the job-related variables, as its aim was to measure both **wage discrimination in the narrow sense** and **job discrimination** (based on qualifications of equal value, inequality of position). During this first step, it was decided to set a **tolerance threshold of +/-5%** in order to take into consideration the fact that other objective and non-discriminatory explanatory factors which are not taken into account by the model could have an impact on wages. The 5% figure was derived from the results of the analysis of the wage difference between women and men in the economy as a whole for 1998: this wage difference stood at 20% when only human capital factors were included in the analysis. The difference fell to 15% once professional position and skill level were taken into account in the analysis as well as human capital factors. 5% of the 20% was thus due to job discrimination. The remaining 15% tallied with wage discrimination in the narrow sense.

This tolerance threshold thus had two main objectives: on the one hand, to provide some degree of legal certainty by avoiding “false accusations” of companies potentially having other internal non-discriminatory variables, and on the other to encourage political acceptance of this new control procedure. For these reasons, the tolerance threshold was set in a relatively generous manner. There is no tolerance threshold provided for under the law.

If the first step revealed a gender-specific wage difference exceeding the tolerance threshold of +/-5%, other explanatory factors such as **job-related factors** (skill level and professional position) were then taken into account in a second step in order to ascertain if the wage differences arose from job discrimination or wage discrimination in the narrow sense. If the second step revealed a gender-specific wage difference which differed significantly from zero, then a third step allowed the auditors to resort to other analyses in order to assess the wage situation and suggest appropriate action. The +/-5% tolerance threshold was thus only intended to be used in the first step in order to detect potential wage or job discrimination. The aim of the two subsequent steps was to ascertain the source of this discrimination and to formulate targeted measures to address it. They therefore did not involve a tolerance threshold. At all stages, however, significance tests were run to verify whether the discrimination coefficient significantly exceeded +/-5% or 0% respectively.

Following this pilot phase, several modifications were made to the tool with a view to setting up equal pay reviews involving bidders for federal government public procurement contracts. The method for the standardised reviews is described more specifically in Strub (2005). The **standardised review** was based on two regression equations: one which measured the share of any wage difference explained by differences in terms of human capital factors (**basic regression**) and a second which considered not only those individual qualification characteristics but also job-related characteristics (**expanded regression**). These two analyses thus provided a way to differentiate between wage discrimination in the narrow sense (unequal wages for equal work or work of equal value) and the presumption of job discrimination. It is thus the results of the **expanded regression** which are decisive for evaluating **compliance with the rules** relating to equal pay in federal government public procurement contracts, and this regression is thus the one retained in the **standard analysis model**. In this model, the tolerance threshold of +/-5% was applied in the two analyses. However, this had initially been set for the human capital factors only. Adding the job-related variables thus did not result in any lowering or abolition of the tolerance threshold in the expanded regression. Furthermore, the significance tests remained unchanged. In addition to this, other

2 Equal pay analyses: methodological foundations and context

in-depth analyses could be conducted subsequently depending on the company. However, it was necessary to ensure that any company-specific factors taken into account in the analysis did not themselves harbour the potential for discrimination.¹²

Table 2 below presents the differences in methodological approach between the pilot phase and the launch of the reviews, in particular with regard to the tolerance threshold. It should be clarified here that both the model and the standard analysis tool were developed as part of the implementation process for equal pay reviews in the federal government's public procurement system. However, from 2004 the tool was made freely available to the public for self-tests. There is thus no difference between the tool used in the reviews and the one available to the public, with the exception of the potential sanctions associated with exceeding the tolerance threshold within the context of a review.

Table 2: Differences in methodological approach between pilot phase and launch of reviews

	Variables taken into account	Tolerance threshold	Objectives
Pilot phase of the tool (2001-2004)			
First step	Human capital factors: education, seniority, potential work experience	+/-5%	Detect suspected discrimination
Second step	Human capital factors + job-related variables: skill level and professional position	0%	Identify source of discrimination (narrow sense vs job), formulate measures to be implemented
Third step	Human capital factors + job-related variables + other non-discriminatory internal variables	0%	Incorporate other company-internal non-discriminatory variables, target measures
Reviews (2006 onwards)			
Basic regression	Human capital factors	+/-5%	Calculate the wage difference with equal personal qualification characteristics
Expanded regression (standard analysis model)	Human capital factors + job-related variables	+/-5%	Calculate the gender-specific wage difference, which is considered discriminatory. Sanctions if the tolerance threshold is significantly exceeded
Analyses at greater depth	Human capital factors + job-related variables + other non-discriminatory internal variables	+/-5%	Clarify the source of the gender-specific wage difference

Source: BASS presentation

¹² For example, one company in the pilot had internal requirements characteristics (in particular physical requirements) which could be incorporated into the analysis.

Interim conclusion 1

Through this review of how the analysis model and tolerance threshold were defined, we can make the following observations:

- The standard analysis model in Logib module 1 is made up of four components: a dependent variable (wage), independent variables (human capital factors and job-related variables as well as gender), a statistical analysis method (semi-logarithmic OLS regression analysis) and a tolerance threshold of +/-5% for the “gender” factor. On top of the tolerance threshold, the standard analysis model provides for statistical significance tests in order to take into account the statistical uncertainty inherent in the regression analysis. Thus, in order to exceed the tolerance threshold, a company must exceed it in a statistically significant manner; that is, with a degree of certainty of at least 95%.
- The +/-5% tolerance threshold was defined during a multi-stage procedure within the context of state reviews conducted within the Swiss Confederation’s public procurement system. There is no tolerance threshold provided for under the law. This +/-5% tolerance threshold was initially set to offset the simplicity of the model, which took into account only three human capital factors, namely education, seniority and potential work experience. The intention was for this threshold to serve as a basis for identifying cases where discrimination was potentially suspected. However, it should be clarified that at the time this analysis with a +/-5% tolerance threshold was only the first step in the three-step procedure envisaged. The plan was to abandon it later on in favour of analyses incorporating job-related variables and in-depth analyses. Strub (2004) thus argued that a +/-5% tolerance threshold was justified when the “simple” model was being used, but not in the context of a “complete” model. Nevertheless, this threshold has continued to be used in the context of the standardised review, as well as in the “expanded regression”, which also includes the “skill level” and “professional position” variables, and the results of which are decisive for determining whether or not an entity is compliant within the context of a review. This is the regression retained in the standard analysis model.

Since 2006 and the start of the reviews, various modifications have been made to the standard analysis tool. These are described in more detail in the next section.

2.4 Developments in the standard analysis tool

One argument put forward to help encourage political acceptance of the tool as a control instrument and to justify the existence of a tolerance threshold was the relative simplicity of the model and the fact that it potentially does not take into account all the pertinent factors for some companies. Nevertheless, this tolerance threshold has remained the same in spite of the “skill level” and “professional position” variables being incorporated into the standard analysis model. Further to this, the variables taken into account in the model, as well as numerous other variables, have undergone detailed evaluations. These evaluations concluded that the variables taken into account were relevant and non-discriminatory, while the other variables considered were discriminatory or required too much effort from companies to be implemented correctly.¹³ Since the integration of the job-related variables, no other variable has been added to the model. However, although the model has not been expanded, the **standard analysis tool** has seen some important developments which need to be taken into account when considering the tolerance threshold. This section describes the main developments.

The first substantial development compared to the pilot phase took place at the time the reviews were introduced. This was to abandon the in-depth analyses element (formerly step 3) of the standardised review procedure and retain only what was considered as the **“expanded regression”** (formerly step 2) as

¹³ See, in particular, Felfe et al. (2015) and Pärli & Oberhausser (2019).

2 Equal pay analyses: methodological foundations and context

the standard analysis model. The final result on wage equality was thus delivered by the regression analysis, taking into account the human capital factors as well as the job-related variables (skill level and professional position). The analysis tool could still be used to obtain the results of the “basic regression” (human capital factors only) and to conduct certain in-depth analyses separately, but the result of the review did not depend on these analyses.

Furthermore, both the standard analysis tool and the instructions for using it have been continuously developed in order to keep up with the latest political, methodological and legal developments in the field of wage equality.

One major development was the integration of **job titles** (or functions performed) into the tool in 2015. Before that there were a limited number of “spheres of activity” which could be selected. According to one authority consulted in the context of the interviews, this could distort the analysis as the company needed to adapt its real functions to make them correspond to the spheres of activity offered by Logib, and this could lead to distortions in the coding, compared to a completely free choice of codes. For that authority, the fact of leaving the field open to companies to enter their own job titles and then code them individually was an important development to allow the tool to better reflect the reality of each company. However, it should be specified here that sphere of activity was not a variable taken into account in the model and that the coding of the functions for the job-related variables was already unrestricted.

However, the major development in this respect was the development of **detailed instructions** on how to code these functions for the two variables “**skill level**” and “**professional position**”. One of the main reasons for developing these instructions was to adapt to national and international norms (CITP-08) which were then adjusted to the reality of companies in Switzerland. The factors which these two variables were intended to measure had been defined in a uniform manner while allowing companies to retain a degree of room for manoeuvre to be able to code their functions in the way that most closely matched the reality and would be pertinent in an internal comparison. Various examples and aids were incorporated into the instructions in this respect. The coding of these two variables is an essential aspect of the analysis method and it is important that this be done correctly in order to guarantee the relevance of the model. Although the model has not been expanded to include other variables, the variables taken into account have been continuously refined and verified to keep them in line with national and international knowledge.

Another essential aspect of the analysis model is the way **wage components are captured**. In this regard, a significant development has occurred with the technical description for the pay specification in the Confederation’s standard analysis model, and in particular the legally compliant pay specification produced by PricewaterhouseCoopers in 2019. This important specification work made it possible to define clear criteria for whether or not to take into account different wage components in the analysis, taking into account the discriminatory potential of each element and the best way to deal with questions relating to the frequency of payments. As a result of this work, a detailed list of numerous existing wage components was drawn up, accompanied by a list of criteria to examine in order to evaluate the relevance of capturing these elements. This development provided a way to harmonise the process for capturing the different wage components found in companies and is thus leading to a more standardised analysis procedure. Furthermore, this pay specification made it possible to spotlight the need to treat compensation in time in the same way as compensation in money; that is, as a wage component. This led to an adaptation in the standardisation of wages to usual weekly working hours in the tool at the time the web version was released.

2 Equal pay analyses: methodological foundations and context

Fittingly, the standard analysis tool was modernised and made available in the form of **a web-based tool** in July 2020 with the aim of ensuring that it would be accessible and simple to use. This modernisation made it possible to integrate various aids, options and filters in order to make it easier to **capture data accurately and correct them directly in the tool**. Both the cockpit incorporated into the tool and its expert mode¹⁴ present results from a number of automatically executed plausibility tests and several in-depth analyses which have already been completed. This interface thus allows users to gain an overview of their data and of a significant number of results which it would only have been possible to obtain previously with the help of in-depth analyses executed separately, outside of the tool. In the view of several of the authorities interviewed, this modernisation allows users to obtain **higher-quality data** more quickly and improves the standardisation of the process, including for in-depth analyses.

Finally, also in 2020, **the estimator proposed by Kennedy** (1981) was integrated into the analysis method in order to estimate the impact of gender on wages. As we have already discussed, the calculation of the discrimination coefficient always involves an element of uncertainty. However, the Kennedy Estimator provides a consistent and virtually unbiased estimator, which thus improves the model's accuracy.

Apart from these many technical developments and clarifications to the tool, it has also been the subject of many external appraisals which have confirmed that it is **scientifically rigorous** and **legally compliant**, as the declaration of conformity attests. These appraisals will be reviewed in more detail in Section 3. However, we can already confirm here that these appraisals have massively reduced the initial uncertainty surrounding the standard analysis model, which was sometimes considered too "simple" or as only allowing an "approximate" measure. The relevance of the model used by the tool and the variables included has been confirmed many times over, as has the choice not to expand the model to include new variables. The lack of empirical data on the tool has also been remedied in the interim.

Furthermore, from 2023 onwards the standard analysis tool will be integrated into ERP (Enterprise Resource Planning) software packages of designers certified by Swissdec, an association which deals with electronic exchanges of financial data within the Swiss e-government system. This impending integration into the Swiss ELM electronic salary reporting system via ERP systems should also help improve the quality of data management and reduce inaccuracies in how the tool is used, in particular at the point of data capture.

These various developments have thus made the tool more accurate, standardised and accessible. They should thus make it possible to limit errors in the use of the standard analysis model and thereby reduce the risk of the analysis producing only an inexact result or one that does not reflect a company's operational reality.

2.5 Developments in control practices

Another pivotal element, directly linked to the analysis method, is **control practices**, which have also seen a great many improvements since the pilot phase began more than 16 years ago.

The study by Trageser et al. (2011) highlighted several problems relating to the **quality** and **completeness of the data** collected in the context of the first reviews: missing data for certain groups of employees (e.g. management), entire variables missing (e.g. actual educational level), capture of allowances and other payments made only once a year, differentiated validation depending on the specialist responsible for the review. Furthermore, in-depth analyses (formerly step 3) were carried out repeatedly, incorporating

¹⁴ Logib's expert mode allows in-depth analyses to be executed and is aimed in particular at equal pay experts conducting analyses.

2 Equal pay analyses: methodological foundations and context

additional, company-specific variables into the analysis. In two cases, this variable was workplace, and in the last it was a variable for shift work. In other cases, these in-depth analyses singled out, for example, two separate employee groups (hourly vs monthly) or carried out matching as the number of women was low. In certain cases, worst case scenarios were established so as to be able to estimate missing education or wage data and thus still be able to conduct the analysis. Finally, the analysis reports were only partly standardised at the time, and consequently less uniform.

Trageser et al. (2011) thus felt that consistency and quality could be further improved by providing **standardised instructions** for data validation and coding of jobs, a better definition of outliers, definition of criteria for more in-depth analyses, a more detailed presentation of the results, as well as greater scope to apply pressure in the event of missing data by means of legal bases and options for sanctions.

These problems have been gradually eliminated by developing control practices at the federal, cantonal and municipal levels. These developments, like those of the tool, have been guided by the goal of **standardising the process**. This consequently translated into the development of a number of documents aimed at harmonising the control practices. In particular, we can cite the following documents:

- **Directive on monitoring of equal pay:** document regulating the legal aspects of control as well as how it is to be conducted and the time frames to be applied to the different steps.

- **Questionnaires 1 and 2:** forms aimed at obtaining information on the data held by the reviewed entity which allow the reviewer to clarify certain questions before starting data validation.

- **Checklist and report for internal validation (IVP):** documents providing the specialists responsible for the reviews a base of the different elements to verify and the validation tests to perform for each of the variables in a detailed and accurate manner in order to guarantee exact and consistent data.

- **Collaboration with specialists and ongoing development of the instructions:** regular exchanges with equal pay specialists allowing knowledge-sharing and guaranteeing continuous improvement and standardisation of the review process, including instructions for use which have developed greatly since the reviews were introduced.

- **Standardised review report:** report to be sent to the entity at the end of the review which presents the key results in a standardised manner.

- **Modernisation of Logib:** Validation aids, detailed analyses, sensitivity analyses, interface with ERP systems (from 2023).

The development of these documents constitutes an important advance in the quality and completeness of the data, as well as in the standardisation of the process. All these elements reduce the potential uncertainties relating to the use of the tool and the review procedure, which was also one of the arguments used to justify the existence of the tolerance threshold in some appraisals (see Section 3.1). We will also look at the experiences of the supervisory authorities with this procedure in Section 4.2.

2.6 Developments of the context

Apart from the developments in the tool and control practices, the **context** surrounding equal pay analyses and reviews plays a crucial role in any discussion of this subject.

As regards the **legal context**, the main developments arise from the revised **Gender Equality Act** (GEA) which entered into force in 2020. The revised GEA requires employers with at least 100 employees to carry out an internal equal pay analysis using a scientific and legal method, such as the standard analysis model provided by the Confederation. Moreover, as spelled out in Article 13c para. 2, the Confederation has the task of providing all employers with an analysis tool free of charge. Furthermore, a new law on

2 Equal pay analyses: methodological foundations and context

public procurement (PPA) entered into force in 2019. This makes it possible to better reflect existing practice in the field of wage equality but is not significantly different to the 1996 version. In particular, the PPA specifies that public procurement contracts may be awarded only to bidders that comply with the requirement for equal pay between women and men and that non-compliance with these conditions may result in them being excluded from the process or having the award revoked, as well as sanctions. The contracting authority also has the right to monitor compliance with wage equality between women and men. The new bidder's declaration notably stipulates that entities with 100 or more employees must provide evidence of having ensured compliance with the requirement for equal pay between women and men.

These recent legal changes cement **the importance of equal pay compliance** by companies as well as the **sensitisation** of the latter to this issue, as all companies with 100 or more employees must carry out an equal pay analysis at least once. These elements also consolidate knowledge of the Logib module 1 analysis tool and how it is used. Companies with 100 or more employees together employ around 46% of Switzerland's workers.

Apart from these legal changes, it is important to note that a large number of cantons (17), municipal areas (128), and semi-public organisations (90) have thus far signed the **Charter for equal pay in the public sector**, introduced in 2016, which marks their commitment to strengthening awareness of wage equality issues and ensuring equal pay compliance in the public procurement and/or subsidies system. In this regard, a report on the strategy for strengthening the Charter on equal pay is currently being drawn up in response to Postulate 20.4263 "Strategy for strengthening the Charter on equal pay" from the National Council's Science, Education and Culture Committees.

In addition to the FOGE, 6 cantonal and municipal authorities (listed in Section 1.3.3), whose experiences will be taken into account in this report, have already put in place **review procedures** within their public procurement and/or subsidies systems. Currently, 3 other cantons are in the process of putting in place a procedure or discussing the possibility of doing so: Jura, Ticino and Basel-Land. The study by Rüegge et al. (2018), which concerned the effects of equal pay reviews in the Confederation's public procurement system, showed that these played an **important role** in increasing companies' awareness, and that most of them had carried out **new analyses** afterwards, the majority of which at **regular intervals**. The proliferation of reviews conducted and of authorities having introduced reviews in their own public procurement and/or subsidies systems are thus undeniably contributing to raising awareness around the issue of wage equality.

Finally, in 2021 the Federal Council adopted the **2030 Gender Equality Strategy**, which counts eliminating wage discrimination as one of its specific objectives. This strategy strengthens political engagement at the federal, cantonal and municipal levels in the fight against wage inequality and identifies as a priority the pursuit of "developing and making available quality tools for analysing wage equality" by the Confederation. This strategy is also part of an international context of various efforts and measures¹⁵ directed towards eliminating all wage discrimination.

Table 3 below provides a summary of the main developments concerning the three aspects reviewed here; namely, the analysis tool, control practices and the context.

¹⁵ Among others, we can cite here the Equal Pay International Coalition (EPIC) and the EU's new Pay Transparency Directive.

3 The tolerance threshold in the literature and other domains

Table 3: Overview of main developments

Aspect	Main developments	Impact
Analysis tool	- Integration of job titles and detailed instructions for coding of job-related variables - Pay specification - Modernisation of the tool and validation aids - Expert opinions guaranteeing the tool's scientific rigour and legal compliance - Kennedy Estimator - Integration into ERP systems from 2023	- Improved accuracy and standardisation in usage of tool, error limitation - Improvement in data quality - Improved accessibility and understanding of the tool
Control practices	- Directive regulating legal aspects - Documents for data collection and validation (questionnaires, checklist, IVP) - Collaboration with specialists and ongoing development of practice as well as of the instructions for use of the tool - Standardised report - Modernisation of the tool, validation aids, sensitivity analyses, integration into ERP systems from 2023	- Standardisation of process - Better quality and completeness of data - Practices up to date with current knowledge
Background	- Revision of the Gender Equality Act: analysis mandatory for companies with 100 or more employees, new PPA (Federal Act on Public Procurement) - Charter for equal pay in the public sector - Introduction of reviews in municipal areas/cantons - Equality Strategy 2030	- Greater sensitisation to the topic - Strengthening of the fight against inequalities

Source: BASS presentation

Interim conclusion 2

■ Since the reviews were introduced, the standard analysis model has not been expanded to include new variables. Several assessments dealing with the question have demonstrated that the variables taken into account were relevant and non-discriminatory, while the addition of new variables was not advised as the majority of these are discriminatory or would require too great an effort on the part of companies to implement in an objective and neutral way. The standard analysis model was deemed to be scientific and legally compliant. Moreover, the variables taken into account have been regularly adapted to correspond to evolving knowledge at the national and international level

■ However, both the tool and the variables taken into account have undergone numerous developments and improvements over the years in order to make them more accurate and standardised. Likewise, control practices have also evolved in this direction, notably with continuous sharing of knowledge and good practices with equal pay specialists. These developments are helping to decrease the room for manoeuvre for correct usage of the standard analysis model and thus increase the probability of obtaining results which concur with the company's operational reality.

■ The developments in the legal, social and political context have contributed to greater awareness of the topic of wage equality, to a growing interest on the part of companies to prove their compliance with this provision, as well as to a strengthened political commitment to combat wage discrimination.

3 The tolerance threshold in the literature and other domains

Since it was defined during the pilot phase of equal pay reviews, the tolerance threshold of $\pm 5\%$ has never been adjusted. However, various studies have yielded interesting insights on this topic. On the one hand, several studies which have looked into the analysis method or the equal pay reviews have assessed the relevance of the tolerance threshold. These are reviewed in Section 3.1. On the other hand, several studies have contributed empirical results, based on different data sources, regarding the tolerance

3 The tolerance threshold in the literature and other domains

threshold and a potential adjustment to it. These results are described and analysed in Section 3.2. Finally, the documentation and literature on tolerance thresholds or limit values used in other countries or fields have also been researched and analysed to determine whether conclusions can be drawn for the tolerance threshold in equal pay analyses in Switzerland. This research is summarised in Section 3.3.

3.1 Evaluations of the tolerance threshold

3.1.1 Evaluation of the pilot review phase (2006-2011)

In 2010, National Councillor Marina Carobbio Guscetti submits a motion entitled “Combating wage inequality effectively” (10.3420). This motion petitions the Federal Council **to evaluate the tolerance threshold** of $\pm 5\%$, taking into account the results of the reviews already conducted. In its response of 18 August 2010, the Federal Council proposes to reject the motion as the FOGE is meant to have already submitted a report to the Federal Procurement Commission (FPC) on equal pay reviews having taken place since 2006 and the tolerance threshold is to be re-examined by the FPC and FOGE after this evaluation. Furthermore, the Federal Council indicates that the tolerance threshold was introduced by the FPC and FOGE mainly to **limit the administrative costs to companies** of reviews, which would be much higher if all the factors with an impact on wages within a given company had to be taken into account in the analysis. The Federal Council adds that the FPC and FOGE “have shown themselves to be quite generous in choosing this value in order to give companies time to take the necessary measures to respond to the requirements. This threshold is used only during reviews conducted in the field of public procurement. This makes it possible to keep the burden represented by these reviews within acceptable limits and to avoid companies being wrongly sanctioned. This tolerance threshold is also intended to have a preventive effect. It should in no way enable wage discrimination. The law itself does not envisage any tolerance threshold at all.”¹⁶ Apart from the aspect aiming to limit the administrative costs and workload for companies, the Federal Council thus brings to the fore the **preventive aspect** of the tolerance threshold as well as the fact that it is **relatively generous** to allow companies to take the necessary measures.

As mentioned in the Federal Council’s response, the FOGE initiated the equal pay review procedure in the public procurement system in 2006, with an initial pilot phase of 5 years.¹⁷ This pilot phase was evaluated by INFRAS in 2011 (Trageser et al.). The main aims of this evaluation were to examine how the method was used in practice during the 14 reviews conducted up to that point, to evaluate the results of those reviews and to evaluate the tolerance threshold of $\pm 5\%$ based on the results.

The evaluation of the tolerance threshold by Trageser et al. (2011) is based primarily on the opinions of experts in the field who were interviewed as part of the study. At the time, the majority of the equal pay specialists interviewed for the study felt that the tolerance threshold was still essentially appropriate as, with the information available to them at the time, they felt that the standard analysis model only allowed for an “approximate measure”. Indeed, although the standard analysis model was deemed appropriate for analysing wage equality at the company level, the specialists identified certain limits or potential improvements at that time: with regard to the variables, some felt that in certain cases adding non-discriminatory variables specific to the company being analysed would provide scope to improve the accuracy of the model and of the result. They also felt that a better specification of the variables already included in

¹⁶ Federal Council opinion of 18.08.2010 on motion 10.3420 “Combating wage inequality effectively”. Consulted online on 30.05.2022: <https://www.parlament.ch/fr/ratsbetrieb/suche-curia-vista/geschaefte?AffairId=20103420>

¹⁷ This pilot phase of the review process between 2006 and 2011 should not be confused with the period during which the review tool was tested on 5 companies between 2001 and 2003.

3 The tolerance threshold in the literature and other domains

the analysis would also contribute to this improvement in the model's accuracy, as would using a heteroskedasticity-robust standard error.

The majority of the specialists thus thought that the tolerance threshold of $\pm 5\%$ was appropriate to the model in the form it was in at the time. Furthermore, the existence of a tolerance threshold was perceived as a good thing from a **strategic** perspective, as it allowed companies to see the tool as **fair** and to be **cooperative**. However, several specialists were of the opinion that the tolerance threshold could be **adjusted under certain conditions**. Thus, one specialist believed that it was already possible to lower the tolerance threshold to $\pm 3\%$, as companies had become familiar with the tool in the meantime. In the case of in-depth analyses, this individual even wondered if the tolerance threshold should not be abandoned completely. Other specialists felt that the tolerance threshold could be lowered for in-depth analyses, or vary depending on economic sector, for example.

Moreover, one statistics specialist interviewed as part of this study saw the tolerance threshold as **problematic**. Indeed, according to this person it cannot be ruled out that, among the companies with results below the tolerance threshold, there are some for which the model is already masking the reality very well. In these cases, the tolerance threshold would represent a "**political choice**" and a "**conciliation value**", as it would mean that a statistically proven instance of wage discrimination would be accepted in such cases. One of the main arguments justifying the introduction of a tolerance threshold was thus not relevant for this specialist.

Based on these considerations, Trageser et al. (2011) felt that the tolerance threshold had the following positive side: it detected the **most significant infringements** of the equal pay principle while also substantially limiting the probability of companies being **wrongly accused** of wage discrimination. However, according to this study the tolerance threshold could have allowed certain companies to be **wrongly exonerated** in the context of reviews and it thus needed to be adjusted, as did the standard analysis model.

For this latter reason, Trageser et al. (2011) notably thought that the **tolerance threshold should be lowered** under certain conditions and gave the following justifications for this:

- The **sensitisation of companies** to wage equality should have increased in the meantime;
- The **additional experiences** relating to wage inequality within companies have shown that the tolerance threshold could also be set lower;
- The reviews conducted up to that point had already led to **in-depth analyses**, which made it possible to obtain a more accurate result.

The authors thus proposed lowering the tolerance threshold to 2-3%, **on the condition that the standard analysis model be adapted** with certain possibilities for improvement and that the additional (in-depth) analyses be introduced as standard (better specification of variables depending on case, calculation of additional variants, separate regression analyses, etc.). According to Trageser et al. (2011), lowering the tolerance threshold would result in fewer companies being **wrongly exonerated** of discrimination. However, expanding the model would entail an **increase in workload** relating to the reviews, and the analysis model would lose its "standard" aspect, as the in-depth analyses would depend on the situation of the companies and the specialists' case-by-case assessments.

This study constitutes one of the most detailed evaluations of the tolerance threshold available to date. However, it is based solely on the results of the 14 reviews carried out up to the time of its writing. A great many insights and experiences have been acquired since. In the meantime, however, other studies focusing primarily on the method used to measure wage equality, the choice of variables, or the impact of the equal pay reviews, have approached the topic in a more concise manner.

3 The tolerance threshold in the literature and other domains

3.1.2 Evaluation of federal government equal pay analyses (2015)

One of these, in particular, is the study conducted jointly by the University of St. Gallen's Schweizer Institut für Empirische Wirtschaftsforschung (SEW) and INFRAS (Felfe et al. 2015) in response to the postulate "Equal pay. Improving the relevance of the statistics" (14.3388) submitted by Rudi Noser on 2 June 2014, which notably tasked the Federal Council with critically examining the **method used to analyse instances of wage discrimination** established by the FOGE, while considering in particular the possibility of taking into account **additional variables** for equal pay analyses, such as actual work experience, management experience, continuing education courses, linguistic skills, or work-time percentage during career. The study by Felfe et al. (2015) concluded that the variables included in the existing model had a high to very high explanatory content and did not have the potential to be discriminatory. In contrast, the majority of the other proposed variables did have discriminatory potential, apart from linguistic skills and difficult mental and physical working conditions, provided they were evaluated in a neutral manner, which would entail a very significant workload for the company. For these reasons, **it was not advisable to expand the standard analysis model.**

As regards the analysis model, the equal pay specialists conducting the reviews in the public procurement system viewed it very positively, although they felt that there was still room for certain **improvements**, in particular in the **specification of the variables**, for example by adding extra levels to the "skill level" and "professional position" variables, using education level as a categorical variable (rather than calculating years of education) or adding an analytical evaluation of functions to the analysis. Furthermore, the equal pay specialists felt that the analyses carried out using the standard analysis model provided less statistically significant results in **small enterprises** (between 50 and 100, or even 200 employees) and that this was leading to a stronger presumption of innocence in these enterprises: for small enterprises, a significant result only appeared in the event of a major violation of the equal pay principle, while for large enterprises the result can already be significant when the gender-specific wage difference exceeds the tolerance threshold slightly.

The wage analysis specialists consulted were more critical of the model: they felt, in particular, that the two variables "skill level" and "professional position" did not go far enough to evaluate the requirements associated with the function performed and that the model was thus too focused on the individual characteristics of the people concerned (education level, years of service) and not enough on the functions. According to these specialists, it was thus necessary to include an **analytical evaluation of functions** in the analysis model to resolve this issue.

Furthermore, the statistical method used in the standard analysis model, namely the OLS regression, was considered **scientific** and **appropriate**. However, in certain cases, it could lead to a distortion in the results of the analysis, in particular when the **distribution** of women and men over the explanatory variables was not proportional (e.g. only women at one professional position level). In these cases, the authors recommended conducting sensitivity analyses to examine how this could impact the result.

To take account of the fact that, in certain individual cases, companies have relevant internal variables (in-house scale for evaluating functions, criteria for difficult physical/mental working conditions) to determine wages which are not discriminatory but are not taken into account in the analysis (omitted variable bias), a +/-5% tolerance threshold and a statistical significance test were introduced. According to Felfe et al. (2015), the tolerance threshold and significance test **thus keep the probability of a penalty at a low level** for a large proportion of companies (above all small and medium-sized enterprises). According to this study, the 5% level of the tolerance threshold has **no theoretical, empirical or legal foundation** and, what is more, it is **not possible to establish the appropriate level for the tolerance threshold**

3 The tolerance threshold in the literature and other domains

in a theoretical and empirical manner. To do this, it would be necessary to know, observe and measure all the other objective factors with an impact on wage inequality and be able to estimate the potential for discrimination of each factor. Broadly speaking, Felfe et al. (2015) felt that the tolerance threshold should be adjusted downwards if the analysis model came to incorporate more variables with explanatory potential or if the use of an alternative statistical method resulted in a lower gender-specific wage difference. The interaction between these elements and the level of the tolerance threshold was unclear, however. They thus concluded that it is not possible to **quantify the adjustment to the tolerance threshold** based on changes made to the model or the method and that the tolerance threshold level needs to be set first and foremost based on the **experts' assessments**.

From the **point of view of the companies** consulted during this study, around a third felt that the tolerance threshold was appropriate, another third felt it was too high, and the final third too low. It is interesting to note that more of the companies which had already completed a wage analysis using the Confederation's standard analysis model than of those who had not yet had that experience judged the tolerance threshold to be too high. However, 85% of the companies stated that they would prefer the analysis model to remain relatively simple and the tolerance threshold to remain the same, rather than the analysis model being expanded and the tolerance threshold lowered as a result.

In the opinion of the specialists carrying out equal pay reviews who were interviewed at the time, the +/- 5% tolerance threshold had proved to be **appropriate enough**, despite having been initially set in a relatively **random** way. Certain specialists, however, felt that the tolerance threshold could be lowered in the light of their experience. All of these specialists agreed that the tolerance threshold would need to be lowered if the model came to be expanded to include new variables, in which case some of them even felt that the tolerance threshold should be scrapped completely. The consultants interviewed also felt that the tolerance threshold was appropriate and that it should be lowered if the model was expanded. However, both the controllers and the consultants said they would prefer to keep a relatively simple model with the current tolerance threshold rather than expanding the model and lowering the threshold.

Based on these assessments, Felfe et al. (2015) thus issued the recommendation that the +/-5% tolerance threshold should be **kept unaltered if the model was to remain in the same state**, and adjusted downwards if additional non-discriminatory variables were to be incorporated. However, they felt that the relevance of the tolerance threshold should be **re-verified in line with new experiences gained** as "the present evaluation is based on only a very small number of experiences."

As we have seen, the evaluation by Felfe et al. (2015) highlights on several occasions that any incorporation of new variables into the analysis model should result in a lowering of the tolerance threshold. However, that same study highlighted the scientific relevance of the existing analysis model and the variables included in it, while demonstrating that the majority of the other variables mentioned, in particular in Noser's postulate, were potentially discriminatory or would require too sizeable an effort from companies to implement in a non-discriminatory way. This apparent contradiction can be explained by the still relatively limited experience available at the time with respect to the analysis tool and by the initial idea that the model applied by the tool was relatively "simple" and could thus still evolve. The experiences accumulated subsequently made it possible to refute this idea and to confirm the scientific relevance of the analysis model used, as we will see in the following sections. Furthermore, the omitted variable bias identified as a potential problem is currently the subject of various discussions and work in progress.

3 The tolerance threshold in the literature and other domains

3.1.3 Assessing the impact of regulation – amendment to the GEA (2015)

That same year, a Regulatory Impact Analysis (RIA) was conducted by Stern et al. (2015) in view of the introduction into the GEA of an obligation to carry out an in-house equal pay analysis. The study reiterates the same conclusions as the report by Felfe et al. (2015) on the tolerance threshold of $\pm 5\%$. However, it offers additional interesting information on the **other countries** with experience of measures for implementing equal pay. The majority of countries with these types of measures use simple direct wage comparisons based on gender-neutral function categories. Tools based on a regression analysis are not prescribed by any country. However, all the countries (England, Finland, France, Germany, Luxembourg and Portugal) which provide a non-compulsory tool for verifying wage equality based on a regression analysis use variants developed/adapted from Logib which are implemented in a similar way to Logib, that is with the same variables. In contrast to Switzerland, these tools similar to Logib do not include **any tolerance threshold**.

In its response to the Noser postulate, the Federal Council, returning to the two studies mentioned above, concluded that the statistical method used to measure wage equality at the company level was scientific and legally compliant, that the choice of variables was appropriate and did not require any expansion, and that the tolerance threshold of $\pm 5\%$ **had proved itself and should be retained**.

3.1.4 Assessments from the legal perspective

Other works have looked into the relevance of the model and its variables from the **legal perspective**, while not addressing the question of the tolerance threshold, or only in a very marginal way. Binggeli et al. (2018) in particular reiterate the argument that the tolerance threshold was set to take into account the potential influence of other non-discriminatory, objective factors specific to the company being reviewed. They also return to the conclusion of Felfe et al. (2015) that the tolerance threshold applied in the public procurement system was appropriate despite not having any scientific foundation. Binggeli et al. (2018) clarify, however, that the tolerance threshold “remains dependent on **political decisions** and should be reconsidered if changes are made to the regression type or the variables.”

Legal specialists Pärli and Oberhauser (2019), for their part, made a detailed study from a legal perspective of the potential for discrimination of the explanatory variables used to measure wage equality between women and men and did not directly discuss the question of the tolerance threshold. However, this study includes several interesting elements with regard to the link between **statistical discrimination** and **equality from a legal point of view**. According to the authors, statistics are ambivalent when it comes to discrimination, as on the one hand they provide a foundation for combating discrimination but on the other they can have discriminatory potential when they generalise certain characteristics.

With regard to the tolerance threshold, the contributions from the legal perspective can be summarised as follows: all wage discrimination, irrespective of its scale, is forbidden by the constitution and by federal law. However, the risk exists that an entity may be wrongly “accused” of wage discrimination according to the statistical analysis and thus be wrongly sanctioned or not obtain a public contract (*type I error*). At the same time, the risk also exists that a competing entity may obtain a public contract despite paying discriminatory wages on the grounds of gender, because the statistical analysis arrives at a result which is still below the tolerance threshold, and that another entity with gender-neutral wages will thus fail to obtain the contract and thus be discriminated against (*type II error*).

3 The tolerance threshold in the literature and other domains

3.1.5 Other analyses and publications on Logib module 1

In addition to the various evaluations carried out on behalf of the FOGE presented above, Logib module 1 has undergone regular external evaluations at the national and international level. However, these evaluations do not provide any guidance on the tolerance threshold.

At the national level, we can refer in particular to the recent work by Felder and Wunsch (2021), who conducted a methodological evaluation of the Logib module 1 standard analysis tool by carrying out analyses based on data from the 2018 ESS.¹⁸ According to the authors, the tool has **two main limitations**: first, it does not take into account the **profession exercised** except through the “skill level” and “professional position” variables. From their point of view, this is not enough as there may be significant differences in terms of requirements and skills between functions in the same combination of skill level/professional position. For this reason, they believe that the 6 criteria taken into account for evaluating functions in Logib module 2¹⁹ are better able to encompass the differences between the professions exercised, although they still feel that the optimal solution would be to take the profession exercised into account directly as a variable within the tool. However, it should be mentioned here that, in the opinion of experts, the proposal by Felder and Wunsch to include the profession exercised in the analysis would only allow measuring of equal pay for equal work, and therefore would no longer meet the legal requirement to compare different work of equal value.

Second, Felder and Wunsch feel that the use of Logib module 1 for **companies with fewer than 100 employees** has little relevance as comparability between women and men is not sufficiently guaranteed: there is not always a counterpart of the other gender with the same individual or function-related characteristics, which lowers the statistical significance of the results. The authors feel that the analysis tool should therefore be complemented by a tool for evaluating comparability between women and men within the company. However, the form this additional tool would take is not specified.

Again at the national level, Vaccaro (2015) highlighted that equal pay analyses using Logib module 1 have helped to **reduce cases of unexplained wage inequality** in companies with at least 50 employees.

Furthermore, at the international level, Logib module 1 has been awarded the UN Public Service Award and the **Good Practice** label issued by the Equal Pay International Coalition (OECD, ILO, UN Women). Other countries, such as Germany or Luxembourg in particular, have used Logib module 1 as a base for developing their own tool for analysing wage equality within companies. In Germany’s case, the **Logib-D** tool (see Section 3.3.1) has also undergone several evaluations since its introduction; however, these are less relevant to the present mandate given the differences in context and usage.

¹⁸ The empirical results from Felder & Wunsch (2021) are briefly presented in Section 3.2.2.

¹⁹ The 6 criteria taken into account in Logib module 2 are: 1) Education/training requirements; 2) Requirements regarding ability to work independently; 3) Requirements in terms of specific expertise and methodological skills; 4) Responsibility-related requirements and demands; 5) Psychological and social requirements and demands; 6) Physical requirements and demands.

3 The tolerance threshold in the literature and other domains

Interim conclusion 3

Broadly, both the authors of the various studies and the interviewed specialists are united in concluding that the tolerance threshold was set in a **prudent** and **generous** way in order to avoid unjustified sanctions, with no real theoretical, empirical or legal foundation. Nevertheless, it was deemed **appropriate** for various reasons. On the one hand, it fostered **acceptance of the process by companies** and gave them time to build **awareness of the topic**. On the other, it provided a way to offset the lack of empirical experience regarding the appropriateness of the model in practice and the fact that certain other potential non-discriminatory company-specific variables are not taken into account in the analysis.

According to these various evaluations, however, this threshold would need to be lowered if the model were to be optimised, expanded or to include in-depth analyses. Nevertheless, various studies have demonstrated that it was not appropriate to expand the standard analysis model as the majority of the variables proposed harboured discriminatory potential or entailed too great an effort for companies. Still, some studies provided scope to identify **potential ways to improve** the standard analysis model, in particular with regard to the specification of the variables included, better evaluation of the functions or the minimum number of data sets required for use of Logib module 1. Some of these recommendations have already been implemented as part of the tool's development, while others are currently being considered.

Furthermore, the studies have highlighted the need to **regularly examine** the appropriateness of the tolerance threshold in the light of new experiences within the context of equal pay reviews or analyses. In the absence of a legal or empirical foundation, the threshold should primarily be set based on **assessments by experts** in the field. Nonetheless, this decision remains to a very large extent a political one. For this reason, it is important to recall that the **law does not provide for a tolerance threshold** and prohibits any discrimination on the grounds of gender. Likewise, from an econometric perspective it is not possible to determine the appropriate value for the tolerance threshold, insofar as any regression analysis already includes a **confidence interval** aimed specifically at limiting the risks of reaching a false conclusion. Finally, other countries using similar tools for analysing wage equality do not apply any **tolerance threshold** at all.

3.2 Empirical results concerning the tolerance threshold

In addition to the various results presented in Section 3.1, which are essentially based on the opinions of equal pay specialists or other stakeholders within the field, a number of reports have provided **empirical results** on the tolerance threshold in equal pay analyses based on different types of data and calculations.

3.2.1 Reviews in the public procurement system

Until now, just one study has been based on the **results of reviews** conducted within the public procurement system, and that is the work by Trageser et al. (2011). This study focused solely on the 14 reviews carried out up to that point. In 10 of these reviews, the gender-specific statistical effect differed significantly from zero. Of these 10 reviews, 2 significantly exceeded the +/-5% tolerance threshold, 4 others exceeded the threshold but not significantly, and 4 were below the threshold. However, lessons to be learned from these initial observations remain relatively limited. One of the objectives of this working paper is thus to update these observations taking into account the results of the numerous other reviews conducted since 2011.

3 The tolerance threshold in the literature and other domains

3.2.2 Swiss Earnings Structure Survey

Several studies have sought to provide empirical results on the tolerance threshold based on data from the **Swiss Earnings Structure Survey (ESS)**, which are representative of the company landscape in Switzerland. These analyses are possible because the variables used in the standard analysis model are similar to those used in the ESS, despite certain differences. Particular attention should be paid to the following points when interpreting the results of these studies.

First, the **definition of wages** in the ESS differs from the legal definition used in the analyses carried out using the Logib standard analysis tool. The dependent variable is thus similar, but specified differently. Then, the ESS variable **“skill level of the profession exercised”**, which corresponds to the international ISCO-08 standard and is used with more than 40 categories in the ESS breakdown analysis, essentially needs to be aggregated on 4 levels to allow for analysis with Logib. From the conceptual perspective, this aggregation does not correspond to the **“skill level”** variable which is used in company-level Logib analyses and which allows for a considerably better explanation of wage disparities than the ISCO-08 code depending on the profession exercised. In principle, the ESS is not designed for analyses at the company level, but the national level. Thus, for example, a qualified computer scientist will always be ranked at the highest level of ISCO-08 as an academic profession, whatever the function the person performs within the company (e.g. even if the person is just starting out in their professional career). These elements can thus explain certain significant differences between the results of the analyses based on the ESS and those based on the reviews, as we will see.

The first of the studies based on ESS data was that of Graf and Garibian (2014), commissioned by the FOGE and presenting the result of a certain number of analyses carried out based on data from **ESS 2010**. The authors of this study simulated Logib analyses on a sample of private companies which had responded to ESS 2010 and which had at least 50 employees, of which at least 10 women and 10 men, and which employed less than 80% managers and supervisors. Only 3,000 companies out of the 42,000 included in the ESS met these criteria and could thus be used for the simulations.

Graf and Garibian (2014) were thus able to establish that the tolerance threshold of a 5% gender-specific wage difference to the disadvantage of women is exceeded in **25% of the companies** in the sample and this moves in an almost linear fashion depending on the tolerance threshold set. Thus, this figure is 33% for a 3% threshold, 38% for a 2% threshold, 44% for a 1% threshold and **50% if no tolerance threshold is applied**. Conversely, if the threshold is set at 9%, 12% of the companies still exceed it. As regards discrimination to the disadvantage of men, the proportions are considerably lower. Thus, this figure is 0.5% for a 5% threshold, and 2% when no threshold is applied.

Graf and Garibian (2014) were therefore also able to establish that certain **sectors** discriminate more against women than others. Thus, in the metallurgy, production of electronic equipment, watchmaking and retail trade sectors, women are discriminated against in approximately 40% of companies with a tolerance threshold of 5%. Conversely, they are discriminated against less frequently in the accommodation, legal or accounting activities, or human health sectors. The proportion of women in the sector does not necessarily affect this, as both the retail trade and health sectors have a high proportion of women. Furthermore, the analyses showed that the discrimination coefficient of the expanded regression was still lower than that of the basic regression, which means that women, all other things being equal, hold less well paid posts.

More recently, simulations by Felder and Wunsch (2021) based on data from **ESS 2018** showed that **24% of private sector companies** and **10% of public sector entities** exceeded the tolerance threshold of a

3 The tolerance threshold in the literature and other domains

+/-5% gender-specific wage difference. This study used a process similar to that of Graf and Garibian (2014) but applied a minimum of 60 employees, of which at least 20 women and 20 men.

The most interesting results, however, came from the recently published study by Chávez-Juárez and Graf (2021), commissioned by the BEFH equality body of the canton of Vaud. These authors carried out an empirical analysis based on data from 2,845 employers taken from **ESS 2018** and meeting the criteria of having at least 50 employees with at least 20 women and 20 men. With these results they were able to show that **80.9% of the entities do not significantly exceed the tolerance threshold when set at +/-5%**, 65% do not with a threshold of 2% and **49.8% do not when the threshold is set at 0%**. These figures thus remain comparable with those of Graf and Garibian (2014) although slightly lower. More than 30% of the entities present a gender effect which is significant but does not exceed the tolerance threshold of +/-5%. Additional analyses show that the results vary depending on sector (public vs private), size of employing entity, economic sector and proportion of women and women managers within the entity.

Thus, more than 90% of the **public sector** entities do not exceed the tolerance threshold of +/-5% significantly, while this share is only around 77% in the **private sector**, which is consistent with the results obtained by Felder and Wunsch (2021). However, this difference falls sharply when there is no tolerance threshold.

The large entities, for their part, are less likely to exceed a tolerance threshold of +/-5%, but this trend reverses when the threshold is set at 0%. This can in particular be explained by greater accuracy in the calculations for larger entities, which thus obtain statistically significant results more frequently, even with relatively small effects.

As in the study by Graf and Garibian (2014), Chávez-Juárez and Graf (2021) also noted some significant differences between **sectors**. The sectors where entities most frequently exceed the tolerance threshold, at both +/-5% and 0%, are the manufacturing industry and trade, which is consistent with the results obtained by Graf and Garibian (2014) and Kaiser and Möhr (2021). Conversely, it is the human health and social welfare sector which has the largest proportion of compliant entities and thus does not exceed the tolerance threshold. It is also one of the sectors with the highest proportion of women. Modifications to the tolerance threshold impact all sectors uniformly.

Chávez-Juárez and Graf (2021) also noted that in entities **with more female employees**, and in particular more female managers, it was less common for the tolerance threshold to be exceeded. With regard to the coefficient of determination (R^2), for which the mean lies at 0.65 and the median at 0.69 for the population under study, this does not seem to be a relevant variable for explaining the proportion of entities exceeding or not exceeding the tolerance threshold as the curves are relatively close together.

Based on their analyses, Chávez-Juárez and Graf (2021) surmise that **the use of a tolerance threshold on top of a significance test is not justified from a methodological perspective** as the tolerance threshold and the confidence interval pursue the same objective: to reduce the risk of arriving at a false conclusion (concluding that there is an effect when there is not, or vice versa). The approach of using a tolerance threshold as well as a significance test would be justified if a certain difference was tolerated by the law, which is not the case. Thus, in their opinion, one sole statistical significance test should be sufficient to determine if an entity is compliant with the law or not. Furthermore, the authors feel that more statistical tests should be added concerning the distribution of residuals from the regression in order to ensure that the standard errors are correct (heteroskedasticity-robust standard errors) and potentially to give non-compliant entities the option to provide objective and legally compatible justifications. Finally, Chávez-Juárez and Graf (2021) recommend **removing the tolerance threshold**, with the option for this to be done in stages via interim thresholds such as +/-2% over a transition phase.

3 The tolerance threshold in the literature and other domains

3.2.3 Analyses in the context of a private certification

The report by consulting firm Comp-On (2021) also provides some interesting empirical data, not based on the ESS on this occasion but on the 193 equal pay analyses carried out within the context of the Fair-ON-PAY **private certification** process, which are thus not representative of the economy as a whole. According to these analyses, just 4% of entities significantly exceed the +/-5% tolerance threshold. However, **90% of the companies show differences to the disadvantage of women, with 25% of these displaying borderline results**, i.e. results which come out at 5% or more, but do not exceed the threshold significantly.

Table 4 below presents a summary of the empirical results concerning the tolerance threshold in the literature.

Table 4: Summary of empirical results concerning the tolerance threshold in the literature

	Trageser et al. (2011) Review pilot phase 2006-2011	Graf and Garibian (2014) ESS 2010	Chávez-Juárez and Graf (2021) ESS 2018	Comp-On (2021) Fair-ON-Pay Report
Number of observations	14	3,000	2,845	193
Tolerance threshold of +/-5% exceeded	14.3%	25.5%	19.1%	4.1%
Tolerance threshold of +/-4% exceeded	-	29.7%	23.5%	-
Tolerance threshold of +/-3% exceeded	-	33.9%	29.6%	-
Tolerance threshold of +/-2% exceeded	-	39.0%	35.1%	-
Tolerance threshold of +/-1% exceeded	-	45.2%	42.1%	-
Tolerance threshold of 0% exceeded	71.4%	52.4%	50.2%	48.2%
No gender effect	28.6%	47.6%	49.8%	47.7%

Source: Studies shown in the table, BASS presentation

3.2.4 Monte-Carlo simulations

Finally, the latest empirical results on the tolerance threshold are provided by Kaiser (2022); these are based on so-called Monte-Carlo simulations to determine the power of the different tests performed by the standard analysis model. These simulations notably showed that the **size of the company**, the **gender ratio** and the **size of wage differences** are the three factors that primarily influence the power of the tests. With respect to the tolerance threshold, Kaiser (2022) came to the conclusion that the effect of the existence of this threshold at the +/-5% level is that companies with no unexplained wage difference almost never generate false positive results (0.2% probability of error). He also highlights that if the tolerance threshold were to be lowered to +/-3%, for example, the probability of gender effects being detected grows, while false positive results only increase slightly in the case of a wage effect of zero. This last finding applies primarily to small and medium-sized enterprises (between 0.3% and 1% probability of error for entities with 100 employees, 0.2% or less for those with 250 employees). The risk of a false result (false accusation) would be around 5% if no tolerance threshold is applied.

3 The tolerance threshold in the literature and other domains

Interim conclusion 4

The different studies which have applied the tolerance threshold to data from the ESS have reached similar conclusions: between **20% and 25% of entities significantly exceed the tolerance threshold** if it is set at $\pm 5\%$, with substantial differences observed according to sector (private/public), line of business, size of entity and proportion of women within the entity. These last two elements are also identified as key drivers influencing the power of the tests performed by the Logib tool.

The report by Comp-On, based on analyses conducted in the context of a certification, presents markedly different results, finding that just 4% of entities significantly exceed the tolerance threshold. This can be explained, in particular, by the differences in how certain variables operate (wage, skill level), the heightened awareness of the topic in companies which are keen to obtain this type of certification and are thus not representative of the economy as a whole, as well as by the data validation process, which is performed by an expert and allows certain errors to be avoided. However, their analyses show that **25% of companies display borderline results**, i.e. results which stand at 5% or more, but which do not exceed the threshold by a significant amount.

The studies by Graf and Garibian (2014) and Chávez-Juárez and Graf (2021) show that **lowering the tolerance threshold** leads to a linear evolution in the proportion of entities significantly exceeding the threshold. This figure stands at close to **50% when no threshold is applied**. According to Kaiser (2022), lowering the tolerance threshold would strengthen the probability of detecting gender effects, while the probability of obtaining false positive results would increase only very slightly (from 0.2% to a maximum of 1% probability of error for a tolerance threshold of $\pm 3\%$). Finally, Chávez-Juárez and Graf (2021) find that, from an econometric perspective, it is not justifiable to have a tolerance threshold on top of a confidence interval, as these two elements have the same objective: to avoid erroneous conclusions that an effect exists when in reality it does not. They would thus recommend **removing the tolerance threshold** while integrating certain **additional statistical tests** to verify the robustness of the model and the results, in particular with regard to the standard errors.

In addition to empirical assessments and results concerning the tolerance threshold in equal pay analyses, the documentation and literature dealing with other thresholds or limit values have also been analysed. The results of this analysis are presented in the next section.

3.3 Tolerance thresholds and limit values in other countries and fields

3.3.1 Equal pay analyses in neighbouring countries

Table 5 below presents an overview of the practices and obligations with respect to equal pay analyses in Switzerland's neighbouring countries: Austria, France, Germany and Italy. The research was based on information contained in the report on wage transparency legislation by the International Labour Organization (2022) and this was supplemented by information provided on the government websites of the different countries.

■ **Germany:** the Transparency in Wage Structures Act, which entered into force in 2017, aims to put into action the principle of "equal pay for equal work or work of equal value". This law is based on 3 main founding principles: first, employees working in a public or private entity with more than 200 employees have the right to ask their employer for information concerning their pay, the criteria and practices applied when setting remuneration, and how an equal-value or equivalent task carried out by workers of the opposite sex in the same establishment is remunerated. Employers must provide this information only if there are at least six comparable employees of the opposite sex, in order to guarantee anonymity. The

3 The tolerance threshold in the literature and other domains

second point concerns entities with more than 500 employees who are required to regularly review their wage structures in order to ensure compliance with wage equality. The entities are free to choose their evaluation method and the government provides the Logib-D tool in Excel format, which is based on Switzerland's **Logib module 1** tool and which thus uses the same regression method with a few differences in how the variables operate. However, Logib-D does **not include a tolerance threshold** and only has one significance test, the aim of which is to determine if the gender-specific wage gap differs from zero in a statistically significant manner. Employees must be informed of the results of the analysis, and if there is a gender-specific wage difference the entity must take action to eliminate it. Furthermore, entities with more than 500 employees which are required to complete an annual activity report must also provide a separate report on equality between women and men and equal pay once every 3 or 5 years. These reports are listed on a federal website and accessible to all.

■ **Austria:** the Equal Treatment Act envisages that public and private sector entities with more than 150 employees will provide a report on the distribution of wages between women and men in the different job classifications or categories established in the collective agreements (which encompass 90% of employees in Austria) or the entity's sector. This report, however, is confidential and there is no obligation for corrective action in the event that a wage gap is found or for sanctions if the report is not drawn up. Employee representative committees may however request information on the report and pass certain information on to the employees. The wage comparison is thus conducted at the level of the job categories defined by the collective labour agreements or internal company classifications which group together jobs of equal value. The mean and median wages between women and men serve as the basis and there is no regression analysis or tolerance threshold. The comparison is thus one of **gross wage differences** within groups of equivalent functions which are defined by the criteria of each collective agreement or internally by the company.

■ **France:** private companies with at least 50 employees and all public entities are required to publish a **professional equality index** every year which contains several indicators including the average wage gap between women and men by age group and equivalent position category. For each of these indicators, a number of points to achieve is set in accordance with various criteria. The points for each indicator are then added up to arrive at an overall professional equality index. If the overall index does not attain a certain minimum score, the company must define progress objectives or even corrective measures which will allow it to attain this score in future.

With regard to **wage differences**, employees are grouped by four age brackets and by equivalent position categories. To define the equivalent position categories, employers may focus on hierarchical level, sector classifications, socio-professional categories (4 categories in total: labourers, employees, technicians and supervisors, and engineers and managers), or on other methods defined with the works council. Only categories containing at least three men and three women are taken into account. The mean wage of women and men is then calculated for each of the groups, as is the wage gap between women and men.²⁰ The differentials at the level of the groups are then added up, taking into account the proportion of employees in each group, in order to obtain the overall wage gap at the company level. The maximum score for the equal pay indicator (40 points) is achieved only **in the event of an overall gap equal to 0%**. The points awarded decrease as the gap widens: a score of 0 is reached when the wage gap is greater than 20%. This is also a comparison of **gross wage differences** filtered by age group and equivalent position category.

²⁰ In the groups made up by socio-professional category, there is a relevance threshold for differentials of 5%. This threshold is set at 2% in the groups made up by hierarchical level. These thresholds are deducted from the differential in the case of a positive differential, and added in the case of a negative differential with a floor of zero.

3 The tolerance threshold in the literature and other domains

■ **Italy:** since the end of 2021, public and private entities with more than 50 employees are required to draw up a report accounting for wage gaps between women and men as well as the terms of employment of their staff. The report must be sent to the employee representative committee and to the gender adviser as well as to various state institutions. Furthermore, companies completing the report and fulfilling certain equality criteria may request a gender equality certificate. Notably, one of the criteria is equal pay for work of equal value. Certified companies benefit from a reduction in their social contributions. This certification system is currently in development, with the aim that it will enter into force in December 2022. We thus do not have any more information on how wage equality is calculated in the context of this certification.

Table 5: Equal pay analyses in neighbouring countries

Country	Obligations	Calculation method	Requirements relative to result	Consequences
Germany	- Entities with 200 or more employees: information on pay practices to employees on request - Entities with 500 or more employees: regular equal pay analyses and report on wage equality every 3 or 5 years	- Free choice of evaluation method - Government provides Logib-D, which is based on Logib module 1 (regression analysis) with certain variables implemented differently and one single significance test to verify if the difference is statistically significant	- For Logib-D: gender-specific wage difference significantly different from zero = difference to be corrected	- Reports published and accessible to all - Employees must be informed - If gender-specific difference exists: measures to take to eliminate it
Austria	- Entities with more than 150 employees: report on the distribution of wages between women and men	- Comparison of mean and median wages between women and men within position categories or in-house classifications for jobs of equal value	- No requirement	- No consequence: confidential report and no requirement for corrective measures
France	- Private companies with 50 or more employees and all public entities: publication of a professional equality index every year containing several indicators including wage equality	- Comparison of average wages between women and men by age group and equivalent job category - Scores added together by group to arrive at an overall wage gap figure	- Overall wage gap: maximum score (40 points) achieved with a gap of 0%. Points deducted as gap grows: 0 points for any gap larger than 20% - Relevance threshold of 5% for comparisons by socio-professional category and 2% for comparisons by hierarchical level or coefficient	- If the index (including all the indicators) is lower than 85 points (out of 100), the company is required to set and publish progress objectives for each indicator - If the index is lower than 75 points, the company must publish corrective and remedial measures
Italy	- Public and private entities with more than 50 employees: report accounting for wage gaps between women and men and terms of employment	To be defined	To be defined	- Companies fulfilling equality criteria (including wage equality) benefit from a reduction in their social contributions

Source: cf. information in the text, research and presentation by BASS

3.3.2 Other tolerance thresholds and limit values

In **Table 6** we present a selection of limit values and tolerance thresholds used in equal pay analyses in other European countries and other fields which fall within the public authorities' execution mandate and which could allow conclusions to be drawn in the discussion of the tolerance threshold for equal pay reviews.

Equal pay analyses

■ In the EU Parliament talks are currently under way on introducing measures for greater **pay transparency in the EU member states** (EP, 2020, 2022b). The introduction of binding measures for greater pay

3 The tolerance threshold in the literature and other domains

transparency is being justified by the argument that despite legal anchoring, equal pay for work of equal value has not yet been achieved, and gender-specific wage discrimination is no longer accepted in society (EP, 2022a). Under a proposal by the relevant Committee, in future all companies with 250 or more employees which have a gender-specific wage difference²¹ of at least 5% for a group of employees performing the same work or work of equal value and for which the wage difference is not explained by objective and gender-neutral factors will be obliged to take appropriate action (COM, 2021).²² The representatives support this course of action, but are calling for a more wide-ranging scope of application as they feel the Committee's proposal is not ambitious enough. The Parliament wishes to oblige all companies with 50 or more employees and a gender-specific wage difference of 2.5% or more to take action. The representatives are also calling for the percentage of the permissible gender-specific wage difference to be lowered to zero eight years after the directive comes into effect, and not to allow an unexplained wage difference of less than 2.5% to serve as proof that the employer is complying with their equal pay obligation (Peter-Hansen & Rafaela, 2022).

■ The **Fair Pay Innovation Lab** (fpi) should also be mentioned here; this organisation conducts private certifications of companies with the aim of lowering the unexplained gender-specific wage difference to +/-1% and the total gender-specific wage difference to +/-10% (fpi, 2022).

Working conditions

■ When reviewing compliance with **pay conditions usual in specific locations, professions and industries** in the context of accompanying measures, cantons use different methods to determine usual pay and several cantons have established malpractice thresholds. However, these are generally not fixed, but are intended to be reviewed on a case-by-case basis and adjusted if necessary. The cantons with no malpractice threshold decide on compliance with the customary wage level on a case-by-case basis (Seco, 2016a, b; Merckx, 2016).

■ The TLV (Threshold Limit Values) value is an example of **limit values in the workplace**. It specifies the maximum permissible concentration of a substance in the form of a gas, vapour or particulates in (breathing) air in the workplace (Suva, 2019; Koller et al., 2013). TLVs are not definite boundaries between hazardous and non-hazardous areas. The limit values are established based on scientific findings and evaluations relating to feasibility (practicability within businesses). The weighting of the factors can vary from one international committee to another, which explains differences in the official limit values between countries. Because TLVs are based on current knowledge, over time adjustments are made or the respective values are lowered as long as implementation remains feasible. When reviews of compliance with the limit values are conducted, no fixed penalty thresholds are applied; rather, each case is judged individually. If limit values are exceeded, improvement measures are defined in consultation with the companies (e.g. installation of a new ventilation system) and a further review may be carried out at a later date (Koller, personal communication of 3.8.2022).

Environmental protection

■ Limit values are also common in the environmental protection field. The example of **noise protection** shows that scientific, social and technical developments can result in a new definition of limit values being imposed in order to fulfil the constitutional obligation. The assessment method used and the exposure

²¹ The indicator for the gender-specific wage difference measures the difference between the average gross hourly earnings of male and female employees as a percentage of the average gross hourly earnings of male employees (Eurostat, 2022).

²² The reason given for setting the threshold at 5% is that it is the same as the standard statistical significance level of 5% used in scientific circles (EMPL Secretariat, personal communication of 19.5.2022). However, the significance level only says something about the certainty of a specific result and cannot justify the level of the threshold in terms of either content or method. Also, the 5% threshold in the EU is intended to be calculated without a significance test.

3 The tolerance threshold in the literature and other domains

limits have increasingly been called into question as these are based on outdated scientific foundations. The Federal Noise Abatement Commission (FNAC) has developed a recommendation for adjusting the limit values which should satisfy the requirements of the Environmental Protection Act (EPA) (EKLb, 2021a). The recommendations (EKLb, 2021b) are based on current scientific knowledge with regard to the health and economic impacts of noise pollution and take into account the years of experience with enforcement of the Noise Abatement Ordinance (NAO) and case law in the field of environmental noise. Relevant to the envisaged lowering of the limit values is reclassifying quietness from a luxury to a (health) necessity. Moreover, pivotal to adjustments in the assessment methods were improved measurement and calculation techniques, changes in the activity patterns of the population (differentiation of day/night) and the character of vehicles. With regard to street noise, reviews of compliance with the regulations are based to a great extent on model calculations and rarely on measurements taken on the ground. Where limit values are exceeded, measures are considered taking into account their economic and political feasibility (Baumann, personal communication of 4.8.2022).

■ The energy consumption and other appliance properties of a range of **electrical appliances** must be declared via the **energy label** (BFE, 2020). Based on preliminary studies, the spectrum of appliance properties (energy consumption and other parameters) is ascertained and an energy efficiency index calculated. The allocation of the index values to the scales of the energy efficiency categories (labels) does not follow a mathematical formula; rather, it is intended to reflect market developments as accurately as possible and has a normative component. Thus, when re-scalings are carried out (every 5-10 years), care is taken to ensure that no currently available appliance is in the best category. During reviews a so-called verification tolerance applies, which is derived from the measurement inaccuracy from the preliminary studies and usually ranges between 5 and 10% depending on the appliance (Stadler, personal communication of 4.8.2022). In 2020 and 2021, 8 (12%) of the total of 65 reviews of the manufacturers' declarations found that they did not satisfy the requirements according to the measurements taken (BFE, 2021, 2022). Because the labelling scale has since become unclear (from A+++ to D) and technological developments have brought all products into the best energy efficiency classes and the labels are thus not very informative, at present the original A-G scale is gradually being reintroduced, but at a higher standard (e.g. EnergieSchweiz, 2021, 2020). The energy consumption of **new cars** also needs to be declared via a label. In order to set the limits for the 7 energy efficiency categories (A to G), all vehicle types on offer are ranked in ascending order and divided into sevenths according to their consumption (primary energy gasoline equivalents). The category limits are recalculated once a year on a specific reference date (Rosser et al., 2021).

■ Also falling under environmental protection are reviews of the **timber declaration**. These have a zero-tolerance threshold; in other words, a breach occurs as soon as a product fails to meet the declaration standards. Depending on the focus of the reviews (e.g. small businesses, furniture retailers or online trade) in the last three years 16-30% of the reviewed businesses had correctly declared all products, and 32-50% had not declared any product correctly (BFK, 2022, 2021, 2020).²³ When a breach is detected the businesses must rectify it and assume the costs of the reviews. However, as no fine has yet been issued since the reviews started 10 years ago, and the implementation thus did not achieve the desired effect, recently as a result of political pressure an adjustment to the legal process has been made (Häne, 2022; Gisin, personal communication of 28.4.22).

Other areas

■ For reviews of compliance with the permissible **maximum speeds in road traffic**, different safety margins apply due to different levels of accuracy in the measuring devices used. In the Swiss Federal

²³ In 2019, 2020 and 2021 between 108 and 130 companies were reviewed (BFK, 2022, 2021, 2020).

3 The tolerance threshold in the literature and other domains

Council's response of 8.9.2004 to the motion by Mörgeli (04.3336) "Road traffic. Tolerance value of 5 instead of 3 kilometres per hour" the Council indicates that the measuring tolerance is not to be understood as a number of kilometres per hour which an individual may drive above the permitted speed limit, but is intended to take into account device- and measurement-related uncertainties. In the wake of technical improvements to radar devices and for reasons of consistent handling, an adjustment to the safety margins for laser and radar devices was discussed (Stellungnahme BR, 2004).

■ The **social insurance field** contains a large number of thresholds or limits which have established themselves without any clear criteria in terms of content or method, and in some cases have been adjusted over time (e.g. lowering of welfare for asylum seekers, lowering of allowable assets for supplementary benefits, raising of the income limit for entitlement to premium reduction in Canton Lucerne). In the discussions on the areas mentioned above, often aspects like public finances, avoidance of supposed disincentives, and definitions of rightful claimants are relevant (e.g. SFH, 2022; AHV/IV, 2020; Piazza, 2019). Other "boundary shifts" follow developments in the cost of living (e.g. raising of maximum rent values and taking into account the regional differences in rent burden for supplementary benefits; cf. AHV/IV, 2020).

The legal perspective

The legal experts consulted were asked, from a legal perspective, to what extent they saw potential comparisons between the areas listed in **Table 6** and the tolerance threshold in equal pay analyses. The only unconditional connecting factors they saw were in the equal pay analyses at the European level. Some limited parallels were also seen with the reviews of the wage level in the context of accompanying measures. One person indicated that the secondments field might also be relevant to the tolerance threshold discussion. Companies that post their employees to Switzerland from abroad must pay the same wage that would be paid to employees based in Switzerland. In cases of direct wage comparisons, zero tolerance applies; a deviation even of only a few francs can result in substantial penalties. However, all the legal experts consulted noted that an in-depth analysis would be needed for them to be able to make more specific statements on potential conclusions for the tolerance threshold in Logib module 1.

3 The tolerance threshold in the literature and other domains

Table 6: Limit values and thresholds for penalties in other areas of the public execution mandate

Area	Context	Level of limit value / penalty threshold	Argument definition / changes
EU pay transparency	The principle of equal pay for men and women performing equal work or work of equal value is enshrined in Article 157 TFEU. However, equal pay has not been achieved, and consequently there are plans to introduce binding measures and establish a threshold for gender-specific wage differences.	Committee proposal: 5% Parliament proposal: 2.5%	Normative (Parliament)
Usual wage level	When reviewing compliance with pay conditions usual in specific locations, professions and industries in the context of accompanying measures, cantons use different methods to determine usual pay and several cantons have established malpractice thresholds. However, these are generally not fixed, but are intended to be reviewed on a case-by-case basis and adjusted if necessary.	If present: as a rule for less than 10% of the usual wage	unknown
Threshold limit values in the workplace	The Threshold Limit Values (TLV) specifies the maximum permissible concentration of a substance in the form of a gas, vapour or particulates in (breathing) air in the workplace. Because MAK values are based on current knowledge, over time adjustments are made or the respective values are lowered. The Federal Council delegated the issuance of directives on limit values in the workplace pursuant to Art. 50 para. 3 APO (Ordinance on the Prevention of Accidents and Occupational Diseases) to Suva.	Varies according to substance (and country)	Striking a balance between potential injuries to health and the risks and costs in manufacturing
Speed limit for road traffic	For road traffic speed checks, as per the FEDRO regulation on these checks a safety margin is deducted from the measured speed.	Depending on the permitted speed limit and measuring device, between 3 and 9 km/h	Measurement accuracy Consistent handling
Noise protection	Pollution limits are enshrined in the Noise Abatement Ordinance (NAO) and based on the Environmental Protection Act (EPA). Currently an adjustment (lowering) of the limit values for road, rail and aircraft traffic noise is being evaluated, as these values are based on outdated scientific foundations and no longer satisfy the requirements of the Environmental Protection Act (EPA).	Depending on region, time (day/night) and limit value type (planning, immission, alarm value) limit values between 40 and 75 dB(A) apply.	Scientific foundations (health and economic consequences of noise) Experience with implementation International guidelines Social change Technological development
Energy labels for household appliances	Since 1 January 2002, energy consumption and other appliance characteristics for a variety of household appliances must be declared using the energy label. The Energy Efficiency Ordinance (EnEO) envisages reviews by the Swiss Federal Office of Energy (SFOE). As the scale has become unclear (from A+++ to D) and technological developments have brought all products into the best energy efficiency classes, the original scale of A to G is being gradually reintroduced, but at a higher standard.	Varies according to appliance	Normative Technological development International guidelines
Timber declaration	The Ordinance on the Declaration for Timber and Timber Products (SR 944.021) has been in force in Switzerland since 2012. This ordinance draws on the Consumer Information Act (SR 944.0) and regulates duties to declare in respect of timber species and place of origin. At the same time as the new articles in the Environmental Protection Act (EPA, 35e-35h), the new Timber Trade Ordinance (TTO) came into effect at the beginning of 2022. As no fine has yet been issued since the reviews started 10 years ago, as a result of political pressure the legal process has been adjusted (now in the event of a breach proceedings are always instigated by the legal service).	Zero tolerance, breach present when one product is not declared correctly	unknown
Social insurance	The social insurance field contains a large number of thresholds or limits which in some cases have been adjusted over time (e.g. lowering of welfare for asylum seekers, lowering of allowable assets for supplementary benefits, raising of the income limit for entitlement to premium reduction in Canton Lucerne).	Varies according to area and region	Public finances Avoidance of disincentives Definition of rightful claimants

Source: cf. details in the text; research and presentation by BASS

4 The tolerance threshold as applied in state reviews

Interim conclusion 5

In Switzerland's neighbour countries, the majority of equal pay analyses are based on **unrefined comparisons of pay** between women and men in equivalent employment categories. The lessons to be drawn from them are thus relatively limited. However, Germany provides Logib-D for equal pay analyses; this tool is based on Logib module 1 with variables which in some instances are implemented differently. Logib-D does **not use a tolerance threshold** and only operates a single significance test to determine if the gender-specific wage gap is significantly different from zero.

Furthermore, discussions are under way at the European Parliament level to compel companies to put in place measures from a certain threshold difference in pay between women and men performing work of equal value, and in cases where the wage difference cannot be justified by objective and gender-neutral factors. The Parliament is looking to set this threshold at **2.5%** for all companies with 50 or more employees.

To sum up, an examination of further reviews carried out by the public authorities shows that due to technical, scientific and social developments as well as experiences with a specific review process, it may be necessary to adjust threshold values, thresholds for penalties, tolerance ranges or the review procedure in order to be able to continue to fulfil the execution mandate. Having said that, direct conclusions from a methodological perspective seem limited, as no area could be identified which is also based on regression analyses (cf. also Infrac, 2011).

4 The tolerance threshold as applied in state reviews

In this section we present the results relating to the use of the tolerance threshold in practice, namely, in the reviews within the public procurement and/or subsidies system at the federal, cantonal and municipal levels. First, we present and analyse the empirical results from the reviews conducted by the authorities at the different levels in Section 4.1. Second, Section 4.2 presents the experiences and assessments of the supervisory authorities with regard to the tolerance threshold, but also regarding developments in practice, and in particular in the quality and completeness of the data.

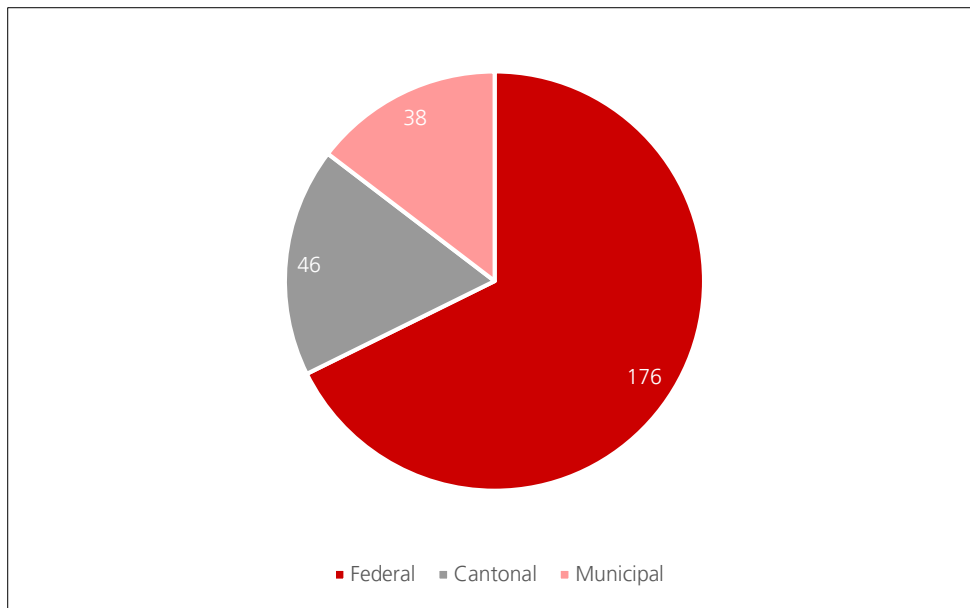
4.1 Results of equal pay reviews

Since the standard analysis model was developed and the first reviews were initiated by the FOG in 2006, other authorities have followed suit and also initiated review procedures within their own public procurement and/or subsidies systems. This section details the results of the reviews conducted at these different levels.

4.1.1 Descriptive data and statistics

We were able to gather the data from 260 reviews carried out by the various authorities, of which the majority were executed at the federal level. **Figure 4** shows the number of reviews conducted by the authorities at the different levels according to data provided in June 2022.

Figure 4: Number of reviews carried out at different levels



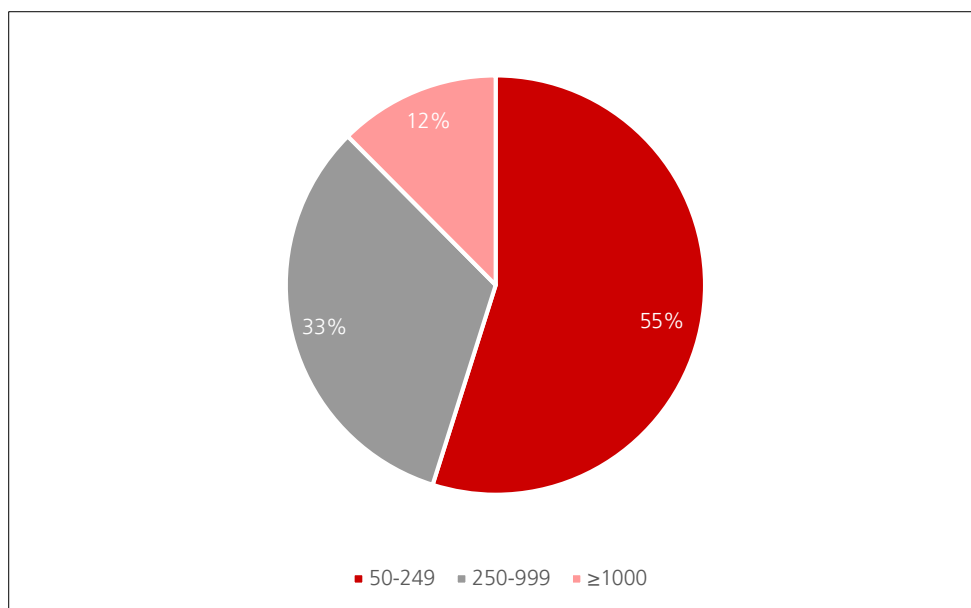
Source: Data from reviews carried out by the authorities (n=260), BASS calculations

It should be clarified here that these reviews were carried out using Logib module 1 only. Furthermore, the data from 21 reviews conducted by one authority could not be obtained for reasons relating to the protection of data specific to the authority in question. In addition, for certain reviews the authorities were unable to supply all of the data requested, in particular with regard to the line of business of the company or the standard error. Thus, some evaluations had to be conducted on a smaller number of cases, which explains the different numbers of observations (n) mentioned within the analyses.

Figure 5 shows the proportion of entities reviewed according to their size. More than half of the reviewed companies had between 50 and 249 employees, one third had between 250 and 999, and 12% had 1,000 or more. The mean size of the reviewed entities was 528 employees, while the median was 204.

4 The tolerance threshold as applied in state reviews

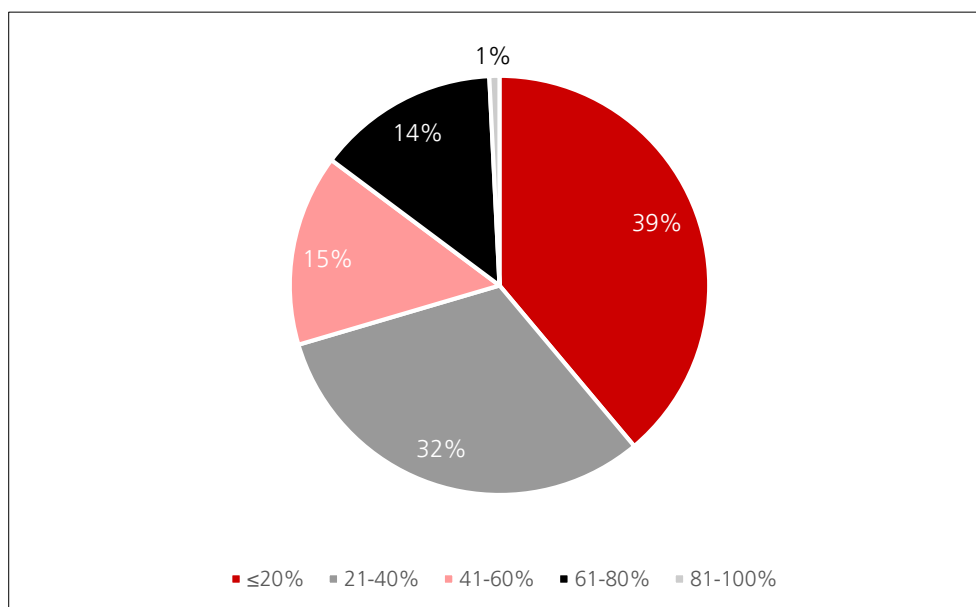
Figure 5: Proportion of reviewed entities according to number of employees



Source: Data from reviews carried out by the authorities (n=260, thereof 3 without information), BASS calculations

These entities employed a total of 137,342 persons, of which almost 30% were women. The mean average of women in the entities themselves was 31%, while the median was 24%. **Figure 6** shows the proportion of women in the different entities reviewed. In more than 70% of cases, the entity has 40% women or less. Only 2 entities had more than 80% women among their employees, while 87 entities had 80% or more men.

Figure 6: Proportion of women in the entities reviewed

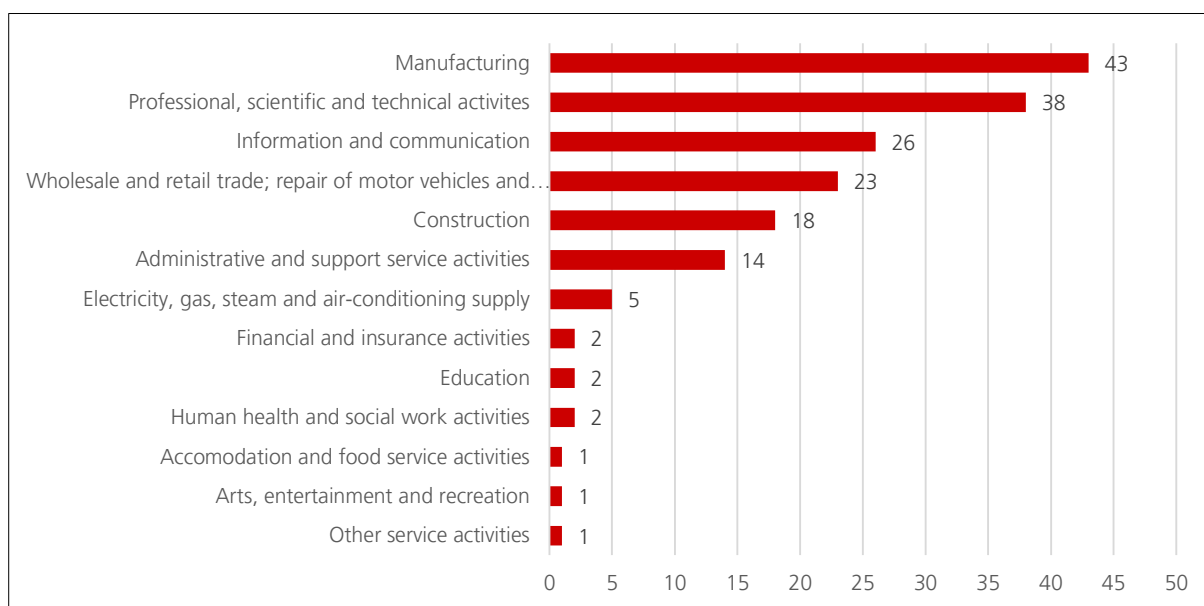


Source: Data from reviews carried out by the authorities (n=257, thereof 3 without information), BASS calculations

4 The tolerance threshold as applied in state reviews

The number of entities reviewed varies quite significantly between the different economic sectors. **Figure 7** below shows the distribution of entities according to economic sector²⁴. On the one hand, these substantial differences can be explained by the very nature of public procurement contracts, which involve certain sectors more than others. On the other, the selection at federal level is also carried out on the basis of risk-weighted samples, and this factor could explain the high proportion of entities from the manufacturing industry or wholesale and retail trade, repair of motor vehicles and motorcycles, which display the most substantial differences in the national statistics (see Section 3.2).

Figure 7: Proportion of reviewed entities according to economic sector



Source: Data from reviews carried out at the federal level (n=176), BASS calculations

4.1.2 Results of reviews and unexplained gender-specific differences

Of the 260 reviews conducted, 226 display an unexplained wage difference to the disadvantage of women, or 86.9% of the cases. Among the reviews showing a result to the disadvantage of women, the unexplained gender-specific wage difference averages -4.9%. Among the 34 reviews showing a result to the disadvantage of men, the unexplained gender-specific wage difference averages 1.9%. **The average unexplained gender-specific wage difference for the whole set of reviews is -4.0%.**

Table 7 below shows the average unexplained gender-specific wage differences in terms of various criteria: supervisory authorities, entity size, sector, proportion of women, R^2 and year of review. This notably allows us to establish that there is a significant gap between the average unexplained gender-specific wage difference in the reviews conducted at the federal level (-4.6%) and those conducted at the cantonal (-2.7%) or municipal level (-2.9%). In particular, this could be explained by the process for selecting entities for review, which is based on risk-weighted samples at the federal level and more on random draw for the other authorities. Another explanatory factor is the fact that the majority of the cantonal and municipal authorities introduced their review procedure only recently. As we shall see, **the reviews conducted recently show unexplained gender-specific wage differences which are not as high on average.**

²⁴ The data concerning line of business could only be obtained for the reviews carried out at federal level.

4 The tolerance threshold as applied in state reviews

Table 7: Average unexplained wage difference according to different criteria

	Average unexplained wage difference
Authorities	
Federal	-4.6%
Cantonal	-2.7%
Municipal	-2.9%
Size	
50-249 employees	-4.1%
250-999 employees	-3.9%
≥1,000 employees	-4.1%
Sector	
Manufacturing industry	-5.7%
Construction	-6.0%
Wholesale and retail trade; repair of motor vehicles and motorcycles	-7.1%
Information and communication	-4.1%
Professional, scientific and technical activities	-3.1%
Administrative and support service activities	-1.6%
Proportion of women	
≤20%	-4.5%
21-40%	-4.2%
41-60%	-4.2%
61-80%	-2.3%
R²	
0.61-0.70	-5.5%
0.71-0.80	-4.2%
0.81-0.9	-4.0%
0.91-1	-3.5%
Year	
2006-2010 (n=15)	-4.9%
2011-2015 (n=63)	-5.0%
2016-2019 (n=161)	-3.8%
2020-2021 (n=21)	-2.3%
Total	-4.0%

Source: Data from completed reviews (n=260), only categories with more than 5 observations are shown, BASS calculations

The average unexplained wage difference varies considerably between sectors. Values range from -1.6% in administrative and support service activities, to -6.0% in construction and -7.1% in wholesale and retail trade, repair of motor vehicles and motorcycles.

Entity size does not seem to have a big impact on the average unexplained wage difference. In contrast, the proportion of women in the entity does play a role: the higher the proportion of women, the more the average wage difference falls. This same effect is observed for the coefficient of determination (R^2).²⁵ The higher this becomes, the more the average unexplained wage difference falls. Relatively significant gaps are also found depending on the group of years being observed. For our analyses, we established four groups of years on the following bases: 2006-2010 represents the pilot phase of the review procedure, as studied by Trageser et al. (2011); 2011-2015 represents the period before job titles and detailed instructions for coding were incorporated; 2016-2019 is the period before the web tool was launched;

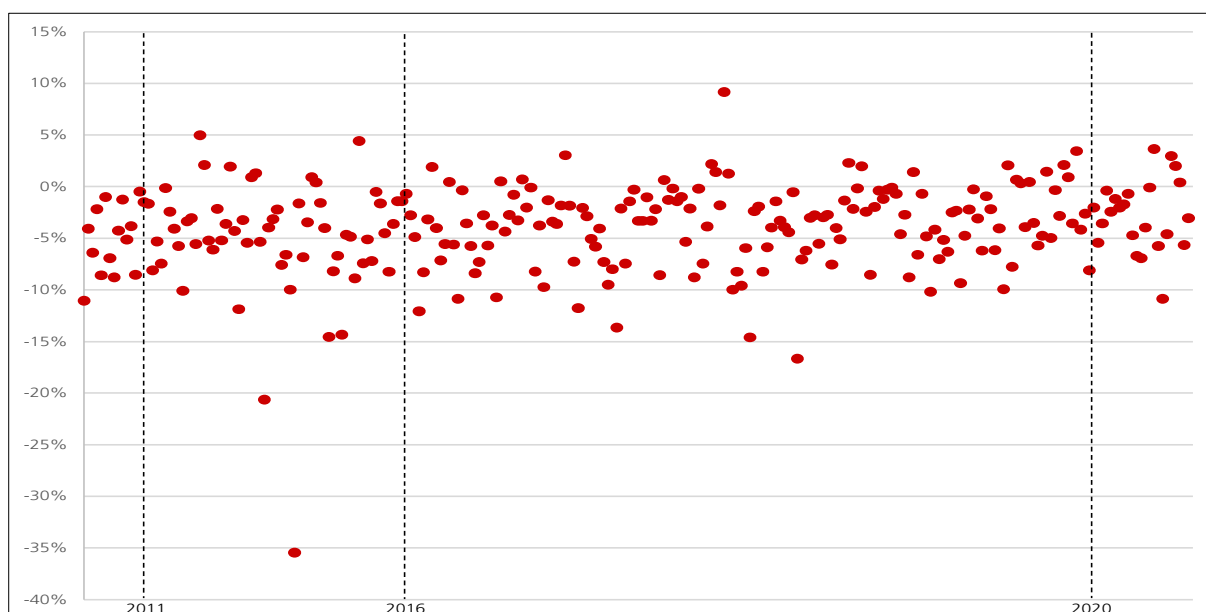
²⁵ This can accept values between 0 and 1 (respectively 0% and 100%) and indicates the share of variance in the wages (dependent variable) which can be explained by the statistical model of the Logib module 1 standard analysis tool. For example, a value of 0.8 indicates that 80% of the wage variation is explained by the variables included in the analysis model and that this therefore has a high explanatory power. Conversely, a value of 0.3, for example, indicates that the model has a low explanatory power.

4 The tolerance threshold as applied in state reviews

and 2020-2021 is reviews conducted with the web tool. We observe that the reviews conducted up until 2015 display an average unexplained wage difference of around -5%. This subsequently fell between 2016 and 2019 (-3.8%), and again in 2020-2021 with an average unexplained wage difference of -2.3%. This can notably be explained by the various developments in the context (in particular greater awareness) and in the tool mentioned earlier.

Figure 8 below shows the development over time of the distribution of unexplained wage differences. The negative values represent unexplained wage differences to the disadvantage of women, while the positive values represent unexplained wage differences to the disadvantage of men. The horizontal axis is divided according to the four periods defined above. The unexplained wage differences to the disadvantage of women are considerably more numerous and higher on average than those to the disadvantage of men. Thus, numerous cases display an unexplained wage difference which goes beyond 5% to the disadvantage of women, two of which showed differences of 20% and more (max. 35%), while there is only one case displaying an unexplained wage difference to the disadvantage of men in excess of 5% (max. 9%). We can also note that there seem to be fewer extreme values as time goes on, in particular in the last few years, during which the differences all lay between +5% and -11%. However, this illustration does not allow an indication to be given of whether the tolerance threshold has been exceeded or not, due to the significance tests implemented by the analysis tool to verify this.

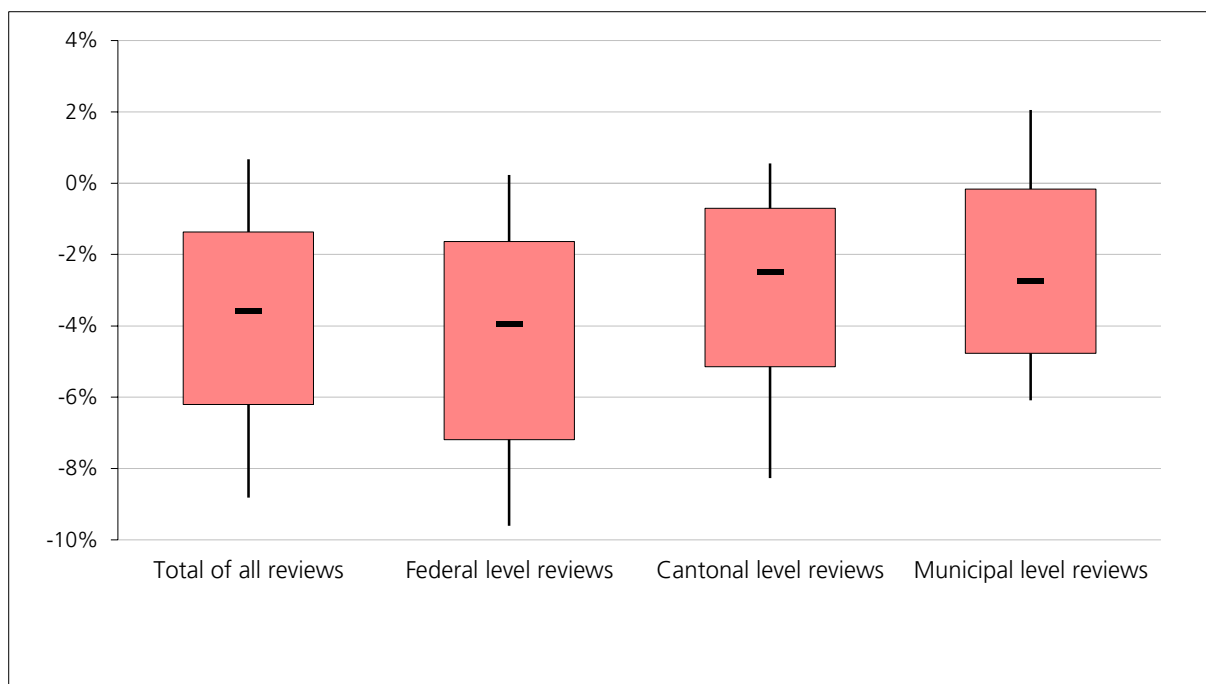
Figure 8: Distribution of unexplained wage differences



Source: Data from completed reviews (n=260), BASS calculations

Figure 9 shows the distribution of unexplained wage differences for the different levels of oversight (federal, cantonal, municipal) in the form of boxplots. The black line in the middle of the boxes represents the median; that is, the value which half of the cases lie above and the other half below. The boxes are delimited by the lower and upper quartiles. This means that half of all values are within the box and 25% each are below or above it. The vertical black lines represent the range between the first and ninth deciles (10% and 90%). We can observe here that half of the reviews display an unexplained wage difference to the disadvantage of women ranging between -1.4% and -6.2%, with a median at around -3.5%. The unexplained wage differences are slightly less significant in the reviews carried out at the cantonal and municipal levels, mainly because these are more recent.

Figure 9: Distribution of unexplained wage differences according to the different levels of oversight

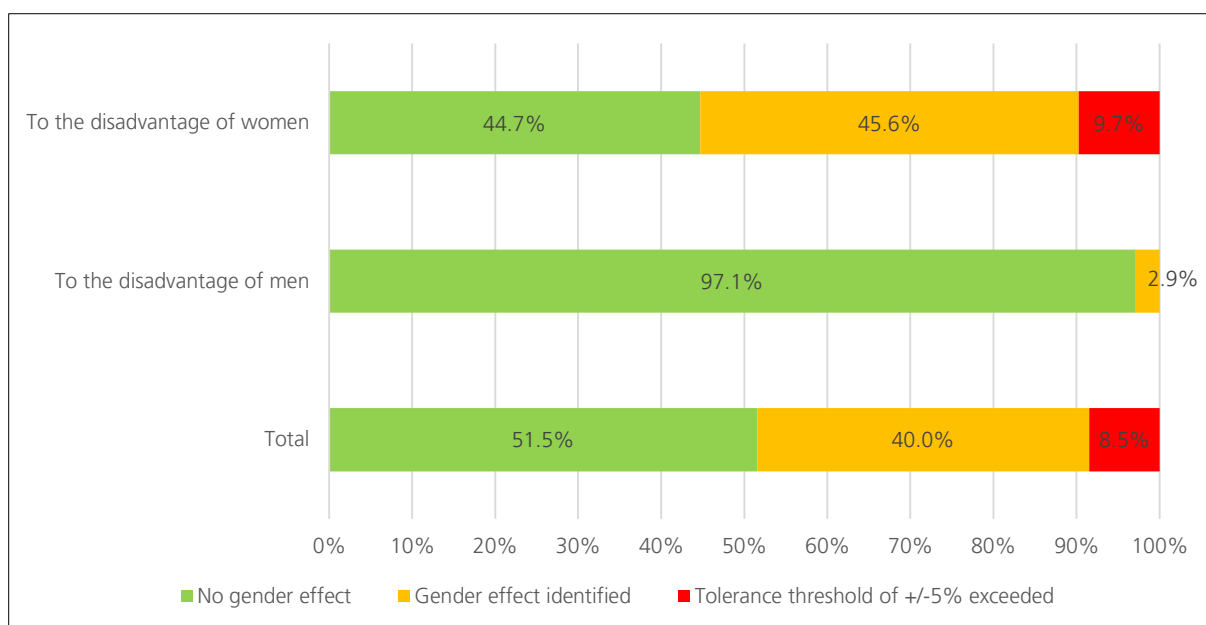


Source: Data from completed reviews (n=260), BASS calculations

Figure 10 below presents the results of the reviews grouped into three levels: “no gender effect”, “gender effect” and “tolerance threshold of +/-5% exceeded”. We can observe here that, among the reviews displaying a wage difference to the disadvantage of women, 44.7% do not display a gender effect, while 9.7% of the cases significantly exceed the tolerance threshold of +/-5%. In the reviews displaying an unexplained wage difference to the disadvantage of men, 97.1% do not display a gender effect and none of the cases significantly exceed the tolerance threshold. **Overall, 91.5% of the reviewed entities do not exceed the tolerance threshold with the standard analysis model when the threshold is set at +/-5%, and 8.5% of entities exceed it significantly.**

4 The tolerance threshold as applied in state reviews

Figure 10: Review results according to the three levels



Source: Data from reviews carried out by the authorities (n=260), BASS calculations

We can also clarify here that in the 21 reviews for which we were unable to obtain data due to reasons of protecting data specific to the authority in question (see Section 4.1.1), the tolerance threshold of $\pm 5\%$ was never exceeded in a statistically significant manner. If we incorporate these 21 reviews into the calculations, the share of entities exceeding the tolerance threshold in a statistically significant manner falls to 7.8%. This share is clearly below those calculated by the various studies based on the ESS (Graf & Garibian, 2014; Felder & Wunsch, 2021; Chávez-Juárez & Graf, 2021). This could be explained by several factors:

- The sample: the reviewed entities are all bidders for public procurement contracts and are thus not representative of Swiss companies as a whole, in contrast to the ESS data. Moreover, for some years now entities bidding for public procurement contracts have had to complete a declaration of compliance with wage equality, and it is therefore probable that they are more sensitised to the topic than the average company.
- The data validation process inherent in every review, which allows certain errors or outliers to be eliminated, to obtain better-quality data and to code functions more accurately and in a way more suited to each entity for the “skill level” and “professional position” variables.
- The differences already mentioned relating to how the wage variable operates and the “skill level” variable.

However, the share of entities significantly exceeding the $\pm 5\%$ tolerance threshold is slightly higher than that presented in the report by Comp-On (2021), which is based on equal pay analyses carried out as part of a certification process.

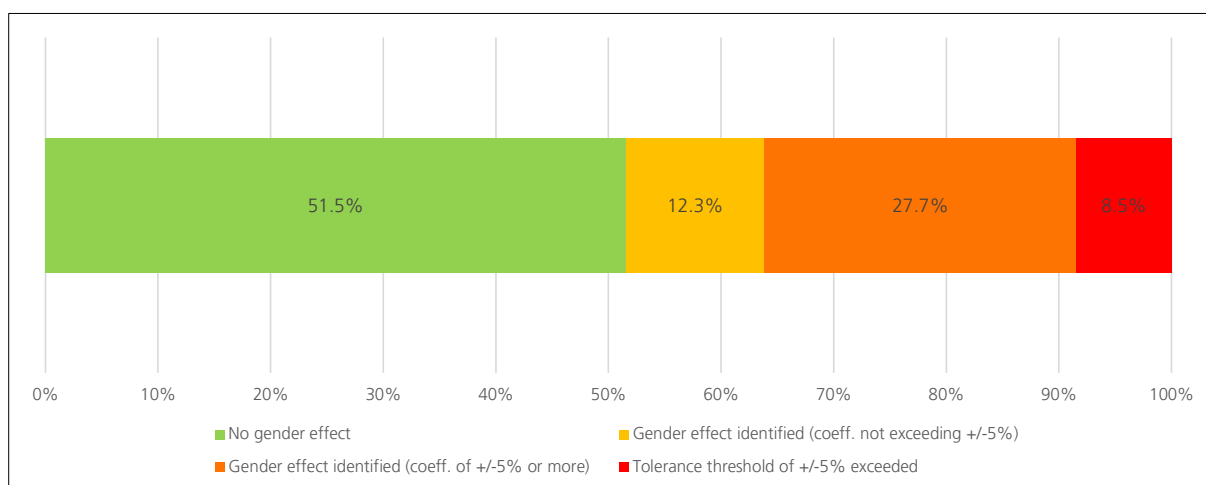
Nevertheless, of the 104 reviews revealing a gender effect²⁶ which does not significantly exceed the tolerance threshold (40% of cases), 72 have a discrimination coefficient of 5% or higher. That means more than 69% of the cases in the “orange” zone show discrimination coefficients of 5% or higher. This proportion is very high and considerably higher than that reported by Comp-On (46% of borderline cases

²⁶ Gender effect which differs from zero in a statistically significant manner.

4 The tolerance threshold as applied in state reviews

among “orange” zone cases). **Figure 11** below shows the distribution of review results according to these four levels.

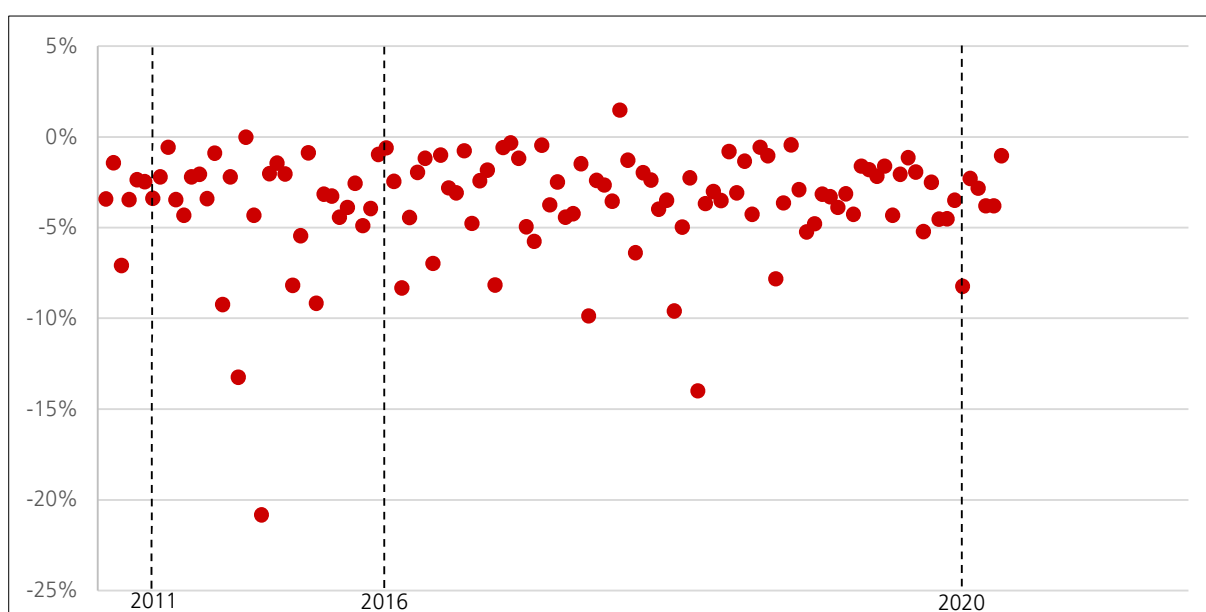
Figure 11: Review results according to four levels (discrimination coefficient: $\pm 5\%$)



Source: Data from reviews carried out by the authorities (n=260), BASS calculations

Figure 12 below shows the evolution over time of the distribution of the minimum significant discrimination coefficient; that is, the value at the lower limit of the 95% confidence interval, for reviews which revealed a verified gender effect. If the value represented in this illustration is higher/lower than $\pm 5\%$, the tolerance threshold has been significantly exceeded. We can observe here that just one completed review shows a verified gender effect to the disadvantage of men with a minimum significant discrimination coefficient of around 1.5%. The cases to the disadvantage of women are clearly more numerous. Furthermore, there are 19 cases which exceed the 5% tolerance threshold in a statistically significant manner to the disadvantage of women, with minimum discrimination coefficient values of up to 20% against women. However, we observe a positive trend in recent years, with fewer extreme results and fewer cases significantly exceeding the $\pm 5\%$ threshold.

Figure 12: Distribution of minimum values of the gender coefficient (lower limit of the confidence interval)



Source: Data from completed reviews with information on standard error, results with a verified gender effect (n=117), BASS calculations

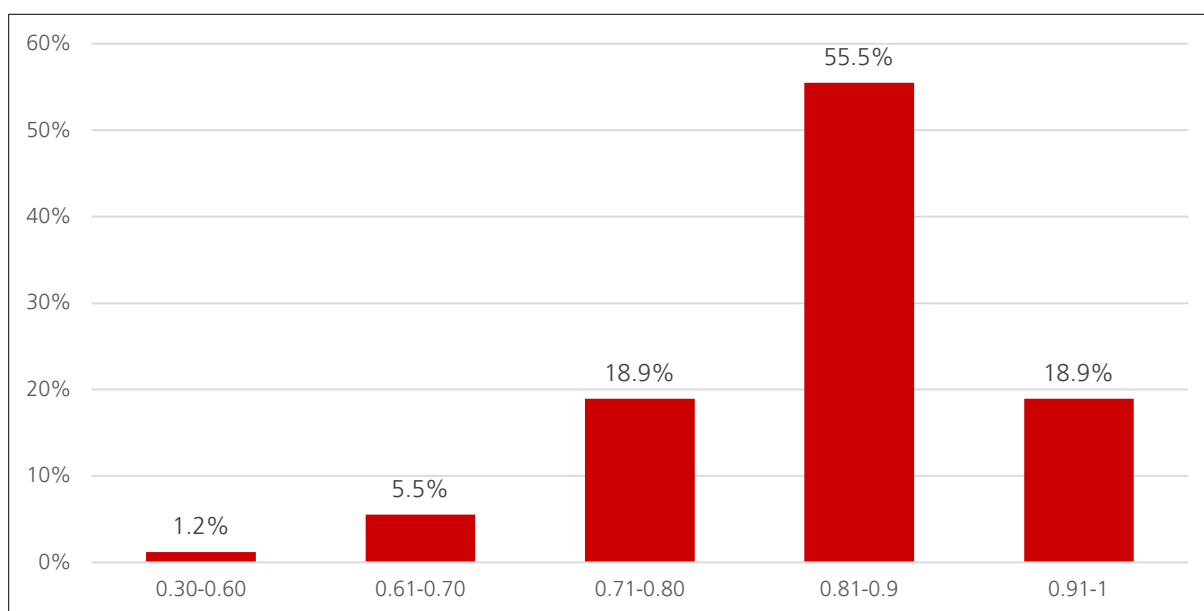
4 The tolerance threshold as applied in state reviews

Among all the reviews conducted, the average coefficient of determination (R^2) is 83%, which is a great deal higher than in the analyses conducted by Chávez-Juárez and Graf (2021). This can be explained partly by the data validation process and the coding of the functions for the “skill level” and “professional position” variables, which can be adjusted to the operational reality of the entity during the review, while the “skill level” variable is coded automatically in the ESS according to the job performed (see Section 3.2.2).

Figure 13 below shows the distribution of R^2 in the different reviews conducted. In the majority of cases, R^2 lies between 0.81 and 0.9. In a little less than 20% of cases, it was between 0.71 and 0.8 or 0.91 and 1. In 5.5% of cases, it was between 0.61 and 0.7, while it very rarely fell below 0.60 (1.2% of cases)²⁷. The explanatory power of the model is thus very high in the majority of the cases reviewed.

In order to carry out the analyses based on years, we returned to the groups formed previously. First, it is interesting to highlight that the average R^2 during these different phases remains stable and moves between 83% and 84%, with the exception of the 2020-2021 phase, where the average R^2 goes above 84% for the first time. The explanatory power of the model has thus remained relatively stable over the years in spite of the various clarifications and changes made to the method. This can be explained by the fact that, since its introduction, the review procedure has guaranteed a quality standard for the data thanks to the validation process performed by an expert and by the supervisory authority.

Figure 13: Coefficient of determination (R^2) in the reviews conducted



Source: Data from reviews carried out by the authorities (n=260, thereof 6 without information), BASS calculations

4.1.3 Adjustments to the tolerance threshold

The scope of this assignment also included conducting analyses on the impact of changing the tolerance threshold on review results. We conducted these analyses using the following criteria: entity size, sector, proportion of women, R^2 , and year of review. For some of these reviews, the data on the standard error could not be supplied and the analyses were thus conducted on a more limited number of observations

²⁷ In 3 cases from the years 2011, 2014 (federal level) and 2018 (cantonal level).

4 The tolerance threshold as applied in state reviews

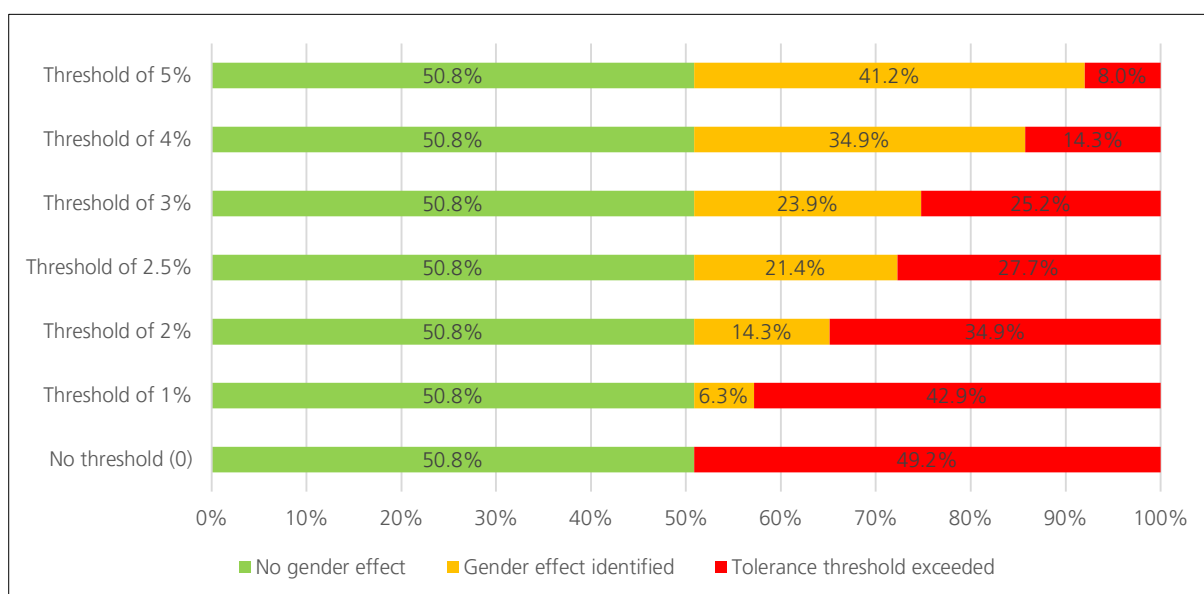
than the various descriptive statistics above. This also explains why there may be differences in the proportions of companies exceeding or not exceeding the tolerance threshold, as the calculation basis is not the same.

In order to carry out these analyses, we needed to calculate the degrees of freedom for each entity in relation to the number of employees in the analysis and the number of variables in the model. With the degrees of freedom, we were then able to define the critical t-value for each review. Then, based on the discrimination coefficient and the standard error²⁸, we were able to carry out t-tests for the different tolerance thresholds we wanted to study. We used the following thresholds for our analyses: 5%, 4%, 3%, 2.5%, 2%, 1% and 0.

Figure 14 below illustrates the impact of adjusting the tolerance threshold on the results of all the reviews. With regard to the reviewed entities for which we obtained information on the standard error, 50.8% do not reveal a gender effect and do not exceed the tolerance threshold whatever the threshold applied.

If a threshold of 0% had been applied, this would have implied that 49.2% of the reviewed entities would have produced a result which differed significantly from zero and thus lay above the tolerance threshold. This percentage decreases in an almost linear fashion in relation to the tolerance threshold applied: 42.9% with a $\pm 1\%$ threshold, 27.7% with a $\pm 2.5\%$ threshold, 14.3% with a $\pm 4\%$ threshold, and finally 8% with a $\pm 5\%$ threshold. Conversely, the share of entities showing a gender effect but one which is not above the tolerance threshold would have fallen as the threshold fell: this share would have been 41.2% with a tolerance threshold of $\pm 5\%$, 21.4% with a threshold of $\pm 2.5\%$, and 6.3% with a threshold of $\pm 1\%$.

Figure 14: Impact of adjusting the tolerance threshold on review results



Source: Data from reviews carried out by the authorities with information on standard error (n=238), BASS calculations

²⁸ As emphasised by Chávez-Juárez and Graf (2021), it is possible that certain standard errors supplied by the standard analysis model are erroneous if the condition of the homoskedasticity of the residuals is not satisfied. Robust standard errors should thus be calculated in certain cases, which has not been done within the present assignment. However, taking into account robust standard errors has only a small impact on the results of Chávez-Juárez and Graf (2021): these would be erroneous for 1.9% of entities with a tolerance threshold of $\pm 5\%$, and 4.8% of entities if no tolerance threshold were applied. Although the impact is limited, this aspect should be considered when interpreting the results presented below as well as in discussions relating to the tolerance threshold.

4 The tolerance threshold as applied in state reviews

Furthermore, we carried out supplementary analyses to ascertain the impact of adjusting the tolerance threshold on review results if these did not incorporate a significance test. This would be in line with the proposals contained in the European Union's wage transparency directive. **Table 8** offers an overview of the number and proportion of entities exceeding the tolerance threshold with and without a significance test. The proportions of entities exceeding the tolerance threshold when the tool does not use a significance test are markedly higher than those presented above. Thus, more than 35% of the reviewed entities have an unexplained wage difference of more than 5% and would exceed this tolerance threshold. This proportion would be around 60% with a 3% threshold, 74% with a 2% threshold, and 86% with a 1% threshold. Furthermore, all the reviewed entities show a result which is different from zero, as this value is very difficult to achieve with a tool based on a regression analysis.

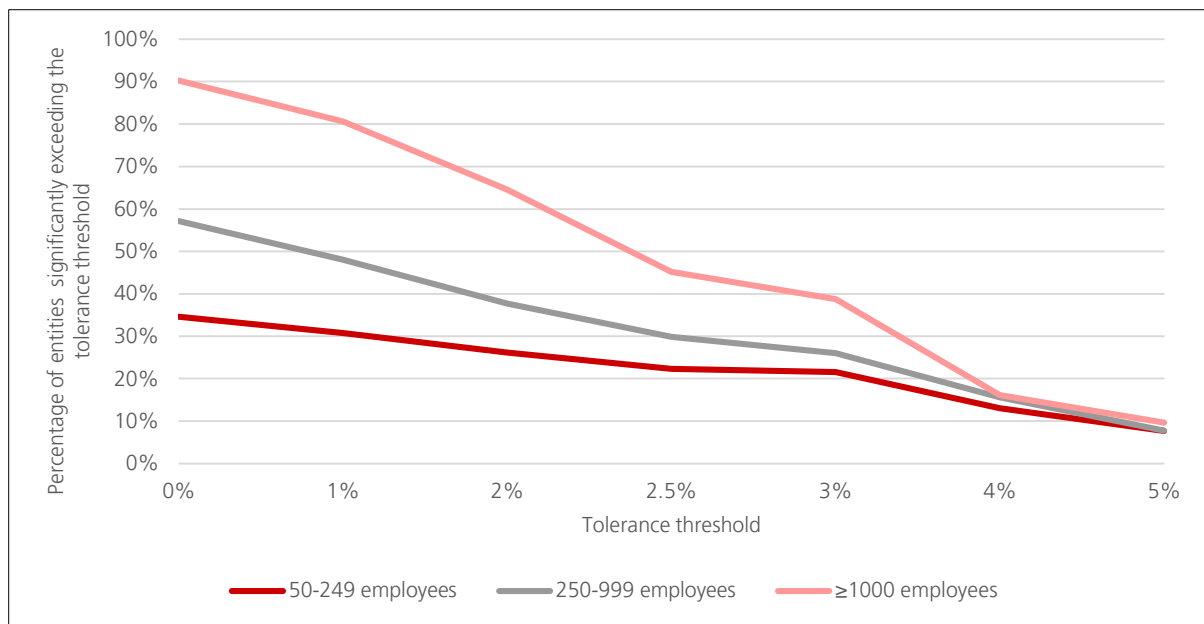
Table 8: Proportion of entities exceeding the tolerance threshold with and without a significance test

Tolerance threshold	With significance test		Without significance test	
	Number	As a % of the total	Number	As a % of the total
+/-5% threshold	19	8.0%	85	35.7%
+/-4% threshold	34	14.3%	108	45.4%
+/-3% threshold	60	25.2%	143	60.1%
+/-2.5% threshold	66	27.7%	154	64.7%
+/-2% threshold	83	34.9%	176	73.9%
+/-1% threshold	102	42.9%	205	86.1%
No threshold (0)	117	49.2%	238	100%

Source: Data from reviews carried out by the authorities with information on standard error (n=238), BASS calculations

Figure 15 below illustrates the impact of adjusting the tolerance threshold on the results of reviews according to entity size. For thresholds of +/-4% and +/-5%, all size categories show a similar proportion of entities significantly exceeding the tolerance threshold. However, it is interesting to note that adjusting the tolerance threshold downwards would above all affect the large entities: with a threshold of +/-2.5%, 22.3% of entities with 50 to 249 employees would have significantly exceeded the tolerance threshold. This share would have been 29.9% for entities with 250 to 999 employees and 45.2% for entities with 1,000 employees or more. The difference would have been greater if a 0% threshold had been applied: in this case, 90.3% of companies with 1,000 or more employees would have significantly exceeded the tolerance threshold, while this share would have been 57.1% for entities with 250 to 999 employees, and 34.6% for those with 50 to 249 employees. However, if we look at the average discrimination coefficient, we observe that it is -4.17% in the entities with 50 to 249 employees, and -3.95% in the entities with 1,000 employees or more. This can be explained by the fact that the calculations are more accurate for the large entities, irrespective of the tolerance threshold: even a relatively small effect can be considered significant, while the smaller entities can more easily benefit from the imprecision of the calculations and thus find themselves below the tolerance threshold even with a larger unexplained wage difference. This is consistent with the conclusions drawn by Kaiser (2022) and Chávez-Juárez and Graf (2021).

Figure 15: Impact of adjusting the tolerance threshold on review results according to entity size

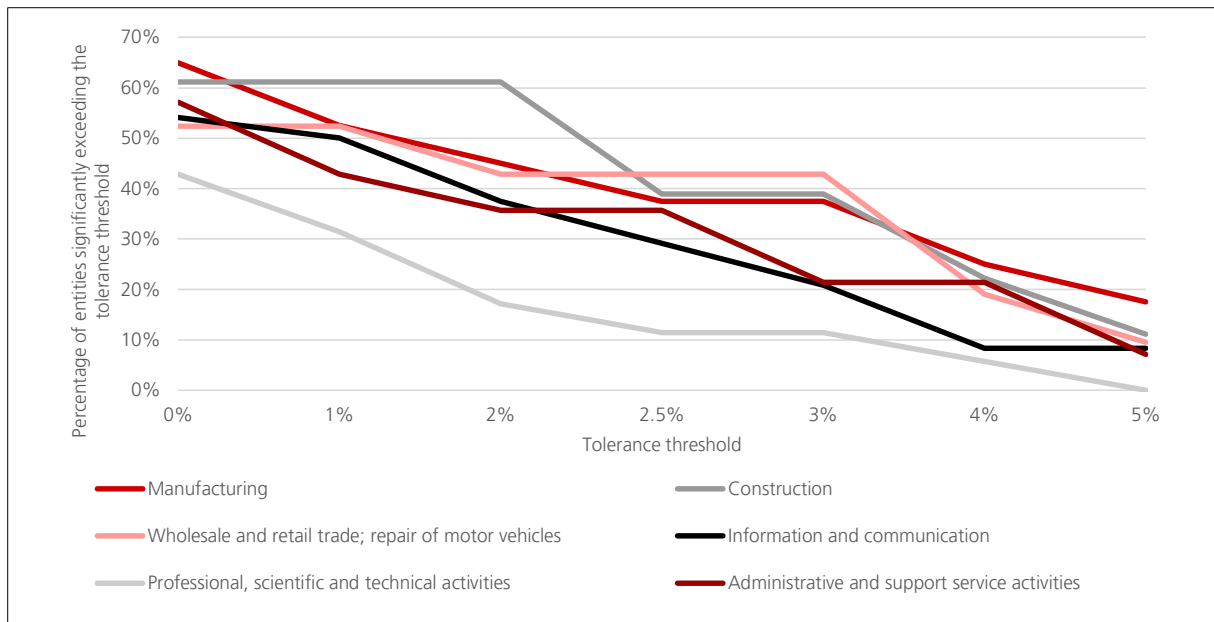


Source: Data from reviews carried out by the authorities with information on standard error (n=238), BASS calculations

Line of business seems to be a less relevant criterion for evaluating the impact of a potential adjustment to the tolerance threshold. **Figure 16** below shows the impact of an adjustment in relation to the line of business of the reviewed entities. We have only represented the sectors with more than 10 reviewed entities here in order to have results with a degree of significance. The professional, scientific and technical activities sector remains the sector with the lowest proportion of entities not significantly exceeding the tolerance threshold, at whatever level it is set. Conversely, the manufacturing industry and construction sectors are those with the highest proportions of entities exceeding the tolerance threshold when it is set at +/-5% or 0%. For the remaining sectors, this varies with the threshold set without affecting the general trends.

4 The tolerance threshold as applied in state reviews

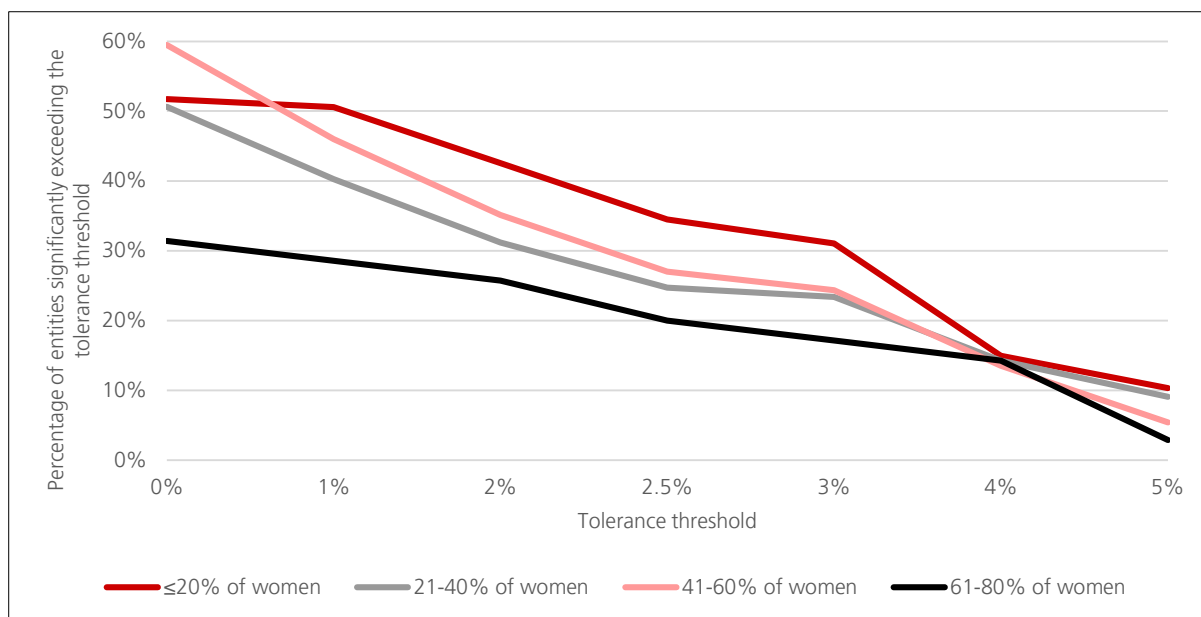
Figure 16: Impact of adjusting the tolerance threshold on review results by sector



Source: Data from reviews conducted at the federal level with information on standard error, only sectors with more than 10 observations (n=152), BASS calculations

Figure 17 shows the impact of adjusting the threshold on results in relation to the proportion of women in the entity. It should be clarified here that only two entities employed more than 80% women. These did not display a gender effect and are not represented on this graph. We can see that the entities employing between 61% and 80% women are those showing the lowest proportion of entities exceeding the tolerance threshold, at whatever level it is set, with the exception of +/-4%, where they show a similar proportion to the other categories. It is when a 0% threshold is applied that the difference is greatest: 31.4% of the entities with 61% to 80% female employees exceed this threshold, while the proportions vary between 50% and 60% for the other categories.

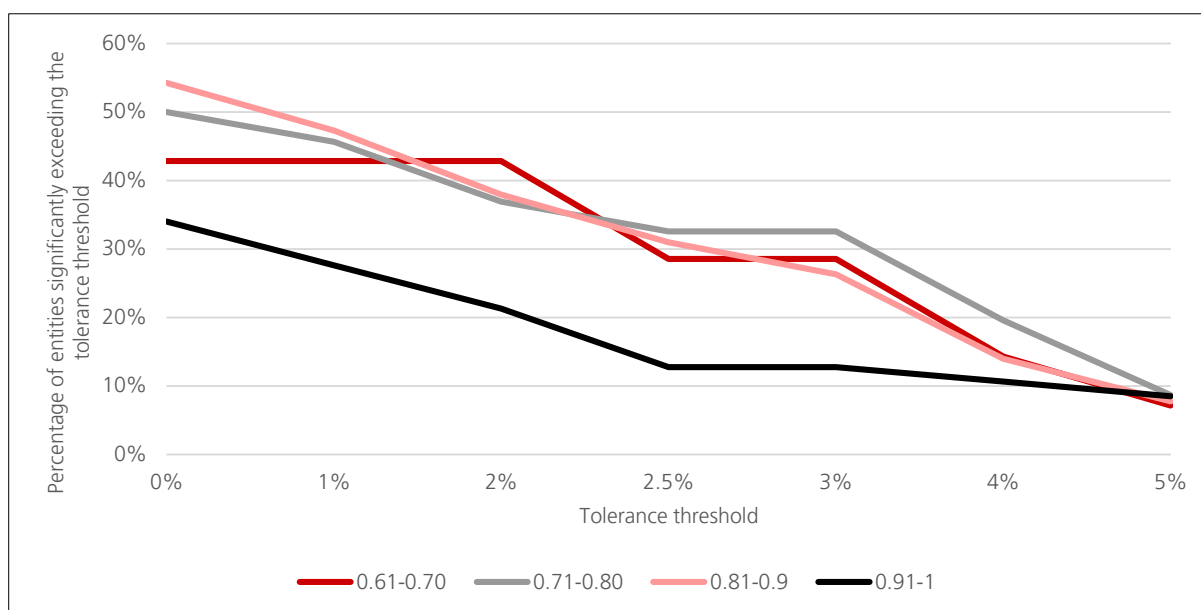
Figure 17: Impact of adjusting the tolerance threshold on review results according to proportion of women



Source: Data from reviews carried out by the authorities with information on standard error (n=238, excluding the 2 cases with more than 80% women), BASS calculations

Figure 18 shows the impact of adjusting the tolerance threshold in relation to the value of R^2 for the reviews conducted. The two reviews with an R^2 lower than 60% are not shown here. The entities with an R^2 higher than 90% have the lowest proportions of entities significantly exceeding the tolerance threshold for all values, with the exception of the +/-5% threshold, for which all categories show similar proportions. With regard to the other categories, they show results which vary depending on the threshold value but remain relatively close together, except in the case of the 0% threshold.

Figure 18: Impact of adjusting the tolerance threshold on review results according to R^2

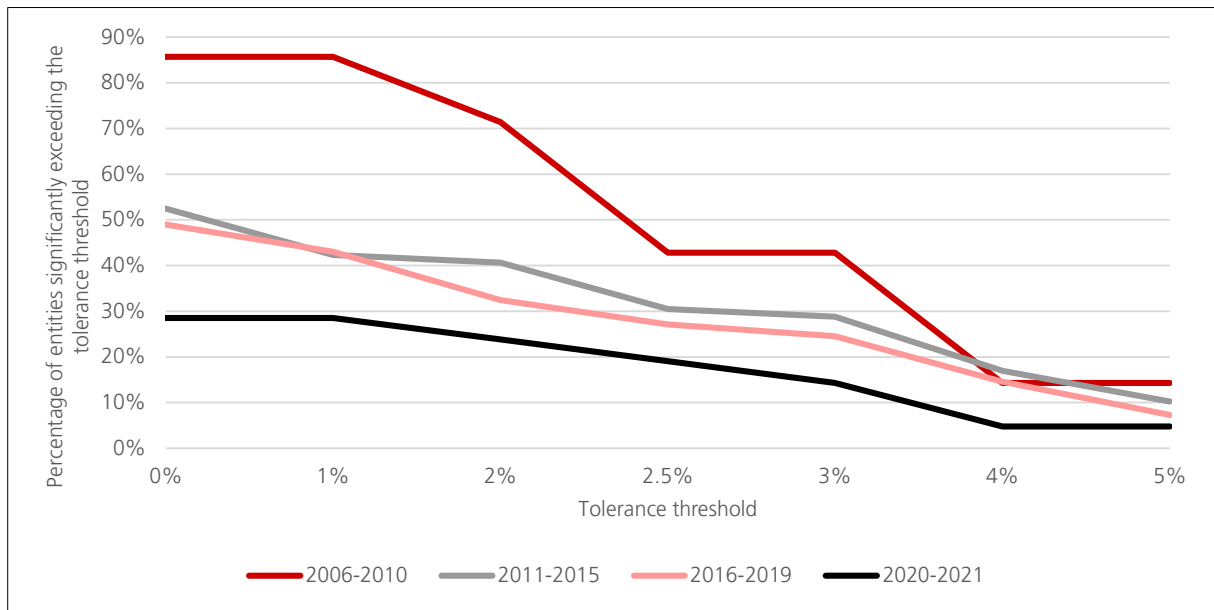


Source: Data from reviews carried out by the authorities with information on standard error (n=238, excluding the 2 cases with R^2 of less than 0.6), BASS calculations

4 The tolerance threshold as applied in state reviews

Figure 19 illustrates the impact of adjusting the tolerance threshold on results in relation to the year of review. We note that the reviews conducted since Logib was made available as a web-based tool display distinctly lower proportions of entities significantly exceeding the tolerance threshold, and this holds for all threshold values. For their part, the reviews conducted in 2011-2015 and 2016-2019 show similar proportions, while those conducted between 2006 and 2010 show clearly higher proportions for tolerance thresholds set between 0% and +/-3%. It should be clarified here that this group is made up of just 7 entities and that the results should thus be treated with caution. Overall, we can speculate that **the entities being reviewed are becoming more and more sensitised to the topic of wage equality** for the reasons cited above.

Figure 19: Impact of adjusting the tolerance threshold on review results according to year of review



Source: Data from reviews carried out by the authorities with information on standard error (n=238), BASS calculations

Interim conclusion 6

In summary, we find that more than **50% of the entities reviewed do not display any gender effect** and would not have significantly exceeded the tolerance threshold even if a 0% threshold were applied. These results are relatively similar to those of the studies based on the ESS. The proportion of entities exceeding the tolerance threshold with a **threshold value of +/-5%** lies at around **8%** and is markedly lower than that of the studies mentioned above. This is explained by the fact that the reviewed entities are not representative of the economy as a whole, and that certain variables are operated differently between the ESS and analyses conducted using the Logib module 1 tool. In addition, there is a data validation process included in the scope of the reviews which is not carried out in the ESS. It is also interesting to note that among the 40% of entities displaying a gender effect which does not exceed the +/-5% tolerance threshold, more than **two out of every three entities** have a discrimination coefficient of 5% or more and can thus be seen as **borderline cases**.

Furthermore, significant deviations are found in the **average unexplained wage difference**, notably in relation to the line of business, the proportion of women, the R^2 and the year of the review.

Adjusting the tolerance threshold downwards would have a considerable impact on the proportion of entities exceeding it. If a 0% threshold had been applied, this would have implied that 49.2% of the reviewed entities would have exceeded this threshold. This percentage decreases in an almost linear fashion in relation to the tolerance threshold: 42.9% with a +/-1% threshold, 27.7% with a +/-2.5% threshold, 14.3% with a +/-4% threshold, and finally 8% with a +/-5% threshold.

These adjustments would have had a differentiated impact on the results of the reviews in relation to various factors. In particular, they would have had a greater impact on the results of entities with at least 1,000 employees, those with 60% women or fewer, and reviews carried out between 2006 and 2010. Conversely, the results for small entities, for those employing more than 60% women, and for reviews carried out since 2020 would have been impacted considerably less by a lowering of the tolerance threshold.

4.2 Experiences of supervisory authorities

In the following section we show how the supervisory authorities assess the quality of the data submitted by companies and control practice, and their experiences with the tolerance threshold (for the methodological procedure cf. Section 1.3).

Assessments of data quality, control practices and review results

According to assessments by the supervisory authorities, since equal pay analyses were introduced in 2006 there has been a marked increase in **data quality** and completeness. This has been helped along by measures within the review process, such as collection of additional information (e.g. job title), provision of a detailed guide on correct data capture and processing, and digitalisation. On the one hand, a web-based analysis tool was developed. On the other, in the wake of digitalisation, the systems for capturing and updating wage data within companies improved. It can therefore be assumed that as digitalisation continues the quality and completeness of the data will continue to improve in the coming years. The revision to the Gender Equality Act in 2020 also contributed to the increase in quality, as it meant that the companies required by the revision to undergo reviews would be working with equal pay experts more frequently.

Significant improvements are also being noted in relation to **control practices**. At the federal level, the process was standardised and consistent control practices established through regular exchanges with

4 The tolerance threshold as applied in state reviews

equal pay experts. The Logib training courses provided by the FOGE also contributed to standardising processes at the cantonal and municipal levels. Most authorities are currently largely following the federal procedure and using the same documents and tools. In some cases, they are working with individuals from the experts' pool for reviews in the public procurement system. According to the authorities' assessments, the dual control principle can also contribute significantly to quality assurance. However, this is only being used at the federal level and in one canton. The other cantons or municipalities are familiar with other quality assurance mechanisms, such as involving a second external expert in complex cases or having a second person double check the audit report. Overall, control practices are thought to be accurate and detailed enough to obtain robust results.

Several of the supervisory authorities consulted were unable to express an opinion on any differences in the **review results** between companies due to the limited number of supervised reviews. Authorities with a broader range of experience tended to notice the following patterns:

- Companies without a (transparent) wage system display larger unexplained wage differences.
- Unexplained wage differences tend to be lower in the subsidies field than in procurement, as the former often relates to government-related organisations which base their wage systems on those of the public sector.
- Results likewise vary from one industry to another. Some industries exhibit problematic results more frequently as analysis of wage policy is less commonplace here.
- Smaller businesses display significant unexplained wage differences less often than larger ones, as the method applied for smaller case numbers (employee numbers) is more generous.

Experiences with the tolerance threshold

Intensification of reviews since 2013/2014, the introduction of reviews at the cantonal and municipal level and revisions to the Federal Act on Public Procurement and to the Gender Equality Act, along with the accompanying political debates, have resulted in greater awareness at company level too, and the tolerance threshold is said to have become established. For many companies it has also become important to comply with equal pay standards. This impression is shared by all authorities able to provide information on the subject. In the case of reviews within administrative bodies, in some places there is now a consensus that zero unexplained wage differences should be the goal irrespective of the tolerance threshold level. This is also shown by a study (Rüegge et al., 2018) according to which the vast majority of companies have observed an effect from their equal pay review (e.g. greater awareness, wage adjustments, modifications to wage systems), irrespective of whether the tolerance threshold had been exceeded. What has come out of the interviews overall is that there is a broad **acceptance of the existence of a tolerance threshold**, but often a lack of awareness regarding its significance. A number of authorities observe that among reviewed companies (and in some cases also administrative bodies) there is an entrenched idea that if pay figures are below the tolerance threshold they have complied with equal pay requirements, and that it is only if the threshold is exceeded that they need to take action. Moreover, one authority found that the 5% threshold is very present at companies and that they are deliberately looking to fall below it, if need be by adjusting roles in minor ways.

Several authorities indicate that it was less the existence of a tolerance threshold than the **significance of the results** that was providing grounds for discussion. Here the question arises as to whether large unexplained wage differences should be examined more closely even if the result is not significant. Also, it appears to cause confusion and legal uncertainty among authorities in the public procurement system when a company exceeds the tolerance threshold but the result is not significant. There is a legal need for clarifi-

4 The tolerance threshold as applied in state reviews

cation here. Furthermore, another source of discussion is that the probability of a significant result increases with the size of the company and the tool works less well for companies with just a few employees of one gender.²⁹

The majority of supervisory authorities consider a **tolerance threshold of 5% wage inequality too generous**, repeatedly referring back to study results (e.g. Trageser et al. 2011) which reach this conclusion. Another widespread argument for lowering the tolerance threshold is the continued development of the tools³⁰, methodology and control practices, which over the years have led to more accurate results. Reports have also shown that the method is in line with current scientific standards and the model takes into account the relevant variables. Moreover, the tolerance threshold is based on neither legal nor scientific foundations. For several authorities, another concern is ensuring that problematic cases cannot slip through the net. For that reason, the tolerance threshold should not be set too high. Two authorities see the low number of cases where the threshold is exceeded as an indicator that it is set too high. A number of discussions also addressed the fact that with Logib module 1 (regression analysis) individual wage discrimination instances cannot be identified and the tool only detects serious problems. One authority also considered it very disturbing that the analysis method allows companies to adapt at a statistical level in order to pass the review without changing anything about their wage policy.

Some authorities were unable or unwilling to provide information on what **level of tolerance threshold** would be appropriate, as this could not be derived scientifically or the methodological skills were lacking. However, all authorities consulted agreed that if the threshold were to be adjusted it should be set lower and not higher. There are arguments in favour of halving the threshold to 2.5%, as this would be easy to communicate and discussions at the European level are going in a similar direction. Individual authorities would also welcome a lower tolerance threshold (2% or 1%) or its complete removal, as the legal basis for wage equality is fundamentally one of zero tolerance. One authority, on the other hand, felt that it would make sense not to set the threshold too low. Lowering the threshold significantly would probably increase review costs and generate more conflicts with the reviewed companies. It was repeatedly noted that when setting the tolerance threshold authorities would also need to take into account the context of the analysis (self-test evidence or review by an expert, subsidies or procurement field) as well as the consequences of exceeding the threshold. For example, a lower threshold would be justified if sanctions were not immediately imposed but corrective action could be taken first. Several respondents also suggested consulting the results of previous equal pay analyses in order to assess how many companies would exceed the tolerance thresholds if adjusted.

At the cantonal and municipal level, the consulted authorities tended to be sceptical with regard to how well **accepted a lowering of the tolerance threshold** would be politically. A few felt that resistance was to be expected at the political level. They also said that it was not within the remit of the cantons/municipal authorities to launch this debate or to change the threshold independently, as they follow the federal approach and would not have the necessary credibility. One person had the impression that companies would be better able to deal with a lowering of the tolerance threshold than the authorities, as any

²⁹ With Logib module 2 an additional tool was developed which was suitable for equal pay analyses at small and medium-sized enterprises (up to 50 employees).

³⁰ However, one authority indicated that while the continued development of the analysis tool meant that the desired data quality was reached more quickly, this was not a good reason for lowering the tolerance threshold as they felt that the review procedure always resulted in good data quality. Having said that, these developments can have a powerful impact on data quality in analyses conducted outside of the review process.

4 The tolerance threshold as applied in state reviews

change would entail a cumbersome political process. Conversely, others expect opposition to come primarily from the company side.

Several authorities feel it is important **not to focus solely on the level of the tolerance threshold**. It would be preferable to spur companies to act in cases where analyses point to wage discrimination, irrespective of whether the tolerance threshold has been significantly exceeded. In this context, the introduction of the traffic light system³¹ is welcomed. What is now called for is to clearly communicate that action is needed even in the case of an “orange” result. If need be, mechanisms could be rolled out to call “orange companies” to account as well. Another authority sees the priority as tools which will be better able to signal discrimination at the individual level and would welcome the use of a uniform review tool, irrespective of company size. Ideally, companies should be informed in all cases that despite having passed the review there are individual inequalities present and the employees can call for an individual review.

Moreover, some respondents mention that the **term** “tolerance threshold” is somewhat unfortunate and misleading, as it is more of a risk score. One person also indicated that companies use the term “margin of error” more often than “tolerance threshold”.

³¹ Since the web-based tool was introduced, a colour-coding system has been used to present results. Green = no gender effect; orange = gender effect not exceeding the tolerance threshold; red = tolerance threshold of +/-5% exceeded.

4 The tolerance threshold as applied in state reviews

Interim conclusion 7

The supervisory authorities consider that numerous **developments in the tool and in control practices** have helped to markedly enhance the quality and completeness of the data, or at least the speed at which these criteria could be achieved. The processes are now more standardised and there is a better guarantee that the results will be robust.

The authorities also agree on the important role that the existence of a tolerance threshold has played in **acceptance** of the review process **by the reviewed entities**.

They feel that awareness of the topic has developed greatly since the reviews were introduced and that it is more and more important for companies to show that they are upholding wage equality. According to one authority, lowering the tolerance threshold would nevertheless still be viewed poorly by entities that could potentially be reviewed.

The majority of the supervisory authorities consider the +/-5% tolerance threshold to be **too generous**. Apart from the developments mentioned which allow for more accurate results to be obtained, many authorities lament the fact that virtually all entities reviewed “pass” the review while in some cases displaying gender effects higher than +/-5%, which suggests that infractions are not being detected. Some authorities feel that the standard analysis model with this +/-5% tolerance threshold only allows them to identify the most serious cases of wage discrimination, as an enormous number of individual instances within the entity would be required to significantly exceed the tolerance threshold. This is considered a problem as there is no tolerance threshold envisaged under the law.

The majority of supervisory authorities agree on the necessity of **lowering the tolerance threshold** in order to better detect potential infractions; however, they are unable to give a precise idea of what an appropriate value might be. Among others, a tolerance threshold of +/-2.5% has been mentioned, as it is easy to communicate (threshold reduced by half) and corresponds in part to discussions at the European level. Other authorities would even argue for a lower threshold, set at 1 or 2%, if not for removing it altogether.

If the majority of the authorities are in favour of lowering the tolerance threshold, they remain sceptical with regard to **acceptance of this move** at the cantonal and municipal levels, whether in certain political circles or by companies.

4.3 Assessments from a legal perspective

All 3 legal experts consulted feel that the Federal Office for Gender Equality (FOGE) is competent to review and if necessary adjust the currently valid tolerance threshold for Logib module 1. They likewise agree with the FOGE's assertions that an adjustment to the currently valid tolerance threshold of 5% would be appropriate (if not essential) for the following reasons:

- Technical progress since the release of Logib module 1 and the setting of the tolerance threshold at 5%;
- Roll-out of various (technical) clarifications for the Logib tool;
- Introduction of numerous guides, information materials, training courses, webinars, training opportunities, clarifications and examples from practice;
- Numerous new validations of the scientific quality and legal conformity of various components of the method by independent bodies;
- Changes to legal and political framework conditions since the release of Logib module 1 and the setting of the tolerance threshold at 5%;

5 Summary and recommendations

- International developments, efforts and tools to eliminate gender-specific wage differences;
- In general, employers and employees are showing greater awareness and consideration for wage equality between women and men since the release of Logib module 1 and the setting of the tolerance threshold at 5%.

One person refers here in particular to the developments in the EU. Another highlights that due to the technical improvements mentioned above and more accurate analytic capabilities, the tolerance threshold should not be set too high in the interests of keeping the tool effective.

Overall, both of the legal experts consulted who were able to give a view on the level of the tolerance threshold felt that lowering it to 2.5% would be appropriate. They again referred to the EU's planned transparency directive in this connection, which in the EU Parliament's version stipulates that only a 2.5% wage difference will still be permissible. From a legal perspective, moreover, an adjustment in (interim) stages would be justifiable (e.g. a first step lowering the threshold to 3% and, depending on subsequent experiences, a second lowering it to e.g. 2%).

On the question of whether, from a legal perspective, further information is required in addition to the basics compiled in the context of this study in order to make a decision on a potential adjustment to the tolerance threshold, the consulted experts were unable to comment. They felt that this would require a more in-depth analysis on their part, which they would not be able to provide within the scope of the written consultation.

Interim conclusion 8

Given that the assessments from a legal perspective provided here are not comprehensive expert reports, the experts consulted could not provide in-depth comments on all the points raised. However, the legal consultations give no indication that the conditions are not in place from a legal perspective to decide on a lowering of the Logib tolerance threshold.

5 Summary and recommendations

The analyses conducted in the context of this study have provided an overview of the existing information and assessments on the tolerance threshold from the methodological, empirical, legal and political perspective. This section sums up the different elements relating to these four aspects and issues recommendations for exploring a potential adjustment to the tolerance threshold by the FOG's equal pay specialists and external experts.

5.1 Methodological aspects

From the point of view of the method, the tolerance threshold was set in the context of a multi-stage procedure. Initially, it was to be applied only within the "basic regression", which only included human capital factors as explanatory variables, and discarded in the context of in-depth analyses. Ultimately, it was also retained within the "expanded regression", which includes job-related variables and constitutes the standard analysis model, also used within the context of the reviews. One of the arguments put forward to justify the tolerance threshold was the fact that the model remained relatively simple and potentially did not take into account all the non-discriminatory variables existing within an entity.

Over time, analyses showed that it was not useful to expand the standard analysis model to include new variables, for a variety of reasons. On the one hand, the model has to allow analyses to be conducted for a limited cost and workload. On the other, it must not take into account variables with discriminatory potential, which is the case for the majority of other variables generally suggested in the context of these

5 Summary and recommendations

discussions (e.g. actual work experience, work-time percentage during career). Above all, the existing model has been externally assessed on several occasions at the national (Felfe et al, 2015; Pärli & Oberhauser, 2019) and international level (ILO, EPIC) and deemed scientifically sound and legally compliant. Finally, various tests have shown that adding variables (e.g. actual work experience) would have only a marginal effect on results. Work is currently under way at the FOG level to establish in a more detailed manner the impact of the omitted variable bias, which has been raised as an important point by some experts and evaluations.

Furthermore, although no new variables have been incorporated into the model, the variables taken into account have been constantly re-evaluated and adjusted in order to ensure that they are being applied in line with the latest national and international insights. The instructions for coding of functions for the “skill level” and “professional position” variables have thus been regularly developed. Discussions are notably under way at the moment on the concept of skill level, and it is possible that this variable will soon be refined via a more detailed definition and the addition of new levels.

In addition to these clarifications regarding the variables and their coding, a number of changes have been made at the level of the analysis tool and control practices over the last fifteen years. According to the experts, these have resulted in a notable improvement in the quality and completeness of the review data. With these gains in accuracy, the majority of experts and supervisory authorities feel that the tolerance threshold of $\pm 5\%$ only allows the most serious cases of wage inequality within companies to be detected. We know today that it is almost impossible for companies to be falsely identified by Logib as practising wage discrimination. Even if the tolerance threshold were to be lowered, the impact on this would be negligible.³² In contrast, due to the methodology used, a great many companies are being “exonerated” even with high values. With the current tolerance threshold, the sensitisation effect is being significantly reduced and has adverse consequences: a high degree of wage discrimination is being tolerated.

Furthermore, according to Chávez-Juárez and Graf (2021), the existence of a tolerance threshold is not justified from an econometric point of view, as the method already includes significance tests with a confidence interval to limit the risks of a false conclusion, and the law does not allow for any tolerance threshold. This opinion is shared by a statistics expert consulted by Trageser et al. (2011), who felt that the tolerance threshold was more a “goodwill value” (*Kulanzwert*). In order to lower or remove the tolerance threshold, however, both Chávez-Juárez and Graf (2021) and several experts consulted by Trageser et al. (2011) feel that it would be necessary to include heteroskedasticity-robust standard errors in the model in order to minimise the possibility of obtaining inaccurate results.

In the other areas of state control, the conclusions which can be drawn regarding the tolerance threshold in equal pay analyses are limited, as the topics and measurement methods are very different and not easily comparable. However, it generally transpires that the different tolerance thresholds or limit values are adjusted in line with developments in the measurement method when this becomes more accurate, as well as in line with gains in knowledge on the topic in question. In particular, gains in knowledge allow for better targeting of situations with a real risk for the factor the threshold is seeking to protect (e.g. health, equal treatment, etc.) and recalibrations are thus a regular occurrence in a number of fields.

From a methodological point of view, it seems appropriate for the FOG’s equal pay experts and external experts to consider adjusting the tolerance threshold. All the more so as, according to several studies, the setting of this threshold must be based primarily on expert assessments, and it is not possible to precisely

³² A BSS study shows that with a tolerance threshold of 5% the probability of false accusations with Logib module 1 is 0.2%. If the tolerance threshold were to be lowered to 3%, the probability would still be vanishingly small, at less than 0.8%.

5 Summary and recommendations

quantify the adjustment required based on the changes made to the method or the tool. To complement these aspects, the following factors should also be considered.

5.2 Empirical aspects

Studies presenting empirical results concerning the tolerance threshold have proliferated in recent years. They are based on different versions of the ESS, on equal pay analyses conducted in the context of a certification after the entry into force of the revised GEA, and on the reviews carried out between 2006 and 2021. Some of these studies, including this one, have also conducted analyses of the impact of lowering the tolerance threshold on the results of equal pay analyses or reviews. **Table 9** below presents a summary of these various empirical results.

Table 9: Summary of empirical results concerning the tolerance threshold

	Trageser et al. (2011) Pilot phase reviews, 2006-2011	Graf and Garibian (2014) ESS 2010	Chávez-Juárez and Graf (2021) ESS 2018	Comp-On (2021) Fair-ON-Pay Report	BASS (2022) Reviews 2006-2021	BASS (2022) Reviews 2019-2021
Number of observations	14	3,000	2,845	193	260	61
Tolerance threshold of +/-5% exceeded	14.3%	25.5%	19.1%	4.1%	8.0%	3.3%
Tolerance threshold of +/-4% exceeded	-	29.7%	23.5%	-	14.3%	8.2%
Tolerance threshold of +/-3% exceeded	-	33.9%	29.6%	-	25.2%	19.7%
Tolerance threshold of +/-2% exceeded	-	39.0%	35.1%	-	34.9%	27.9%
Tolerance threshold of +/-1% exceeded	-	45.2%	42.1%	-	42.9%	39.3%
Tolerance threshold of 0% exceeded	71.4%	52.4%	50.2%	48.2%	49.2%	39.3%
No gender effect	28.6%	47.6%	49.8%	47.7%	50.8%	60.7%

Source: Studies shown in the table, BASS presentations

All the studies except that by Trageser et al. (2011), which dealt with only a reduced sample of 14 reviews, present relatively similar proportions of entities not showing any gender effect, ranging between 47.6% and 50.8%. Thus, almost half of the entities display gender effects which differ significantly from zero, whether this relates to all Swiss entities, those reviewed within the context of the public procurement and/or subsidies system, or entities seeking to obtain equal pay certification. However, the proportion of entities significantly exceeding the +/-5% tolerance threshold varies considerably between the different sources: this figure stands at 25% and 19% in the studies based on the ESS, but at only 8% in the context of reviews and 4% in the context of certification analyses. These substantial differences between the ESS on the one hand and reviews/analyses on the other can be explained by the differences in terms of samples, operationalisation of the variables and data validation process.

Furthermore, we observed a significant development in review results over time: the proportion of entities not exceeding the tolerance threshold of +/-5% has decreased substantially. If the proportion was already relatively low on average (8%) for all reviews, it has only been 3.3% for the reviews carried out since 2019 (64 reviews over this period). Among these reviews, 60.7% do not show any gender effect. This development seems to confirm the greater sensitisation of entities to the topic of wage equality and a better understanding of the analysis tool. At the same time, the figures from the FSO show an increase in instances of wage inequality between 2014 and 2018 at the national level and underpin the need to act to decrease these inequalities. Lowering the threshold would thus allow for better detection of equal pay infractions while also being consistent with developing awareness and knowledge.

In terms of empirical data, it thus seems that the FOG specialists now have enough information to be able to judge the impact of a potential adjustment to the tolerance threshold on the results of equal pay

5 Summary and recommendations

reviews and analyses at different levels. Moreover, this information should soon be supplemented by the results of tests concerning the effects of the omitted variable bias, which has been mentioned repeatedly as an important factor to take into account. The data in their current state seem to indicate that the FOGÉ should consider moving towards a lowering of the tolerance threshold.

5.3 Legal aspects

From a legal perspective, no gender discrimination whatsoever is tolerated by the law, no matter how small. The $\pm 5\%$ tolerance threshold thus has no basis in law.

In the opinion of the legal equal pay experts consulted within the context of this assignment, the FOGÉ is the competent body to define the appropriate tolerance threshold for equal pay analyses using the standard analysis model. Furthermore, the legal experts share the opinion that not only the various technical and methodological developments, but also those relating to sensitisation and social and political context require the FOGÉ to look into possibly adjusting the tolerance threshold. What is more, they feel that lowering the tolerance threshold to $\pm 2.5\%$ would be appropriate. In addition to this, the current discussions at the European level are heading in the same direction.

However, the legal experts were unable to answer the question of whether additional information (in addition to that gathered in the context of this paper) would be necessary to allow a decision to be made on a potential adjustment to the tolerance threshold. Answering this question would have required an in-depth analysis on their part, which was not the purpose of the current assignment. Nevertheless, the legal experts acknowledged that it would be interesting to know what impact such an adjustment might have on Logib module 2.

Based on the information obtained, it does not seem necessary to have any additional bases for a legal assessment. Defining a threshold is first and foremost a political task, and the legal experts have confirmed the competence of the FOGÉ to deal with this matter and to define the appropriate threshold level.

5.4 Political aspects

At the time the tolerance threshold was introduced as part of the standardised review procedure, it had two objectives: to provide some degree of legal certainty while avoiding “false accusations” and to encourage political acceptance. The Federal Council, in its response to the motion by Carobbio Guscetti (10.3420), as well as certain specialists consulted by Trageser et al. (2011), argued that the tolerance threshold gave time to companies to adjust and to raise awareness of the topic in order to justify its existence. In the meantime, almost 300 reviews have been conducted at different levels and a great many self-tests performed, in particular following the revision of the GEA, which in July 2020 introduced an analysis obligation for all companies with 100 or more employees.³³ Furthermore, many equal pay labels and certifications have sprung up in recent years, with growing success. The report by Comp-On (2021) confirms this trend: the demand for certification has increased substantially since 2020, with more than 100 certifications in 18 months, while the list of companies receiving such a certification was still very short just a few years ago. The argument for a period of adjustment and the need to build awareness of the topic in order to justify the existence of a tolerance threshold thus seems hard to accept today.

In addition, the Federal Council and the experts consulted at the time felt that this threshold was set in a rather generous manner. Today, the supervisory authorities are of the same opinion. Many of them, who have been conducting reviews for a number of years, feel that certain infractions are not being detected

³³ According to STATENT figures, there were 5,083 companies with more than 100 employees in Switzerland in 2019. However, we do not have figures for the number of analyses performed.

5 Summary and recommendations

with this tolerance threshold, as there is a significant gap between the low number of reviewed entities exceeding the threshold set and the number of those showing a statistically significant gender effect.

Furthermore, in recent years there have been efforts at the political level to move towards greater wage equality between women and men at both the Swiss and international level. At the Swiss level, these efforts have taken the form of the revised GEA, the new PPA, development for the Charter for equal pay in the public sector, and the 2030 Gender Equality Strategy. In this context, a report on the strategy for strengthening the Charter on equal pay is currently being drawn up in response to Postulate 20.4263 “Strategy for strengthening the Charter on equal pay”. At the international level, the ILO, OECD and UN Women have established a coalition for wage equality (the Equal Pay International Coalition, or EPIC). The UN has also introduced an international Equal Pay Day (18 September). In addition, discussions are currently under way in the European Parliament to improve pay transparency. In particular, a tolerance threshold of $\pm 2.5\%$ deviation in gross pay between women and men holding positions with comparable requirements is being discussed.

In view of the various developments, deliberations regarding a possible adjustment to the tolerance threshold also seem necessary from a political point of view and are pertinent to the current context.

5.5 General conclusion

In conclusion, all the aspects considered within the context of this working paper justify giving some thought to adjusting the tolerance threshold. Moreover, the 2.5% level seems appropriate from the perspective of the various developments in the tool and the context, and would be in line with discussions at European level. However, this is primarily a normative decision which must be taken by the competent authorities.

In order to provide the best conditions for making this decision, this working paper aims to provide a documentary base that is as wide-ranging as possible, bringing together an analysis of the literature, empirical data, experiences and assessments by experts, supervisory authorities and legal equal pay specialists. Taking into account this working paper as well as the other work recently carried out on the subject, in particular by Chávez-Juárez and Graf (2021) and Kaiser (2022), we believe that the FOGÉ now has a sufficiently robust base of information to be able to launch the discussion process on a potential adjustment to the tolerance threshold.

Moreover, other work is currently being carried out in parallel and will need to be taken into account in this process. In particular, we can cite the work on the effects of the omitted variable bias, the drafting of a report on the strategy for strengthening the Charter for equal pay in the public sector, deliberations regarding the definition and number of tiers in the “skill level” variable, and the upcoming integration of Logib module 1 into the Swiss ELM electronic salary reporting system through ERP systems. In addition, the FOGÉ had to prepare a report by the end of 2022 on the status of European legislation and possible effects on Switzerland.

Based on the results of this document, it would seem appropriate to move towards a lowering of the tolerance threshold. If such a move is decided, one of the major challenges will be communicating it, in particular to companies and the general public. For this reason, this working paper provides an overview of the available information which can be leveraged for this communication. Finally, we agree with the conclusions of Felfe et al. (2015), who feel that establishing a tolerance threshold must be based mainly on expert assessments. The upcoming discussions between FOGÉ equal pay specialists and external experts will thus constitute a vital part of this process and the final decision should thus arise mainly from these discussions.

6 Bibliography

- AHV/IV (2020). Ergänzungsleistungen (EL) 2021: Was ändert? Stand am 1. Januar 2021[Merkblatt]. <https://www.ahv-iv.ch/de/Merkbl%C3%A4tter-Formulare/Merkbl%C3%A4tter/Erg%C3%A4nzungsleistungen-zur-AHV-und-IV> (abgerufen am 16.5.2022).
- Bauer, T., Strub, S., Füglistner, M. (2001). Instrumente zur Kontrolle der Lohnungleichheit bei Beschaffungen des Bundes. Interner Bericht über die Möglichkeiten der Umsetzung von Art. 8 Abs. 1 Bst. c des Bundesgesetzes über das öffentliche Beschaffungswesen. Im Auftrag des EBG. Bern.
- Baumann, P. (2022, 4. August). Bundesamt für Umwelt BAFU [Persönliche Kommunikation].
- BAFU Bundesamt für Umwelt (2022, 1. Februar). Belastungsgrenzwerte für Lärm. <https://www.bafu.admin.ch/bafu/de/home/themen/laerm/fachinformationen/laermbelastung/grenzwerte-fuer-laerm/belastungsgrenzwerte-fuer-laerm.html> (abgerufen am 23.5.2022).
- BFE Bundesamt für Energie (2022, 1. März). Marktkontrolle Elektrogeräte 2021. <https://www.bfe.admin.ch/marktueberwachung> (abgerufen am 4.8.2022).
- BFE Bundesamt für Energie (2021, 22. Juni). Marktkontrolle Elektrogeräte 2021. <https://www.bfe.admin.ch/marktueberwachung> (abgerufen am 4.8.2022). BFE Bundesamt für Energie (2020, 15. Mai). Marktüberwachung. <https://www.bfe.admin.ch/bfe/de/home/effizienz/energieetiketten-und-effizienz-anforderungen/marktueberwachung.html> (abgerufen am 12.5.2022).
- BKF Eidgenössisches Büro für Konsumentenfragen (2022, 25. April). Durchgezogene Ergebnisse bei den Kontrollen der Holzdeklaration [Medienmitteilung]. <https://www.konsum.admin.ch/bfk/de/home/holzdeklaration/holzdeklarationspflicht.html> (abgerufen am 8.8.2022).
- BKF Eidgenössisches Büro für Konsumentenfragen (2021, 27. April). Weiterhin Verbesserungspotenzial bei der Holzdeklaration [Medienmitteilung]. <https://www.konsum.admin.ch/bfk/de/home/holzdeklaration/holzdeklarationspflicht.html> (abgerufen am 8.8.2022).
- BKF Eidgenössisches Büro für Konsumentenfragen (2020, 3. März). Verbesserungen bei der Holzdeklaration festgestellt [Medienmitteilung]. <https://www.konsum.admin.ch/bfk/de/home/holzdeklaration/holzdeklarationspflicht.html> (abgerufen am 8.8.2022).
- Bureau fédéral de l'égalité entre femmes et hommes BFEG (2021). Modèle d'analyse standard de la Confédération. Descriptif méthodologique Logib Modules 1 et 2. Version 2021.1.
- COM Europäische Kommission (2021, 4. März). Vorschlag für eine Richtlinie des Europäischen Parlaments und des Rates zur Stärkung der Anwendung des Grundsatzes des gleichen Entgelts für Männer und Frauen bei gleicher oder gleichwertiger Arbeit durch Lohntransparenz und Durchsetzungsmechanismen. <https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX:52021PC0093> (abgerufen am 12.5.2022).
- Comp-On AG, HR Solutions (2021). Fair-ON-Pay Report. Erkenntnisse zur Lohnungleichheit zwischen Frauen und Männern in der Schweiz aus über 190 Fair-ON-Pay Analysen.
- Chávez-Juárez, F. & Graf, R. (2021). Réflexion empirique sur le seuil de tolérance utilisé lors des contrôles de l'égalité salariale. Bureau de l'égalité entre les femmes et les hommes BEFH, Lausanne.
- Ecoplan (2010). Forschungskonzept Lärm - Konzeptstudie zur Aktualisierung der Grundlagen für die Lärmbeurteilung, im Auftrag der Eidgenössische Kommission für Lärmbekämpfung und des Bundesamts für Umwelt. Altdorf.
- Eurostat (2022, 25. Februar). Geschlechtsspezifischer Lohnunterschied ohne Anpassungen. https://ec.europa.eu/eurostat/de/web/products-datasets/product?code=SDG_05_20 (abgerufen am 3.8.2022).

6 Bibliography

- EKLB Eidgenössischen Kommission für Lärmbekämpfung (2021a). Eidgenössische Kommission für Lärmbekämpfung empfiehlt Anpassung der Grenzwerte für Strassen-, Eisenbahn- und Fluglärm [Medienmitteilung]. <https://www.admin.ch/gov/de/start/dokumentation/medienmitteilungen.msg-id-86339.html> (abgerufen am 12.5.2022).
- EKLB Eidgenössischen Kommission für Lärmbekämpfung (Hrsg.) (2021b). Grenzwerte für Strassen-, Eisenbahn- und Fluglärm. Empfehlungen der Eidgenössischen Kommission für Lärmbekämpfung EKLB. Bern.
- EMPL Secretariat des Europäischen Parlaments (2022, 19. Mai). Ausschuss für Beschäftigung und soziale Angelegenheiten [persönliche Kommunikation].
- EnergieSchweiz (2021, 1. März). Energieetikette für Geschirrspüler ab 1. März 2021 [Faktenblatt]. <https://www.bfe.admin.ch/bfe/de/home/effizienz/energieetiketten-und-effizienzanforderungen/haushaltsgeraete/geschirrspueler.html> (abgerufen am 12.5.2022).
- EnergieSchweiz (2020, 15. Mai). Energieetikette für Kühl- und Gefriergeräte ab 1. März 2021 [Faktenblatt]. <https://www.bfe.admin.ch/bfe/de/home/effizienz/energieetiketten-und-effizienzanforderungen/haushaltsgeraete/kuehl-und-gefriergeraete.html> (abgerufen am 12.5.2022).
- EP Europäisches Parlament (2022a, 21. Februar). Geschlechtsspezifisches Lohngefälle: Verbindliche Maßnahmen für Lohntransparenz [Briefing]. [https://www.europarl.europa.eu/think-tank/de/document/EPRS_BRI\(2022\)698934](https://www.europarl.europa.eu/think-tank/de/document/EPRS_BRI(2022)698934) (abgerufen am 12.5.2022).
- EP Europäisches Parlament (2022b, 5. April). Geschlechtsspezifisches Lohngefälle: Verbindliche Maßnahmen für Lohntransparenz [Medienmitteilung]. <https://www.europarl.europa.eu/news/de/press-room/20220401IPR26532/geschlechtsspezifisches-lohngefalle-verbindliche-massnahmen-fur-lohn-transparenz> (abgerufen am 12.5.2022).
- EP Europäisches Parlament (2020, 3. April). EU will geschlechtsspezifisches Lohngefälle verringern. <https://www.europarl.europa.eu/news/de/headlines/society/20200227STO73521/eu-will-geschlechtsspezifisches-lohngefalle-verringern> (abgerufen am 12.5.2022).
- Felder, R. & Wunsch, C. (2021). Evaluation des Analyse-Tools Logib Modul 1 für den Lohnvergleich zwischen Frauen und Männern. WWZ Insights.
- Felfe, C., Trageser, J., Iten, R. (2015). Studie zu den statistischen Analysen der Eidgenossenschaft betreffend die Lohnungleichheit von Frau und Mann.
- FPI Fair Pay Innovation Lab (2022). Universal Fair Pay Check. <https://www.fpi-lab.org/universal-fair-pay-check/> (abgerufen am 23.5.2022).
- Gisin, D. (2022, 28. April). Eidgenössisches Büro für Konsumentenfragen BFK [Persönliche Kommunikation]
- Graf, R. & Garibian, V. (2014). Seuil de tolérance et discrimination salariale : statistiques basées sur l'Enquête suisse sur la structure des salaires 2010. Document de réflexion.
- Häne, S. (2022, 6. Mai). Parmelin kündigt neue Kontrollen bei IKEA an. *Der Bund*. <https://www.der-bund.ch/parmelin-kuendigt-neue-kontrollen-bei-ikea-an-445916199985> (abgerufen am 23.5.2022).
- Kaiser, B. & Möhr, T. (2021). Analyse des différences salariales entre femmes et hommes sur la base de l'enquête suisse sur la structure des salaires (ESS) 2018. B,S,S Volkswirtschaftliche Beratung. Sur mandat de l'Office fédéral de la statistique (OFS). 2021.
- Kaiser, B. (2022). Monte-Carlo-Simulation der Lohnungleichheitsanalyse in Logib Modul 1, EBG, Bern.
- Koller, M. (2022, 3. August). Prozessverantwortlicher Grenzwerte der Suva [Persönliche Kommunikation].

- Koller, M., Pletscher, C., & Jost, M. (2013). Schweizer Grenzwerte am Arbeitsplatz [Factsheet]. <https://www.suva.ch/de-CH/material/Factsheets/grenzwerte-am-arbeitsplatz-grundlagen-und-anwendung> (abgerufen am 2.5.2022).
- Kennedy, Peter E. (1981). Estimation with Correctly Interpreted Dummy Variables in Semilogarithmic Equations, *American Economic Review*, 71(4), p. 801.
- Lempen, K. & Voloder, A. (2017). Analyse der kantonalen Rechtsprechung nach dem Bundesgesetz über die Gleichstellung von Frau und Mann (2004-2015), Bern.
- Marti Whitebread, C. (2016). « Egalité salariale entre femmes et hommes: état des lieux » in: Dunand, Lempen, Mahon « L'égalité entre femmes et hommes dans les relations de travail ».
- Merckx, V. (2016). Vollzug der flankierenden Massnahmen: Verbesserungen sind möglich. *Die Volkswirtschaft*. <https://dievolkswirtschaft.ch/de/2016/07/merckx-07-2016/> (abgerufen am 27.4.2022).
- Organisation internationale du Travail (2022). Législation sur la transparence des salaires : Implications pour les organisations d'employeurs et de travailleurs. Bureau international du Travail – Genève : BIT, 2022.
- Pärli, K. & Oberhauser, C. (2019). Document de réflexion : analyse juridique du potentiel de discrimination des variables explicatives servant à mesurer l'égalité salariale entre femmes et hommes, BFEG, Berne.
- Peter-Hansen, K.M., & Rafaela, S. (2022, 22. März). Bericht über den Vorschlag für eine Richtlinie des Europäischen Parlaments und des Rates zur Stärkung der Anwendung des Grundsatzes des gleichen Entgelts für Männer und Frauen bei gleicher oder gleichwertiger Arbeit durch Lohntransparenz und Durchsetzungsmechanismen [Plenarsitzungsdokument]. https://www.europarl.europa.eu/doceo/document/A-9-2022-0056_DE.html (abgerufen am 12.5.2022).
- Piazza, M. (2019, 1. Februar). Prämienverbilligung: Die Rüge des Bundesgerichts an den Kanton Luzern interessiert auch in Nid- und Obwalden. *Luzerner Zeitung*. <https://www.luzernerzeitung.ch/zentral-schweiz/nidwalden/praemienverbilligung-die-ruege-des-bundesgerichts-an-den-kanton-luzern-interessiert-auch-in-nid-und-obwalden-Id.1090234> (abgerufen am 16.5.2022).
- PricewaterhouseCoopers AG (2019). Descriptif technique pour la spécification du salaire dans le modèle d'analyse standard de la Confédération. Spécification du salaire conforme au droit. Sur mandat du BFEG, Berne.
- Rosser, S., de Haan, P., & Gehwolf, G.L. (2021). Energieetikette für Neuwagen: Anpassung der Kategoriengrenzen und mittlerer CO₂-Wert der Neuzulassungen per 1.1.2022. Bern: Bundesamt für Energie BFE.
- Rüegge, B., Petry, Ch., Stern, S. (2018). Wirkungen der Lohngleichheitskontrollen des Bundes, Schlussbericht, Zürich.
- SECO Staatssekretariat für Wirtschaft (2016a). Erfolgsfaktoren beim Vollzug der flankierenden Massnahmen flankierenden Massnahmen auf Grundlage der Erfahrungen der Audits (Oktober 2012 – Mai 2015). Bern.
- SECO Staatssekretariat für Wirtschaft (2016b). Massnahmen zur Konkretisierung des Aktionsplans. Bericht der Arbeitsgruppe zum Verbesserungsbedarf von Vollzug und Missbrauchsbeämpfung der FlaM zuhanden des Bundesrats. Bern.
- SFH Schweizerische Flüchtlingshilfe (2022, 29. April). Änderung des AIG: Anpassung der Sozialhilfeleistungen für Ausländerinnen und Ausländer aus Drittstaaten. Vernehmlassungsantwort der Schweizerischen Flüchtlingshilfe. Bern.
- Stadler, P. (2022, 4. August). Bundesamt für Energie BFE [Persönliche Kommunikation].

- Stellungnahme des Bundesrats vom 8.9.2004 zur Motion Mörgeli (04.3336) «Strassenverkehr. Toleranzwert von 5 statt 3 Stundenkilometern».
- Strub, S. (2004). Contrôle du respect de l'égalité de salaire entre femmes et hommes dans les marchés publics de la Confédération. Rapport sur la phase-pilote portant sur la mise en œuvre de l'art. 8, al. 1, let. c, de la loi fédérale sur les marchés publics, BFEG, Berne.
- Strub, S. (2005). Approche méthodologique relative au contrôle de l'égalité salariale entre femmes et hommes dans les marchés publics de la Confédération, BFEG, Berne.
- Stutz, H., Schär Moser, M., Freivogel, E. (2005), Evaluation der Wirksamkeit des Gleichstellungsgesetzes, Synthesebericht, im Auftrag des Bundesamts für Justiz.
- Suva (2019). Grenzwerte am Arbeitsplatz. Luzern
- Trageser, J., Stern, S., Iten, R. (2011). Evaluation der Kontrollen im Beschaffungswesen.
- Vaccaro, G. (2015). How to Reduce the Unexplained Gender Wage Gap ? Evidence from a Regression Discontinuity Design, Thèse de doctorat, Université de Genève, Genève, 2015.