Economics School of Louvain - ESL

Economics School of Namur - ESN

Field of Study Differences and Gender Wage Inequality in Belgium

A decomposition using Gelbach's method

Author : Packo Edouard

Thesis Director : Vandenberghe Vincent

Academic Year 2025-2026

Master in Economics – 60 credits

Université catholique de Louvain - Economics School of Louvain - Place Montesquieu 3, 1348 Louvain-La-Neuve
Université de Namur – Economics School of Namur – Rempart de la Vierge 8, 5000 Namur

Abstract

This thesis analyses the extent to which differences in field of study between men and women may contribute to gender wage inequality among higher education graduates in Belgium, both directly and indirectly through occupation, sector, and working time. The analysis uses the 2012 Vacature Salarisenquête dataset and focuses on a sample of 25,795 employed higher education graduates. Ordinary Least Squares regressions and Gelbach's (2016) decomposition method are used to quantify the contribution of each factor to the gender wage gap. The results indicate that women earn approximately 21.09% less than men, conditional on age. Observable characteristics explain 68.78% of this gap. Field of study alone is associated with 16.81% of the total gap, while an additional 11.41% appears to be associated with the gap indirectly through occupation, sector, and working time. Overall, educational segregation is associated with approximately 28.23% of gender wage inequality, suggesting that a substantial share of wage disparities may originate before labour market entry through gendered educational choices.

Acknowledgments

Before going any further in this thesis, I would first like to thank everyone who has helped me, in one way or another, with the completion of this master's thesis.

First of all, I would like to express my sincere gratitude to Professor Vincent Vandenberghe, my supervisor, who guided me in researching my topic, provided me with the necessary resources to start with my analysis, gave me his time whenever I had questions, and, above all, placed his trust in me to carry out this research even if I had limited experience in applied econometrics.

I would also like to thank my parents, my brother and my friends for their constant support, as well as for their advice. I also want to specifically thank Lena Aroud for her encouragement and understanding, but also for her help in proofreading my work.

This year, and this thesis on a tangible and topical subject, have reminded me why I have always liked economics and why I decided to study it when I started university. For this, my gratitude goes to the Economics School of Louvain and to all the lecturers who have contributed to my enlightenment on this broad discipline.

Finally, to everyone I have not mentioned who contributed, either directly or indirectly, to the writing of this thesis, I would like to thank you.

Preliminary Note

In order to improve the readability of this paper by correcting spelling and grammar mistakes, the AI tools DeepL Write (DeepL SE) and ChatGPT (OpenAI) were occasionally used, to a limited extent, during the writing and proofreading of the following pages. These tools were not involved in reflections, methodological choices, research, analyses, or interpretations, and have not been used to generate ideas or explanations.

Table of Content

Introduction

In spite of important progress in the research about gender equality, substantial wage gaps between men and women persist in many countries, and Belgium is no exception. Although the Belgian institutional and legal framework promoting equal treatment is relatively well developed, women continue to earn less than men on average. Understanding the mechanisms linked to these disparities remains an important issue in labour economics and public policy.

Gender wage inequality is a complex phenomenon that cannot be attributed to a single factor. Instead, it results from a combination of mechanisms, including differences in labour market characteristics and working time arrangements, as well as educational, occupational and sectoral segregation. Measuring the overall wage gap alone provides only limited information. A more informative approach consists in identifying and quantifying the contribution of the different factors associated with this gap in order to better understand its origins and to identify potential areas for policy intervention.

Within these various determinants, educational choices have a particularly important position because they operate before individuals even enter the labour market. Men and women are not evenly distributed across educational specialisations, and these differences may have significant consequences for wages. Some fields of study are associated with higher wages on the labour market than others. This implies that educational segregation may contribute to gender wage inequality if women are overrepresented in fields of study that are characterised by relatively lower wage premia and underrepresented in those which are associated with higher earnings. In the sample used in this thesis, for example, women represent 83.9 percent of graduates in secretary studies, but only 7.8 percent of graduates in computing science.

The contribution of the field of study in which individuals obtained their degree to gender wage inequality may also extend beyond its direct effect on wages. Educational choices can influence the occupations individuals enter, the sectors in which they work, and their working time arrangements. Consequently, part of the effect of educational segregation may operate indirectly through these labour market characteristics. Differentiating the direct and indirect roles of the field of study is therefore essential to more precisely assess its contribution to the gender pay gap.

The international literature provides substantial evidence that the field of study matters for labour market outcomes. Thomas Breda and Clotilde Napp (2019) show that gender differences in educational specialisation arise early in the educational process. Livanos and Pouliakas (2012) estimate that the field of study variable explains an important share of the gender wage gap in Greece. Other authors such as Hägglund (2024), Sánchez-Mangas and Sánchez-Marcos (2021), as well as Michelmore and Sassler (2016), confirm that educational segregation can play a significant role in the wage gap, even if its extent may differ across countries and study disciplines. However, in Belgium, while the literature has documented gender wage inequality and wage-setting mechanisms, the specific contribution of the differences in fields of study between men and women among higher education graduates has not been quantified in a systematic way.

This thesis aims to fill this gap by addressing the following research question:

To what extent do differences in field of study between men and women contribute to gender wage inequality among higher education graduates in Belgium?

The thesis makes three main contributions. First, it provides a Belgium-specific estimate of the role of educational segregation in explaining gender wage inequality. Second, it focuses specifically on higher education graduates, for whom study discipline constitutes a meaningful and clearly defined educational choice. Third, it makes a distinction between the direct contribution of the field of study and its indirect contribution through other labour market factors, helping to connect the literature on educational segregation with the literature on the sorting of the labour market and wage inequality.

The empirical analysis is based on an individual-level dataset from Vacature Salarisenquête, which contains detailed labour market information for Belgium. The analysis focuses on the most recent available year, 2012. After restricting the sample to employed individuals who obtained higher education degrees, and with complete information on all relevant variables, the final dataset contains 25,795 observations. The dependent variable of the analysis is the logarithm of gross monthly wages. The main determinants that are included in the models are gender, age, field of study, occupation, sector of activity, and working time percentage.

The methodology combines Ordinary Least Squares (OLS) wage regressions with the decomposition method proposed by Gelbach (2016). A baseline model including only gender and age estimates a reference measure of the total gender wage gap (conditional on age), and the full model then adds field of study, occupation, sector, and working time to assess how much of this gap may be linked to these components. Gelbach's decomposition is then used to attribute the difference between the two models' estimated gender coefficient to each group of variables. After this, additional auxiliary regressions are used to estimate the extent to which the contributions of occupation, sector, and working time to the gender wage gap are associated with differences in field of study. The results should be interpreted as descriptive and associative rather than causal. The objective of the analysis is not to measure discrimination, but rather to assess the role of segregation in generating wage inequality.

The empirical results report that the total gender wage gap conditional on age amounts to approximately 21.09% to the detriment of women. After controlling for field of study, occupation, sector, and working time, the gap falls to approximately 6.59%, implying that these observable characteristics are associated with approximately 68.78% of the total gender wage gap. The Gelbach decomposition shows that the field of study alone may be associated with 16.81% of the overall gap. Additional indirect effects operating through occupation, sector, and working time appear to be associated with a further 11.41%. Taken together, these estimates suggest that educational segregation may be associated with approximately 28.23% of the gender wage gap among higher education graduates in Belgium. Among the other determinants considered, working time represents the largest individual contribution.

These findings suggest that an important share of gender wage inequality may already arise before labour market entry, through differences in educational specialisations. They also suggest that educational choices contribute to shape later labour market outcomes that further reinforce wage disparities. As a result, policies aimed at reducing gender segregation across fields of study may contribute to narrowing gender wage inequality over the longer term.

The structure of this thesis is organised as follows. Chapter 1 presents the motivations of the study. Chapter 2 reviews the literature on field of study and gender wage inequality. Chapter 3 describes the methodology (and its limitations) used in the analysis, the dataset, the econometric models, and the decomposition. Chapter 4 presents and interprets the empirical results of the analysis. The conclusion summarises the main findings, discusses their implications and limitations, and suggests areas for future research.

1. Motivations

Understanding the determinants of gender wage inequality is an important issue in labour economics. This remains true in developed countries, such as Belgium where, even if the already existing policies and frameworks addressing equality between men and women are relatively strong, gender wage gaps persist.

A challenge related to the study of these gaps is that they are not due to a single cause. In fact, it may result from multiple factors that can be linked to each other, closely or not, like working time, types of jobs, gender segregation, etc. This means that studying the overall gap does not explain much by itself. However, understanding and quantifying the factors lying behind can help to better understand the causes of this overall wage gap and to identify areas where we can act in order to reduce it.

Among these various factors, there is one that could explain a part of the pay gap even before the moment people enter the labour market: the field of study. Indeed, some diplomas are less rewarded on the labour market. This means that an educational gender segregation may contribute to the observed gap if women are overrepresented in fields of study characterised by lower incomes on the labour market for example.

Besides, the diplomas could not necessarily only have a direct impact on wages. The role of the fields of study on the gender wage gap could also be linked to the fact that some of them influence the type of occupations people will eventually enter, as well as the sector in which they will work. This means that a part of the wage gap that may be associated with educational segregation could be indirect and operate through other determinants.

In that context, this paper aims to better understand and quantify the direct and indirect roles played by field of study in explaining the gender wage gap inequality in Belgium. This would provide a quantitative base people can rely on to identify if it is necessary to encourage a more balanced distribution of men and women within different fields of study to reduce the wage inequality, and in which ones. A better understanding of these determinants can be useful to help design new policies aiming to reduce the pay gap.

2. Literature review

The literature on gender wage inequality shows that educational choices matter well before workers reach the labour market. Among graduates, men and women are not evenly distributed across fields of study, and these differences shape access to occupations, sectors, contract types, working time arrangements, and firms. For that reason, the field of study should not be treated as a marginal control variable. It is one of the main early sorting mechanisms which later leads to gender wage inequality. This point is central in the work of Breda and Napp (2019), Livanos and Pouliakas (2012), Hägglund (2024), as well as Sánchez Mangas and Sánchez Marcos (2021).

At the same time, the literature does not support a simple claim that field of study alone explains the gender wage gap. Its effect is partly direct, through differences in wage returns across fields, and partly indirect, through later sorting into jobs with different pay structures. This distinction is highlighted by Leuze and Strauß (2016), Petrongolo (2004), Ponthieux and Meurs (2005), as well as Matteazzi, Pailhé and Solaz (2013). This distinction is central to a thesis on the case of Belgium. The Belgian literature has produced useful evidence on firm-level wage discrimination and labour market wage setting, but it has said much less about how educational segregation by field of study contributes to the wage gap among graduates. This gap is clear in the Belgian contributions of Vandenberghe (2011, 2016). Therefore, the research gap is not whether gender wage inequality exists in Belgium, but rather how much of it can be linked to field of study directly, but also indirectly through channels such as occupation, sector, and working time.

2.1. Field of study as an upstream mechanism

A first strand of the literature explains why men and women enter different educational tracks in the first place. Breda and Napp (2019) show that gender differences in comparative advantage between reading and mathematics help explain why girls are less likely to choose math intensive studies. Using PISA data for 64 countries, they find that once the difference between performance in mathematics and reading is taken into account, the gender gap in intentions to pursue math related studies falls by roughly three quarters. Their argument is important because it links field segregation to earlier educational processes rather than to labour market events alone. In this view, gender wage inequality among graduates begins partly through gendered academic specialization.

This upstream perspective is really interesting in order to analyse wages because, as mentioned by Breda and Napp (2019), math intensive fields tend to lead to more highly rewarded labour market positions. The implication is not that women are less able in highly paid fields, but rather that educational choices reflect relative skills, expectations, and social norms in ways that, later, can affect earnings. The literature then suggests that, when analysing the gender wage gap among graduates, the field of study distribution should be considered as a primary factor and should be accounted before examining labour market variables.

2.2. Evidence that the field of study contributes to the wage gap

Several studies do quantify the role of the field of study, but the estimated magnitude varies substantially across settings, samples, and field classifications.

Livanos and Pouliakas (2012) provide one of the clearest estimates of the direct role of the field of study on gender wage inequality. Using Greek Labour Force Survey data, they show that women are overrepresented in subjects such as education and humanities, which are fields of study linked to lower wage returns. The results of their decomposition indicate that field of study explains 8.4 percent of the overall wage gap. This figure rises to 9.8 percent in the private sector, and to 23.6 percent among new entrants. This makes their paper particularly relevant, because it shows that educational segregation matters strongly even before career breaks or promotions can change wage profiles.

Sánchez Mangas and Sánchez Marcos (2021) provide a similar conclusion for young European graduates. Their study shows that gender wage inequality is not the same across fields, not only when entering the labour market, but also later in terms of wage growth. In economics, business, and law, their research estimates that there is a penalty for women's annual wage growth which is equal to 1.1 percentage points among childless graduates and 2.5 percentage points among those who become parents. This suggests that the field of study can influence future career paths as well as their associated initial wage levels.

The recent literature also shows that the effects of differences in study disciplines are highly heterogeneous and should not be treated as uniform. Michelmore and Sassler (2016) find that within STEM studies, the role of the observable elements differs significantly across subfields. In life and physical sciences, the observed characteristics account for most of the gender gap, whereas in computer science and engineering, a residual gap remains after controls, reaching about 8 to 12 percent in computer science. Zając et al. (2024) similarly find large differences across STEM fields in Poland, with women earning more than 20 percent less than men in the first year after graduation, with a gap above 25 percent in mathematics but below 3 percent in chemical and Earth sciences. These papers suggest that broad field categories may hide important variation and understate the contribution of educational segregation if the classification is too aggregated. In order to be more precise, fields of study should be analysed separately without grouping them in broader categories such as “STEM” for example.

Chevalier (2006) adds an important graduate level result by showing that later labour market expectations already matter at the start of the career. In the preferred model, career break expectations explain about 10 percent of the graduate wage gap. This does not quantify the contribution of the field of study directly, but it does quantify a closely related mechanism through which educational and occupational sorting may later affect wages.

Hägglund (2024) adds an important comparative dimension. Studying Finland and Germany, this paper shows that the field of study matters in both countries, but not exactly in the same way. In Finland, women appear to benefit more than men from fields that are strongly linked to the occupation, while in Germany the pattern differs. The broader implication is that the field of study does not carry a fixed wage effect across contexts. The importance of this issue depends on the connection between national education systems and labour markets, as well as the manner in which occupations reward different forms of specialisation.

Breda and Napp (2019), working upstream of labour market outcomes, show that controlling for comparative advantage between mathematics and reading reduces the gender gap in intentions to pursue math-related studies by about 75 percent. Their result is important because it quantifies an early mechanism that helps generate later segregation by field of study.

2.3. From field of study to labour market sorting

The literature also shows that field of study matters because it channels graduates into different labour market positions. In that sense, occupation, working time, and contract type are not alternative explanations to field of study. They are part of the mechanism through which educational segregation becomes wage inequality.

Chevalier (2006) provides an important contribution by showing that career expectations and anticipated interruptions already shape the graduate wage gap in the United Kingdom. In his preferred specification, career break expectations explain about 10 percent of the wage gap. This does not mean that expectations are purely individual choices. On the contrary, the paper suggests that gendered expectations about careers and family roles help produce different labour market paths even among similarly educated graduates. For a thesis focusing on the impact of the field of study, this is important because educational choices and later labour market behaviour are closely linked.

Leuze and Strauß (2016) make the channel argument more concrete. Their analysis of higher education graduates shows that lower pay in women-dominated occupations is tied to women typical work tasks and to working time arrangements. The implication is that some of the apparent wage penalty associated with educational specialization may actually reflect the characteristics of the jobs to which graduates from those fields are sorted. Field of study, occupation, and working time should therefore be analysed together rather than as isolated determinants.

Contract type is another important channel. Petrongolo (2004) shows that across 15 European countries, women are overrepresented in part-time work everywhere and are also more likely to hold fixed-term contracts in several countries, especially in Southern Europe. Part-time and temporary contracts are not equivalent. Temporary contracts are broadly associated with lower job satisfaction and weaker job quality, while part-time work is more voluntary in Northern Europe and more constrained in Southern Europe. This distinction matters because graduates from different fields may face different contract structures at labour market entry, and these structures can themselves contribute to wage inequality.

The literature treating the working time points in the same direction. Ponthieux and Meurs (2005) show that differences in hours worked are one of the most common contributors to the monthly gender wage gap in ten European countries. Matteazzi, Pailhé and Solaz (2013) extend that conclusion by showing that the mere prevalence of part time employment explains only a share of the gap, while segregation and the specific nature of part time jobs matter more. These findings are directly relevant for this paper. They imply that working time should not just be added mechanically as a control, because it may capture one of the routes through which the field of study affects wages.

2.4. Belgian literature and its limits

The Belgian evidence in the selected corpus is informative, but it does not directly answer the thesis question. Vandenberghe (2011) studies the Belgian private sector with firm level productivity data and finds little evidence of generalized within firm gender wage discrimination. The later Belgian analysis by the same author reaches a similar conclusion and suggests that any statistically significant discrimination is concentrated in particular groups, such as women under worker contracts, rather than being spread evenly across all employees.

These papers are important because they shift the focus of explanation. If the wage gap in Belgium is not mainly driven by broad within firm pay discrimination, then a larger share of inequality may arise through sorting across firms, occupations, sectors, and working time arrangements. That conclusion fits well with the broader European literature discussed above. However, the Belgian papers do not quantify the specific role of the field of study, nor do they focus on graduates. As a result, they leave open the question of whether educational segregation contributes to the Belgian gender wage gap before later labour market processes reinforce it.

This is the main gap the thesis can address. Belgium is not absent from gender wage gap research, but the educational origins of the gap among graduates remain underexplored in the selected literature. A study that quantifies the contribution of differences in field of study between men and women to the gender wage gap in Belgium, and that examines how this contribution changes once occupation, sector, and working time are introduced, would therefore connect two literatures that are usually studied separately: the literature on educational segregation and the literature on labour market wage inequality.

2.5. Methodological implications for the thesis

The methodological literature helps clarify how such a contribution can be made. Gelbach (2016) proposes a decomposition framework that attributes the change in a coefficient between a baseline model and an expanded model to specific covariates. This is especially appropriate in the context of this paper's research question. A standard wage regression can show that the gender coefficient falls when field of study, occupation, sector, and working time are added. A Gelbach decomposition can go further by showing how much of that fall is associated with each variable.

This feature is important because the thesis is not only interested in whether the field of study matters or not, but in how it matters relative to other variables. If occupation and sector reduce the estimated role of the field of study, that does not necessarily mean that its impact is not important. It may mean that the contribution of educational segregation to wage inequality also operates through those channels. Gelbach provides a transparent way to separate the direct contribution of study discipline from the contribution carried through later labour market sorting.

Tverdostup and Paas (2019) show the usefulness of this logic in a broader European setting. Using Gelbach decomposition on data from 17 countries, they find that work experience, occupation and industry specific experience, and numeracy all materially reduce the measured wage gap. Their study is not graduate specific and is not centered on the field of study, but it demonstrates how decomposition can uncover the relative importance of grouped explanatory factors.

At the same time, the literature about decomposition also shows an important limitation. Kunze (2008) explains that consistency and interpretation are central concerns in wage gap decomposition. Variables such as occupation, hours, contract type, and even skills are not necessarily exogenous. They may themselves reflect earlier gendered constraints and choices. For that reason, the decomposition results should be interpreted as descriptive attribution rather than definitive causal mediation. This caveat does not weaken the thesis. It clarifies the kind of claim that can be made responsibly.

2.6. Position of this thesis

The literature reviewed here supports several conclusions:

- Field of study has already been identified as a contributor to the gender wage gap in some countries, but the estimated magnitude varies widely. For instance, existing estimates range from 8.4 percent of the gap overall in Greece to 23.6 percent among Greek new entrants.
- There is a part of the effect of the field of study that likely operates indirectly through other channels, such as occupation, sector, contract type, and working time.
- The effect of field of study is heterogeneous across countries and across fields.
- In Belgium, even if the literature has rather well documented wage inequality, the impact of educational segregation on the gender wage gap has not really been quantified yet.

This thesis therefore takes a clear place in the literature. Its contribution is not simply to test whether the field of study matters in Belgium, since the broader literature already shows that it often does in Europe. Its contribution is to produce a Belgium specific estimate of how much field of study contributes to the graduate gender wage gap and to distinguish between its direct role and its indirect role (through other variables such as occupation, sector, and working time). A Gelbach decomposition is particularly appropriate for this objective because it allows the contribution of each group of variables to be assessed in a transparent and structured way, following Gelbach (2016).

3. Methodology

3.1. Empirical Strategy

As previously mentioned, this thesis aims to analyse the impact of the field of study on the gender wage gap in Belgium. In this context, the analysis focuses on two dimensions: the direct effects of the field of study on the wages, but also its indirect impact through three other determinants, which are:

- The occupation that people are engaged in
- The sector in which they work
- The working time

For that matter, the empirical strategy of this paper is based on a linear regression framework, as well as a decomposition using the method of Gelbach (2016) which allow us to decompose the wage gap in additive contributions so that it becomes possible to identify each determinant's role in this gap.

3.2. Dataset and Sample Selection

The empirical analysis made in this paper is based on a dataset provided by Vincent Vandenberghe, coming from *Vacature Salarisenquête*, which contains various labour market information for Belgium, such as the wage, the gender, the educational background including the field of study (represented by the certificate obtained by the person), the occupation and sector of activity, as well as other employment characteristics including the working time percentage. This information is contained in 158,579 entries at an individual level, covering three years: 2008, 2010 and 2012. As it is the most recent year in the dataset, the analysis only focuses on the year 2012 in order to use the latest data available.

In addition to limiting the dataset to the year 2012, it is necessary to apply some other restrictions in order to obtain an economically meaningful sample for the analysis.

First, the analysis only contains people who are currently employed and whose information about wage is not missing.

Then, the entries containing missing information on important variables such as gender, field of study, occupation, sector or working time, are excluded so that the final sample doesn't contain missing values.

Eventually, the sample is restricted to people who have completed higher education, which means people who possess either a bachelor's degree (or more) or a master's degree (or more). This is a key restriction in this analysis as it mainly focuses on the field of study. Individuals whose education level is lower than that do not meaningfully choose a field of study as secondary school is more general. Including them would introduce a bias in measuring the effects of the studies on the gender wage gap.

The implementation of this set of restrictions ensures that the final sample, containing 25,795 observations, only represents employed individuals who have attained higher education and possess complete information on all variables necessary for conducting decomposition analysis.

3.3. Econometric Models

3.3.1 Baseline Model

The starting model is a linear regression estimating the overall gender wage gap, conditional on age:

$$Y_i = \mu + \lambda^b D_i + \gamma Age_i + \epsilon_i \qquad (1)$$

Where Y_i is the logarithm of monthly wages, D_i is a gender dummy which is equal to one for women and zero for men, and Age_i is used to capture age effects to measure the gap at a given age.

The coefficient λ^b measures the gender wage gap at a given age, which is more meaningful to compare male and female individuals at similar stages of their life.

3.3.2 Full Model

Then, the model is extended to include the field of study, but also the other characteristics, i.e. occupation, sector and working time:

$$Y_i = \alpha + \lambda^f D_i + \gamma Age_i + Z_i \Gamma + \eta_i \qquad (2)$$

Where Z_i includes all the characteristics cited above, which means the field of study, the occupation, the sector of activity and the working time. The effect of variable Z_i on wages is represented by Γ, the "price effect".

The coefficient λ^f measures the gender wage gap at given age, field of study and at given labour market characteristics.

In this model, no interactions between gender and the other variables are included. This implies that potential shadow prices are not taken into account or analysed separately in the analysis, which means that potential differences in return to the characteristics between men and women are not considered.

3.3.3 Gelbach Decomposition

The objective is now to decompose the difference between the baseline model and the full one. In order to do this, this thesis applies the decomposition method proposed by Gelbach (2016), which is built on the omitted variable bias framework. An advantage of this approach is that the decomposition is independent of the order in which variables are introduced, which ensures that the contribution of each variable is interpreted consistently.

Consider the baseline model previously mentioned, written as follows:

$$Y_i = X_i \beta^b + \epsilon_i \tag{3}$$

Where X_i includes the gender indicator and the age, and β^b then corresponds to the gender wage gap conditional on age.

Consider also the full model including additional covariates, the vector Z omitted in the baseline model:

$$Y_i = X_i \beta^f + Z_i \Gamma + \eta_i \tag{4}$$

Where β^f represents the gender wage gap after controlling for the additional factors that Z includes, and Γ corresponds to their effect on wages.

The observation of Gelbach is that the difference between β^b and β^f can be expressed as follows:

$$\widehat{\beta^b} = \beta^f + (X'X)^{-1} X' Z \Gamma \tag{5}$$

Where $(X'X)^{-1}X'Z$ captures the extent to which the variables in Z is related to the variables already included in the baseline model. This term is denoted by $\hat{\rho}$, which gives:

$$\widehat{\beta^b} = \widehat{\beta^f} + \hat{\rho}\hat{\Gamma} \tag{6}$$

This formula can be further decomposed into the contribution of each variable Z, written as follows:

$$\widehat{\beta^b} = \widehat{\beta^f} + \sum_j \widehat{\delta_J} \tag{7}$$

Where each $\widehat{\delta_J}$ represents the contribution of the variable Z_j to the difference between the coefficient of the baseline model and the one of the full model, defined as $\widehat{\delta_J} = \widehat{\rho_J}\widehat{\Gamma_J}$.

In this formula, it is then possible to compute $\widehat{\beta^b}$ by estimating the baseline model, to compute $\widehat{\beta^f}$ and $\widehat{\Gamma_J}$ by estimating the full model, but also to compute $\widehat{\delta_J}$ by estimating auxiliary regressions of each Z_j on the baseline variables X, which provides the coefficients $\widehat{\rho_J}$ such that $\widehat{\delta_J} = \widehat{\rho_J}\widehat{\Gamma_J}$.

In this case, occupation, sector and working time are different sets of dummy variables, for which Gelbach's method proposes, for each set, to first construct a heterogeneity term $\widehat{H^g} = \sum_{k \in g} Z_k \widehat{\Gamma_k}$ and then to estimate an auxiliary regression of H^g on X, so that the coefficient associated to the gender represents the contribution $\widehat{\delta^g} = \sum_k \widehat{\rho_k}\,\widehat{\Gamma_k}$ of the set of dummy variables g to the gender wage gap.

3.4. Definition of Decomposition Channels

The variables included in Z_I are organised according to the main economic mechanism through which wage inequality may arise. Instead of grouping variables into broad categories, this thesis focuses on distinct channels that correspond to particular labour market processes. As previously mentioned in this paper, four main components are considered:

- Variable used to capture the direct effect of the field of study on wage inequality:
    - *Field of study*, which represents the central variable of interest, and reflects the differences in educational specialisation. It captures educational segregation.

- Variable used to capture the indirect effect of the field of study on the gap:
    - *Occupation*, which reflects the allocation of individuals across different jobs. It captures occupational segregation.
    - *Sector*, which captures the distribution of workers across different sectors of activity. It represents sectorial segregation.
    - *Working time*, which reflects the differences in labour supply intensity, particularly the percentage of working hours compared to a full-time position.

These variables are each treated separately in the decomposition process, so that their specific contributions to the wage inequality can be isolated and quantified individually. Then, it is possible to estimate the way educational segregation can operate indirectly through these components on the gender pay gap.

Some variables, such as educational level or area of work for example, are not included in order to maintain a better focus on field of study and its potential main transmission channels.

3.5. Identification of Indirect Effects of Field of Study

Although the Gelbach decomposition provides a breakdown of the gender wage gap into the contributions of different covariates, it does not really distinguish directly between direct effects and indirect ones. This section therefore focuses on identifying the extent to which labour market variables are associated with field of study.

As previously said in the paper, for each variable, a heterogeneity term $\widehat{H^g}$ is constructed, which represents the portion of wage explained by the variable. The total contribution of this component is then obtained by estimating an auxiliary regression on X, which means estimating:

$$\widehat{H^g} = \alpha_0 + \alpha_1 D_i + \alpha_2 Age_i + \varepsilon_i \qquad (8)$$

Where α_1 represents the contribution of the variable to the gender wage inequality after controlling for age, as mentioned above.

In order to evaluate the extent to which this contribution is related to the field of study, the regression can be extended by adding the field of study FOS to X, in order to estimate:

$$\widehat{H^g} = \beta_0 + \beta_1 D_i + \beta_2 Age_i + \theta FOS_i + \omega_i \tag{9}$$

Where β_1 captures the contribution of the variable to the gender wage gap after controlling for age, but more importantly, field of study.

The difference between the two coefficients α_1 and β_1 measures the share of the contribution of the variable that is associated with the field of study. In other words, it can provide an empirical assessment of the effect of differences in terms of educational specialisation operated indirectly through this other component. This step is, and should be interpreted as, a complementary extension of the Gelbach decomposition used to assess the extent to which the contributions of occupation, sector, and working time are associated with the field of study. It is not a new separate decomposition of the gender wage gap.

3.6. Estimation Procedure

The empirical implementation is conducted in accordance with a structured sequence that is consistent with the framework that has been described above.

Firstly, the baseline wage equation is estimated, only incorporating gender and age, in order to get a "reference measure" of the gender wage gap. Then, the full equation is estimated by adding the field of study and the other labour market variables, allowing the corresponding coefficients to be recovered.

After that, the Gelbach decomposition is implemented by constructing heterogeneity terms, for each channel, and estimating the associated auxiliary regressions. This step is intended to determine the contribution of each variable to the gender wage gap.

Finally, in order to assess the indirect role of field of study, the auxiliary regressions are extended by including educational specialisation. A comparison of the gender coefficients across these specifications enables the proportion of each channel associated with field of study to be determined.

All regressions are estimated using Ordinary Least Squares (OLS) on a consistent sample across specifications.

3.7. Limitations of the Methodology

The analysis is based on a number of methodological choices which, whilst appropriate and offering many advantages, also have certain limitations.

Firstly, the approach is based primarily on OLS regressions, which is a standard methodological choice in labour economics that is highly consistent with the objective of decomposition, and which allows a clear and straightforward interpretation of the results. However, one of the main limitations of this method is that it deals with conditional, rather than causal, relationships. Typically, certain variables, such as preferences or abilities, for example, are not taken into account in the data analysis, even though they also could explain and be correlated with wages, gender or field of study. Furthermore, whether for the field of study or other labour market variables, these are treated here as exogenous, whereas they could themselves be partly determined by wages, for example. Although these limitations are common in most cross-sectional analyses, the use of OLS is justified here given the objective of the analysis in terms of decomposition.

Secondly, the decomposition is based on Gelbach's method (2016) and is used to determine the contribution of each variable to the gender pay gap. An advantage of this method is that it is particularly useful for analyses in which variables are strongly correlated, as might be expected in this case, as it gives order-invariant estimates which do not depend on which controls are introduced before the others. It means that the order of the variables does not need to be taken into account. Nevertheless, once again, the results obtained cannot really be interpreted as causal effects. Indeed, Gelbach's decomposition method is associative and reflects the way in which the variables are correlated with gender and how they impact wages within the model, which does not imply a causal relationship. Furthermore, unlike the Oaxaca-Blinder method, Gelbach's method does not separate explained wage differences from unexplained ones, which therefore does not allow us to quantify discrimination, for example. In this thesis, the analysis focuses on quantifying the effects captured by the Gelbach decomposition, the shadow prices are not considered nor analysed. This means that potential differences in return between men and women for a same characteristic are not analysed, the paper only focuses on assessing how differences in characteristics may contribute to the gender wage gap. Nevertheless, Gelbach's decomposition method remains appropriate for this analysis, given that the objective is not to isolate a residual gap or analysing shadow prices, but rather to understand the significance of certain mechanisms on the wage gap. The aim here is not to study discrimination, but rather segregation.

Thirdly, the method used to identify the indirect effects of the field of study on the pay gap also has certain limitations. It relies on an econometric method based on auxiliary regressions, which allow occupation, sector and working hours to be interpreted as channels through which the field of study may affect wages. This therefore makes it possible to measure the extent to which these components are associated with the field of study, but it does not allow us to know for certain whether the field of study directly affects these variables. The results of this method depend on the direction in which the analysis is conducted. Here, it is assumed that the field of study has an effect on the other variables, whereas intermediate variables may be missing or the effects may work both ways: an individual may well choose their field of study with the aim of entering a particular sector or securing a particular job; in this case, the field of study would itself be influenced by these variables. The results should therefore be interpreted as providing indications on the role of observable mechanisms rather than as actual evidence of a causal effect.

Finally, the dataset may also have certain limitations. Apart from the fact that errors may creep into the sample, as is the case with many survey-based datasets, the analysis is based on a dataset restricted to individuals who have completed higher education. This choice was made deliberately in order to measure more accurately the impact of the choice of field of study on the gender pay gap, but it significantly limits the number of observations in the dataset. Furthermore, in order to use the most recent year possible, the dataset is further restricted to include only observations from 2012, which significantly reduces the sample size compared to the base dataset and may make it less representative. The use of a single year also prevents us from observing trends and potential changes in the results over time.

In conclusion, these limitations of course do not invalidate this analysis, but they rather define its scope. The results should be interpreted as an econometric decomposition based on observable characteristics rather than as a prediction of causal relationships.

4. Results and Interpretations

4.1. Descriptive statistics

This section presents some characteristics of the sample used for the decomposition as well as some information about gender differences observed in the distribution across fields of study, as well as differences in wage premia associated with these educational specialisations.

4.1.1 Sample Composition and Raw Gender Wage Gap

As previously mentioned in the methodology section, the final sample contains 25,795 entries of employed individuals who have completed higher education and for whom information is available on wages, gender, age, field of study, occupation, sector and working time. These observations were made in Belgium in 2012.

The share of women in the selection is about 42.59% of all the entries and their average gross monthly wage is 2,968.67 €, while men represent about 57.41% of the sample and earn on average 3,906.90 € per month, which gives a wage gap of 938.23 € to the detriment of women. This gap is raw and has not been controlled with any variable yet.

4.1.2 Gender Segregation Across Fields of Study

Men and women are far from being evenly distributed across the different fields of study, as it can be noticed in the following table representing the share of women who obtained a diploma in each academic branch, according to the sample. For example, the most feminised field is the bachelor's degree in secretary studies, where women represent almost 84% of the graduates, while the least feminised one is the master's degree in computing science, where they represent barely 8% of the graduates. Only 14 educational specialisations out of the 40 in the sample have a share of female graduates between 40% and 60%, which shows that most fields of study are characterised by a considerable asymmetry regarding the gender of the degree holders. This segregation could contribute to the gender wage inequality if women are overrepresented in fields that are associated with a lower salary on the labour market.

FIELD OF STUDY	WOMEN SHARE (%)
bac-Secretary	83.9
bac-Tourism, entertainment	75.9
mast-Journalist, translator	75.5
mast-Psychology	74.6
bac-Social assistant	71.7
mast-Languages, Philology	71.1
bac-Nurse, midwife	71.0
mast-Biomedical sciences	70.2
bac-Medical assistant	69.7
mast-Interior design	69.7
mast-Pharmacists	68.2
mast-Medical studies (Physicians, Dentists)	64.3
bac-Teaching (primary, kindergarten)	63.9
mast-Social Sciences	61.6
mast-Criminology	59.7
mast-History, History of art, Musicology, Archaeology	59.2
mast-Veterinarians	59.1
mast-Medico-social sciences (nurses)	58.9
bac-Architecture	55.9
mast-Fine arts (theater, design, music, visual arts)	53.6
bac-Media, publicity	52.6
bac-Teaching(low secondary)	50.2
mast-Law	50.1
mast-Philosophy, Theology	50.0
bac-Business assistant (bookkeeping, tax, marketing, insurance)	46.7
mast-Other	45.1
mast-Physical education, physiotherapy, speech therapy	44.0
mast-Accountant	42.1
mast-Natural sciences (biology, maths, chemistry, physics, geology, geography)	38.2
bac-Accommodation, restauration	37.8
mast-Business, applied economics	37.7
mast-Economics	36.8
mast-Bio Engineer, agricultural engineer	36.5
mast-Sales engineer	32.8
bac-Construction, wood	12.1
mast-Engineer	10.8
mast-Industrial engineer	10.6
bac-Computing IT(A1)	9.9
bac-Technician (audio, photo, electricity)	9.7
mast-Computing science	7.8

Figure 1 – Educational segregation: representation of women across fields of study

While the share of women within each field of study already tells the extent to which each discipline is feminised or not, it does not indicate the relative importance of each field among male and female graduates, which could also bring useful insights. The figure below therefore shows the distribution of fields of study within the total population of men and women in the sample.

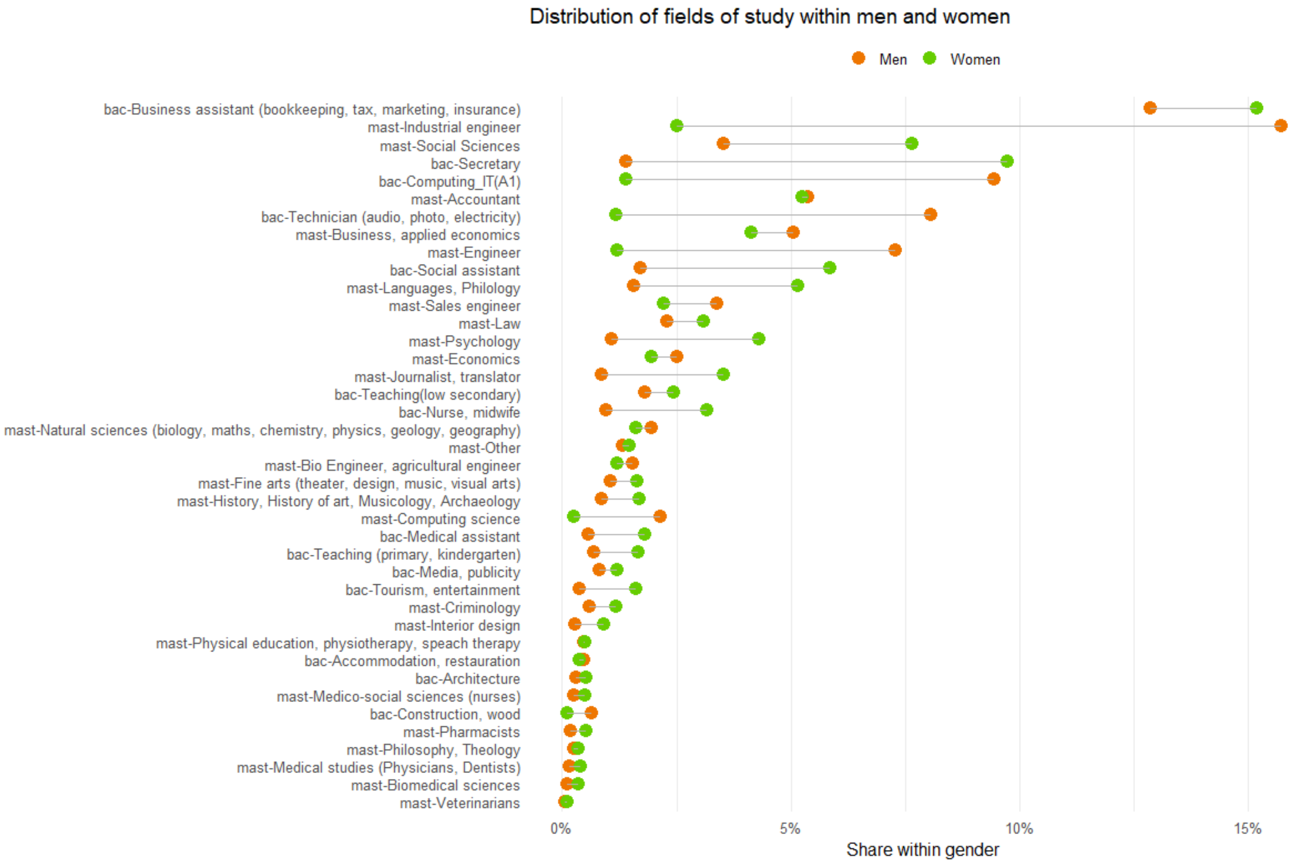

Figure 2 – Distribution of fields of study among men and women

This distribution gives another perspective on educational segregation. It indicates that some fields of study are popular among men but not particularly attractive for women, such as industrial engineering, information technology and technical studies for example. On the other hand, some other fields seem to be more common among women than among men, such as social sciences, secretary studies or studies in social work for example.

4.2. Baseline and Full Regression Results

4.2.1 Baseline Model

This section presents the results of the baseline regression, in which the logarithm of gross monthly wages is regressed on gender and age only, as previously mentioned in the methodology section.

These results indicate that, in the baseline model, the estimated coefficient β_b associated with the female dummy is about -0.21089. As the wage variable is expressed in logarithms in the model, it is possible to convert this coefficient into approximate percentage terms, which indicates that women approximately earn 21.09% less than men when only conditioning on the age. This β_b coefficient is the reference measure of the total gender wage gap that will be used for the rest of the analysis. It is the wage inequality that the Gelbach decomposition tries to explain by quantifying the part of it that can be attributed to the field of study, as well as the occupation, the sector and the working time percentage.

	Regression Results				
Variable	**Coefficient**	**Std. Error**	**t-value**	**p-value**	**Significance**
(Intercept)	7.97008	0.30861	25.82597	0.00000	***
fem	-0.21089	0.00391	-53.92935	0.00000	***
Age	Included				
Note:					
Dependent variable: log gross monthly wage.					
Age fixed effects are included in the regression but not individually reported.					
Significance levels: *** p < 0.001, ** p < 0.01, * p < 0.05, . p < 0.10.					

Figure 3 – Baseline regression results

4.2.2 Full Model

This section presents the results of the full regression which, as explained in the methodology section, extends the baseline model by including the following variables: field of study, occupation, sector of activity, and working time.

After controlling for these components in addition to those included in the basic model, the coefficient β_f associated with the gender dummy drops to about -0.06585, which gives a conditional wage gap of approximately 6.59% to the detriment of women, when converting the coefficient into percentage terms.

This residual gender wage gap remaining after controlling for all these variables can be interpreted as the part of the wage inequality that is not captured by the observable characteristics included in the model. It may be due to other relevant factors that are not observed here and does not necessarily reflect discrimination.

Variable	Regression Results				
	Coefficient	Std. Error	t-value	p-value	Significance
(Intercept)	6.77254	0.22838	29.65437	0.00000	***
fem	-0.06585	0.00346	-19.02576	0.00000	***
Age	Included				
Working time	0.01331	0.00017	78.50668	0.00000	***
Occupation	Included				
Sector	Included				
Diploma	Included				
Note:					
Dependent variable: log gross monthly wage.					
Age, field of study, sector, and occupation fixed effects are included in the regression but not individually reported.					
Significance levels: *** p < 0.001, ** p < 0.01, * p < 0.05, . p < 0.10.					

Figure 4 – Full regression results

4.2.3 Differences Between the Models and Reduction of the Gap

The difference between the baseline model and full model coefficients associated to the female dummy is equal to -0.14504 log points, which represents the part of the baseline coefficient that can be associated with the field of study, sector, occupation and working time. It represents an important part of the wage inequality observed between men and women. Indeed, considering the baseline coefficient of -0.21089, this reduction in the gap represents 68.78% of the total gender wage gap. It other words, this means that 68.78% of the gender wage gap can be associated to the variables mentioned above, while 31.22% of the gap remains after controlling for these components.

Even if this comparison gives a first insight of the role of the observable characteristics on the wage gap, it does not quantify how much each variable contributes individually. This will be addressed in the section about Gelbach's decomposition.

4.3. Gelbach Decomposition of the Gender Wage Gap

This section contains the results of the Gelbach Decomposition, which attributes the difference between the coefficients associated with the gender dummy in the baseline and full models to the different variables included in the analysis. In other words, it decomposes the reduction of 0.14504 log points into several contributions associated with the field of study, the sector, the occupation and the working time, in order to quantify the extent to which these components are associated with the total gender wage gap in Belgium.

As previously stated in the methodology section, for each variable, the decomposition combines two elements:

- ρ, which assesses how men and women are distributed differently across the categories of the variable
- Γ, which measures the wage premia that is associated with these categories

The contribution of each variable is then given by $\delta_j = \rho_j \Gamma_j$.

4.3.1 Wage Premia by Field of Study

After estimating the full model, it is possible to determine the coefficients Γ of the different variables. For the field of study, these coefficients represent the wage premia associated to each educational specialisation. The following figure represents the conditional relative gross wage premia of the different disciplines.

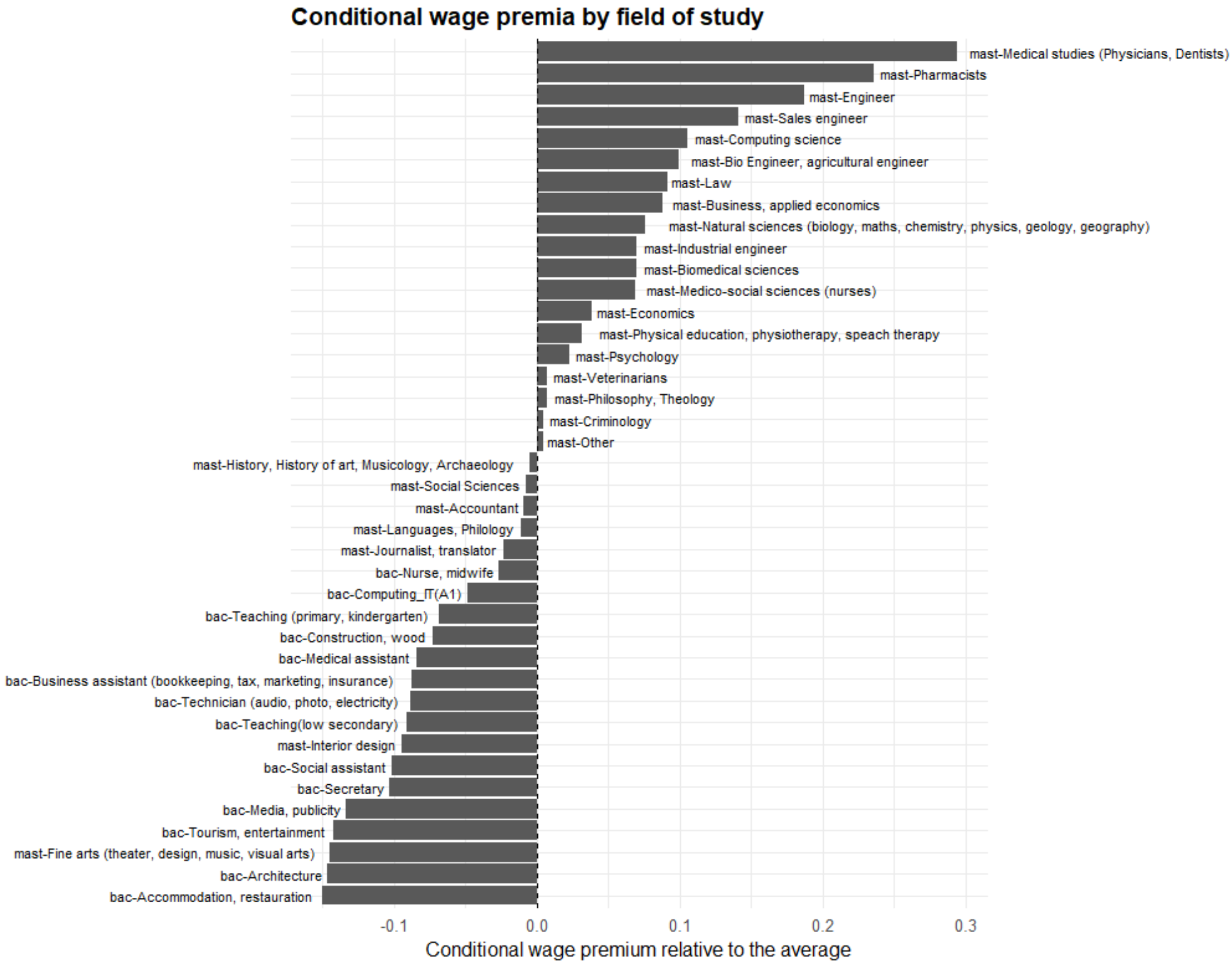

Figure 5 – Wage differential by field of study

The graphic indicates important differences in wage returns across the different educational specialisations. For example, some fields such as medical or pharmacy studies are characterised by an important positive wage premium of up to almost 0.3 log points (approximately 30%) compared to the average for the medical studies. Conversely, other degrees such as in catering or fine arts are characterised by a negative wage premium of up to -0.15 log points (approximately -15%) compared to the average.

In order to have a first insight of the potential contribution of the field of study to the gender wage gap, this figure can be compared to the table in section *4.1.2* of this paper (*Figure 1)*, reporting the share of women by fields. For example, secretary studies, which is a discipline were women are strongly overrepresented with a women share of 83.9%, are characterised by a negative wage premium of approximately -0.10 log points (approximately -10%) relative to the average. On the other hand, engineering studies, a field where men are heavily overrepresented and where women account for only 10.8% of the graduates, are characterised by a positive wage premium of about 0.18 log points (approximately 18%) above the average.

4.3.2 Contribution of Each Component

This section reports the estimated contributions of each variable that have been introduced in the full model to extend the baseline regression.

a) Field of Study

The estimated contribution of the field of study on the gender wage gap is δ_{dipl} = -0.03546, which means that the educational segregation is associated with a wage penalty of approximately 3.55% for women. Compared to the baseline coefficient β_b of -0.21089, this represents:

$$\frac{-0.03546}{-0.21089} * 100 = 16.81\%$$

of the total gender wage gap. In other words, the difference in distribution of men and women across fields of study is associated to approximately 16.81% of the gender wage inequality. These results show that educational segregation accounts for an important part of the gender wage inequality, even when controlling for labour market characteristics such as working time, sector and occupation.

The figure below offers a graphical illustration of this contribution δ_{dipl}:

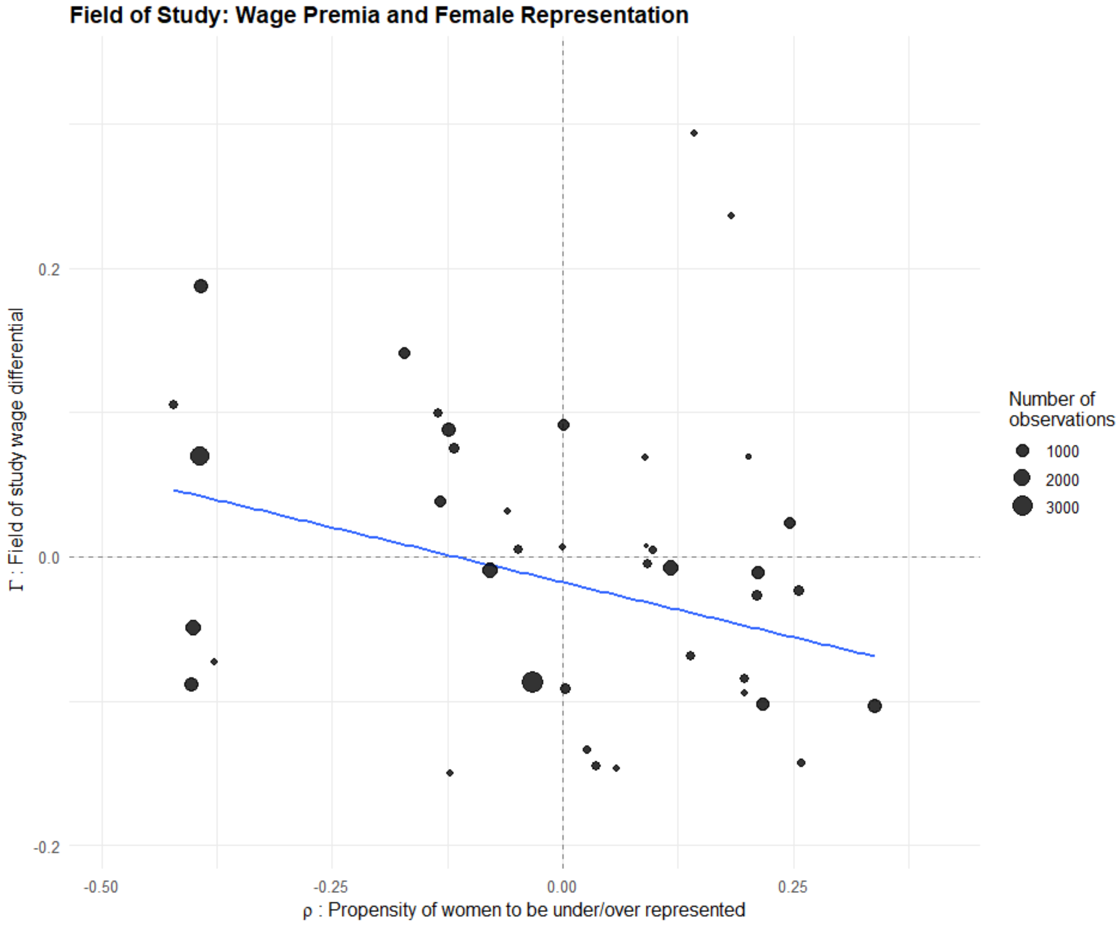

Figure 6 – Contribution of educational segregation to the gender wage gap: relation between women's representation and wage premia across fields of study

As previously explained, the contribution of the field of study combines two elements, ρ and Γ. Each point in this scatter plot represents a specific study discipline, with the horizontal axis showing its ρ and the vertical axis giving its Γ, and can be interpreted as an individual contribution $\delta_k = \rho_k \Gamma_k$. The overall contribution of the fields of study to the gender wage gap δ_{dipl} is given by summing all these individual contributions. The size of each point is proportional to the number of observations of this field of study in the sample. Most of the observations are located in the top-left and bottom-right quadrants of the graph, which gives a clear illustration of the way educational segregation can be associated with gender wage inequality. Indeed, the upper-left quadrant represents the fields in which women are underrepresented and which are associated with positive wage premia, while the lower-right quadrant indicates disciplines in which women are overrepresented and which are associated with negative wage premia. To confirm this pattern, the regression line has a downward slope, which shows a negative relationship between the representation of women in each study discipline and the corresponding wage premia.

b) Other Variables

The contribution of occupation to the gender wage inequality is estimated at δ_{occ} = -0.04009, representing about 19.01% of the total wage gap, when comparing with the baseline coefficient. This means that occupational segregation is also associated with a substantial part of the gender pay gap, corresponding to a reduction in women's salaries of approximately 4.01%.

Besides, sectoral segregation contributes to the gender wage gap as well, even if its impact is more limited. Indeed, the contribution of the sector variable is estimated at δ_{sect} = -0.01540. Compared to the baseline coefficient, it represents a share of nearly 7.30% of the total gender wage inequality, reducing women's wages by approximately 1.54% in comparison to men.

Finally, the working time seems to be the component contributing the most to the gender wage gap with an estimated contribution of δ_{wkpc} = -0.05409, which corresponds to about 25.65% of the total wage inequality between men and women. In percentage terms, it represents a reduction in women's wages of approximately 5.41% compared to men.

These results suggest that women and men are unevenly distributed across the different occupations, sectors and working time arrangements. More precisely, women are more concentrated in jobs and industries that offer a lower remuneration and are also more likely to work part-time. By contrast, men are more concentrated in higher-paying occupations and sectors and are more likely to work full-time.

c) Overview

The following table summarises the main results of the Gelbach decomposition. Overall, the sum of the estimated contribution coefficients (δ_j) is equal to -0.14504, which exactly corresponds to the previously computed difference between the coefficients associated with the female dummy in the baseline model and in the full model. This confirms that the decomposition is consistent with the properties of Gelbach's method, and that the estimated contribution coefficients are coherent.

Component	Estimated coefficient	Approx. effect on women's wages (%)	Share of the total gender wage gap (%)
Beta baseline model	-0.2108949	-21.089491	100.000000
Beta full model	-0.0658510	-6.585100	31.224559
Working percentage	-0.0540905	-5.409054	25.648099
Sector	-0.0154015	-1.540152	7.302935
Occupation	-0.0400912	-4.009115	19.010014
Field of study	-0.0354607	-3.546070	16.814393

Figure 7 – Results of the Gelbach decomposition

It is also important to state that the working time, sector and occupation variables are not independent of the degree obtained. Depending on the field of study followed by the individuals, it is reasonable to assume that these variables can differ, as some educational specialisations can lead to specific occupations or sectors for example. As a result, a part of the contributions of these components to the gender wage gap may also reflect an indirect effect of educational segregation. The next section explores and addresses this by assessing the extent to which the impact of the field of study on the gender wage inequality may occur through these channels.

4.4. Direct and Indirect Effects of the Field of Study

The previous Gelbach decomposition estimated the direct contribution of the field of study on the gender wage gap, as well as the contributions of the other variables included in the model. However, since the degree obtained by an individual may later influence the job and sector in which they work, as well as their working time, a part of the contributions of these variables on the wage gap may be associated with differences in the study discipline as well.

As mentioned in the methodology section, in order to assess this, the heterogeneity terms are regressed on the baseline variables as well as the field of study. The coefficient associated with the female dummy estimated by this regression is then compared to the contribution coefficient estimated in the previous section. The difference between the two coefficients reflects the part of the variable's contribution that can be associated with differences in field of study, which are interpreted as "indirect effects" of field of study on the gender wage gap.

In order to later verify the coherence of the results, an additional regression of the logarithm of the gross monthly wage is estimated, in which the field of study variable is added to the baseline components. The coefficient associated with the gender dummy is equal to -0.15136 in this model, compared to -0.21089 in the baseline regression. The difference between the two coefficients is therefore equal to -0.05953 log points.

4.4.1 Indirect Contribution

The contribution coefficient of occupation after controlling for the field of study in addition to the baseline variables is equal to -0.03296. By subtracting this value from the initial contribution of occupation δ_{occ} = -0.04009, a difference of -0.00713 log points is obtained. This difference can be interpreted as the part of the occupation's contribution to the gender wage gap that may be associated with differences in field of study. In relative terms, it implies that approximately 17.78% of the contribution of occupation to the gender wage gap may be associated with field of study differences.

As for the sector, after controlling for the field of study, the coefficient of contribution to the gender wage gap is equal to -0.00638. Compared to the initial contribution of the sector variable δ_{sect} = -0.01540, it implies a reduction of -0.00902 log points, which can be interpreted as the part of the sector's contribution to the wage inequality between men and women that may be associated to the field of study. In relative terms, this implies that approximately 58.57% of the contribution of the sector to the gender wage gap may be associated with field of study differences.

By applying the same reasoning to the working time variable, after controlling for the field of study, the contribution coefficient associated with this component equals -0.04617, compared to an initial contribution of δ_{wkpc} = -0.05409. The difference between these coefficients indicates that there is a part of -0.00792 log points of the working time contribution that may be associated with differences in fields of study. In relative terms, this implies that approximately 14.64% of the contribution of the working time to the gender wage gap may be associated with field of study differences.

Labour market mechanism	Contribution to gender wage gap in the full model	Contribution controlling for field of study	Indirect effect of field of study
Working time	-0.05409	-0.04617	-0.00792
Sector	-0.01540	-0.00638	-0.00903
Occupation	-0.04009	-0.03296	-0.00713

Figure 8 – Indirect contribution of the field of study variable to the gender wage gap through other variables

The sum of the contributions found after controlling for the field of study is equal to -0.08551, which is equal to the difference between the coefficient of the additional regression extending the baseline by adding the field of study (-0.15136), and the coefficient of the full model (-0.06585). This confirms the internal consistency of this extension of the decomposition.

Taken together, these results imply that a total of -0.02407 log points of the gender wage gap may be indirectly associated with differences in field of study through occupation, sector and working time. Relative to the baseline coefficient (-0.21089), it represents approximately 11.41% of the gender wage gap.

Given that the combined contribution of the occupation, sector and working time accounts for about 51.96% of the total gender wage gap, this means that approximately 11.41 percentage points may be associated with differences in study disciplines. In other words, about 21.96% of the combined contribution of these three factors to the total gender wage gap appears to be associated with educational segregation.

4.4.2 Total Contribution

The total contribution of the field of study to the gender wage gap can be obtained by combining its direct contribution δ_{dipl} and its indirect effects estimated in the last section. This gives a total contribution of:

$$\delta_{total\ dipl} = -0.03546 + (-0.02407) = -0.05953$$

log points, suggesting that differences in distribution across field of study between genders may be associated with a wage penalty of approximately 5.95% for women relative to men. Compared to the baseline coefficient, this contribution represents approximately 28.23% of the total gender wage gap, including a "direct effect" representing approximately 16.81% of the gap, and an "indirect effect" which accounts for approximately 11.41% of the gap through the influence of educational segregation on occupation, sector and working time.

However, this total contribution should be interpreted with caution as it does not represent an exact causal measure of the effect of the field of study on the gender wage gap. Instead, it gives an extended measure of the share of the gender wage gap that may be associated with field of study differences, both directly and indirectly through their association with occupation, sector, and working time. Therefore, this estimate should be understood as an indication of the broader role of educational segregation, rather than as a precise causal effect. Also, it is important to mention that this total contribution should not be added to the individual contributions of occupation, sector, and working time that have been reported in the main Gelbach decomposition, since it already includes the parts of these labour market channels that are associated with educational specialisation. It should instead be interpreted as an alternative estimate of the overall contribution to the gender wage inequality that may be associated with educational segregation.

Overall, these results suggest that the role of educational segregation in gender wage inequality seems to extend beyond its direct contribution to wages, by influencing other labour market characteristics that contribute to this wage gap.

Conclusion

The goal of this thesis was to answer the following research question: To what extent do differences in field of study between men and women contribute to gender wage inequality among higher education graduates in Belgium? More specifically, the objective was to quantify both the direct contribution of educational segregation as well as the extent to which this contribution operates indirectly through the occupation, the sector of activity, and working time arrangements. The purpose of the analysis was to assess how differences in educational specialisations may be associated with the gender wage gap.

The empirical results indicate that, among the graduates in Belgium, the total gender wage gap conditional on age amounts to approximately 21.09% to the detriment of women. After controlling for field of study, occupation, sector, and working time, this gap falls to approximately 6.59%, which is considered as the residual wage gap that remains unexplained by the observable components included in the model. This residual gap may reflect the influence of unobserved factors. These figures suggest that the observable characteristics mentioned above may be associated with a difference between the wages of men and women of about 14.5%. This represents approximately 68.78% of the total gender wage gap.

The sample reveals a substantial gender segregation across the different fields of study, and the Gelbach decomposition shows that this mechanism is associated with 16.81% of the total gender wage gap, while occupation, sector and working time are in turn associated with 19.01%, 7.30% and 25.65% of the overall gap, respectively. A part of these contributions may itself be linked to differences in field of study between men and women, as the analysis suggests that an additional 11.41% of the gender wage gap may be indirectly associated with educational segregation through occupation, sector, and working time. More specifically, educational segregation appears to be associated with approximately 17.78% of the contribution of occupation, 58.57% of the contribution of sector, and 14.64% of the contribution of working time. Taken together, these results suggest that differences in field of study may be associated with approximately 28.23% of the gender wage gap among higher education graduates in Belgium, of which 16.81% corresponds to a direct contribution and 11.41% to an indirect contribution through occupation, sector, and working time.

These findings give clear elements to answer the research question. Differences in field of study between men and women appear to play an important role in gender wage inequality. Moreover, the contribution of educational segregation extends beyond the direct wage premia associated with different degrees. Fields of study also seem to participate in shaping subsequent labour market trajectories, by influencing the occupations individuals enter, the sectors in which they work, and their working time arrangements. In this way, a substantial share of the gender wage gap may be influenced even before entering the labour market, through educational choices and the unequal distribution of men and women across study disciplines.

This conclusion has important implications for public policy. Policies aimed at reducing gender wage inequality should not only focus on labour market outcomes such as occupational segregation, sectoral sorting, or part-time employment. Some interventions could occur earlier in the educational process in order to encourage a more balanced representation of women and men across the different fields of study that are highly segregated. For example, organising talks with women engineers and computer scientists in schools may highlight the contribution of women in the STEM sector and encourage girls and young women to undertake studies in those fields. In Belgium, where women remain strongly underrepresented in disciplines such as engineering, computer science, and technical studies, increasing female participation in these relatively highly rewarded fields could contribute to reducing wage disparities over the longer term. Such interventions would not replace labour market policies aimed at promoting equal pay, improving access to full-time employment, or reducing occupational segregation, but rather complement them by acting on one of the upstream mechanisms through which gender wage inequality appears to be generated.

At the same time, the results should be interpreted with caution. First, the analysis is based on Ordinary Least Squares regressions and Gelbach's decomposition, which identify descriptive and associative relationships rather than causal effects. Indeed, for example unobserved factors such as preferences, abilities, family expectations, or social norms may influence both educational choices and wages. Second, the analysis does not incorporate interactions between gender and the explanatory variables. This means that potential differences in returns to the same observable characteristic between men and women (shadow prices) are not examined. Third, the estimation of indirect effects relies on an extension of the decomposition that provides an approximate measure of the extent to which the contributions of occupation, sector, and working time are associated with field of study, but it should not be interpreted as a formal causal mediation analysis. Finally, even if the field of study is considered as given in this analysis, in reality, it cannot be considered as exogenous, as educational choices are the outcome of a long process shaped by earlier social, educational, and institutional influences (family background, school experiences, stereotypes, or academic performance for example). They may also even reflect anticipated labour market outcomes, as some students choose their field of study partly on the basis of the occupations or sectors they hope to enter.

However, these limitations do not undermine the relevance of the findings, but they help define their scope. The results provide an econometric decomposition of the gender wage gap based on observable characteristics and offer a clear indication of the relative importance of educational segregation within this framework.

Several research paths for future research emerge from this study. First, it would be interesting to include interactions between the gender and the other components, in order to examine whether men and women receive different returns to the same characteristics. That would allow to analyse potential discrimination rather than focusing on segregation. Then, the analysis could be extended using more recent data to assess whether the contributions of educational segregation, as well as the other components, have evolved over time. A future work could also assess the upstream determinants of gender differences in educational choices, or it could evaluate the contribution of other determinants to the gender wage gap, such as firm characteristics, contract type, or career interruptions.

Overall, this thesis suggests that educational segregation is an important mechanism associated with gender wage inequality in Belgium. The results indicate that differences in field of study between men and women may be associated with gender wage inequality both directly, through differences in wage premia across fields of study, and indirectly, through later labour market sorting into different occupations, sectors, and working time arrangements. Reducing gender wage inequality may therefore require attention not only to what happens in the labour market, but also to the educational pathways that can shape individuals' opportunities before they even enter employment.

Bibliography

- Breda, T. and Napp, C. (2019), "Girls' comparative advantage in reading can largely explain the gender gap in math-related fields", *Proceedings of the National Academy of Sciences*, vol. 116, pp. 15435–15440, DOI: 10.1073/pnas.1905779116.

- Chevalier, A. (2007), "Education, Occupation and Career Expectations: Determinants of the Gender Pay Gap for UK Graduates", *Oxford Bulletin of Economics and Statistics*, vol. 69, pp. 819-842, DOI: 10.1111/J.1468-0084.2007.00483.X.

- Gelbach, J.B. (2016), "When Do Covariates Matter? And Which Ones, and How Much?", *Journal of Labour Economics*, vol. 34, pp. 509-543, DOI: 10.1086/683668.

- Hägglund, A.E. (2024), "Same degrees, different outcomes? Fields of study choices and gender wage inequality in Finland and Germany", *Social science research*, vol. 122, pp. 103029, DOI: 10.1016/j.ssresearch.2024.103029.

- Kunze, A. (2008), "Gender wage gap studies: consistency and decomposition", *Empirical Economics*, vol. 35, pp. 63–76, DOI: 10.1007/S00181-007-0143-4.

- Leuze, K. and Strauss, S. (2016), "Why do occupations dominated by women pay less? How 'female-typical' work tasks and working-time arrangements affect the gender wage gap among higher education graduates", *Work, Employment & Society*, vol. 30, pp. 802–820, DOI: 10.1177/0950017015624402.

- Livanos, I. and Pouliakas, K. (2012), "Educational segregation and the gender wage gap in Greece", *Journal of Economic Studies*, vol. 39, pp. 554–575, DOI: 10.1108/01443581211259473.

- Matteazzi, E., Pailhé, A. and Solaz, A. (2013), "Does Part-Time Employment Widen the Gender Wage Gap? Evidence from Twelve European Countries", *ECINEQ, Society for the Study of Economic Inequality*, working paper 293, URL: http://www.ecineq.org/milano/WP/ECINEQ2013-293.pdf.

- Michelmore, K. and Sassler, S. (2016), "Explaining the Gender Wage Gap in STEM: Does Field Sex Composition Matter?", *RSF: The Russell Sage Foundation Journal of the Social Sciences*, vol. 2, pp. 194-215, DOI: 10.7758/RSF.2016.2.4.07.

- Sánchez-Mangas, R. and Sánchez-Marcos, V. (2021), "Wage Growth across Fields of Study among Young College Graduates: Is There a Gender Gap?", *CESifo Economic Studies*, vol. 67, pp 251–274, DOI: 10.1093/CESIFO/IFAA021.

- Petrongolo, B. (2004), "Gender Segregation in Employment Contracts", *Journal of the European Economic Association*, vol. 2, pp. 331-345, URL: https://www.jstor.org/stable/40004908.

- Ponthieux, S. and Meurs, D. (2005), “The Gender Wage Gap in Europe: Women, Men and the Public Sector”, *INSEE*, *Direction des Statistiques Démographiques et Sociales*, working paper F0502, URL: https://www.bnsp.insee.fr/ark:/12148/bc6p06zqtdv.pdf.

- Tverdostup, M. and Paas, T. (2019), “Valuation of Human Capital and the Gender Wage Gap in Europe,” DOI: 10.2139/ssrn.3460494.

- Vandenberghe, V. (2011), “Firm-level Evidence on Gender Wage Discrimination in the Belgian Private Economy”, *LABOUR*, vol. 25, pp. 330-349, DOI: 10.1111/J.1467-9914.2011.00524.X.

- Vandenberghe, V., “Women Make a Fraction of Every Euro Earned by Men...”, *Reflets et perspectives de la vie économique*, vol. 4, pp. 11-22, DOI: 10.3917/RPVE.554.0011.

- Vandenberghe, V. (2025-2026), “LECON2608: GENDER and MIGRANT labour inequalities”, *Economic School of Louvain, UCLouvain*.

- Zając, T., Magda, I., Bożykowski, M., Chłoń-Domińczak, A. and Jasiński, M. (2024), “Gender pay gaps across STEM fields of study”, *Studies in Higher Education*, vol. 50, pp. 126–139, DOI: 10.1080/03075079.2024.2330667.

Appendices

A: Summary of Regression Results

Model	Intercept	Intercept std. error	Gender coefficient	Gender coef.td. error	t-value	p-value	Approx. wage gap (%)	R-squared	Adj. R-squared	Observations
Baseline model	7.970082	0.3086073	-0.2108949	0.0039106	-53.92935	0	-21.08949	0.3761551	0.3748220	25795
Baseline + field of study	7.887808	0.2818528	-0.1513595	0.0040769	-37.12568	0	-15.13595	0.4819478	0.4800530	25795
Full model	6.772540	0.2283825	-0.0658510	0.0034611	-19.02576	0	-6.58510	0.6715483	0.6682818	25795

Figure 9 – Results of the regressions

	Baseline model	Baseline + field of study	Full model
Coefficient on female indicator (log points)	-0.21089	-0.15136	-0.06585
	(0.00391)	*(0.00408)*	*(0.00346)*
Intercept	7.97008	7.88781	6.77254
	(0.30861)	*(0.28185)*	*(0.22838)*
Included covariates:			
Age	Yes	Yes	Yes
Field of study	No	Yes	Yes
Working time	No	No	Yes
Sector	No	No	Yes
Occupation	No	No	Yes
Note:			
Dependent variable: log gross monthly wage.			
Standard errors are reported in brackets.			

Figure 10 – Results of the regressions and list of included controls

B: Model with Interactions Between Gender and Field of Study (Shadow Prices): Exploratory Results for Further Research

Model with gender–field of study interactions (shadow prices): main results

Term	Value/Coefficient	Std. error	t-value	p-value	Significance
Intercept	6.75908	0.22860	29.567	0.00000	***
Female dummy	-0.04949	0.01741	-2.843	0.00447	**

R-squared	Adjusted R-squared	Observations
0.6729	0.66914	25795

Figure 11 – Results of the regression including interactions between gender and diploma

Field of Study	Interaction coefficient	Difference in returns (%) between men and women	Std. Error	t-value	p-value	Significance
mast-Biomedical sciences	-0,109120025	-10,9120025	0,067474039	-1,617214944	0,10584426	*
bac-Construction, wood	0,10848628	10,84862803	0,069056512	1,57097828	0,1162001	*
mast-Interior design	-0,10671367	-10,67136698	0,044848594	-2,379420626	0,017347188	***
mast-Other	-0,101207758	-10,12077579	0,029628834	-3,415853513	0,000636815	****
mast-Physical education, physiotherapy, speach therapy	0,072194046	7,219404562	0,044237544	1,631963235	0,102699579	*
bac-Accommodation, restauration	0,071064568	7,106456777	0,047414123	1,498805902	0,133936372	*
bac-Technician (audio, photo, electricity)	-0,064811493	-6,481149323	0,027487154	-2,357882987	0,018386993	***
mast-Engineer	-0,062103346	-6,210334571	0,027216585	-2,281819948	0,022508174	***
mast-Law	-0,059103201	-5,91032011	0,024577549	-2,404763837	0,016190001	***
bac-Teaching (primary, kindergarten)	0,055202355	5,520235464	0,032896682	1,678052373	0,093349129	**
mast-Fine arts (theater, design, music, visual arts)	0,055042996	5,504299603	0,030410709	1,809987263	0,070309529	**
mast-Economics	-0,04943432	-4,943431971	0,025972688	-1,903319372	0,057010133	**
mast-Veterinarians	-0,04926831	-4,926831009	0,099236511	-0,496473625	0,619564571	
bac-Social assistant	0,047517551	4,751755113	0,024072698	1,973918761	0,048401757	***
mast-Business, applied economics	-0,045259886	-4,525988577	0,02194549	-2,062377541	0,039181911	***
mast-Languages, Philology	-0,041822043	-4,182204268	0,024716411	-1,692075877	0,090643724	**
mast-Medical studies (Physicians, Dentists)	-0,041171573	-4,117157261	0,058844455	-0,699667832	0,484141152	
mast-Philosophy, Theology	0,038161348	3,816134757	0,054578447	0,69920178	0,484432315	
bac-Medical assistant	-0,037731357	-3,773135672	0,034115613	-1,10598503	0,268743432	
bac-Computing_IT(A1)	-0,036096613	-3,609661348	0,025930688	-1,392042282	0,163921734	
mast-Pharmacists	0,034674831	3,467483126	0,056710426	0,611436621	0,540916016	
mast-Journalist, translator	0,031347062	3,134706198	0,028875669	1,085587381	0,277671847	
mast-Industrial engineer	-0,026777407	-2,677740667	0,022527268	-1,188666392	0,23458204	
mast-History, History of art, Musicology, Archaeology	0,02620336	2,620336008	0,031127271	0,841813588	0,399900199	
mast-Natural sciences (biology, maths, chemistry, physics, geology, geography)	-0,023354788	-2,335478772	0,027620416	-0,845562504	0,397804784	
bac-Media, publicity	0,019578142	1,957814209	0,033365312	0,586781321	0,557355778	
bac-Business assistant (bookkeeping, tax, marketing, insurance)	-0,017067724	-1,706772412	0,018878723	-0,904071967	0,365965782	
mast-Accountant	-0,01668326	-1,668325982	0,021248094	-0,785164996	0,43236413	
bac-Nurse, midwife	-0,015919844	-1,591984392	0,028620964	-0,556230185	0,578058387	
mast-Social Sciences	-0,014883474	-1,488347394	0,021413445	-0,69505275	0,487028577	
bac-Architecture	0,013166813	1,316681276	0,048821201	0,269694571	0,78739742	
mast-Criminology	0,012957243	1,295724322	0,035923768	0,360687201	0,718336282	
bac-Teaching(low secondary)	-0,012189748	-1,218974758	0,026127597	-0,466546829	0,640828121	
mast-Computing science	-0,011029452	-1,102945204	0,048428734	-0,227746034	0,819845508	
mast-Medico-social sciences (nurses)	0,010711026	1,071102578	0,050115406	0,21372721	0,8307615	
mast-Bio Engineer, agricultural engineer	0,010196113	1,019611264	0,030180229	0,337840793	0,735485948	
bac-Tourism, entertainment	-0,009525453	-0,952545327	0,038850123	-0,245184633	0,806315447	
mast-Psychology	0,008714656	0,871465592	0,026942392	0,323455165	0,74635321	
mast-Sales engineer	-0,005482078	-0,54820783	0,024702987	-0,221919652	0,824378209	

Figure 12 - Estimated gender differences in field-of-study returns (shadow prices)

These results suggest the existence of shadow prices, meaning that the wage returns associated with some fields of study may differ between men and women. This could constitute an interesting avenue for further research.