Norwegian School of Economics

Bergen, Autumn 2023

Gender Pension Gap in Norway: An Empirical Study

"What are the key contributing factors affecting the pension gap?"

Juni Holberg and Sara Sadeghi Khosroshahi

Supervisor: Siri Isaksson

Master thesis, Economics and Business Administration

Profile: Financial Economics

NORWEGIAN SCHOOL OF ECONOMICS

Acknowledgements

We would like to express our sincere gratitude to our supervisor, Siri D Isaksson, for her valuable guidance, encouragement and insightful feedback throughout the process that has been crucial in the development and completion of this thesis. Our appreciation also extends to Statistics Norway (SSB) for providing access to essential microdata, which has been the cornerstone of our analysis. We would also like to thank our family and friends for their never-ending support.

Juni Holberg and Sara Sadeghi Khosroshahi

Norwegian School of Economics

Bergen, December 2023

Abstract

This thesis investigates the gender disparities in pension income in context of Norwegian population, focusing on the socio-demographic factors that contribute to this gap. More specifically, this study is an empirical investigation on whether there is a difference in average pension income between men and women among the Norwegian retirees in the age range of 67 to 74 in year 2017, as the latest year with available data. Utilizing microdata provided by Statistics Norway and employing multiple linear regression analysis, this research offers a comprehensive understanding of the factors influencing pension income disparities between genders.

Our findings reveal a significant gender pension gap. The disparity is around 28% on average against females and decreases after adjusting for sociodemographic variables, highlighting the possibility of systemic issues contributing to this inequality. The disparity after controlling marital status and parenthood, particularly motherhood, is reduced to 2.9% highlighting the importance of these variables in explaining the gap which can provides valuable insights for policymaking aimed at reducing the gender pension gap.

While the methodology offers clarity in interpretation, the study acknowledges limitations due to data constraints, particularly in fully capturing the nuances of gender-specific employment patterns and their impact on pension outcomes.

Keywords_ Pension Gap, Gender Differences, Norwegian Pension System,

Contents

1. Introduction

An increase in gender equality has been a significant development in modern Western societies, and Norway, as a Nordic country, has been instrumental in this progression (OECD, 2018). The nation has been proactive in encouraging female participation in higher education and in reforming labor market structures to enhance gender equality (World Economic Forum, 2021). Such measures, including state-supported childcare and paid maternity leave, have been pivotal in enabling women to actively participate in the workforce and pursue careers (Nav, 2023). These advancements are deeply intertwined with the evolution of the welfare state, helping women break free from traditional gender roles and achieve financial independence. However, despite these strides in gender equality, financial disparities remain evident, particularly in terms of pension savings. In 2021, the average wage gap in Norway was reported to be 4.25% (OECD, 2021), a noticeable disparity in earnings between men and women. Yet, this gap is significantly more pronounced in the context of pension savings, where it stands at 28.1% (NTB, 2020). This considerable disparity in pension income is more than just a statistic; it reflects the deep-seated societal and economic factors that continue to affect the financial well-being of individuals, especially women, in their post-retirement years. This situation underscores the need for ongoing efforts to address these inequalities and ensure a more equitable distribution of financial resources in retirement.

The Norwegian pension system, while supportive in some aspects, still presents challenges in achieving complete gender equality. The Norwegian pension system does not necessarily offer increased benefits for periods of maternity leave, and its design to redistribute wealth is complex, with areas needing improvement. The system faces specific challenges, such as non-mandatory pension accrual during parental leave, the absence of pension points for caregiving periods, and a lack of pension indexation for retirees. These issues disproportionately affect women, especially considering their longer life expectancy (NTB, 2021). Addressing these complexities in the Norwegian pension system is crucial for advancing gender equality further, as well as considering a larger aging population. This includes reevaluating policies related to pension accrual during caregiving periods and ensuring that pension schemes are inclusive and equitable for all, irrespective of gender (NTB,2021).

In terms of changes in the pension system, it is important to mention that the Norwegian pension system is currently adapting to changes after the pension reform came into effect 1. January 2011. The goal of Norway's pension reform, with its new model for accruing and withdrawing old-age pension from the National Insurance Scheme, is to ensure an economically and socially sustainable pension system that has a strong distribution and equality profile and is based on simple and understandable main principles (Regjeringen, 2023). Despite recent advancements, the Norwegian pension system still faces challenges in achieving gender equality. It is difficult to measure the outcome of these improvements as the new rules have not yet been fully implemented. Section 2 will provide further information regarding the structure of the pension system and the implications of the reform.

The motivation behind this thesis is firstly to lift pensions into the spotlight as one significant challenge for gender equality in old age. Additionally, it strives to uncover what can explain these gender differences in pension income in Norway and how different characteristics could potentially explain the pension gap. This leads to the research question of the thesis:

"Is there a gender gap in pension income among the Norwegian population, and what are the key contributing factors affecting the pension gap?"

This research question will be answered by first examining the generally accepted theories of how gender relates to pension income. The method is fundamentally empirical and deductive through prior literature. Secondly, an empirical analysis of the gender pension gap in Norway is conducted. To address this critical issue, the study utilizes microdata from Statistics Norway, employing linear regression analysis as the primary methodological approach. This analysis is complemented through various linear regression models, each providing insights into how different socio-demographic factors contribute to the gender pension gap. The research focuses on variables like employment history, marital status, and the impact of parenthood, with a particular emphasis on 'potential experience' as a proxy for labor market engagement and the role of 'children' as a control variable. By examining these factors, the thesis aims not only to contribute to academic discourse but also to inform policymaking and inspire further research in this field.

Our research uncovers a striking disparity in pension income based on gender in Norway. We find that there is a significant difference in pension income between men and women, with an average gap of 28.46%. This disparity in pension income is noteworthy, given the context of Norway's social and economic framework. It's crucial to understand that these findings are derived from an in-depth analysis of comprehensive microdata from the entire Norwegian population, provided by Statistics Norway (SSB). This data, while extensive, has its own set of limitations, particularly in terms of completeness and privacy constraints, which impact the depth of longitudinal analysis possible.

Besides all the limitations, this study is contributing to the relevant literature since very few country-based research has been done on differences in pension income. Authors believe that deep investigation on pension system of each country, and how it behaves with men and women, can be a crucial step to address equality and potential discrimination.

1.1 Outline

To lay the foundation for the rest of the thesis, we provide further information about the Norwegian pension system and the importance of the Norwegian pension system in terms of gender in section 2. Section 3 summarizes economic explanations and prior research, which are the basis of the derived hypotheses of this thesis. The underlying methodology for the analysis as well as is presented in section 4. Section 5 describes the data applied and includes descriptive statistics on the relevant variables for explaining the gender pension gap. Section 6 presents the empirical findings from the linear regression analysis and robustness check for our models. In section 7, the findings and limitations of the thesis are discussed, and suggestions for further research are presented. Lastly, the main findings are presented in the conclusion of the thesis in section 8.

2. The Norwegian Pension System

Norway is renowned for its robust social welfare system, characterized by substantial annual social expenditures, which often exceed 20,7% of the country's GDP, a little below OECD average of around 21,1 % (OECD, 2023). One of the fundamental components of Norway's social welfare system is its public pension system.

The Norwegian pension system follows a model designed to prevent poverty and ensure economic security for all citizens during their retirement years (Pensjonsutvalget, 2020). It shares similarities with the Beveridgean model, where a fundamental objective is to guarantee a basic level of pension benefits to eligible individuals, regardless of their occupation or income throughout their lifetime (PPHR, 2017). The Norwegian pension system consists of multiple pillars and aims to provide comprehensive coverage for retirees (Pensjonsutvalget, 2020).

2.1 Implications of the Pension Reform

For the relevance of the thesis, it is important to emphasize that the Norwegian pension system underwent a reform that was officially implemented on January 1, 2011 (Regjeringen, 2023) and the reform has over the last 10 years been slowly adapted. The reform has introduced several new and improved models for earning and accessing old-age pensions in the National Insurance Scheme to ensure a pension system that it is possible to pay for in the future as well (Regjeringen, 2023).

Considering this, the master's thesis will look at pension payments occurring in 2017. Since the new system is currently being adapted in Norway, only in 2025 will the first individuals exclusively covered by the new old-age pension in the National Insurance Scheme be able to claim their pension at the age of 62. As many of the new pension schemes do not apply yet, we see it appropriate that the data used for the analysis will be collected from a sample of people between the ages 67-74. Further details on sample selection will be described in section 6.

2.2 Overall View

The pension system in Norway is divided into three parts: The national insurance scheme, employment-based pensions, and individual pension savings (Statens Pensjonskasse, 2023). This thesis aims to explain the calculations and adjustments involved in pension payments for individuals in Norway who received support in 2017 and were aged 67-74.

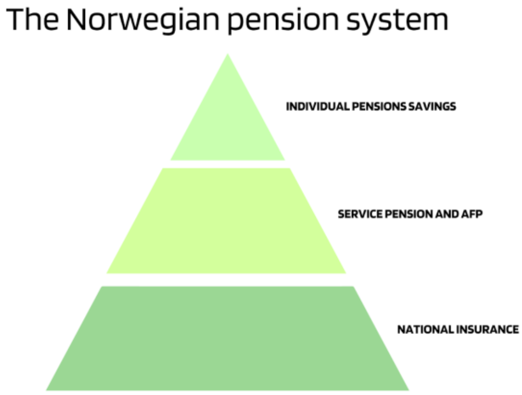

Figure 2.1: Brief overview of the Norwegian pension system

2.3 First Pillar

Pension from the National Insurance is lifelong and ensures that you have an income when you retire. “Folketrygden” is the basic pension scheme in Norway and is the major component of most people's pension payment. It covers everyone who resides in Norway or is an employee and the size of the pension depends on the income as a worker (standard insurance). All are nevertheless guaranteed a pension of a certain minimum level through the guaranteed pension (Nav, 2023).

Norway's recent pension system reform, specifically the Folketrygden program, has significantly changed the way pension payments are calculated (Regjeringen, 2023). This thesis specifically focuses on individuals born before 1953 to ensure that our analysis is conducted under consistent conditions for pension calculation. Focusing on this demographic allows for a detailed examination of the reform's subtle impacts. Under the old system, the old-age pension had two parts: a basic pension that was guaranteed regardless of previous income, and a supplementary pension that depended on pensionable income and accumulated pension points. By focusing on individuals born before 1953, we can analyze the implications of the pension system's changes in a consistent way.

2.3.1 Eligibility for Retirement Pension

To qualify for retirement pension, individuals generally need to have a minimum of five years of National Insurance coverage. National Insurance coverage is accrued through living and/or working in Norway (Nav, 2023). Those born before 1953 can start receiving their retirement pension the month after they turn 62. This option lets individuals choose when and how they want to start their pension while potentially continuing to work, without any decrease in pension benefits. To start withdrawing your pension at age 62, your earnings must be at least equal to the minimum pension level or guarantee pension for your cohort. If your spouse receives an old-age or contractual pension from the public sector, you are eligible for a reduced rate of 187,801 (Nav, 2023). You can work and receive a retirement pension at the same time without any reduction in the pension amount.

2.3.2 Accumulation of Retirement Pension

The amount of retirement pension you receive in Norway depends on your income and how long you have lived there. An individual's pension size is based on their highest 20 years of income during their career. Earnings beyond the National Insurance basic amount contribute to pension points, which determine pension rights. If you were born in or after 1943, you can start accumulating pension rights at age 17 and continue until you turn 75. Unemployment benefits and disability benefits are among the income sources that contribute to one's pensionable income. Pension rights in Norway can be accumulated through periods of residency, known as National Insurance cover, from age 16 to 66. To receive a full retirement pension, you need to accumulate 40 years. If you have fewer years, your pension will be reduced proportionally (Nav, 2023).

2.3.3 Components of Retirement Pension

The retirement pension from the National Insurance Scheme consists of two main components: the basic pension and the supplementary pension. The Basic Pension is calculated without considering previous work income. To qualify for a full basic pension, a minimum of 40 years of National Insurance coverage is required. The pension amount is being calculated from the basic amount (G) and the rules on life expectancy. From the period of 2017 the amount of G was calculated to 93, 281 NOK. However, if you are married or living with a partner, the amount is reduced with 90% of what you would have received as single (Nav, 2023).

Supplementary pension is accumulated through pensionable income and pension points. At least five years with pension points are needed to be entitled to a supplementary pension. Pensioners without a supplementary pension, or with a low supplementary pension, are guaranteed a minimum pension level. The minimum pension level consists of multiple rates and depends on your marital status as well as the income of any spouse or cohabitant. For example, if you live with a spouse who receives an old-age pension or contractual pension in the public sector, you are entitled to a low rate which is equivalent to 187,801 NOK (in 2023). For comparison, if you are single, you are entitled to a special rate for single people. Spouses who do not live together are also considered single in this context and correspond to a sum of 257,040 (Nav, 2023).

2.3.3 Life Expectancy Adjustments

The pension amount is adjusted based on the life expectancy of the individual's age group, known as life expectancy adjustment. If people live longer, each age group may have to work for a longer period to receive the same pension amount.

People born before 1942 or those born in 1943 who began receiving retirement pension before January 1, 2011, are exempt from life expectancy adjustments. This system helps individuals born before 1953 in Norway plan for retirement by taking into account factors like pensionable income, years of National Insurance coverage, and life expectancy adjustments. It aims to support their financial well-being in old age (Nav, 2023).

2.4 Second Pillar

The second pillar of the Norwegian pension system consists of service pension and contractual pension. The rules for these pensions vary depending on whether you work in the public or private sector.

2.4.1 Service Pension

Service pension is a pension scheme offered by employers in Norway as part of the employment contract. It is a pension you earn through your professional career with an employer (Norsk Pensjon, 2023).

Public Service Pension

Public occupational pensions are occupational pension schemes that have been established for those who are employed by the state, municipalities, healthcare organizations or companies with public connection (Norsk Pensjon, 2023).

In the public sector, for those born up to and including 1962, the pension system operates as a defined benefit scheme, ensuring a fixed level of retirement income derived from the combined contributions of the occupational pension scheme and the National Insurance Scheme (Folketrygden). The pension amount is calculated based on the salary earned in the occupational pension scheme of which an individual was most recently a member prior to retirement, referred to as the final salary. Initially, the gross occupational pension is computed, typically amounting to 66 percent of the pensionable income based on the final salary, up to a certain income threshold (12 times the National Insurance basic amount or 12 G), before accounting for adjustments due to changes in life expectancy (Nav, 2023).

Full accrual of pension rights typically occurs after 30 years of membership, and the payouts are lifetime. With shorter periods of membership, the pension entitlement decreases proportionally. The net occupational pension is then determined by offsetting the National Insurance old-age pension payments according to specific coordination rules. If an individual holds pension entitlements from multiple public sector occupational pension schemes, the "last scheme" in which membership was held becomes responsible for disbursing the pension, encompassing all periods of membership within public sector schemes (Norsk Pensjon, 2022).

Private Service Pension

The Act on Mandatory Occupational Pension (OTP) mandates most private sector employers to establish pension schemes for their employees and outlines specific requirements for the design of these pension schemes (Skatteetaten, 2023). The legislation permits the establishment of pension schemes in the form of defined benefit plans, defined contribution plans, or hybrid plans (introduced post-2015).

Benefit plans, often referred to as enterprise pensions, aim to provide plan members with a predetermined annual pension. This contrasts with defined contribution plans, where the contributions are fixed, and the ultimate pension amount is variable. Typically, defined benefit pensions are structured such that the annual pension is calculated as a percentage of the salary at the retirement age, subtracting an estimated National Insurance pension. On the other hand, defined contribution plans operate as savings schemes, with the employer, and occasionally the employee, making annual contributions to the pension scheme. Accumulated amounts are subject to potential annual returns, and the annual pension is determined by dividing the accumulated sum by the number of years over which the pension is to be received (Norsk Pensjon, 2023).

In this analysis, we investigate the gender gap in both private and public service pensions. For private service pensions, individuals can start receiving payments from the age of 62, continuing at least until age 77, with the condition that the duration of payment spans a minimum of ten years. Unlike public service pensions, which are typically lifelong, private service pensions are designed to be time-limited initially (Norsk Pensjon, 2023). The analysis will therefore consider the impact of both the longevity of payments and the type of service—private or public—on the gender gap in pension income.

2.4.2 Contractual Pension, AFP

AFP is short for contractual pension, which is a pension scheme included in a collective agreement between employers and employees. It is considered part of the second pillar of the Norwegian pension system (Norsk Pensjon, 2023). The right to AFP requires that you work in a company that has a collective agreement in which AFP is included. Originally, the AFP was an early retirement scheme that paid a pension to older workers who stopped working between the ages of 62 and 67. Since 2011, the scheme in the private sector has been transformed into a supplementary occupational pension scheme with the payment of benefits for life, while the old rules still apply in the public sector if you were born before 1963. AFP in the public sector involves a fixed supplement of NOK 1,700 per month when withdrawing full AFP. If you choose to work during the period, the supplement is reduced depending on the income (Nav,2023).

2.5 Third Pillar

In Norway, the private pension savings, often referred to as the third pillar of the pension system, are structured as voluntary individual pension schemes. These schemes are typically established by insurance companies, private pension funds, or banking institutions. The ultimate pension benefits are contingent upon an individual's savings and the returns generated by their specific scheme. Consequently, private pension savings can exhibit considerable variation among individuals. These individual schemes commonly take the form of capital pension plans or rate pension plans, although they may also encompass life-long annuities or age-based pensions. (Norsk Pensjon, 2023)

For example, IPS is a special type of private pension savings that the state has implemented to encourage private individuals to save more for their own pension. The pension scheme means that you can save up to NOK 15,000 per year for a pension. A deduction is given in ordinary income for deposits, and withdrawals must be taxed as ordinary income. Deposited funds in the new scheme are exempt from wealth tax and current income tax on the return. The scheme can be established as a pure savings scheme or as an insurance scheme (Norsk Pensjon, 2023).

2.6 Overview of Composition of the Pillars

Below, you will find a comprehensive overview of the distribution of pension assets within the Norwegian pension system. This visual representation illuminates how pension funds are allocated across the various pillars, offering valuable insights into the weight and significance of each component in the retirement security of Norwegian citizens.

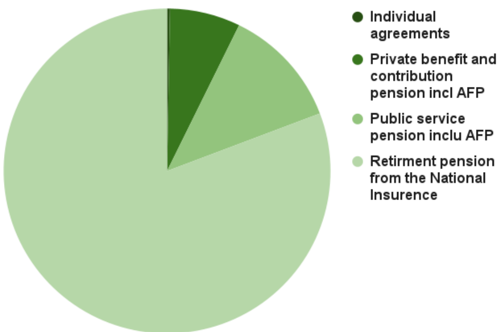

Figure 1.1: Composition on pension assets in 2017 Source: Adapted from SSB (2017)

The pie chart clearly shows how the Norwegian pension system's assets were distributed in 2019. The National Insurance retirement pension accounted for 81% of the total. The majority of pensions are public service pensions, including AFP, at 12%, followed by private benefit and contribution pensions, also including AFP, at 7%. Individual agreements account for only 0.29% of total pension agreements, indicating minimal direct individual involvement. The chart shows that state-sponsored schemes are important for ensuring retirement security in Norway.

2.7 Norwegian Pension System in the Context of Gender

In this section, we will discuss the differences between men and women in terms of demographics and how these differences are related to the Norwegian pension system. This highlights the importance of studying the gender pension gap. Pension systems are supposed to be gender-neutral, and the Norwegian pension system aims to achieve equal payments for men and women (Pensjonsutvalget, 2022).

There is a significant difference in life expectancy between men and women. Women, on average, live about 3.43 years longer than men (Haug, 2023). Life expectancy differences directly affect pension systems, especially when converting pensions to annuities. Annuity rates are based on expected remaining lifetime. This means that women, who generally have longer life expectancies, tend to receive lower annual pension payouts compared to men, assuming equal pension savings. The contrast is most noticeable in the second pension pillar (SSB, 2023). Occupational funds often calculate annuity entitlements based on gender-specific life expectancy measures. Women, due to their longer life expectancy, accumulate more public pension entitlements in the first pension pillar. This highlights that women rely more on public pensions than men (Halvorsen, E. & Hetland, 2023).

Second, having children has different impacts on men and women’s labor market attachment, which affects supplementary pension from “Folketrygden", and occupational pension contributions, since they are earnings-related. In 2022, 36.6% (16,6%) female (male) labor force participants in Norway were part-time (SSB, 2020). As we are looking at the people being born before 1953, it is reasonable to believe the old rules apply for parental leave as most people have an average childbearing age for the birth of their first child of 24,1 years (Only date back to 1961, SSB). Looking at the old system, studies on Norwegian parental leave show that the average days on parental leave for children being born in the first half of 2009 for mothers was ~218, compared to ~37 for fathers (Engvik, M., & Pettersen, M. (2021). However, the design of the pension system does not compensate for lower attachment in the labor market arising from childcare other than that parental leave money up to a certain amount of 6 G, an amount equal to 560 000 NOK in 2017. Since mothers on average take longer parental leave, it is possible to claim that women could possibly have been receiving in a period less credit points due to home care of children, and that this should be encountered in the system.

Furthermore, as mentioned above the Norwegian pension system does not include care credits beyond the normal parental leave period (SSB, 2020). Additional pension entitlements can exclusively be built up through labor market participation or private investment and you are only entitled to care credits if you care for a person who is ill, elderly or has a disability, (Nav, 2023). This can also contribute to occupational segregation, where women who plan to have children tend to gravitate toward employment in the public sector rather than the private sector. The public sector offers more family-friendly regulations and benefits, and that the “wage penalty” by having children for women is markedly lower in the public than in the private sector (Hardoy & Schøne 2008). Moreover, contributions to occupational pensions are typically more substantial and higher in the public sector compared to the private sector (SSB, 2022). However, it's important to note that wages in the public sector are generally lower and may not experience as significant growth as those in the private sector, thereby highlighting the complex nature of the Norwegian pension system concerning gender.

These observations align with the findings of a study conducted by Pål Schone (2015), titled 'Women, children, and choice of sector: Is the public sector still attractive?' The study, which pertains to the public sector in Norway—a sector that is predominantly female-dominated—investigates the influence of children on mothers' decision to pursue careers in the public sector. It reveals that the public sector remains an attractive option for women during the first phase of motherhood. Women with children are more inclined to seek employment in the public sector compared to their childless counterparts, and this inclination becomes more pronounced as the number of children increases (Schøne, P. 2015).

3. Economic Explanations and Hypotheses

This section aims to give a summary of well-known economic literature that explains the gender pension gap by introducing theoretical and empirical labor market findings. Examining empirical data and conventional economic theories will help determine the relevant variables that can possibly contribute to the gender pension gap in Norway. Furthermore, the economic explanations found in relevant literature will prepare a foundation for the development of theoretical hypotheses that will be tested empirically later in the study.

3.1 Gender Pension Gap vs Wage Gap

Researchers examining the gender inequalities in income and labor markets typically focus on the gender wage gap. The difference between the average yearly earnings of men and women is known as the gender wage gap. The wage gap focuses on how factors like labor market attachment variables, resource access, and societal standards can result in inequalities in salaries between men and women. In the western world, the gender pay gap is frequently estimated to be between 10% and 20% (Lane, 2018). However, the data indicates that pension payments for individuals 65 and over were, on average, 28% lower for women than for men in the United States and the European OECD countries (OECD, 2015).

The pension gap between men and women can be considered as a natural extension of the gender wage gap. It applies the same concept to a relatively older population in which their explanatory variables are largely historical. But these factors still have a significant impact on their retirement financial well-being. Therefore, the greatest difference between wages and pensions is the natural retrospectivity of pension determinants. Pension entitlements are a result of accumulated lifetime earnings, where the process is fourfold. Wage is transformed into annual earnings, annual earnings constitute lifetime earnings through the length of the career, which in the end results in pension entitlements (Bettio et al., 2013).

As mentioned, the estimated average gender pension gap in the world in the studies is quoted to be around 30–40%. This significantly wider gap compared to the wage gap can be caused by the cumulative impact of the gender wage difference over the course of a person's life, impacting pensions later in life. Therefore, it is crucial to investigate this aggregated impact in economic studies.

3.2 Economic Explanation of Gender Pension Gap

The focus of this section is on introducing the main explanatory variables mentioned in well-known previous literature in the field of gender inequality in the labor market and specifically in pension income. Although there is extensive literature on gender differences in the labor market with respect to the wage gap (Blau & Kahn, 2017), less research has been done on differences in pension income, especially country-based research. Therefore, the major resource for this study is cross-country studies on gender differences i.e. Möhring (2014) and Dewilde (2012). Since many of the explanatory factors contributing to the gender pension gap are the same as the ones in the gender wage gap, the literature will be combined in this section.

3.2.1 Labor Market Discrimination

Labor market gender discrimination has been extensively studied in recent decades literature (Kunze, 2015; Blau & Kahn, 2017) and the majority of them find the similar results of the existence of a direct link between gender and wage for the employees that are equal in other factors. Thus, gender discrimination in the labor market is an established phenomenon, even in modern society (Blau & Kahn, 2017).

As the Norwegian pension system is based on lifetime contributions in the labor market, it seems reasonable to assume that there is a link between employment background and pension income (Kuivalainen et. al., 2020). Thus, the primary cause of gender disparities in the labor market can be attributed to pay gaps between men and women, which are also reflected in the gender pension gap. (Fasang et. al., 2013). Since women usually have more irregular employment careers with years of discontinuity in the labor market, lower lifetime earnings also translate into lower additional resources for private pension schemes and fewer contributions to occupational pension schemes. Due to lower savings in employer and private pensions, this increases the pension gap in retirement. (Fasang et. al., 2013).

3.2.2 Human Capital

To explain why men and women experience different outcomes in the job market, several conventional economic hypotheses have been presented. One of the most well-known hypotheses was created by Becker (1985) and is based on the concept of human capital. The concept of human capital offers reasons why some people decide to invest in human capital while others do not. It can therefore be interpreted as a supply-side explanation of gender differences in wages and therefore also pension income (Becker, 1985). The two primary factors affecting human capital are “Work Experience” and “education” (Blau & Kahn, 2017). These two human capital factors are based on the premise that productivity increases with time spent acquiring education and experience. This productivity-enhancing effect of education and work experience results in higher wages, and therefore also pension income (Becker, 1985). Differences between men and women in these two human capital variables can consequently lead to large pension gaps in retirement (Veremchuk, 2020).

In the context of this study, when referring to work experience, it indicates the number of years of employment in the labor market. This is a crucial factor when assessing the pension incomes of retired people. Reviewing the literature, it is found out that work experience is a key factor in explaining the differences in pension income between the two genders (Sefton et. al. 2011; Dewilde, 2012). Due to social gender norms, throughout history, men have had an advantage over women in terms of job experience, leading to longer years spent in the labor market for men. However, over time, this pattern has undergone a significant shift. The difference in work experience between men and women decreased from roughly 7 years in 1981 to just 1.4 years in 2011 in US (Blau & Kahn, 2017). This may be a result of institutional variables that are more commonly available and offered, including childcare services or maternity benefit programs (Veremchuk, 2020). These institutional elements may have an impact on the number of years of employment a woman can accumulate, which may increase pension income upon retirement because of continuous lifelong earnings. In addition, gender norms affect how much women participate in the job market (Akerlof & Kranton, 2000). The word "Gender identity" is used to introduce this idea, and it is discovered that an individual's behavior is influenced by cultural models. Cultural norms influence how men and women behave differently, which also has an impact on their utility. This is experimentally shown, where it is discovered that women typically perform more household duties than men because doing "women's work" reduces men's utility (Akerlof & Kranton, 2000). As a result of their gender identity and traditional gender roles, women lower their labor market connection in favor of dedicating themselves to housework, which is reflected in their years of work experience (Veremchuk, 2020).

Along with work experience, education is a classic human capital factor used to account for gender variations in pension income. Higher levels of education can be reflected in the level of productivity of employees and as a result, associated with greater wages and pension incomes. According to Sefton et al (2011), less educated women are less likely to have private pension funds, which lowers their overall pension income in older age. According to a different study by Bettio et al (2013), higher educational levels are linked to greater pension income. This supports the idea that education is a crucial element in the field. These two studies contend that a person's greater degree of education favorably affects their pension income.

Over history, men have had a competitive advantage in years of education, which has widened the gender wage gap, but this trend seems to have been reversed in modern society starting in 2010. According to the study of Blau and Kahn (2017), women currently have a somewhat higher degree of education than males do in US. Literature is divided on whether education still matters when contrasting men and women because the impact seems to have faded in contemporary culture. According to Blau & Kahn (2017), the importance of education in the cumulative work career has decreased over time, which reduces the explanatory power of the variable when calculating the gender pension gap. On the other hand, Veremchuk (2020) discovers that when examining gender disparities in retirement, education level is crucial. However, based on the human capital model, this analysis anticipates that if women spend more on the human capital variable education, the gender pension gap will narrow down.

3.2.3 Occupational Types

When analyzing what leads to gender inequalities in pension income, occupational type, or the diversity of employment kinds, is a key explanatory factor. Since higher-level jobs pay more than lower-level occupations, the choice of occupation establishes the basis of compensation (Kuivalainen et al, 2020).

In terms of representation across different employment kinds, men and women are remarkably diverse. Women often work more than males do in administrative, service support, and low-status manual tasks (Kuivalainen et al., 2020). They are also more prevalent in low-wage sectors like nursing and teaching, whereas males are more likely to pursue management and blue-collar employment (Blau & Kahn, 2017). It has also been shown that women choose jobs that allow for the care of children, which are often found in the public sector. Therefore, this is consistent with the research that shows that women are more likely to work in professions like nursing and teaching (Flyer & Rosen, 1997).

The hypothesis of the "glass ceiling" is introduced because there are fewer women in positions of leadership compared to men. According to the "glass ceiling" idea, there are unspoken barriers to women's advancement in top positions that prevent them from doing so (Blau & Kahn, 2017).

Fortunately, as more women have started to work in high-ranking positions and professions over the past few decades, they have made significant progress in overcoming the issue (Tinios, Bettio, & Betti, 2015). However, this does not eliminate the reality that there are still career inequalities depending on gender (Blau & Kahn, 2017).

3.2.4 Children

Several studies have been studying the effect of having children on pension income for men and women (e.g. Dewilde, 2012; Sefton, 2011; Kuivalainen, 2020). Also, there have been studies to estimate the effect of children on other labor market factors like employment status. Traditional economics (Becker, 1985; Gronau, 1988) explain that women can be less productive in the labor market when they have children since they traditionally have responsibilities in society to spend more time on unpaid work like taking care of children and household work, so they have less time to focus on self-development factors for job market like education and experience. Kleven, Landais, & Søgaard (2019) also discovered that having children is a significant contributor to the wage gap between men and women. According to Blau & Kahn (2015), the phrase "motherhood penalty" refers to the association between having more children and earning less money. They discover that when the number of children rises, the maternal penalty does as well. Moreover, according to Möhring (2015), having children significantly lowers women's pension income. An empirical study shows that despite having similar capabilities, women who are mothers are perceived as being less competent and productive than women who are not mothers (Correll, Benard, & Paik, 2007). In light of the widespread perception that mothers will prioritize their children above their careers, women who have children often struggle to advance in their careers due to perceptions of their inadequacy. The term "status-based discrimination competencies" is used in literature to describe this (Correll, Benard, & Paik, 2007). Dewilde (2012), shows that having children can be associated with opportunity cost if the mother needs to get a break from a paid job to take care of her child. In such situations, according to Möhring (2014), the welfare system of the country can play a crucial role in mitigating the problem.

3.2.5 Participation in Labor Market

In addition to the previous section and the effect of having children on the level of income, women with children are more likely to change their labor market engagement (Liu & Marois, 2023). This means that women who have children often choose to leave the workforce altogether or just work part-time (Burkevica et al., 2015). Labor market involvement is strongly impacted by salaries and, consequently, pension income in later life (Dessimirova, 2019). Dewilde (2015) in his study reveals that males have more consistent job histories than women do, who frequently alternate between employment, inactivity, full-time work, and part-time work, generally being less connected to the labor market. As mentioned in the previous section, family obligations are the primary cause of women's irregular job histories. It is challenging to balance full-time job and home duties due to the traditional, yet controversial gender roles that persist in modern culture. These roles suggest that women should do a greater portion of housework and childcare. As Becker (1985) revealed, women usually pick jobs that require less work to do but provide greater flexibility. As a result, there is a decrease in productivity, human capital investment, salaries, and pension income.

3.2.6 Marital Status

It has also been observed that marital status can affect the level of pension income among retired people, but not in the same way among the two genders. According to several studies (Fasang et al., 2013; Möhring, 2014), single or divorced women receive more pension income than married women. This gap between married and unmarried women can be substantial, to the extent that in a study, Bettio et al. (2015) show that the average pension gap between married and unmarried women in European Union countries is 54%. The reason can be rooted in the traditional gender roles and responsibilities that are expected by society from married women compared to men.

Interestingly, the studies reveal the exact opposite effect of marital status on the income of men. According to Baradasi and Taylor's (2008) research, married men make more money overall than single men. It's noteworthy to consider that they also discover that married men receive a higher salary premium if their wives do not work 40 hours per week as opposed to if they do. This can be explained as the wife spends more time taking care of the home, and the husband can more easily find the opportunity to devote more of his time to developing human capital, improving productivity, earnings, and retirement income. They have called this phenomenon the “marriage premium”.

3.2.7 Investment Behavior

Investment practices play a critical role when considering private pension funds. In these funds, capital is allocated based on an individual's risk appetite. Studies indicate that women are generally more cautious with their investment decisions compared to men (Dawson, 2023). Mercer, a consulting company, noted that European women often exhibit more financial hesitancy than their male counterparts, which is evident in their retirement planning. As a result, women tend not to opt for high-growth, long-term investment strategies for their retirement funds, hindering their ability to achieve financial stability comparable to men in their later years (Lane, 2018). Furthermore, a 2020 Forbes article highlighted the significance of financial knowledge in enhancing women’s financial well-being upon retirement, potentially narrowing the existing gender pension discrepancy. The article asserts that by enhancing their financial literacy, women can strengthen their confidence in making financial choices, subsequently setting and achieving ambitious long-term savings targets, thereby ensuring a comfortable retirement income (Forbes, 2020). Similarly, Van Rooij and colleagues (2012) found evidence suggesting a direct relationship between financial know-how, retirement preparations, and augmented pension assets. Their research indicated that individuals well-versed in financial matters are more prone to diversify into stocks and exhibit superior retirement savings habits (Van Rooij, Lusardi, & Alessie, 2012).

Furthermore, marital status emerges as a factor affecting investment tendencies. Research by Jianakoplos and Bernasek (1998) revealed that single women exhibit greater risk aversion than both single men and married pairs. By examining U.S. demographic data, they validated prevalent findings that indicate a reduction in risk aversion with increasing household wealth. However, they noticed that this reduction isn't as pronounced for single women. They observed that as single women have more children, they tend to decrease the risky components in their investment portfolio. This heightened caution among single women can lead to less optimal asset distribution, potentially yielding lower returns. This observation helps to understand the diminished risky financial assets held by women, which may affect their financial comfort during retirement (Jianakoplos & Bernasek, 1998). It anticipates that enhancing financial literacy among women can lower the pension income gap observed between genders.

3.3 Hypothesis Development

Drawing from the economic explanations discussed in Section 3.2, the following hypotheses have been developed to provide a deeper understanding of the factors influencing gender disparities in pension income. This approach allows for an investigation into the overarching research question of this thesis.

These hypotheses address both the unadjusted and adjusted gender pension gaps. The unadjusted pension gap quantifies the raw average disparities in pension income between genders, while the adjusted pension gap considers various observable factors that may contribute to gender-based differences in pension income.

Hypothesis 1: “When accounting for sociodemographic factors, the average pension income for males is consistently higher than that of females."
Hypothesis 2: "A portion of the disparity in average pension income can be attributed to differences in marital status and the number of children."
Hypothesis 3: “The gender pension gap can be partially explained by disparities in human capital, occupational choice, labor market participation, and investment behavior between males and females."

These hypotheses establish the anticipated outcomes for the empirical analyses.

4 . Methodology

The scientific method employed in this thesis is fundamentally empirical and deductive, meaning that it relies on real-world observations to test and support its theories. This process is illustrated in Figure 4.1. It begins with a comprehensive review of relevant scientific literature and theories, as detailed in section 3.2. From this literature review, hypotheses are derived to understand how theories and literature relate to the specific research question at hand. Section 3.3 outlines how the research model should be structured and defines hypotheses for the model's results.

Moving on to the third step, data is collected, primarily from Statistics Norway as secondary data for this thesis. From this data, the necessary variables for testing the hypotheses are constructed, as explained in section 5. In the fourth and fifth steps, the findings are presented, and the hypotheses are rigorously tested using appropriate models.

Finally, the last step involves revision of the theory including challenges and limitations of our current analyses. In other words, it entails a discussion of how the findings and their limitations have implications for the overarching theory, as discussed in section 7. This deductive and quantitative approach aligns with an epistemological framework associated with the natural sciences and positivism. Furthermore, the ontological orientation is rooted in objectivism (Bryman & Bell, 2011).

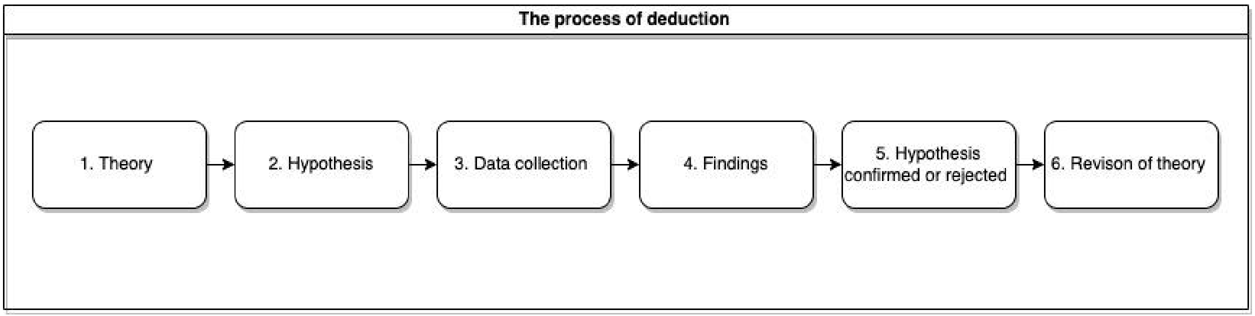

Figure 4.1: Modified from (Bryman & Bell, 2011)

The following sections aim to delve into the relevant methodology for examining the gender pension gap from a statistical perspective. We will commence by outlining the methodology of linear regression analysis. This approach allows us to estimate the partial association between various explanatory factors and the dependent variable.

In this context, we use linear regression to examine how gender (male or female) affects pension income while keeping other characteristics constant. By employing this model, we can establish, on a reasonable basis, whether there exists a statistically significant relationship between gender and pension income. It's important to note that while a significant statistical correlation does not imply causation, its absence can be a basis for further investigation (Field, A. 2018).

4.1 Linear Regression

Linear regression models are well-suited for estimating the influence of gender on pension income (Woodrige, 2019). This method stands as one of the most widely employed techniques in econometric analysis, offering a straightforward and intuitive approach. Linear regression models are commonly estimated using Ordinary Least Squares (OLS), which determines model coefficients by minimizing the sum of squared errors between the observed sample values and the model's predicted values (Wooldridge, 2019). The main goal is to measure how gender affects pension income while considering other factors and determining if gender has a statistically significant impact on pension income.

4.1.1 The Multiple Linear Regression Model

The general population multiple linear regression model has more than one independent variable and can be formally defined as an equation (4.1).

$$Yi = \beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + \ldots + \beta_k X_{k,i} + u_i \quad for\ i = 1, \ldots, n \qquad \text{Eq. 4.1}$$

Where Y_i is the dependent variable of observation i, and $X_{1,i}$, $X_{2,i}$, …, $X_{k,i}$ represent the k independent variables. Moreover, β_0 is the intercept, which can be thought of as a slope coefficient on a vector, $X_{0,i}$ which equals 1 for all i. Correspondingly, $\beta_1 \ldots \beta_k$ are the slope coefficients for each of the remaining independent variables. These can be interpreted as the partial impact on Y_i for a one-unit change in the corresponding variable X_k. Lastly, ui represents the error term for observation i (Wooldridge, 2019).

In investigating the thesis question, we find it appropriate to use multiple regression as our regression tool. This decision is based on the recognition that the simple linear regression, in its most basic form, includes only one regressor. In this case, employing just a gender dummy variable (X_1) to predict pension income allows us to identify an unadjusted gap in pension income. This unadjusted gap represents the effect that gender, in isolation, has on pension income and is identified by studying $\hat{\beta}_1$, the coefficient associated with the gender dummy variable. However, the simple regression model falls short in accounting for other pertinent factors that may influence the effect of gender on pension income. In essence, it leaves out variables that could be relevant in determining pension income. In this framework, these omitted variables are incorporated into the error term, u, which results in the coefficient being biased. This phenomenon is commonly referred to as omitted variable bias (Woodridge, 2019).

To address and mitigate the problem of omitted variable bias and gain a more comprehensive understanding of the factors contributing to the gender-based disparity in pension income, we employ a multiple linear regression. This approach involves the inclusion of control variables that are not only determinants of the dependent variable (pension income) but also relate to the independent variable (gender). Using a multiple regression model helps us understand how different factors interact and gives us a better way to identify and evaluate the main factors that contribute to the gender gap in retirement payments (Woodridge, 2019).

As mentioned previously, the coefficients in equation (4.1) can be estimated using the Ordinary Least Squares (OLS) method. This traditional approach hinges on the fundamental principle of minimizing the sum of squared errors by employing first-order derivatives concerning each coefficient and setting them equal to zero. Assuming b_0, b_1, and so forth to be the general estimators of β_0, β_1, and the remaining coefficients, the ensuing equation (4.2) is minimized:

$$\sum (Y_i - b_0 - b_1 X_1, i \ - \cdots - b_{1k} X_{ki}\ )^2$$

Eq.4.2

Under the fundamental assumptions of linear regression analysis, the OLS methodology assembles what are known as the Best Linear Unbiased Estimators (BLUE). Should $\hat{\beta}_0$, $\hat{\beta}_1$, and $\hat{\beta}_k$ signify the OLS estimators of the elusive true population coefficients β_0, β_1, and so forth, the OLS sample regression function (SRF) and residuals are expressed in equations (4.3) and (4.4), respectively:

$$\hat{Y}i \ = \ \beta_0 \ + \ \widehat{\beta_1} X_{1,i} \ + \ \widehat{\beta_2} X_{2,i} \ + \ \ldots \ + \ \widehat{\beta_k} X_{k,i} \ + \ u_i$$
$$for\ i \ = \ 1, \ldots, n$$

Eq. 4.3

$$\hat{u} = Yi\text{-}\hat{Y}i \quad for\ i\ =\ 1,\ldots,n$$

Eq. 4.4

The OLS method is important for our research. It helps us estimate coefficients and create a sample regression function to understand the relationships between variables. We are using it to assess gender-related disparities in retirement payments. The residuals, as indicated in equation (4.4), play a vital role in evaluating the model's accuracy and its ability to capture the nuances of these relationships.

4.1.2 Economic Question of Interest

Building upon the general linear regression model discussed earlier, the Sample Regression Function (SRF) can be tailored to address the specific economic inquiry of interest. Given the focus of this paper on examining gender disparities in pension income, the model incorporates the variable Female i. This variable is a dummy variable, taking the value 1 if the individual is female and 0 if the individual is male in all regression analyses. It serves as the primary independent variable of interest, aligning with the central research question.

Furthermore, as elaborated in section 5.3.1, the dependent variable underestimation is represented by the natural log-transformed pension income, ln (Pension Income). This transformation leads to the formulation of the following SRF, as derived from equation (4.5).

$$\ln \widehat{pensionincome}_{\,i} = \beta_0 + \beta_1 Female_1, i + \beta_2 X_2, i + \ldots + \beta_k X_k, i + u_i \quad for\ i\ =\ 1,\ldots,n$$

Eq. 4.5

As we have already established that our main variable of interest is Beta 1, the focus of this section of our study revolves around ensuring that this coefficient estimator is Best Linear Unbiased (BLUE). Our ultimate objective is to determine whether this estimator is an outcome of random sampling variation or if, with a certain level of confidence, we can conclude that β_1 differs significantly from zero. This concept is formulated as expressed below:

Null Hypothesis (H0): $\boldsymbol{\beta_1 = 0}$
Alternative Hypothesis (HA): $\boldsymbol{\beta_1 \neq 0}$

This constitutes a two-sided statistical t-test, and it holds relevance in addressing our economic question. Within the context of the sample regression function (SRF) utilized in this paper, this test allows us to assess whether gender-based disparities in pension income carry statistical significance. Although the t-test is useful for analyzing specific limitations, it is important to emphasize that our models strive for a comprehensive specification. Including joint hypothesis tests is important in this context. These tests are important for determining if the control variables help explain gender disparities in pension income. The null hypothesis states that the coefficients for these variables are all zero, as shown in equation (4.6). This approach helps us evaluate the overall impact and importance of these variables in our analysis.

$$H_0: \beta_1=\beta_2=\ldots=\beta_k=0 \qquad \text{Eq 4.6}$$

The statistical framework allows us to use the linear regression model to determine if being female has a significant impact on pension income while keeping all other factors constant. Equations (4.5) and (4.6) are mathematical representations that serve as a statistical framework for addressing our main research question. Statistical tests are important for evaluating assumptions in the linear regression model and choosing the best model for studying gender-based differences in pension income. The next section will discuss the basic assumptions of the classical linear model. Then, we will provide an overview of statistical testing.

4.1.3. Gauss-Markov Assumptions for Simple Regression

To achieve the Best Linear Unbiased Estimation (BLUE) for the coefficient on $Femalei$, certain assumptions must be met. In this section, we will discuss the theory behind these assumptions. In section 6.1, we will conduct a multiple linear regression analysis to determine if these assumptions are held in the context of this thesis.

The consequences of violating these assumptions vary depending on which assumption is being breached. When the assumptions are not met, there are three possible outcomes for the estimator $\hat{\beta}_1$ (Woodridge, 2019).

(i) Bias: Estimation bias, where $\hat{\beta}_1$ systematically deviates from the true population value.

(ii) Inconsistency: Inconsistent estimates, where $\hat{\beta}_1$ fails to converge to the true value as sample size increases.

(iii) Inefficiency: Inefficiency in estimation, characterized by larger standard errors and reduced precision in estimating β_1.

It is crucial to evaluate the validity of assumptions to ensure accurate parameter estimates. OLS is proven to be unbiased and efficient based on a set of simple assumptions. The following section will provide a brief overview of each assumption.

1. Linear in parameters

Equation (4.5) introduces the multiple linear regression model, which is commonly used to analyze statistical relationships with linear parameter dependencies. This model allows for the inclusion of curvilinear associations and does not require linearity among the explanatory variables (Woodridge, 2019). This flexibility is important when considering the population model, where the dependent variable 'y' is related to the independent variable 'x' and the error term 'u' in the equation:

$$y = \beta_0 + \beta_1 x + u_i \qquad \text{Eq 4.7}$$

By carefully choosing values for ‘y’ and ‘x,’ one can investigate interesting nonlinear relationships, such as constant elasticity models (Wooldridge, 2019). We will further explain the implications of this non-linearity about Gauss-Markov Assumption 4.

2. Random sampling

Random sampling is essential to obtain an unbiased estimate of β. The data used in the analysis must come from a random sample of the population. If the sample does not accurately represent the population, this assumption may be violated. If the sample is not representative of the population, it would introduce bias into the estimate of β_1. This bias occurs because the Simple Random Sample (SRF) and the Population Random Sample (PRF) have different partial effects. Therefore, the sample for the PRF must consist of n random observations (Woodridge, 2019).

3. Sample variation in the explanatory variables

The sample’s independent variables (xi) have diverse observed values. Diversity is crucial for understanding the impact of changes in x on y. Our understanding of the relationship would be limited without this variability.

Additionally, independent variables in both the sample and the broader population are not constant, and there are no perfect linear relationships among them. This condition is violated when there is a pair of values (a, b) where $x_j = a + bx_i$.

4. Zero conditional mean

The error term, denoted as 'u,' is characterized by having an expected value of zero for any given value of the explanatory variable:

$$E(u|x) = 0 \qquad \text{Eq 4.9}$$

This key is crucial for interpreting causal relationships as it can be violated in multiple ways. If the zero conditional mean assumption is violated, we cannot determine if changes in x cause changes in y, or if it is u. (Somville, 2022).

The omitted variable bias occurs when two conditions are met: (i) the excluded variable is correlated with the included explanatory variable, and (ii) the omitted variable has a significant impact on the dependent variable. The assumption is that a negative correlation between pension income and gender does not necessarily mean there is a negative cause-and-effect relationship. The correlation we observed may be misleading because gender could be associated with other factors, like choosing different occupations with varying pay scales. To ensure an unbiased estimation of β_1, it is imperative to account for this variable. In this context, the omission of an independent variable, denoted as Xk, which significantly influences pension income ($\beta k \neq 0$) and is simultaneously correlated with the gender variable ($\rho XFemale, Xk \neq 0$), introduces bias into the estimation of β_1.
An alternative perspective suggests that omitting a relevant variable can lead to a correlation between the error term and the variable of interest. Mathematically, the omitted variable bias, assuming there are only two determinants of Yi, can be expressed as below:

$$\beta_1\ p \overset{p}{\to} \rho\ \beta_1 + \rho X\ u\ \frac{\sigma \mathrm{u}}{\sigma \mathrm{x}} \qquad \text{Eq. 4.10}$$

This implies that as the sample size increases, β_1 does not converge toward β_1 but instead approaches below:

$$\beta_1 + \rho X\ u\ \frac{\sigma \mathrm{u}}{\sigma \mathrm{x}} \qquad \text{Eq. 4.11}$$

This mathematical expression reveals two important insights that are directly relevant to the empirical analysis in this thesis. It is important to include variables in our statistical model that affect pension income and are related to gender. If factors like place of residence do not correlate with gender, they can be excluded from the model, even if they have an impact on pension income. There are two scenarios where the estimates from the complete model and a simplified model are the same.

1. When the omitted variable's estimated coefficient is 0 ($\hat{\beta} omit$ = 0), indicating that xk has no impact on $\hat{y}$ and thus should not be included in the model.

2. When the omitted variable is not correlated with the variable of interest, its exclusion from the model doesn't introduce omitted variable bias, even if the coefficient on the omitted variable is not 0.

These findings underscore the importance of including control variables in the empirical analysis, specifically those that influence pension income and are correlated with gender.
Additionally, a violation of this assumption can be due to a misspecification of the functional form. According to Assumption 1, the linear regression model does not require the explanatory variables to be linear, only the parameters need to be linear. This flexibility allows for incorporating different types of relationships, such as curvilinear patterns, between the independent and dependent variables.

Functional form misspecification is strongly linked to the bias caused by omitted variables. If a squared term is not included in the model to represent a curvilinear relationship, the error term in the regression accounts for this effect. Excluding the squared term can lead to omitted variable bias (Woodridge, 2019).

Although eliminating all potential sources of omitted variable bias is challenging, this thesis benefits from having data that helps reduce the most significant sources of bias. Section 5.3.2 will discuss the selection of relevant variables for the empirical analysis to reduce the risk of omitted variable bias.

5. Homoscedasticity

Homoscedasticity means that the variability of the error term is not dependent on the values of the explanatory variables. In simpler terms, it means that the variance stays the same and does not depend on the independent variables:

$$Var\ (u_i \mid x_1, \ldots, x_k) = \sigma^2 \qquad \text{Eq. 4.12}$$

The assumption of homoscedasticity allows for making inferences across observations without having to consider variations in variance based on different values of $X_{k,\ i}$. In the context of this paper, it would mean that the variance of u_i is the same for both males and females, implying that the variance of pension income is consistent for both genders. However, this assumption can be challenging to support, as the data could exhibit heteroskedasticity.

To address the issue of heteroskedasticity, robust standard errors are reported, as further discussed in Section 4.2.3. These robust standard errors are used in statistical hypothesis tests. Nevertheless, it's important to note that while the estimator of $\hat{\beta}_1$ remains unbiased and consistent, the standard error calculated in the hypothesis test may exhibit bias unless the dependency is properly addressed, a topic explored in more detail in Section 4.2.

4.2 Statistical Testing

Statistical tests play a crucial role in assessing the robustness of the estimator β_1 to variations within the sample. They enable us to determine whether the estimated coefficient can be statistically generalized to the broader population. As discussed in section 4.1.2, these statistical tests are instrumental in addressing the research question of this thesis from a statistical standpoint. The subsequent section provides an overview of the specific statistical tests employed in the empirical analysis within this thesis. To ensure clarity, significance tests will be presented using p-values, critical values, and test statistics. The tables will include significant stars to show the p-value thresholds for the t-test.

4.2.1 T-test of Coefficients

In the pursuit of answering the economic question outlined in Section 4.1.2, a two-sided test is conducted for each coefficient estimator of interest, specifically focusing on the 'female' variable from Equation (5.5). Assuming a normal distribution for the sampling distribution of $\hat{\beta}_1$ in large samples, a straightforward t-test is employed:

$$t\text{-}statistic = \frac{\beta_1 - \beta_{1,0}}{SE(\beta_1)} \qquad \text{Eq. 4.13}$$

Here, SE ($\hat{\beta}_1$) represents the standard error of $\hat{\beta}_1$, and $\beta_{1,0}$ is set to 0 based on the null hypothesis (Eq. 5.6). In large samples, the t-statistic adheres to a normal distribution. Despite the acknowledgment in Section 4.1.3 that the assumption of homoscedasticity might be unrealistic, employing robust standard errors in the denominator mitigates potential issues, ensuring the validity of the t-test results.

4.3.2 Testing for Mulitcollinariy

In econometric modeling, it is crucial to prioritize the reliability and interpretability of coefficient estimates. Multicollinearity requires careful examination. Multicollinearity occurs when independent variables in a regression model are strongly correlated (Woodridge, 2019). The interdependence between variables presents challenges in econometric analysis. It can increase the variability of coefficient estimates, resulting in less precise and potentially misleading results.

Multicollinearity does not impact the model's predictive accuracy but weakens the ability to determine the individual effect of each predictor (Gujarati & Porter, 2019). Multicollinearity is important to address in situations where it is crucial to understand the specific impact of each variable, such as in pension income analysis. It ensures that the influence of each variable is accurately interpreted without being affected by its correlation with other predictors in the model.

This study will use a correlation matrix to detect multicollinearity. A correlation matrix shows the relationships between all pairs of independent variables. By analyzing the correlation coefficients, we can identify variables that may have strong linear relationships (Woodridge, 2019). It is an important initial step in identifying potential issues, although it cannot capture more complex multicollinearity involving three or more variables. Section 5.4 will further discuss multicollinearity and the correlation matrix.

4.2.3 Addressing Heteroskedasticity

Heteroskedasticity is a common concern in econometric models, especially with large and complex datasets like the one used in this study. It occurs when the variance of the error terms in a regression model is not constant across observations (Wooldridge, 2019). This violation of the classical linear regression assumption can lead to inefficiencies in the Ordinary Least Squares (OLS) estimates, particularly affecting the reliability of standard error estimations.

In practical terms, heteroskedasticity can arise from a variety of sources in empirical data. For instance, in the context of pension income analysis, the variability in income and other socio-economic factors across individuals can contribute to heteroskedastic errors. Such variations often become more pronounced in datasets that span a wide range of demographics and economic conditions, as is the case in our study (Zumbach, 2009).

The presence of heteroskedasticity does not bias the coefficient estimates themselves; however, it does lead to biased standard error estimates (Wooldridge, 2019). This bias in standard errors can mislead inferences made about the significance of the coefficients, as traditional statistical tests assume homoscedasticity (constant variance of errors).

To address this issue, econometricians often turn to robust standard errors, also known as heteroskedasticity-consistent standard errors. These are designed to provide more accurate standard error estimates that are not biased by the presence of heteroskedasticity (Wooldridge, 2019). By accounting for the possibility of varying error variances, robust standard errors enable more reliable hypothesis testing and confidence interval construction. Considering these aspects, our study has chosen to report robust standard errors as described in the subsequent section. This choice is grounded in the recognition of potential heteroskedasticity within our dataset, and it aligns with best practices in econometric analysis to ensure the accuracy and reliability of our findings (Wooldridge, 2019).

Part 5.4 will demonstrate additional static testing and model diagnostics to show the robustness of our models. This section will focus on multicollinearity. Additionally, section 6.5 will cover model diagnostics.

5. Data

This section details the methodology applied in our empirical analysis, utilizing data from Norway's Statistics Central Office (SSB). It provides an overview of the dataset that forms the basis of the research presented in this thesis.

First, we provide a detailed overview of the dataset, including how different variables have been combined to create a comprehensive dataset with all the variables of interest. Next, we will discuss the process of sample selection and the important assumptions that guide these choices. This section will provide a detailed description of both the dependent and explanatory variables. The study will analyze the selection and modification of these variables using relevant literature. To better understand the data, we will include descriptive statistics. This will help us gain insights into the variables before we analyze them empirically.

5.1 Data Description

Secondary data is used in the empirical analysis to answer the research question of this thesis. SSB has provided access to administrative data on the entire Norwegian population for this thesis, using their analysis tool called microdata (Microdata, 2023). SSB is the main authority for Norwegian statistics, making it a reliable and objective source of data. SSB collects microdata that is longitudinal, meaning it gathers data on the same individual at different time points. The study of pension income relies on the historical factors of everyone, making this information valuable for empirical analysis (Rabe-Hesketh, S., & Skrondal, A. 2012).

This paper examines how various factors, including education, occupation, children, and other variables, impact an individual's pension income. The focus is on understanding the gender differences in pension income (see section 3.2). The registers are connected through personal identification numbers (PNR) and years, creating a dataset that is relevant to the research question of this paper (Microdata, 2023).

5.2 Sample Selection

We analyzed pension data from individuals aged 67 to 74 in Norway in 2017 to study gender differences in benefits. This year was selected because it contains all the necessary data about the three main components of the Norwegian pension system. To ensure a valid comparison of pension incomes, it is important to use a representative sample with consistent data. Our focus on this age group in 2017 is because we have comprehensive data and a need for reliable economic analysis.

Our study focuses on analyzing gender differences in pension payments in Norway using pension data from the year 2017. By focusing on this year, we can provide a snapshot of pension outcomes for individuals aged 67-74. This approach ensures that our sample has similar economic conditions during their working lives, which could impact pension accumulation (Peris-Ortiz et al., 2020). We analyze data from 2017 to account for variables. This ensures that individuals in our chosen age group transition into the pension phase under similar economic conditions. Being specific is important to identify the gender gap in pension benefits. It helps to reduce the impact of different economic conditions on the results. Our analysis focuses on a single year to avoid the complexities of fluctuating economic conditions over time (Sánchez Serrano & Peltonen, 2020).

The age of 67 is important in Norway because it is the earliest age at which someone can start receiving pension benefits from all three pillars (Nav, 2023). Pension systems and regulations change over time. This research focuses on individuals aged 67 to 74 in 2017 to ensure that the participants experienced retirement under similar regulatory frameworks. Limiting the age range helps prevent inaccuracies when comparing retirees from different years who may have experienced different pension regulations, contribution rates, or benefit calculations.

Limiting the age of the cohort has practical advantages, in addition to economic considerations. By narrowing down the year and age range, we can better understand the details of pension incomes for this specific group, instead of dealing with the complexities of a broader age or time range. This precision helps us better understand how pensions work for this specific group, without the complications that could arise from external factors like major regulatory changes or macroeconomic events affecting different age groups in different ways.

In summary, we chose to focus on the age group of 67 to 74 in 2017 because it strikes a good balance between economic reasoning and practical analysis. This approach ensures that our findings are both relevant and reliable.

The data was obtained from microdata and included multiple variables. Using a "left join" method to integrate these variables means that data will only be included if there is a match with individuals in the primary dataset (Microdata, 2023). The initial sample selection is refined and reduced as variables are added. To start, import a variable with the least missing values from a hypothetical population, such as gender, country of origin, or birthdate. The sample selection process reduces the sample size to 403,871. Out of the total observations, 199,628 (about 49.42%) are male and 204,208 (about 50.57%) are female. Gender data balance is crucial for accurate statistical analysis in this thesis (Stokes, 2022).

5.3 Variable Description

This section will explain the variables used in the analysis. To analyze the variables, we need to consider the construction, economic justifications from section 3, methodology from section 4, and data from SSB Norway. A brief analysis of the descriptive statistics for the variables will be provided.

5.3.1 Dependent Variable

The main goal of this study is to estimate the differences in pension between men and women. The dependent variable of the model is an index that shows the level of annual pension for each individual in the sample. Section 2 provided a detailed explanation of the Norwegian pension system and identified its pillars. Analyzing each pension pillar separately may provide a misleading representation of gender disparities in annual pension entitlements. To accurately analyze gender differences, it is important to take a holistic view of the pension system, considering that evidence suggests women rely more on public pensions (see section 2.5).

Pension wealth and pension income are commonly used metrics for calculating pensions in academic literature. Studying pension wealth gives an overall view of savings in old age, without considering specific payment methods like lump sums, annuities, and life rates (OECD, 2023). However, using pension wealth does not account for public pension entitlements for annual annuities, which may result in an inaccurate representation of the first pillar of the Norwegian pension system. When using pension wealth as the dependent variable, government assistance is not considered. Pension wealth can create bias when studying gender inequalities and may distort the perception of financial well-being in old age for men and women (Cordova, K., Grabka, M. M., & Sierminska. E, 2022). A significant portion of the Norwegian pension system, specifically the public pensions of the first pillar, would not be covered by it. However, when looking at annual pension income, this issue does not arise because it includes pension components from all three pillars of the Norwegian pension system, including the annuity payment linked to public pensions. Using pension income as the dependent variable would provide a more accurate approach to address the study question.

The dependent variable is the annual Pension Income of individuals aged 67 to 74 in the sample. The study's dependent variable is created by combining multiple variables to include all three pillars of the Norwegian pensions system. This variable represents the annual income pension for 2017, using the most recent data.

To understand gender differences in pension income, it is important to examine the distribution of pension income in the sample. Figure 5.1 shows the average pension income received by individuals aged 67 to 74 in 2017, categorized by gender. The pension income in this study is based on the three pillars of the Norwegian pension system.

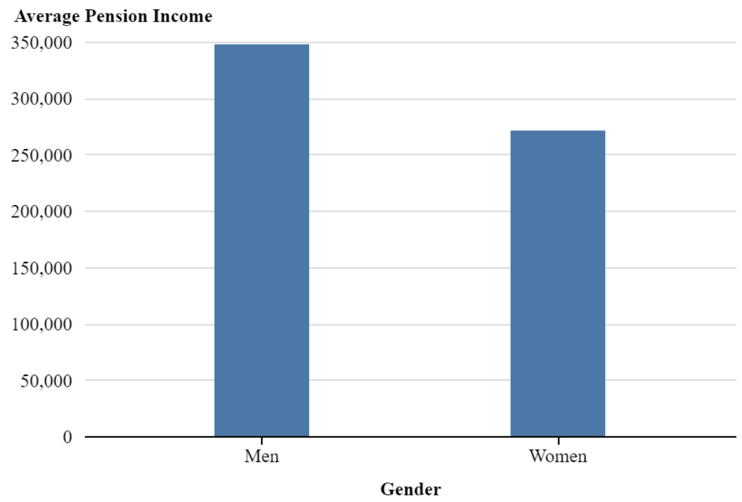

Figure 5.1: Average Pension Income Divided by Gender (2017)

The data shows clear gender differences in the average pension income. In 2017, men earned an average of 251,471 NOK per year, while women earned 179,785 NOK per year. This represents a difference of 71,686 NOK annually. Many studies use equation (5.1) to calculate the gender pension gap. This equation determines the percentage of women's average pension income compared to men's (Bettio et al., 2013).

$$\left(1-\frac{\text{women' average pension income}}{\text{Men's average pension income}}\right)*100 \qquad \text{Eq.5.1}$$

However, this method only accounts for the absolute difference in average pension income; it does not consider variations in personal factors, such as educational background and individual life choices, which can significantly influence pension outcomes (cf. section 3.2).

Equation 5.2 converts the dependent variable 'Pension Income' to its natural logarithm in our empirical analysis. As per Wooldridge (2019), it is common practice to transform income data to stabilize variance and normalize the distribution. This is done to better align with the statistical assumptions of regression analysis.

$$\text{Pension income} = \ln(\text{Pension Income}) \qquad \text{Eq. 5.2}$$

Adapting the dependent variable into natural logarithms, as shown in equation (5.2), is beneficial because it allows us to interpret the coefficients of the independent variables as the percentage change in pension income. Additionally, this transformation can mitigate the impact of outliers.

5.3.2 Explanatory Variables

This section will explain how the explanatory variables were created for the empirical study. The variables were chosen based on the economic justifications discussed in section 3.2.

Female

The main variable of interest in this thesis is gender. The gender variable from the register data provided by SST is converted into a dummy variable. The binary encoding of the dummy variable represents females as 1 and males as 0.

$$Female \ \epsilon \ \{0,1\}$$

Education Level

Higher levels of education are associated with higher pension income, as mentioned in section 3 and supported by previous studies (Bettio et al., 2013; Sefton et al., 2011). Recent academic studies have found that as women's education surpasses that of men in modern society, this variable becomes less important in determining gender inequalities in wage and pension income (Blau & Kahn, 2017). Contradictory research suggests that education level is important when studying gender gaps in retirement. Although there are differing opinions on the impact of education on the labor market and gender gaps (cf. 3.2.2.), it is still included in the model to analyze the potential outcomes of the gender gap in pension.

This study uses the variable NUS-code provided by SST to measure the level of education for each individual (SSB, 2023). The NUS-code represents the highest completed education. The variable has 5546 categories, which can be grouped into 9 categories based on the first digit of the code. Categories starting with zero represent no education or basic education, including kindergarten or preschool programs. Categories starting with 1 represent primary school, while categories starting with 8 represent PhD programs and Doctorates. Categories 4, 5, and 6 are considered equivalent to a bachelor's degree, so they are grouped together. Therefore, there is a categorical variable for education level.

$$\text{Education Level} \in \{1,2,3,4,5,6,7\}$$

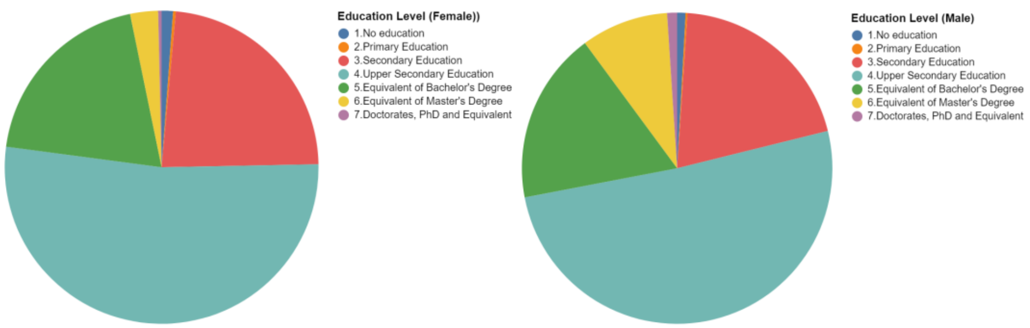

Figure 5.2: Share of Education level Between Men and Women in 2017

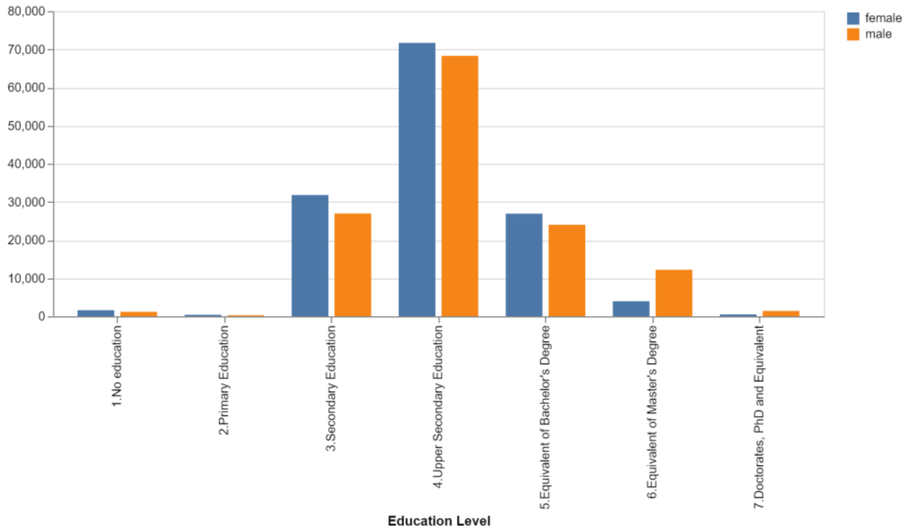

Figure 5.3: Distribution of educational level by gender

Figures 5.2 and 5.3 examine the distribution of education levels by gender. In general, women are slightly more educated than men across most levels. However, in the last two categories, men have a larger share. Overall, women have higher levels of education than men. However, a larger proportion of men have higher education compared to women. This observation supports the theory discussed in section 3, which suggests that higher levels of education should lead to higher pension income in old age. Men with higher education are expected to receive more pension income. However, women are generally more educated, but since we study the average pension income, the average education level is what matters. The average pension income for females is slightly lower than that of males, indicating potential underlying factors. Women may be more represented in teaching professions, which require significant education but do not offer equal financial rewards. Additionally, this difference may be due to gender-based discrimination in pension income and educational opportunities. Even if women have the same level of education as men, they may not receive equal pension income (Belingheri, P., Chiarello, F., Fronzetti Colladon, A., & Rovelli, P. 2021).

Experience

Section 3 of the economic theory explains the gender pension gap by considering experience as a key factor. Due to limited statistics on real job experience, accurately quantifying it is challenging. One way to gain experience, as mentioned in literature (Möhring, 2014), is by calculating the total number of years spent working. Zveglich et al. (2019) suggest adding job probabilities for each year of working age to calculate projected work experience.

However, this proxy has limitations, and the findings cannot be applied to different types of data. Additional scholars (Filer, 1993; Garvey & Reimers, 1979; Moulton, 1986) have developed more indicators of experience. Most empirical investigations follow Mincer's (1974) suggested approach. Mincer (1974) developed an equation to analyze the wage gap, which relates experience to age and education. The function assumes that job experience is calculated by subtracting the number of years of schooling from a person's present age, and then subtracting six years to account for the typical age at which children start school.

$$Experience = age - educ - 6 \qquad \text{Eq 5.3}$$

It's important to note that this index represents potential job experience, not actual work experience. One may argue that it is illogical to assume that every sample member has worked continuously since completing their education. This index is only reliable for individuals who have strong connections to the labor market. In section 3, it is explained that women's contribution to the job market is more discontinuous than men's because they typically have more household and childcare responsibilities in societies. This measurement mistake is more likely to impact women than men, resulting in a negative bias of returns in pension income (Miller, 1993). However, despite these drawbacks, the use of prospective experience is now common in previous literature (Zveglich, Rodgers, & Editha, 2019).

To create this variable according to equation 5.3, a number is assigned to each level of education (from 1 to 9 as explained in the previous section), which represents the number of years usually needed to complete that level of education. For example, a normal child spends 7 years in primary school according to the Norwegian education system. Then, having the age of each person in the sample, the potential years of experience are calculated. Again, to make the interpretation of the coefficient easier for the variable, 4 categories are created which respectively represent the potential work experience of 30-40, 40-50, 50-60, and 60-70 years.

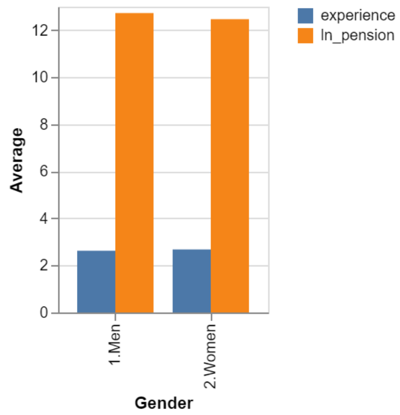

Figure 5.4: Average level of experience and Average Logarithm of Pension income across Genders (2017)

By looking at Figure 5.4, which presents data on the average level of experience for both women and men, it is evident that contrary to our initial expectations, females possess a slightly higher level of experience compared to males. On average, women have an experience level of approximately 2.68, while men have an average experience level of approximately 2.62. Further analysis will be discussed in section 6.

Financial Literacy

Financial literacy is used to infer an individual's investment tendencies in empirical research. Section 3.2.7 emphasizes that understanding financial concepts affects a person's inclination to invest their savings. Women often seek expert advice before making investment decisions, suggesting they may be less assertive than men in this area. One reason for this behavior is that women generally have less financial knowledge than men (Doe, J., & Smith, A. (2020).

A binary variable was created to measure financial literacy, based on whether an individual has an educational background in economics or finance. This approach includes people who have knowledge of economics or finance. This includes specialized education in finance, economics, pensions, business administration, accounting, insurance, and auditing. Additionally, it has significance to include degrees that are not solely focused on economics, such as combined courses in Marketing Management and Economics. This binary variable assigns a value of 1 to individuals with a finance or economics background, and 0 to others.

$$\text{Financial Literacy} \in \{0,1\}$$

The choice of educational fields used as benchmarks for financial knowledge was guided by the Norwegian standard for educational grouping (NUS2000). The SSB uses this standard to classify people's educational achievements and backgrounds in their educational statistics and other metrics that involve education as a factor (SSB, 2023). As detailed above, the variable for financial literacy is created from individuals who have undergone any form of economic education during their lifetime. Economic principles suggest that people who have a better understanding of finances tend to have higher pension income in their later years. (cf. 3.2.7) This is due to better investment decisions and savings management over time.

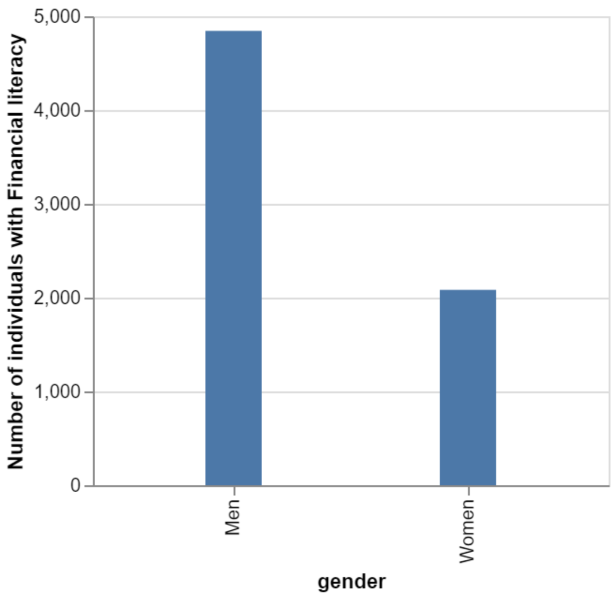

Figure 5.5: Financial Literacy by Gender (2017)

Figure 5.7 illustrates the relationship between gender and financial literacy in the chosen sample. The initial data analysis supports economic theories, indicating that men in this sample have higher financial literacy than women, on average. Section 6 will provide a detailed analysis of the specific effects.

Labor market

There is a clear difference in the number of men and women in the labor market. Women generally participate in part-time work more frequently than men (see section 3.2.5). Traditional gender roles, like caregiving, contribute to this trend. After giving birth, many women switch from working full-time to working part-time to accommodate their family responsibilities (Veremchuk, 2020).

Historical labor market participation is important for analyzing pension income. The dataset provided by Statistics Norway (SSB) does not include the complete work history of the individuals in our sample. Despite this limitation, our methodology aims to analyze the impact of current employment patterns, specifically part-time work, on the transition into retirement. We analyze the impact of work attachment in 2017 on individuals aged 67-74. This group includes people who have become eligible for pension at different times with diverse backgrounds. Our analytical framework combines working hours and job percentages to create four dummy variables that evaluate labor market engagement. These variables represent different categories of labor participation. '0-10%' refers to individuals working up to one-tenth of full-time equivalence. '10-50%' refers to individuals working between one-tenth and half-time. '50-80%' refers to individuals working between half-time and four-fifths of a full-time schedule. We aim to study labor engagement among those transitioning to pension and its impact on pension outcomes. These categories serve two purposes. Firstly, it helps us understand the level of employment during a critical time when people are about to retire or have recently started their retirement. Additionally, it allows us to compare full-time and part-time employment during this important period.

While valuable, it's important to acknowledge the limitations of our analysis. Our approach is missing historical labor market participation data. The dataset from SSB does not include longitudinal information about our sample, which would provide a complete view of an individual's entire work history. This context is necessary for fully understanding pension income. Our current method only considers the immediate impact of employment before receiving a pension, but it may not account for the overall impact of an individual's entire career on their pension. As we interpret the findings, it is crucial to consider this limitation and understand that historical labor market participation remains a missing piece in our analytical framework. Section 7 will provide more details about our limitations.

The average work percentages (also known as "Stillingprosent") for men and women in various work % categories are shown in the figures below. Men typically work longer hours than women do, however women are overrepresented in the 10–50% and 50%–80% categories. However, in comparison to categories 1 and 2, these groups only comprise a minor portion of the population. In section 6, there will be a presentation of additional analysis of these findings.

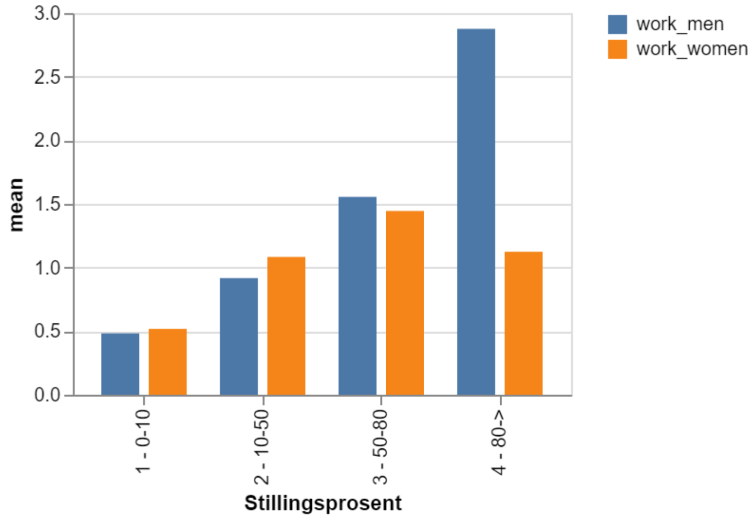

Figure 5.6: Average work percentages by gender in 2017

Industrial Sectors

Because of the inherent relationship between occupations and incomes and the earnings-based design of the Norwegian pension system, an individual's occupational status has a major impact on their pension income (see section 3.2.3). Notably, there are differences in how men and women are distributed among different sectors and occupations. Generally speaking, women are more likely than men to work in manual labor, service support, and lower-paying administrative jobs (see section 3.2.3). As a result, women often contribute less to private pension plans and accrue less in employment pensions. Consequently, it is expected that the industry a person works in would be a major factor in the gender pension gap due to differences in job titles (more on this in section 3.2.3).

The variable on industrial sectors includes 9 dummy variables and is classified according to the standard occupational classification (SSB,2011) The dummies represent the occupations within the 10 fields: i) Military and non-given, ii) Managers iii) Academic professions, iv) College occupations v) Office occupations, vi) Sales & Service vii) Farmers, fishermen, etc., viii) Craftsmen, ix) Process, Machine, Transport x) Cleaners, Auxiliary, etc.

However, our analysis faces a fundamental challenge. We use an analytical approach to understand how individuals are connected to the labor market as they near retirement. We aim to select a year close to retirement age but still early enough to include people who are actively working. We chose the year 2015 as our data anchor because it is the earliest year where we have data for every individual, even though it may not align perfectly with everyone's retirement age.

Although limited, the 2015 data provides valuable insights into the work situation of the people in our sample near retirement. It is important to note that post-2015 data may show a decrease in labor market activity because of retirement. Although the dataset has weaknesses, it is important to highlight that it still has significant value due to its substantial size. Our substantial and varied sample ensures the significance of our findings. The data provides valuable insights into how different occupations at the very end of the career affect pension outcomes, highlighting gender disparities and their causes.

The sector of employment is important in determining gender-based differences in pension earnings, as mentioned in section 3.2.3. This research clearly shows the differences in male and female representation across professional sectors. Figures 5.7 and 5.8 display the breakdown of gender-specific distribution across sectors and the corresponding average pension earnings.

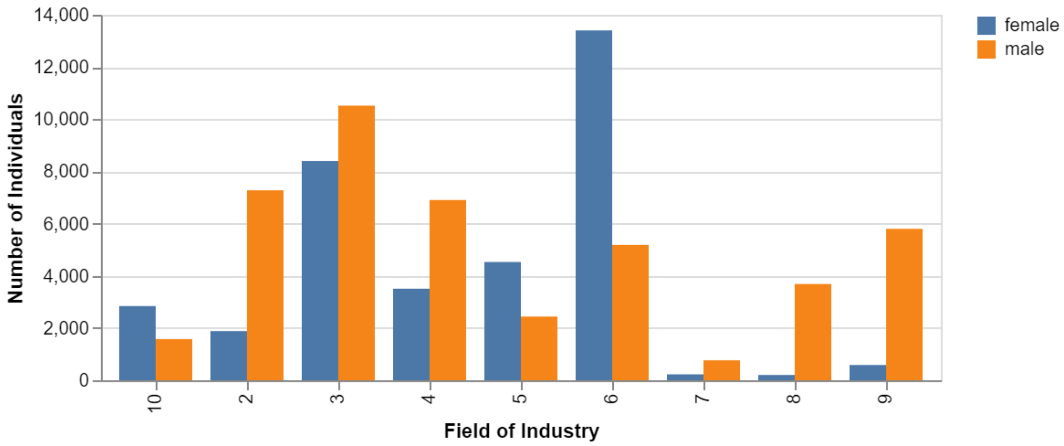

Figure 5.7: Breakdown of employment in industry sectors, by gender in 2015

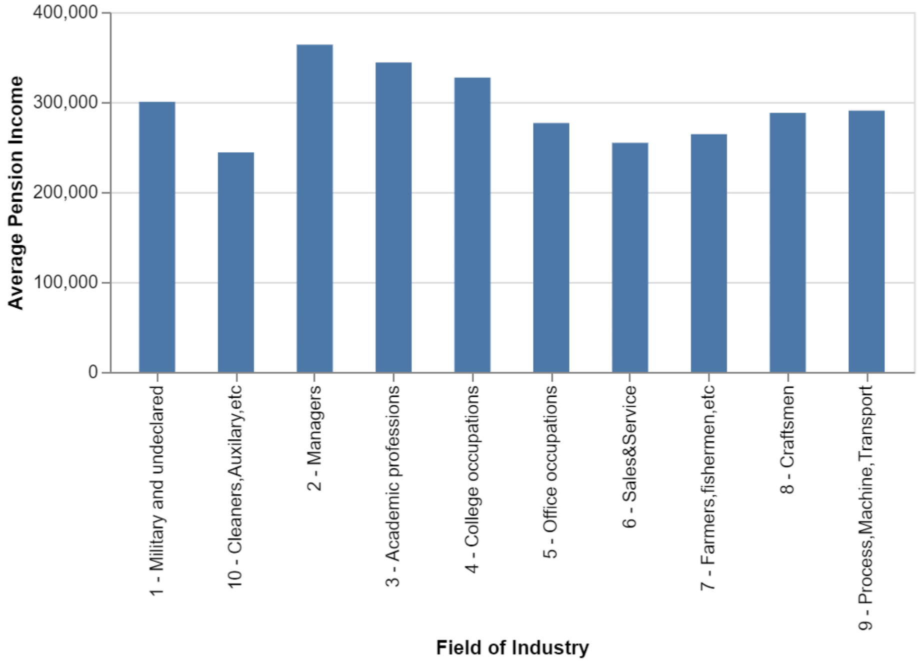

Figure 5.8: Average income pension by work of field

Figure 5.8 shows clear differences in the representation of males and females in various industrial sectors. Women dominate the military, office and service, sales, and cleaning industries in this sample. Managerial occupations have the highest average pension income compared to other industries. Regarding the pensioners in Military occupations, they can work in the private sector and earn unlimited income without any reduction in their pension, which can also affect their overall results (Forsvaret, 2023).

The graphs clearly show that men are the majority in certain fields. Males are more numerous in the "Managers" and "College occupations" sectors. Interestingly, the sectors shown in the lower graph also have higher average pension incomes. This suggests that managers and college professionals may earn more money over their careers. These professions may naturally lead to higher pension incomes because they are usually linked to lifetime earnings. The graphs show that men dominate in these sectors, with around 70% in "Managers" and over 60% in "Academic Professions". Furthermore, these two sectors together make up a significant portion of the entire sample. Based on this data, males are more likely than females to be in sectors with higher pension payouts.

Marital Status

The individual's marital status for the year 2017 is relevant for determining their pension incomes. The variable is constructed by considering retrospective observations, including marital histories and divorces. In 2017, a person's marital status is determined by considering their past marriages. The variable used is Sivilstand, which is provided by SSB. It has 9 categories: Single, Married, Widow/Widower, Divorced, Separated, Registered Partner, Divorced partner, and Survived partner. The study combines similar categories in their logic and creates three dummy variables: i) single and never been married or in a partnership, ii) single, but previously married, and iii) married or in a partnership.

According to the literature, marital status has different effects on men and women, as discussed in section 3.2.6. Single or divorced women have higher pension income (about 50%) compared to married women in all countries, whereas men benefit from marriage. Research shows that married men tend to have higher earnings compared to unmarried men. Additionally, married men experience a greater increase in wages if their wives do not work, compared to when their wives also work full-time. This is also reflected in their pension income, as discussed in section 3.2.6.

To clarify the impact of gender on pension income, we include interaction terms between the gender variable and different marital status categories. Interaction terms are used to examine gender-related variations in married status. Including this additional factor is beneficial because it allows for variations in the coefficient on the female variable, which supports the chosen methodology for this study.

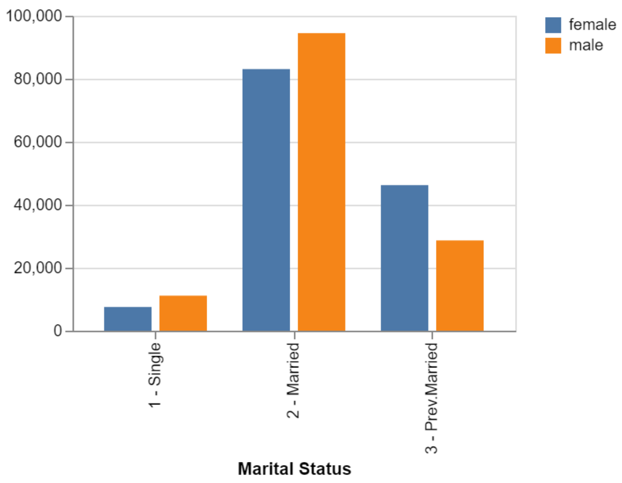

Figure 5.9: Distribution of marital status by gender (2017)

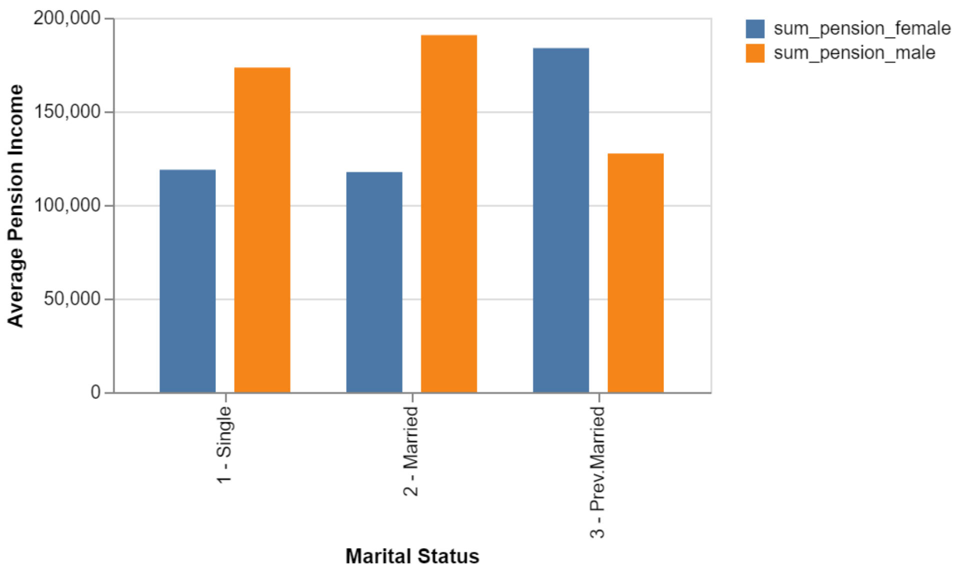

Figure 5.10: Distribution of marital status and average pension income, by gender (2017)

Figures 5.9 and 5.10 show that married men receive higher pension income than married women. However, in the previously married category, women receive more income than men.

Children

In this thesis, the number of children is used as a measure of labor market interruptions. This helps us understand how childbearing affects pension incomes, and also allows us to examine potential gender disparities as highlighted in section 3.2.4. We will investigate how the number of children interacts with gender, as this can reveal gender-specific differences related to this variable. This is important because we are focusing on gender disparities in pension incomes, rather than general pension effects.

In this study, the child's variable is merged with the family identification number in SSB, attributing it to both parents. According to SSB's classification, a child is considered independent from their parents when they have a separate permanent address from the family residence or parent if they have different addresses. SSB's methodology for tracking child count may not accurately represent the overall population. When people retire, the number of children they have may change if some of their children have already moved out.

To account for possible measurement errors, we have used the maximum number of children reported by everyone as of 2005, the earliest data available for our group. Adopting this approach introduces two specific constraints:

i) Instances where siblings have a wide age gap, leading to scenarios where the older child has already established an independent residence by 2005.

ii) Tragic instances of child mortality, which under this methodology, become indistinguishable from cases where children simply move to a separate address.

Acknowledging these limitations is crucial. However, these scenarios are rare in our dataset. The large sample size makes it unlikely that these exceptional cases will greatly affect our findings.

Empirical evidence consistently demonstrates that having children can affect labor market participation, which in turn affects pension income after retirement (see section 3.2). The association between the number of children and pension income is shown in Figure 5.11. The data shows a pattern: the average pension income tends to decline with the number of children. This is consistent with research that indicates having more children generally results in a lower pension in later life. This decrease might be ascribed to higher childcare costs, which may result in irregular employment (see section 3.2 for more information).

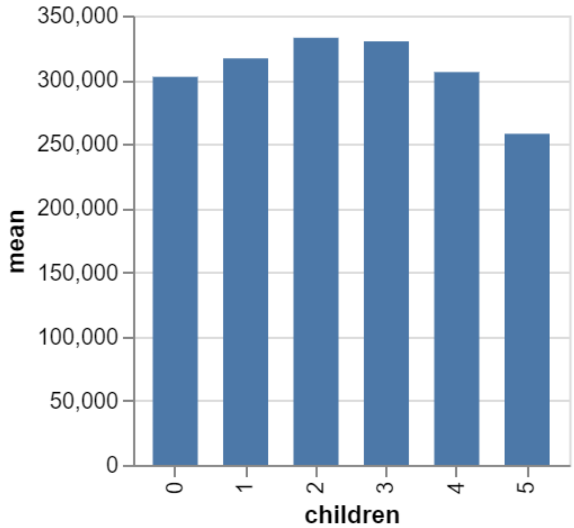

Figure 5.11: Distribution of the number of children and average pension income for the whole population in the sample (2005)

5.4 Collinearity

Multicollinearity is a statistical phenomenon that arises when two or more independent variables in a regression model exhibit high correlations with one another (Woodridge, 2019). Determining the distinct impacts of every variable on the dependent variable becomes challenging as a result. This phenomenon can lead to inflated standard errors and imprecise calculations, and it may make it difficult to discern the true links within the model. Verifying the correlation between variables is crucial since high correlations may point to multicollinearity issues. Strong correlations might make the model sensitive to small changes in the data and make it challenging to identify the specific effects of each variable on the dependent variable. The identification and handling of multicollinearity ensures that the coefficients accurately reflect the relationships between the variables and helps to prevent false or misleading conclusions in statistical studies, which is essential for the reliability and interpretability of regression results (Woodridge, 2019).

Thus, in this section, following a thorough overview of the explanatory factors included in the research, the correlation matrix is supplied and displayed in Table 5.12. The correlation matrix acts as a guide, helping us navigate through the complex web of factors under examination. By investigating the direction and intensity of relationships, we can learn more about the dynamics present in our dataset. Additionally, the discovery of multicollinearity—a phenomena that may slightly affect the results of later analyses—is made possible through this exploration. Low correlation ($|r| < 0.3$) between variables is generally seen as acceptable and denotes a weak association, raising no urgent concerns. In general, moderate correlation ($0.3 \leq |r| < 0.7$) is appropriate as well since it provides information about moderately strong relationships, which should be carefully interpreted considering the goals of the study. When dealing with strong correlation ($|r| \geq 0.7$), especially between independent variables, care should be taken because it could indicate multicollinearity (Woodridge, 2019).

Examining Table 5.12, which displays the matrix of correlations among the variables included in the research, one can observe that Education Level and Experience have the highest degrees of correlation. The relationship between the variable Experience and years of schooling is highly dependent, which is not surprising. Assuming that an individual's experience is a function of their age and years of schooling (eq. 5.3), a high degree of correlation between those two variables is already reasonable. Section 6 will address the problem by including only one variable of each type in the model. The findings will be presented separately for each variable.

	Female	Education Level	Experience	Financial Literacy	Industry Field	Labor Market	Marital Status	Children
Ln (Pension Income)	-0.21	0.12	0.05	0.05	-0.12	-0.08	0.02	-0.01
Female		-0.11	0.05	-0.07	0.04	-0.23	0.17	-0.11
Education Level	-0.11		-0.64	0.17	-0.37	0.11	-0.06	0.16
Experience	0.05	-0.64		-0.15	0.23	-0.2	0.06	-0.18
Financial Literacy	-0.07	0.17	-0.15		-0.1	0.05	-0.02	0.03
Industry Field	0.04	-0.37	0.23	-0.1		0	0.03	-0.07
Labor Market	-0.23	0.11	-0.2	0.05	0		-0.02	0.1
Marital Status	0.17	-0.06	0.06	-0.02	0.03	-0.02		-0.05

Table 5.1: Correlation Matrix

6. Analysis and Discussion

This section outlines the primary results derived from the empirical study. Initially, the results of the multiple linear regression will be presented, and subsequently, a subsample for the purpose of robustness check will be conducted.

Evaluating disparities in retirement payments between genders often begins with analyzing the mean differences between male and female retirees. The primary aim of this thesis is to delve deeper into the underlying explanations for such disparities in the Norwegian context. So, to achieve this goal, first, Model 1 is presented as an unadjusted view of the gender pension gap. Then, Model 2 will be presented to offer a more comprehensive insight. This model illustrates the potential shortcomings of relying solely on Model 1 and seeks to provide a more unbiased representation of gender differences in retirement payments.

The first two assumptions relating to the linear regression model are linearity in parameters (1) and random sampling (2) (cf. section 4.1.3). They relate to the applied model's properties and the sample selection, respectively, and are discussed before running a regression model.

(1) In this section, linear regression analysis is utilized to recognize the impact of gender on pension income while accounting for several control variables. The foundational premise of the model is its linearity. This denotes that there's a linear connection between pension income and its influencing variables, paving the way for the use of multiple linear regression. The majority of our regressors are classified as dummy variables and the rest are categorical variables (see section 5.3.2). This naturally satisfies the linearity prerequisite.

(2) This study uses the gross sample of the Norwegian population given data provided by SSB. From this population data, the sample has been reduced into a smaller net sample to ensure only including relevant observations and avoiding measurement errors. As explained in section 5, this study is anchored around data regarding individuals aged 67-74 who were receiving pension income in 2017. The significance of the year 2017 is highlighted by its status as the most recent year with comprehensive data encompassing all the essential variables that form the trifold pillars of the Norwegian pension system. Drawing from the vast pool of data provided by SSB, we narrowed it down to a more precise net sample. The objective was to include only the most relevant observations (as detailed in section 5.2).

In this study, robust standard errors have been reported. This decision is made to recognize the potential heteroscedasticity within the dataset. Robust standard errors provide a more reliable estimation of the standard errors in the presence of heteroscedasticity, offering results that are less sensitive to violations of the classical assumptions. This method enhances the robustness and reliability of our statistical inferences, contributing to a more accurate interpretation of the findings (Angrist, & Pischke, 2009).

6.1 MODEL 1: Simple Regression

Model 1 is the initial model in this analysis, as it includes solely the gender indicator without incorporating any control variables to the dependent variable, which is the logarithmic transformation of pension income. This model serves as the benchmark within this study, isolating the gender difference in pension incomes. Essentially, this model quantifies the average pension income disparity between genders on a percentage basis, representing the 'unadjusted' gender pension gap. The formula for this model is as follows, with the empirical outcomes presented in Table 6.1.

$$\text{Ln (Pension Income)} = 0 + 1\text{Femalei} + \text{ui} \qquad \text{Eq. 6.1}$$

In this model, the inclusion of a constant term designates males as the reference category. The coefficient 1 reflects the average percentage difference in pension incomes observed for females compared to males. In Table 6.1, the coefficient associated with the female variable stands at -0.2846. This indicates that women, on average, have a 28.46% lower pension income than their male counterparts. The significance of this coefficient at the 1% level is derived from the t-values, which rely on the robust standard errors presented in the model. This outcome aligns with our expectations according to previous literature explained in section 3.1.

The model's R-squared value is 0.0758, suggesting that the variable female accounts for 7.58% of the variability in the dependent variable, ln(pension). However, the absence of control variables in the model means that it overlooks sociodemographic factors influencing pension income and differing by gender. This omission means that various elements influencing both pension income and gender are encapsulated within the error term of Model

1. This oversight pertains to the zero conditional mean assumption, making the model prone to omitted variable bias in its estimated coefficient.

	Model 1 (Baseline for gender differences)	
	Coef	**Robust Std. Err**
Female	-0.28463	0.001562 ***
Constant	12.68489	0.001097 ***
Adjusted R-square	7.587 %	
R-square	7.587 %	
Sample size	403,870	
Note: $* \ p < 0.10, ** \ p < 0.05, *** \ p < 0.01$		

Table 6.1: Simple Regression

To sum up, Model 1 highlights that females, on average, receive 28% less pension income compared to their male counterparts. However, this gap and the influence of the gender variable is not reliable yet, since the unadjusted disparity doesn't incorporate other gender-related factors that could influence pension income. In the next step, Model 2 is introduced which concludes other significant determinants of pension income that might also be associated with gender, as control variables.

6.2 MODEL 2: Multiple Regression with Control Variables

In Model 2, the relevant variables are identified based on previous literature that are explained in section 5. Most of the control variables are categorical variables, each with k mutually exclusive categories, and since the model is estimated to include an intercept, the no perfect multicollinearity assumption (3) will not hold. Thus, only $k-1$ dummies of every given categorical variable are included. This implies that the omitted category of the given dummy is represented in the intercept. Hence, Model 2 is specified in the following form:

$$\ln(PI) = \beta_0 + \beta_1 Female_i + \sum_{k=2}^{8} \beta_k \, \varphi_{k,i} + \sum_{k=9}^{12} \beta_k \, \chi_{k,i} + \sum_{k=13}^{16} \beta_k \, \vartheta_{k,i} + \sum_{k=17}^{26} \beta_k \, \zeta_{k,i} + \beta_{27} Financial\ Literacy_i + u_i$$

E.q. 6.2

Where $\varphi_{k,i}$ represents 6 dummy variables for ‘Educational Level’, where the first category for education (refer to 5.3.2) is the omitted dummy. Furthermore, $\chi_{k,i}$ represents 3 dummy variables for experience while the first category (30-39 years of experience) is omitted dummy. Additionally, $\vartheta_{k,i}$ represents ‘Activity in the labor market’ as another categorical variable that includes 3 other dummies in which the first category (0-10% activity) is omitted dummy. The next categorical variable is $\zeta_{k,i}$ in which presents the 'Industrial sectors' including 10 categories, of which the first category "Military and undeclared" is omitted dummy in this case. The last explanatory variable is another dummy representing 'Financial Literacy' that as explained in section 5.3.2, holds the value of 1 in case the individual has an education background in Economics, Finance, or similar fields, and is holding a value of zero otherwise.

	Model 2 (Control variables included)		Model 2 (Excluding 'Work Experience')	
	Coef	**Robust Std. Err**	**Coef**	**Robust Std. Err**
Female	-0.26035	0.001492 ***	-0.26154	0.00151 ***
Education Level				
1. No Education	Omitted		Omitted	
2. Primary Education	-0.13092	0.047477 *	-0.20621	0.027407 ***
3. Secondary Education	0.46033	0.05467 ***	0.37087	0.013814 ***
4. Upper Secondary Education	0.66967	0.054657 ***	0.53361	0.013769 ***
5. Bachelor's Degree or Equivalent	1.00310	0.054705 ***	0.76973	0.013852 ***
6. Master's Degree or Equivalent	1.17315	0.054849 ***	0.90417	0.0143 ***
7. Doctrates, PhD or Equivalent	1.20616	0.05684 ***	0.84025	0.021934 ***
Work Experience				
1. 30-40 years	Omitted			
2. 40-50 years	0.85803	0.078135 ***		
3. 50-60 years	1.04360	0.078178 ***		
4. 60-70 years	1.13412	0.094396 ***		
Involvement in Labor Market				
1. 0- 10%	Omitted		Omitted	
2.10-50%	-0.04819	0.003923 ***	-0.04596	0.003979 ***
3.50-80%	-0.15085	0.00669 ***	-0.16695	0.006782 ***
4.Higher than 80%	-0.25084	0.005082 ***	-0.27078	0.005134 ***
Industrial Field				
1.Military and non-given	Omitted		Omitted	
2.Managers	0.16377	0.006792 ***	0.13828	0.006812 ***
3.Academic professions	0.03474	0.004835 ***	0.00846	0.004822 *
4.College occupations	0.12635	0.005753 ***	0.08772	0.005732 ***
5.Office occupations	0.09466	0.006423 ***	0.05878	0.0065 ***
6.Sales&Service	0.01725	0.004116 ***	-0.00711	0.004174 ***
7.Farmers, fishermen, etc	-0.08167	0.013087 ***	-0.09052	0.013236 *
8.Craftsmen	0.06101	0.007622 ***	0.02791	0.007675 ***
9.Process, Machine, Transport	0.08656	0.005642 ***	0.06852	0.005703 ***
10.Cleaners, Auxiliary, etc	-0.03518	0.008979 ***	-0.05185	0.009139 ***
Financial Literacy	0.04017	0.005944 ***	0.03752	0.005958 ***
Constant				
Adjusted R-square	21.037 %		19.234 %	
R-square	21.041 %		19.238 %	
Sample size	403,870		403,870	
Note: $* \ p < 0.10, ** \ p < 0.05, *** \ p < 0.01$				

Table 6.2: Model 2 Table of Results

Current control variables in Model 2 are chosen relying on the reviewed literature in section 3 (cf. section 3.2), to mitigate omitted variable bias. Thus, Model 2 is a better-specified model to explain the gender influence on pension income and provides a fairly unbiased estimate of the coefficient on variable ‘Female’. It is important to note that according to the possibility of revealing confidential information, Microdata analysis tool does not provide the constant in the table of results in this multiple regression in Model 2 (Microdata, 2023).

As mentioned in section 5.4, the variables ‘Experience’ and ‘Education Level’ show a high level of correlation due to their structure. This issue can bring up the possibility of collinearity. Thus, in Model 2, two different regressions are conducted, one including all explanatory variables and one excluding ‘Experience’ to avoid the potential collinearity with ‘Education Level’. The results of Model 2 are reported in Table 6.2 and show an estimated coefficient on ‘Female’ of −0.2603 which is statistically significant at the 1% level in both versions.

The R-squared of the model has increased remarkably compared to Model 1 and shows that the included control variables in combination with ‘Female’, explain 21% of the variation in pension income, while the coefficient has not increased drastically. The gender difference in pension income is found to be about 2% smaller than the unadjusted pension gap in Model 1 after accounting for the variables included in Model 2. This suggests that there is still a significant gender effect, with women receiving an average of 26% less in pension income than men, even for individuals who are more similar in terms of sociodemographic and economic factors. It demonstrates that the differences between men and women are difficult to explain, and the gap appears to remain even after controlling variables are included.

The control variables added in Model 2 contain useful information on pension income, making it important to develop a general intuitive understanding of the variables' signs. This is especially true given that the majority of the control variables are statistically significant at the 1% level as seen in Table 6.2. Looking at the coefficient of Financial Literacy, 0.04, one can interpret it as having a background education in financial literacy can have a positive effect on pension income. This is in line with the intuitions from the literature mentioned in section 3.2.7. A positive effect was expected since having more financial literacy helps individuals to make better decisions regarding their investments, and those good investment decisions will be reflected in their future payments. In this case, this effect is around 4%.

Furthermore, as highlighted in the literature (cf. section 3.2.2), it can be demonstrated that having a higher education has a partially favorable impact on pension income at a relative cost to having no education. Though this seems counterintuitive, the results claim that having a basic education lowers pension income relative to not having any education. One can explain it as people who only complete the required education have more years of work experience and can therefore accumulate pension contributions over longer periods. It is also observed that the higher the education, the higher the coefficient for the variable 'Education Level', in which the biggest jump is from category 2 to 3 which means having secondary education can be very effective in changing the level of payment for individuals. The interesting fact is that the more an individual is educated, the effect of it on their pension payment is not the same and is increasing at a diminishing rate. This seems correct since getting a PhD, for individuals holding a master's degree is not very favorable regarding their payments in the Norwegian job market according to intuitive observations.

Furthermore, the sign and amount of the coefficients for the variable 'Stillingprosent' (which indicates the percentage of individuals' activity in the job market after retirement) are in line with expectations relying on the mentioned literature in section 3.2.5. According to the Norwegian pension system, the more the individual is active based on their interest, they will receive less pension income, but this income cannot be less than a certain amount. We can see that the coefficients are all negative in this case, meaning that people with 10-50% activity, will receive 4% less compared to the people who are inactive after retirement or have very small contributions (less than 10%), and ultimately people who are active more than 80% in the labor market will receive approximately 25% less than inactive group.

The variable experience can also since we observe a positive coefficient which shows a diminishing increase among categories. Meaning that compared to the group with the least experience (30-39 years) which can be the base experience for an individual to retire, the more experience, the pension income would be higher. Also, it is interesting to note that this increase in income is not substantial, moving from the 3rd category (50-59 years) to the 4th category (60-69 years). This means that experience cannot be very effective after some point since it cannot add very substantial knowledge and skill to the employee after a certain point.

The last control variable in this Model is the 'Industry sectors' which represents the type of industry that each individual was active in. As mentioned, and discussed in section 3.2.3, one reason for the gender difference in pension income may be potentially from the fact that men are usually more active in certain fields and industries which is not common among women. By adding this control variable, we try to measure this potential effect. As explained in Section 5.3.2, this variable is categorical with 9 different categories each including a set of activities. Looking at the coefficients of this variable, it is recognizable that categories 2 and 3 are receiving respectively 16% and 3% more pension income compared to the first category. Considering that categories 2 and 3 are respectively present Managerial Professions and Academic professions, it can make intuitive sense. On the other hand, categories 7 and 10 show a negative effect on pension income compared to the first category. Because these categories represent "Farmers, fishermen, etc." and "Cleaners, auxiliaries, etc." respectively, it can make intuitive sense.

To investigate more about the factors explaining the gender differences, in the next step, Model 3 is introduced, that is including additional control variables and the interaction terms with the variable of interest ‘Female’. Previous economic research indicates that pension income is significantly influenced by children and marital status in relation to gender.
Model 3 is developed considering the literature to determine if gender influences on marriage and having children exist and whether they may partially account for the gender pension disparity reported in Model 2.

6.3 MODEL 3: Multiple Regression with Additional Control Variables

In a pooled regression, the interaction of the female with the marital status and child variables enables the analysis of gender differences in these control variables and the determination of whether the pension difference is dependent on the values or distinct categories of these two variables. Thus, the specifications of Model 3 are as follows, and Table 6.3 presents the findings. It is important to note that according to the possibility of revealing confidential information, Microdata analysis tool does not provide the constant in the table of results in this multiple regression in Model 2.

$$\ln(PI) = \beta_0 + \beta_1 Female_i + \sum_{k=2}^{8} \beta_k\, \varphi_{k,i} + \sum_{k=9}^{12} \beta_k\, \chi_{k,i} + \sum_{k=13}^{16} \beta_k\, \vartheta_{k,i} + \sum_{k=17}^{26} \beta_k\, \zeta_{k,i} + \beta_{27} Financial\ Literacy_i + \sum_{k=28}^{32} \beta_k\, \tau_{k,i} + \sum_{k=33}^{35} \beta_k\, \delta_{k,i} + \beta_{36} Pre.Married * Female_i + \beta_{37} Married * Female_i + \beta_{38} Women\ with\ children_i + u_i$$

E.q. 6.3

Where $\varphi_{k,i}$ represents 6 dummy variables for educational level, where no education is the omitted dummy. Furthermore, $\chi_{k,i}$ represents 3 dummy variables for experience while the first category (30-39 years of experience) is omitted dummy. Additionally, $\vartheta_{k,i}$ represents activity in the labor market as another categorical variable that includes 3 other dummies in which the first category (0-10%) is omitted dummy. The categorical variable $\zeta_{k,i}$ presents the industrial sectors including 9 categories, in which the first category "Military and undeclared" is omitted dummy in this case. The variable $\tau_{k,i}$ represents the variable 'Marital Status' including 3 categories of Single, Single but previously married, and married individuals, where single is the omitted dummy. The variable $\delta_{k,i}$ represents children which as explained in section 5 includes 6 categories from 0 to 5, where 0 children are omitted, dummy. Lastly, an interaction term is included for married women and previously married women. Of course, the interaction effect of Single women would be aggregated in the constant coefficient.

Like Model 2, to account for the potential collinearity of the variables 'Experience' and 'Education Level', two different regressions are conducted in Model 3, one including all explanatory variables and one excluding 'Experience'. Results are presented for both versions in Table 6.3.

As for Models 1 and 2, the estimated coefficient on 'Female' is negative and significant at the 1% level, but interestingly, it has substantially decreased in absolute value. The coefficient for the main independent variable in Model 3 has reduced to -0.029 implying that the average pension income is 2.9% lower for single females with no children compared to males with the same specified characteristics, holding all other control variables constant. This effect on the dummy variable female, including the significance of the interaction terms, indicates that gender matters in the case of having children and marital status when investigating pension income.

The model's R-squared is 23.5%, a minimal rise over Model 2's R-squared of 21%. This makes sense because the interactions aim to find impacts in the independent variable, female, and as a result, they don't add much to the explanatory power of the dependent variable's variance in pension income.

Looking at the coefficients of the interaction terms with 'Female', it is observable that they all hold negative signs, meaning that having children and being married have a negative effect on a female individual's pension income. Having a negative coefficient of -0.33 for the interaction term of Married Women implies that married women receive 33% less than the equivalent married men. This is supported by the theory mentioned in section 3, implying that Married women rely on their husband’s pension income because they dedicate more time to family responsibilities than labor market attachment (cf. section 3.2.6). On the other hand, it is also in line with the theory that married men earn more compared to single men. Also, this is noteworthy to mention the coefficient for the previously married women. The effect is negative but interestingly, it is much less (-0.084) in absolute value compared to the coefficient for married women. It can potentially be interpreted as women trying to put their focus back on their profession after getting a divorce (since they cannot rely on their husband’s pension anymore), they achieve more and receive higher pension income after retirement. The effect cannot be easily reversed as they have never been married, but still, a high improvement can be seen as they only receive 8.4% less than single women.

Furthermore, by adding the interaction term between ‘Children’ and ‘Female’, we are looking at the effect of being a female and having children on the individuals' pension income compared to men who have children. The coefficient indicates that females with children are on average receiving 1.1% less pension income than men with children, which is in line with the intuitions covered in previous literature mentioned in section 3.2.4.

	Model 2 (Control variables included)		Model 2 (Excluding 'Work Experience')	
	Coef	**Robust Std. Err**	**Coef**	**Robust Std. Err**
Female	-0.26035	0.001492 ***	-0.26154	0.00151 ***
Education Level				
1. No Education	Omitted		Omitted	
2. Primary Education	-0.13092	0.047477 *	-0.20621	0.027407 ***
3. Secondary Education	0.46033	0.05467 ***	0.37087	0.013814 ***
4. Upper Secondary Education	0.66967	0.054657 ***	0.53361	0.013769 ***
5. Bachelor's Degree or Equivalent	1.00310	0.054705 ***	0.76973	0.013852 ***
6. Master's Degree or Equivalent	1.17315	0.054849 ***	0.90417	0.0143 ***
7. Doctrates, PhD or Equivalent	1.20616	0.05684 ***	0.84025	0.021934 ***
Work Experience				
1. 30-40 years	Omitted			
2. 40-50 years	0.85803	0.078135 ***		
3. 50-60 years	1.04360	0.078178 ***		
4. 60-70 years	1.13412	0.094396 ***		
Involvement in Labor Market				
1. 0- 10%	Omitted		Omitted	
2.10-50%	-0.04819	0.003923 ***	-0.04596	0.003979 ***
3.50-80%	-0.15085	0.00669 ***	-0.16695	0.006782 ***
4.Higher than 80%	-0.25084	0.005082 ***	-0.27078	0.005134 ***
Industrial Field				
1.Military and non-given	Omitted		Omitted	
2.Managers	0.16377	0.006792 ***	0.13828	0.006812 ***
3.Academic professions	0.03474	0.004835 ***	0.00846	0.004822 *
4.College occupations	0.12635	0.005753 ***	0.08772	0.005732 ***
5.Office occupations	0.09466	0.006423 ***	0.05878	0.0065 ***
6.Sales&Service	0.01725	0.004116 ***	-0.00711	0.004174 ***
7.Farmers, fishermen, etc	-0.08167	0.013087 ***	-0.09052	0.013236 *
8.Craftsmen	0.06101	0.007622 ***	0.02791	0.007675 ***
9.Process, Machine, Transport	0.08656	0.005642 ***	0.06852	0.005703 ***
10.Cleaners, Auxiliary, etc	-0.03518	0.008979 ***	-0.05185	0.009139 ***
Financial Literacy	0.04017	0.005944 ***	0.03752	0.005958 ***
Constant				
Adjusted R-square	21.037 %		19.234 %	
R-square	21.041 %		19.238 %	
Sample size	403,870		403,870	
Note: $* \ p < 0.10, ** \ p < 0.05, *** \ p < 0.01$				

Table 6.3: Model 3 Table of Results

6.5 Robustness Check

The validity and reliability of empirical findings are crucial in economic research, especially when exploring issues of significant social concern such as gender gaps in pension savings. As mentioned in Section 4 and earlier in this section, to accommodate the issue of heteroskedasticity in this study, robust standard errors are reported. All the models mentioned in the previous section are run using the "Robust" option in the analysis tool introduced by SSB. This option will as a result present regression estimates with adjusted standard deviations for the estimated coefficients. Associated t-, z-, and p-values are also affected. Other values are not affected compared to standard estimation.

Moreover, a subsample analysis is performed as a robustness check to evaluate the stability and consistency of our results across various dataset segments. This further analysis checks whether the observed patterns hold up in various conditions, which strengthens the robustness of our findings. Subsampling confirms that conclusions drawn from the entire sample are not unduly influenced by particular subsets and enables us to investigate potential heterogeneity within the data. Comparing the results of the subsample models with results in the previous sector, it is notable that the patterns are mostly still recognizable. Most of the coefficients are still statistically significant at the 1% level. However, some of the categories in the categorical variables are assigned with higher p-values than 0.05, and thus it is not safe to interpret the corresponding coefficients regarding those lines.

6.5.1 Findings from Subsample Analysis

This study on the gender pension gap was conducted based on the data from the year 2017, with a specific focus on individuals aged 67-74. The results revealed significant disparities in pension earnings among genders. To deepen our insights and validate these findings, we expand our analysis by introducing a sub-sample. This section will delve into the gender pension gap among a subsequent age cohort, specifically individuals aged 75-94 in 2017. The aim here is not to replicate our earlier analysis but to reinforcing the robustness and credibility of our research outcomes. It is plausible that the pension gap dynamics in this age group may show significant differences or similarities, influenced by factors specific to their life stage. We will use a methodological approach that is similar to our initial study but customized for this specific group to explore these subtleties.

In the following sections, we will utilize the models outlined in Sections 6.1 to 6.3, applying them to our new subsample. This approach will enable us to directly compare the results with those from our main sample, providing a thorough robustness check and deeper insights into pension disparities. The regression results regarding the subsample will be presented in Table 6.4 for Model 1, table 6.5 for Model 2, and lastly in table 6.6 for Model 3.

Model 1

In our analysis of the gender pension gap, two distinct age groups were examined to ascertain the robustness of our findings. For the main sample, comprising individuals aged 67-74, the model's R-squared value was 0.0758, indicating a moderate fit by explaining around 7.58% of the variability in the natural log of pension (ln_pension). The coefficient for 'female' was -0.2846, a statistically significant result suggesting that being female is associated with a lower ln (pension income). The constant, or intercept, was 12.68, which represents the average expected log pension for males. Turning to the subsample, which focused on the older age group of 75-94, we observed an improved model fit, with an R-squared value of 0.1285. This indicates that the model explains approximately 12.85% of the variability in ln (pension income). The coefficient for 'female' in this group was -0.2891, slightly larger in magnitude than that of the main sample. This finding points to a somewhat stronger negative association between being female and having a lower ln (pension income). in this older cohort. Additionally, the constant in the subsample was 12.693, marginally higher than in the main sample.

Overall, the consistency in the direction of the gender coefficient across both age groups, alongside the increased magnitude of this effect in the older cohort, underscores the robustness of our findings in Model 1. It suggests that the gender pension gap not only persists but may also widen as individuals age, highlighting an important consideration for economic and social policies targeted at the elderly population.

Model 2

In the analysis of individuals aged 67-74, the model shows that approximately 19.24% of the variability in ln (pension income). is accounted for, with an R^2 value of 0.192. The coefficient of -0.261 indicates that being female has a significant negative impact on ln (pension income). Financial literacy has a positive impact on pension levels, with a coefficient of 0.037. Higher education generally correlates with higher ln (pension income), but the impact of other factors like field of work and employment percentage can vary. In the analysis of the older age group (75-94 years), the model shows a stronger fit with an R^2

value of 0.3395. This means that the model explains about 33.95% of the variability in ln (pension income). In this case, the impact of gender is still significant but slightly less noticeable. The coefficient for females is -0.2346, indicating a persistent but slightly smaller gender pension gap in this group. The importance of financial literacy has decreased, as shown by a small and statistically insignificant coefficient of 0.005493. Educational level and field of work have a significant impact on ln (pension income), but the extent of their influence varies.

The comparison of these models for the robustness check shows consistent findings and some noticeable variations. Both models show that there is a gender pension gap, which is slightly smaller in the older cohort. This implies that the disparity may decrease as people get older. Financial literacy is important for the main sample, but its impact decreases for older individuals. This suggests that the benefits of financial literacy on pension outcomes may diminish with age. A higher R^2 value in the subsample model indicates that variables such as education level and field of work may have more consistent effects on pension outcomes in older age groups. The different impacts of these factors on pension outcomes for different age groups show that their influence can change as people get older. This highlights the importance of considering age-specific factors in policies and planning.

Model 3

Implementing Model 3 to the new subsample focuses on the effect of variables of marital status and number of children and the interaction terms on the pension payment among genders. The comparative analysis between the main sample and the older cohort subsample provides insightful results. The main sample, which includes the general population, indicates that Model 3 explains about 21.95% of the variation in ln (pension income), as shown by an R^2 of 0.2195. Being female has a small but significant negative impact on ln (pension income). In this model. Having financial knowledge is linked to higher pension levels while having children is associated with a decrease in pension.

Model 3 implemented to the new sample shows a better fit for the older group, with an R^2 of 0.4151, explaining approximately 41.52% of the variance in ln (pension income). The gender gap is less pronounced here compared to the main sample. As observable in Table 6.6, the coefficient for the dummy variable ‘Female’ is -0,01466. One can interpret it as if we control the socio-demographic factors of ‘Marital status’ and ‘Children’, the pension income gap will be reduced to only 1.4% among men and women. The impact of financial literacy on pension levels is not significant as people get older which can be reasonable since for the older age cohort, there might not be very significant differences in financial literacy. The impact of having children on pension outcomes is greater for older individuals, indicating that family dynamics have a changing effect on pensions as people get older. Interestingly, the coefficient for the interaction term presenting ‘Married Women’ is significantly higher in the older age cohort (-0.415 compared to -0.342). This is a reasonable result since one can argue that in the older times, the traditional responsibilities were stronger for a married woman and society is slowly changing regarding its expectations of a married woman. Also, the pattern mentioned in section 6.3 regarding the previously married women is still observable in subsample analysis. Model 3 results confirm the presence of a gender pension gap across various age groups. The stronger model fit in the older cohort suggests that the variables in Model 3 are more predictive of pension outcomes in this age group.

In summary, these models collectively support the robustness of our thesis, highlighting consistent trends and surprising shifts in factors like financial literacy and family dynamics. The consistent gender pension gap across samples and the changing influence of other variables could emphasize the need for age-specific approaches in addressing pension disparities.

7. Challenges and Limitations

The previous chapter provided the analysis, and the statistical findings. As mentioned in the previous section, the empirical analysis reveals a significant gender disparity in pension income, with women receiving an average of 28.46% less than men. However, it is crucial to critically evaluate these findings. This section will provide a critical evaluation of the results of this study, mainly by discussing the limitations in data, the challenges and shortcomings faced through the investigation. Next, we will focus on conceptualizing the perfect data scenario. An ideal data set to measure the pension income gap will be explained, without the constraints of our current data. This explores how a dataset could improve our understanding of gender disparities in pension outcomes. This step is important for evaluating our findings and understanding the complexity of the gender pension gap issue. Also, this may shed light on future studies to improve our current work to achieve more precise results.

7.1 Limitation of the approaches

In this study microdata has been the main source of data which covers data for the entire Norwegian population. The data from Statistics Norway (SSB) is easily accessible and can be used to include many relevant factors in the analysis. However, there are many constraints when handling and examining register variables in SSB's Microdata, which will be explained in the following sections.

7.1.1 Considerations of Data Selection

Our study delves into the gender pension gap in Norway, an issue deeply rooted in and propelled by historical factors (Bazilchuk, N. 2018). The gender pension gap is not simply a snapshot of the present, but it is a set of complex interconnections of generational patterns of labor participation, societal norms, wage discrepancies, and changes in pension policy (OECD, 2021). To accurately dissect and understand this gap, it is essential to employ historical data, tracing the professional and financial journeys of individuals throughout their lives.

Nonetheless, we confront a substantial limitation in this respect: the microdata provided by Statistics Norway (SSB) does not offer a complete historical archive for each individual. The datasets available are rich with information but fall short of spanning entire lifetimes, thereby introducing an inherent constraint to our longitudinal analysis. This means that for our specific sample—individuals aged 67-74 in the year 2017—we are only able to observe a segment of their employment and pension accumulation history.

Further complexities arise when we consider the constraints imposed by data privacy laws, which are rigorously upheld by the SSB. While these regulations are crucial for safeguarding individual privacy, they also restrict access to certain detailed variables that could enhance the quality and depth of our analysis. For example, the lack of access to granular data on individuals' exact pension contributions and accruals over their working lives poses a challenge in constructing a full picture of the gender pension gap.

One of the main shortcomings is the variable 'employment histories' in the labor market, which is one of the main explanatory variables when interpreting the gender pension gap. The limitation of SSB is the lack of longitudinal labor market data for individuals in our sample. The data from Statistics Norway (SSB) for 2017 does not include complete employment histories of individuals. This shortfall is significant because it hinders our ability to fully understand and consider the various employment patterns that impact pension outcomes. In other words, we don't know how big the gender gap in pension would be, had they worked the same amount. Employment history details like full-time versus part-time work, career breaks, and employment contracts affect pension accumulation and the gender pension gap, especially since women often have career interruptions or work part-time due to caregiving, resulting in lower pension savings. (Bravo, J., & Herce, J., 2020)

Our approach attempts to mitigate this limitation by incorporating the concept of "Stillingprosent" — the percentage of full-time employment — into our model as a proxy for labor market engagement. We have categorized employment levels into four discrete categories based on working hours, which allows us to analyze the immediate pre-retirement employment patterns. However, this method has its drawbacks which are explained below.

Incomplete Employment Histories: By focusing solely on the year 2017, we capture only a snapshot of employment status, potentially missing the cumulative effects of earlier employment experiences on pension savings.

Assumption of Continuity: The assumption that the "Stillingprosent" for 2017 reflects an individual’s typical working pattern throughout their career may not hold. Individuals may have fluctuated between different employment levels, which would not be reflected in our data.

Unaccounted Variables: The lack of detailed historical data also means that we cannot control for other relevant variables that might have changed over time, such as periods of unemployment, changes in the sector of employment, or variations in part-time work over an individual’s career span. However, for the variable ‘Marital Status’ since we had access to historical data, we tried to mitigate this issue, by considering the previously married status for individuals instead of only their current status, so we could capture the effect of their historical status if they have been married before and are single in 2017.

Gender-Specific Employment Patterns: Women's labor market participation is particularly prone to underrepresentation in terms of historical full-time equivalent work (Bravo, J., & Herce, J., 2022), leading to potential underestimation of the gender pension gap.

Sample Representation: The "Stillingprosent" categories, while informative, may not evenly represent the labor force, with the possibility that some categories have a limited number of observations, which could bias the analysis toward the more populous categories.

Cross-Sectional Constraints: The cross-sectional nature of our data inherently limits our ability to draw conclusions about the impact of life-long labor market pathways on pension outcomes.

While the data enables a detailed examination of the gender pension gap, it lacks a source of exogenous variation to establish causality. This study, therefore, provides a descriptive analysis that highlights correlations within our sample, without claiming causative conclusions. We fully utilize the available data to shed light on the gender pension gap, cognizant of the study's descriptive nature and the limitations inherent in our dataset.

7.1.2 Consideration on Variable Construction

In our analysis, relying on the method presented by Mincer (1974), 'potential experience' has been utilized as a proxy for actual experience, implicitly assuming continuous engagement in the labor market. Such an approach, by design, does not account for labor market interruptions like parental leave. Since women, particularly mothers, are more likely to take prolonged breaks from work for caregiving, 'potential experience' might not accurately reflect the true labor market attachment for this group (Carmichael F., Ercolani M. 2016). To mitigate this, we introduced the control variable ‘children’. Our findings show that women with children experience a -1.2% significant effect of pension income penalty compared to men, which is consistent with labor market expectations mentioned in section 3.2.1. This indicates that motherhood, as predicted, adversely affects women’s pension income. The ‘children’ variable does not mitigate the complex reality of women's work interruptions, such as differing parental leave durations. It likely reflects a combination of factors, including societal norms, childcare obligations, and the structure of Norway's welfare system, which cumulatively impact mothers' careers and pensions.

Said, R. (2020) also highlights that childcare-related employment interruptions have a notable impact on women's pension income. On the other hand, variables like childcare accessibility and work flexibility are important variables to consider when analyzing pension outcomes to avoid the potential omitted variable bias. These factors may offset the negative impact of motherhood on pensions, resulting in more neutral outcomes for women with children compared to men.

Despite its limitations, 'potential experience' is a flawed yet useful indicator. We can make inferences about labor market engagement, but we need to consider that our analysis doesn't fully account for the career interruptions related to motherhood that women experience. Although we can estimate the impact of work history on pension outcomes, our findings may not fully reflect the actual gender disparities.

7.1.3 Omitted Variable Bias

This study conducted a multiple regression analysis to examine how gender affects pension earnings. Our model includes control variables from section 5.3.2 to thoroughly investigate this dynamic. However, it is important to note that the concern of omitted variable bias cannot be completely disregarded. It is difficult to account for every possible factor that could affect pension outcomes and gender disparities. The literature often discusses how to measure work effort and talent in the labor market (Becker, 1993). Human capital factors are difficult to measure because they are elusive (Abraham, K. G. 2022). Unmeasured attributes could influence the results.

Additionally, factors like employment trends in different industries, the quality of job opportunities over time, and access to professional networks may also have a significant impact on pension income Manyika, J. (2017). However, these variables are not directly included in our model due to lack of availability and difficulty of measurement. Some of these factors could be only measured perfectly in a controlled social experiment.

7.2 Conceptualizing the Perfect Data Scenario

The analysis of gender differences in pension payouts is limited due to the available data and the way variables were constructed. These issues are directly impacting our findings. In this section, we explore the connection between gender and pension disparities, considering the limitations of our data and variable construction. We examine the impact of gender, specifically the female experience, on disparities in pension income. Our goal is to explain a dataset that includes all gender-specific variables without any current limitations. This is important for getting accurate results in linear regression analysis, especially in dealing with omitted variable bias and following the Best Linear Unbiased Estimators (BLUE) standards. We will examine the factors that contribute to gender disparities in pension outcomes to better understand the complexity of these issues. This analysis explores gender disparities in pension income and sets the stage for future research in this area.

The main challenge in this study was the process of collecting high-quality data. We recognize that Statistics Norway (SSB) datasets have limitations, particularly in reflecting historical employment trends and their impact on pension outcomes, with a significant focus on gender disparities. The ideal dataset should go beyond current limitations, enabling clearer insights into the gender pension gap. It would ideally be a comprehensive, long-term study with these key features:

Comprehensive Employment and Pension Histories: Detailed, lifetime data on employment and pension accumulations, particularly focusing on gender-specific pathways, to facilitate a robust analysis of how these factors differentially impact pension outcomes for men and women.

Consistency with Gauss-Markov Assumptions: A dataset structured to ensure adherence to the Gauss-Markov assumptions, thereby enabling our regression models to provide unbiased, efficient estimates that truly reflect the underlying dynamics of the gender pension gap.

Mitigation of Omitted Variable Bias: A careful selection of variables and data that effectively captures the complexity of gender dynamics in employment and pension systems, thereby reducing the risk of omitted variable bias which is crucial for obtaining BLUE estimates in our linear regression analysis.

Longitudinal Perspective and Exogenous Variation: To accurately measure the effect of gender on pension income, we need to find a source of variation that is not influenced by other factors affecting pension outcomes. We want to create a dataset that shows current and past pension and employment trends. Enriched data with exogenous variation would help us understand the actual impact of gender on the pension gap and improve the quality of gender-focused economic research.

7.3 End of Discussion

In summary, the previous chapters have examined the gender pension gap in Norway, specifically highlighting a significant difference of 28.46% in pension income between men and women, a difference of 71,686 NOK annually. We used microdata from the entire Norwegian population, provided by Statistics Norway (SSB). Although the data was detailed, it had limitations in terms of handling and examination. The constraints of incomplete historical data for individuals and data privacy laws posed challenges in creating a thorough longitudinal analysis.

The main issue we faced was the limited historical data availability. This limitation affected our understanding of employment patterns that greatly affect pension incomes, particularly for women with caregiving responsibilities who may have career interruptions or work part-time according to historical literature (Carmichael F., Ercolani M. 2016)

To address these limitations, we used "Stillingprosent" as a measure of labor market engagement in the model. However, this approach had drawbacks like incomplete employment histories and assumptions about employment continuity, which may not accurately represent a person's career path. Furthermore, our analysis lacked other important factors such as periods of unemployment or part-time work due to the limited historical data available.

Our study aimed to understand the gender pension gap using the available data despite all the limitations. We used 'potential experience' as a measure for labor market engagement, but it didn't consider labor market interruptions like parental leave, which are more common among women (IU International University of Applied Sciences, 2023). We introduced the control variable 'children' to show that motherhood significantly affects women's pension income, demonstrating the 'motherhood penalty'.

We acknowledge the possibility of omitted variable bias in this study and this awareness prompts a careful consideration of the potential impact on the validity and precision of the results. It means we need to consider factors like childcare accessibility and work flexibility in pension outcome analyses. This highlights the complexity of the gender pension gap and the need for a more comprehensive dataset.

In the ideal data scenario, we explained a dataset that includes employment and pension histories, follows Gauss-Markov assumptions, reduces omitted variable bias, and offers a longitudinal perspective. This dataset would provide a better understanding of the gender pension gap over time. In summary, having a broad dataset, closer to the explained ideal scenario, is important for better understanding the gender pension gap. Understanding this issue is crucial for supporting future research and informing policy decisions and efforts aimed at addressing gender disparities in pension benefits.

7.4 Suggestions for Further Research

By employing linear regression methods, the investigation of the gender pension gap in Norway has revealed substantial socio-economic disparities, indicating a substantial pension gap of 28.46%. The primary focus of the analysis was centered on uncovering the factors that contribute to differences in retirement payments between men and women in Norway, specifically identifying the key elements that account for the pension gap. However, there is a wide range of areas that might be further examined in a more detailed and segmented approach, which should be considered for future research.

Using the Oaxaca-Blinder (OB) decomposition in conjunction with recentered influence function (RIF) regressions, a methodological combination that was highlighted by Fortin, Lemieux, and Firpo (2011), could be considered as a substantial extension of our work. This approach would make it possible to conduct a detailed analysis of the disparity in pension benefits across different income quantiles, which could lead to the discovery of intricate patterns and trends within particular income brackets. An investigation of this kind could shed light on whether the pension difference is more evident at different income levels and investigate the structural prejudices that lie beneath the surface, such as the glass-ceiling effect, which may have a disproportionate impact on women (Wickwire, K. S., & Kruper, J. C.,1996).

Furthermore, it would be intriguing to explore pension wealth as the dependent variable in addition to pension income as the dependent variable. If similar patterns emerge when comparing pension wealth and pension income, this consistency across different pension metrics could offer a more robust basis for addressing the pension gap, potentially leading to more effective policy interventions.

8. Conclusion

The primary aim of this thesis was to analyze gender differences in pension income in Norway, with a focus on socio-demographic characteristics. To address this objective, The data was retrieved from microdata under the provision of Statistics Norway (SSB). The linear regression analysis was applied as the core methodology and the analyses were done by the internal analysis tool provided by microdata.no due to the confidentiality of the data. This approach was augmented by conducting several variations of the methodology to ensure robustness across different models. Figure 8.1 presents the overall process of linear regression conducted in this study.

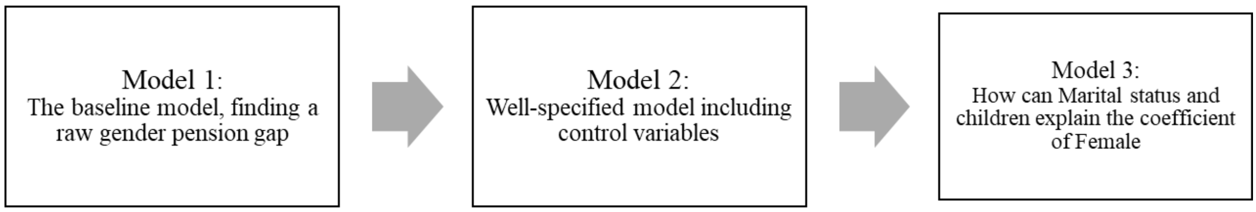

Figure 8.1: The process of multiple linear regression

First, an unadjusted simple regression was implemented implying a raw gender gap of 28% among men and women, meaning that women on average receive 28% less than men in the selected sample of the Norwegian retired population. In the next step, the multiple linear regression analysis identified a significant disparity in pension income along gender lines, confirming the predictions when accounting for sociodemographic factors, the average pension income for males is consistently higher than that of females (Hypothesies 1). This gap persisted even after factoring in sociodemographic variables. In a more detailed analysis incorporating interaction terms, it became apparent that married women are indeed at a disadvantage in terms of pension income when compared to married men, which is consistent with economic theories regarding unpaid work and dependency on a spouse's pension. After controlling for socio-demographic factors, especially 'Marital status', 'Children', and the interaction terms with the variable 'Female', this gap is reduced to 2.7% less pension income among women in the original age cohort and this pattern was consistent in the sub-sample, showing 1.4% less pension income for women compared to men.

Contrary to the expectations outlined in Hypothesis 2, the analysis further revealed that women with children tend to have lower pension incomes than men, thereby suggesting that the presence of children contributes negatively to the pension income of females, which is a finding in line with the overarching discussions in the thesis. This aspect inverts the anticipated positive relationship between the number of children and pension income for females suggested in Hypothesis 2, indicating that the impact of having children on pension income for women is indeed in the opposite direction than initially expected.

Despite following the methodology rigorously, a number of limitations were faced due to the data availability. The datasets from Statistics Norway were not able to fully capture the complex dynamics of gender-specific employment patterns and their impact on pension outcomes. Our use of 'potential experience' and the control variable 'children' did not fully capture the complex realities of women's labor market engagement and its impact on pension accumulation.

As highlighted in section 7, the thesis emphasizes the importance of improving data collection and analysis methods to better understand the gender pension gap. A dataset that includes detailed labor market data and control variables, following strict econometric principles, would allow for a more accurate and comprehensive analysis. Implementing this method is key not only to exploring the intricate relationship between gender, labor market behavior, and societal norms but also to achieving our primary objective: obtaining the Best Linear Unbiased Estimator (BLUE) in line with the Gauss-Markov assumptions. This approach is crucial for understanding how gender, labor market behavior, and societal norms come together to influence pension outcomes.

In summary, this research sheds light on the gender disparities in pension income in Norway and paves the way for future investigations. It emphasizes the need for better data and methods to understand gender disparities more effectively. The insights from this thesis are valuable for both academics and policymakers working to reduce the gender pension gap. It is important to continue researching this area to gain insights that can help achieve a more precise distribution of pension income between genders.

References

- Abraham, K. G., & Mallatt, J. (2022, August 1). Measuring Human Capital. *Journal of Economic Perspectives*, *36*(3), 103–130. https://doi.org/10.1257/jep.36.3.103 Accessed 30.10.2023

- Akerlof, G., & Kranton, R. (2000). Economics and Identity. *Quarterly Journal of Economics*, 715- 753.

- Angrist, J. D., & Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press. https://doi.org/10.2307/j.ctvcm4j72

- Bazilchuk, N., & Amundsen, B. O. A. A. B. B. (2018, October 31). *Gender gap in retirement: Norwegian men retire earlier than women*. Science Norway. https://www.sciencenorway.no/forskningno-norway-retirement/gender-gap-in-retirement-norwegian-men-retire-earlier-than-women/1459980 Accessed 08.11.2023

- Becker, G. S. (1985, January). Human Capital, Effort, and the Sexual Division of Labor. *Journal of Labor Economics*, *3*(1, Part 2), S33–S58. https://doi.org/10.1086/298075

- Belingheri, P., Chiarello, F., Fronzetti Colladon, A., & Rovelli, P. (2021). *Twenty years of gender equality research: A scoping review based on a new semantic indicator*. PloS one, 16(9), e0256474. https://doi.org/10.1371/journal.pone.0256474

- Bettio, F., Tinios, P., & Betti, G. (2013). The Gender Gap in Pensions in the EU. Luxembourg: European Union. https://www.social-protection.org/gimi/Media.action;jsessionid=wcBD4JtLysoSFq1EEQ92NOATLS62Ak_xqzHUnbeaD4p5VZR7wIpw!1945465934?id=13384 Accessed 15.10.2023

- Blau, F. D., & Kahn, L. M. (2017, September 1). The Gender Wage Gap: Extent, Trends, and Explanations. *Journal of Economic Literature*, *55*(3), 789–865. https://doi.org/10.1257/jel.20160995

- Bravo, J. M., & Herce, J. A. (2020, July 22). Career breaks, broken pensions? Long-run effects of early and late-career unemployment spells on pension entitlements. *Journal of Pension Economics and Finance*, *21*(2), 191–217. https://doi.org/10.1017/s1474747220000189

- Bryman, A., & Bell, E. (2011). *Business Research Methods*. New York: Oxford University Press.

- Burkevica, I., Humbert, A., Oetke, N., & Paats, M. (2015). Gender gap in pensions in the EU: Research note to the Latvian Presidency. Vilnius: European Institute for Gender Equality (EIGE)

- Carmichael F., Ercolani M. (2016). Unpaid caregiving and paid work over life-courses: Different pathways, diverging outcomes. *Social Science & Medicine, 156, 1–11*.

- Cordova, K., Grabka, M. M., & Sierminska, E. (2022). Pension Wealth and the Gender *Wealth Gap. PubMed*. https://pubmed.ncbi.nlm.nih.gov/36237300/ Accessed 10.09.2023

- Correll, S., Benard, S., & Paik, I. (2007). Getting a Job: Is There a Motherhood Penalty? *American Journal of Sociology, 1297-1304.*

- Dawson, C. (2023, June 9). Gender differences in optimism, loss aversion and attitudes towards risk. *British Journal of Psychology*, *114*(4), 928–944. https://doi.org/10.1111/bjop.12668

- Dessimirova, D., & Audera Bustamante, M. (2019) The gender gap in pensions in EU.https://www.europarl.europa.eu/RegData/etudes/BRIE/2019/631033/IPOL_BRI(2019)631033_EN.pdf Accessed 12.09.2023

- Dewilde, C. (2012). Life course determinants and incomes in retirement. *The Journal of the Centre for Policy on Ageing and the British Society of Gerontology, 587-615.*

- Doe, J., & Smith, A. (2020). *Gender Differences in Investment Behavior*: The Role of Assertiveness and Advice Seeking. *Journal of Financial Psychology, 15(2)*, 154-168.

- Engvik, M., & Pettersen, M. (2021). *Pappaperm blant lærere, men langt fra en likedeling*. SSB analyser (2021/14) Published 20. september 2021

- Fasang, A., Aisenbrey, S., & Schömann, K. (2013). Women's Retirement Income in Germany and Britain. *European Sociological Review*, 968-980.

- Field, A. (2018). Discovering Statistics Using IBM SPSS Statistics (5th ed.). *SAGE Publications.*

- Flyer, F., & Rosen, S. (1997). The New Economics of Teachers and Education. *Journal of labor economics, 104-139.*

- Forbes. (2020, December 16). H*ow The Gender Income Gap Impacts Women's Retirement*.https://www.forbes.com/advisor/retirement/retirement-gender-incomegap/ Accessed 31.10.2023

- Forsvaret (2023) *Lønn og skatt* https://www.forsvaret.no/soldater-og-ansatte/ansatt/lonn Accessed 05.12.2023

- Fortin, N., Lemieux, T., & Firpo, S. (2011). Decomposition Methods in Economics. In O. Ashenfelter & D. Card (Eds.), *Handbook of Labor Economics, Vol. 4A* (pp. 1-102). Elsevier.

- Gronau, R. (1988). Sex-Related Wage Differentials and Women's Interrupted Labor Careers - The Chicken or The Egg. *Journal of Labor Economics*, 277-301.

- Halvorsen,E. & Hetland, A. (2023) *Pensjonsformue i Norge 2019-2020.* https://www.ssb.no/inntekt-og-forbruk/inntekt-og-formue/artikler/pensjonsformue-i-norge-2019-2020/_/attachment/inline/a17d5db8-fba9-4850-be61-0dd6d2833a12:549732467a7b6b8429991a333d40a2ffc7e5d611/RAPP2023-21.pdf Accessed 23.10.2023.

- Hardoy, I., og Schøne, P. (2008). Hvor mye betyr barn for lønnsforskjeller mellom kvinnerog menn? *Tidsskrift for samfunnsforskning*, 49: 3–36

- Haug, M (2023) *Forventa levealder falt i 2022* https://www.ssb.no/befolkning/fodte-og-dode/statistikk/dode/artikler/forventa-levealder-falt-i-2022 Accessed 13.10.2023.

- Isaacs, K. (2010). Women's Retirement Insecurity Across U.S. Birth Cohorts. *Dissertation, Duke University*. https://dukespace.lib.duke.edu/dspace/handle/10161/2408 Accessed 23.09.2023.

- IU Internationale Hochschule (2023, December 1). *Study on kids and careers: Traditional role models still largely exist*. Phys.org. https://phys.org/news/2023-12-kids-career-traditional-role-largely.html Accessed 23.09.2023.

- Jeffrey M. Wooldridge (2019): *Introductory Econometrics:* A Modern Approach, 7th edition

- Jianakoplos, N., & Bernasek, A. (1998). Are Women More Risk Averse? *Economic Inquiry, 620-630.*

- Kleven, H., Landais, C., & Søgaard, J. (2019). Children and Gender Inequality: Evidence from Denmark. American Economic Journal: Applied Economics, 181-209.

- Kuivalainen, S., Nivalainen, S., Jarnefelt, N., & Kuitto, K. (2020). Length of working life and pension income: empirical evidence on gender and socioeconomic differences from Finland. *Cambridge University Press*, 126-146.

- Kunze, A. (2015). The Family Gap in Career Progression. *Research in Labor Economics*, 115-142.

- Lane, S. (2018). *The scary facts behind the gender pension gap:*https://www.weforum.org/agenda/2018/03/retired-women-less-money-pensions-than-men Accessed 25.11.2023.

- Liu, S., & Marois, G. (2023, March 23). The effect of motherhood on the labor force participation of married women in China. *Asian Population Studies*, 1–17. https://doi.org/10.1080/17441730.2023.2193518

- Manyika, J. (2017). Technology, jobs, and the future of work. McKinsey & Company. https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/technology-jobs-and-the-future-of-work Accessed 05.12.2023.

- Microdata (2023). *How to define and create populations*: https://www.microdata.no/en/eksempel/how-to-define-and-create-populations/ Accessed 08.11.2023.

- Microdata (2023). *Om microdata.no*: https://www.microdata.no/om-microdata-no/ Accessed 09.11.2023.

- Mincer, J. A. (1974). The Human Capital Earnings Function. In J. A. Mincer, Schooling, Experience, and Earnings (pp. 83-96). NBER.

- Möhring, K. (2014). Employment Histories and Pension Incomes in Europe. *European Societies*, 3- 26.

- Norsk Pensjon (2023) *Offentlig tjenestepensjon* https://norskpensjon.no/om-pensjonssystemet/tjenestepensjon/tjenestepensjon/ Accessed 20.11.2023.

- Norsk Pensjon (2023) *Om pensjonssystemet* https://norskpensjon.no/om-pensjonssystemet/individuelle-pensjonsavtaler/ Accessed 20.11.2023.

- Norsk Pensjon (2023) *Individuelle pensjonsavtaler*: https://norskpensjon.no/om-pensjonssystemet/individuelle-pensjonsavtaler/ Accessed 20.11.2023.

- NTB (2021)- *Pensjonsgapet mellom menn og kvinner i Norge er betydelig* *https://www.dnb.no/dnbnyheter/no/bors-og-marked/betydelig-pensjonsgap*

- OECD. (2015). New OECD data and analysis revealing the wide gap in pension benefits between men and women. Data - OECD Accessed 10.11.2023.

- OECD (2018), *Is the Last Mile the Longest? Economic Gains from Gender Equality in Nordic Countries*, Gender Equality at Work, OECD Publishing, Paris, https://doi.org/10.1787/9789264300040-en.

- OECD (2021), Understanding the gender pension gap beyond labor market drivers through a literature review", in Towards Improved Retirement Savings Outcomes for Women, OECD Publishing, Paris, https://doi.org/10.1787/2c610b86-en.

- OECD (2023), *Social spending* https://data.oecd.org/socialexp/social-spending.htm#indicator-chart Accessed 02.10.2023.

- OECD (2023), *Pensions at a Glance 2023: OECD and G20 Indicators*, OECD Publishing, Paris, https://doi.org/10.1787/678055dd-en.

- Pensjonsutvalget (2022). *Et forbedret pensjonssystem*: Utredning fra et utvalg oppnevnt ved kongelig resolusjon 12. juni 2020. Avgitt til Arbeids- og inkluderingsdepartementet https://www.regjeringen.no/contentassets/de89ff1348ab4754a64d08b9ce77a584/no/pdfs/ Accessed 02.10.2023.

- Peris-Ortiz, M., Álvarez-García, J., Domínguez-Fabián, I., & Devolder, P. (2020, March 20). Economic Challenges of Pension Systems. Springer Nature

- PPHR (2017, December 2). *Health Care Reform:* Learning From Other Major Health Care Systems – Princeton Public Health Review. https://pphr.princeton.edu/2017/12/02/unhealthy-health-care-a-cursory-overview-of-major-health-care-systems Accessed 30.11.2023.

- Rabe-Hesketh, S., & Skrondal, A. (2012). Multilevel and Longitudinal Modeling Using Stata (Volume I: Continuous Responses). *Stata Press*.

- Said, R. (2020). *The motherhood penalty on retirement income.* University of Malta. https://www.um.edu.mt/library/oar/handle/123456789/66686 Accessed 20.11.2023.

- Sánchez Serrano, A., & Peltonen, T. (2020). Pension Schemes in the European Union: Challenges and Implications from Macroeconomic and Financial Stability Perspectives. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3723347

- Sefton, T., Evandrou, M., Falkingham, J., & Vlachantoni, A. (2011). The relationship between women's work histories and incomes in later life in the UK, US and West Germany. *Journal of European Social Policy*, 20-36.

- Schøne, P (2015) *Søkelys på arbeidslivet 2015*, vol. 32 (4), 360-376 https://www.spk.no/en/about-us/the-norwegian-pension-system/ Accessed 20.10.2023.

- Skatteetaten (2023) *OTP - Mandatory occupational pension scheme - The Norwegian Tax Administration*. (n.d.). https://www.skatteetaten.no/en/business-and-organisation/employer/otp---mandatory-occupational-pension-scheme/ Accessed 20.09.2023.

- Smogli, P. O., & Halvorsen, E. (2019) *Pensjonsformue i Norge 2017* https://www.ssb.no/inntekt-og-forbruk/artikler-og-publikasjoner/_attachment/399044?_ts=16d81612778 Accessed 23.11.2023

- Soville, V (2022) *The Simple Linear Regression Model:* Retrieved from Canvas ECN4002: Spring 2022.

- SSB (2011) *Standard for yrkesklassifisering* https://www.ssb.no/klass/klassifikasjoner/7 Accessed 10.09.2023.

- SSB (2022) *Indikatorer for kjønnslikestilling i kommunene* https://www.ssb.no/befolkning/likestilling/statistikk/indikatorer-for-kjonnslikestilling-i-kommunene Accessed 06.10.2023.

- SSB (2022) *Pensjonsformuen i Norge 2019-2020:*https://www.ssb.no/inntekt-og-forbruk/inntekt-og-formue/artikler/pensjonsformue-i-norge-2019-

2020/_/attachment/inline/a17d5db8-fba9-4850-be61-0dd6d2833a12:549732467a7b6b8429991a333d40a2ffc7e5d611/RAPP2023-21.pdf Accessed 06.10.2023.

- SSB (2023) *Standard for utdanningsgruppering NUS*: https://www.ssb.no/klass/klassifikasjoner/36 Accessed 06.10.2023.

- Statens Pensjonskasse (2023) *The Norwegian pension system* https://www.spk.no/en/about-us/the-norwegian-pension-system/ Accessed 23.09.2023.

- Stokes, S. L. (2022, October 2). Sampling: Design and Analysis, 3rd ed. *Journal of the American Statistical Association*, 117(540), 2287–2288.

- Van Rooij, M., Lusardi, A., & Alessie, R. (2012). Financial literacy, retirement planning and household wealth. *The Economic Journal, 449-478.*

- Veremchuk, A. (2020). Gender Gap in Pension Income: Cross-Country Analysis and Role of Gender Attitudes. University of Tartu - *Faculty of Economics & Business Administration Working Paper Series*, Issue 126, p3-64.

- Wickwire, K. S., & Kruper, J. C. (1996). The glass ceiling effect: An approach to assessment. *Consulting Psychology Journal:* Practice and Research, 48(1), 32–39. https://doi.org/10.1037/1061-4087.48.1.32.

- World Economic Forum (2021) *Global Gender Gap Report 2021*: Insight Report https://www3.weforum.org/docs/WEF_GGGR_2021.pdf Accessed 07.12.2023.

- Wooldridge, J. M. (2019). Introductory Econometrics: A Modern Approach (7th ed.). Cengage Learning.

- Zumbach, G. O. (2009). Characterizing Heteroskedasticity. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1420162

Appendix

	Model 1 (Baseline for gender differences)	
	Coef	Robust Std. Err
Female	-0.28915	0.001234 ***
Constant	12.69315	0.000916 ***
Adjusted R-square	12.856 %	
R-square	12.856 %	
Sample size	365,898	
Note: $*\ p<0.10, **\ p<0.05, ***\ p<0.01$		

Table 6.4: Simple Regression on Sub-Sample

	Model 2	
	Coef	Robust Std. Err
Female	-0.23462	0.001094 ***
Education Level		
1. No Education	Omitted	
2. Primary Education	-0.38442	0.035661 ***
3. Secondary Education	0.44643	0.015193 ***
4. Upper Secondary Education	0.59490	0.015195 ***
5. Bachelor's Degree or Equivalent	0.88506	0.015278 ***
6. Master's Degree or Equivalent	1.07075	0.015491 ***
7. Doctrates, PhD or Equivalent	1.14684	0.019738 ***
Involvement in Labor Market		
1. 0- 10%	Omitted	
2.10-50%	0.04782	0.005463 ***
3.50-80%	0.07812	0.012004 ***
4.Higher than 80%	0.05476	0.008082 ***
Industrial Field		
1.Military and non-given	Omitted	
2.Managers	0.16065	0.01524 ***
3.Academic professions	0.07779	0.014533 ***
4.College occupations	0.09116	0.015066 ***
5.Office occupations	0.06747	0.016928 ***
6.Sales&Service	0.03865	0.008364 ***
7.Farmers, fishermen, etc	-0.12599	0.021306 ***
8.Craftsmen	-0.01782	0.01853
9.Process, Machine, Transport	0.04112	0.011534 ***
10.Cleaners, Auxiliary, etc	-0.03252	0.019296 *
Financial Literacy	0.00549	0.00720
Constant		
Adjusted R-square	33.949 %	
R-square	33.953 %	
Sample size	365,898	
Note: $*\ p<0.10, **\ p<0.05, ***\ p<0.01$		

Table 6.5: Model 2 results on Sub-sample

	Model 3	
	Coef	**Robust Std. Err**
Female	-0.01467	0.004396 ***
Education Level		
1. No Education	Omitted	
2. Primary Education	-0.37253	0.036171 ***
3. Secondary Education	0.43477	0.015248 ***
4. Upper Secondary Education	0.59445	0.015252 ***
5. Bachelor's Degree or Equivalent	0.89590	0.015326 ***
6. Master's Degree or Equivalent	1.07072	0.015549 ***
7. Doctrates, PhD or Equivalent	1.14993	0.019775 ***
Involvement in Labor Market		
1. 0- 10%	Omitted	
2.10-50%	0.05141	0.005249 ***
3.50-80%	0.08100	0.011799 ***
4.Higher than 80%	0.04732	0.008011 ***
Industrial Field		
1.Military and non-given	Omitted	
2.Managers	0.16349	0.014981 ***
3.Academic professions	0.09931	0.014482 ***
4.College occupations	0.10027	0.014782 ***
5.Office occupations	0.09990	0.016012 ***
6.Sales&Service	0.06329	0.008141 ***
7.Farmers, fishermen, etc	-0.08765	0.019502 ***
8.Craftsmen	-0.01090	0.01819
9.Process, Machine, Transport	0.04437	0.011509 ***
10.Cleaners, Auxiliary, etc	-0.02333	0.01863
Financial Literacy	-0.00947	0.00708
Children		
1. zero	Omitted	
2. one child	-0.03419	0.002353 ***
3. two children	-0.06409	0.004495 ***
4. three children	-0.10717	0.00996 ***
5. four children	-0.16021	0.024012 ***
6. five or more children	-0.33747	0.063515 ***
Interaction: Children*Female	-0.00008	0.00314
Marital Status		
1.Single	Omitted	
2. Previously Married	0.17023	0.003341 ***
3.Married	0.18083	0.003529 ***
Interaction: Single*Female	Omitted	
Interaction: Pre.Married*Female	-0.13413	0.004742 ***
Interaction: Married*Female	-0.41539	0.004642 ***
Constant		
Adjusted R-square	41.511 %	
R-square	41.516 %	
Sample size	365,898	
Note: $*$ p <0.10, $**$ p <0.05, $***$ p <0.01		

Table 6.6: Model 3 results on Sub-Sample