

Modalsli, Jørgen

**Working Paper**

## Female labor force participation in historical census microdata

EHES Working Paper, No. 282

**Provided in Cooperation with:**

European Historical Economics Society (EHES)

*Suggested Citation:* Modalsli, Jørgen (2025) : Female labor force participation in historical census microdata, EHES Working Paper, No. 282, European Historical Economics Society (EHES), s.l.

This Version is available at:

<https://hdl.handle.net/10419/331498>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



European  
Historical  
Economics  
Society

**EHES Working Paper | No. 282 | September 2025**

# **Female labor force participation in historical census microdata**

**Jørgen Modalsli<sup>1</sup>**

Oslo Business School at Oslo Metropolitan University

<https://www.ehes.org/>

# Female labor force participation in historical census microdata

Jørgen Modalsli<sup>1</sup>

Oslo Business School at Oslo Metropolitan University

## Abstract

How reliable is historical microdata? Understanding historical labor force participation is crucial for assessing long-term trends in economic development and intergenerational mobility. Most existing historical studies are, however, limited to men, and little is known about how reliable quantitative historical sources are when studying labor market outcomes for women. This paper documents that the measurement of women's economic activity in the 1910 Norwegian population census had a high level of consistency. There is extensive discussion of measurement issues in historical census reports, micro data can reproduce historical census tables with a high degree of accuracy, and other contemporary reports such as industrial censuses and tax statistics confirm the results found in the census. In addition, a double-enumeration feature of the Norwegian census is leveraged to assess consistency between enumerators, finding no indication that precision in the occupational classification of women is any lower than for men. Some potential sources of downward bias are found in the historical census microdata set provided by IPUMS. Based on the results in this paper, historical census data appears well suited to study economic activity using modern econometric methods, for women as well as men. A slight upward revision of the 39% female labor force participation in Norway in 1910 might be in order.

**JEL Codes:** N33, J21, J16

**Keywords:** Labor force participation, historical census data, gender, occupational measurement, Norway

<sup>1</sup> Corresponding Author: Jørgen Modalsli ([jorgenmo@oslomet.no](mailto:jorgenmo@oslomet.no))

I am grateful to Elien Dalman, Vincent Delabastita, Suvi Heikkuri, Jakob Molinder, Benjamin Schneider, Paul Sharp, Gunnar Thorvaldsen and participants at workshops and conferences for helpful comments and suggestions, and to Lars Nygaard for access to National Archives data.

## Notice

The material presented in the EHES Working Paper Series is property of the author(s) and should be quoted as such. The views expressed in this Paper are those of the author(s) and do not necessarily represent the views of the EHES or its members.

# 1 Introduction

Administrative data on individuals' economic and demographic characteristics is a valuable source of information for researchers in economics and other social sciences. The existence of large historical data sets makes it possible to use these types of data also for earlier historical periods. But is it straightforward to impose modern definitions on data collected more than a hundred years ago? This is a challenging question to answer. An important dimension of social change is the evolution of the role of women in the labor market. How should we understand census enumerators' decision to report women as working, and how has this been interpreted in data sets compiled a century after the fact?

Research on long-run trends in labor economics, intergenerational mobility, and inequality typically relies on census data sets, made available for research through large-scale public or private investments. The most prominent source of data is IPUMS, which again builds on data sets made available through public projects and by cooperation with organizations that provide services to genealogists. For many countries, microdata sets including demographics and variables denoting economic activity are available from the mid-nineteenth century onward. Historical studies of occupational patterns and social inequality have traditionally focused on outcomes for men, but recently improved linking methods and a richer set of data sources have made it possible to examine the entire adult population in the economy.

In this paper, a set of approaches to evaluate the enumeration of female labor force participation (LFP) in historical micro data censuses will be presented. They will be applied to the Norwegian census of 1910, a high-quality census from a country which had started to industrialize but where small-scale agriculture was still a substantial part of the economy. A thorough understanding of historical reporting standards is crucial when using historical data to draw inferences about long-term changes. For example, studies of the role of family background in shaping labor market outcomes usually compare characteristics of parents and children at separate times. Although recent studies have emphasized the role of data linkage (Bailey *et al.*, 2020; Abramitzky *et al.*, 2021), less work has been done to evaluate how to accurately understand the economic variables reported at the individual level, such as participation in the labor force and occupational classification. This is particularly important when studying the role of mothers, as views of women's labor force participation have changed substantially over time. Indeed, information on female participation in the labor force is sometimes claimed to be more problematic to use than that of men, as enumerators would be prejudiced against recording women as working (Humphries & Sarasúa, 2012; Maas & Van Leeuwen, 2023).

This paper documents that the information on female labor force participation in Norway in 1910 is well suited for modern analysis using descriptive statistics or econometric techniques. This claim is based on four observations. First, it is evident from the results published in the years after

the census that the question of female labor force participation was seen as important by the people compiling the statistics. Contemporary analysis uses elaborate classification schemes and discusses the challenges with consistent measurement. Women, both married and unmarried, are shown to inhabit a wide range of occupations and social statuses.

Second, a comparison of the original census tabulations with tabulations based on micro data disseminated by IPUMS and the Norwegian National Archives shows that it is largely possible to reproduce the original census tables. This micro data also makes it possible to revisit some of the classification choices made when coding historical occupations. Despite some cases of potential downward bias in the classification of labor force status found in the IPUMS data set (to be detailed below), the overall classification of labor force participation appears to be sound.

Third, other evidence from the early twentieth century, where available, corroborates the rates of female labor force participation found in the census. Industrial statistics find gender ratios among industrial workers that are comparable to those found in the Census, and tax records do not indicate any systematic under-enumeration of women compared to men. Fourth, an examination of individuals enumerated twice in the census because they were traveling away from their location of residence shows that mismeasurement is not systematically different between men and women.

The main source of potential bias is found when comparing the text strings from the original enumeration to occupation codes in the IPUMS data. Here, some of the choices that have been made lead to an under-enumeration of female labor force participation. Retired men and women are treated differentially, housework is systematically coded as non-participation and farmers and fishermen's wives are not considered to be participating in the labor market. However, such comparisons are challenging, and there is likely to have been substantial heterogeneity in the underlying economic activity that is not captured by the short enumeration strings. Although the focus of this paper is on the reliability of census information in individual-level data, the methodology also provides information on the aggregate participation rate. The reported labor force participation rate of 39 per cent for women 15 years and older appears reliable, though, as we will see, some revision of the interpretation of occupation strings can give rise to higher reported participation rates.

Ideally, we would like to come as close to the common use of labor force definition as possible (see e.g. ILO, 2013): we want the share of the population who work for pay, profit, compensation in kind or in a family business or farm. In practice, we will be limited by the available source material, as will be detailed below. To illustrate the methodology, an illustration of the data generation process that is being analyzed is given in Figure 1.

The box at the top of the figure illustrates the economic activity that takes place: whether or not a person is gainfully employed. Directly below is a box named “enumerator”: The enumerator observes the economic activity and decides how to classify it, typically by writing the desired economic status and/or activity on a paper form. The original census publications then result from the paths in the left leg: First, individual records are processed into a set of census publications

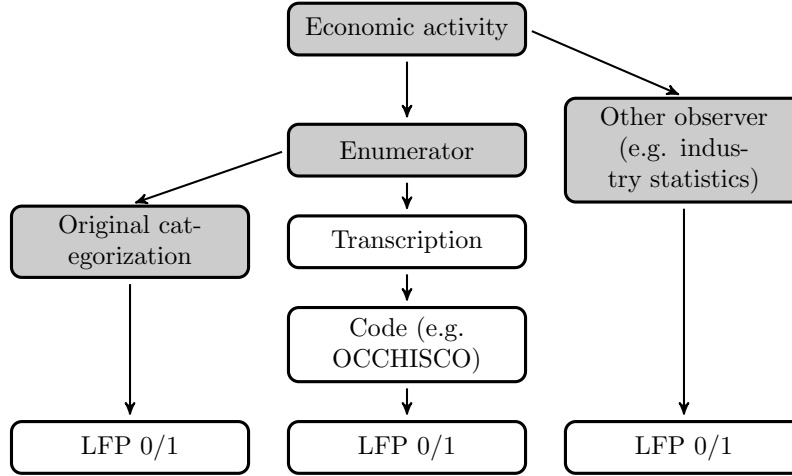


Figure 1: Model of the data generating process involved in analysis of the 1910 Census. Shaded boxes denote activity/analysis in the early 1900s; white boxes denote analysis performed in recent decades/today

(“original categorization”), which the modern researcher can then interpret with an eye to modern definitions of what constitutes labor force participation (“LFP 0/1”, left box). The first three boxes have a shaded color, to illustrate that these processes took place roughly contemporaneously with the census, in the 1910s, while the final box is white as the interpretation is performed in the present day.

The middle path in Figure 1 denotes the way research micro-data sets (e.g., IPUMS) are typically constructed. The enumerators’ text strings are transcribed, and then these strings are evaluated according to a coding system, such as HISCO. This then serves as the basis for a measurement of, for example, the labor force participation. A comparison of these two pathways is one way in which the reliability of economic variables such as labor force participation can be measured.

However, economic activity was not only observed in the census; there are also other contemporary reports. Although narrative and qualitative data can often be illuminating, the analysis here will be restricted to quantitative analyses. In 1910, censuses of manufacturing, employee insurance, and tax authorities all addressed the issue in various ways. These reports can be compared to learn more about the reliability of each of the sources.

Finally, there is a way to assess the reporting quality within the census itself. External evaluation of the enumerators’ evaluation can be provided in the cases where individuals are enumerated twice: both at home (the municipality where they were legally resident) and away (the municipality where they were present on census day).

The paper proceeds as follows. In Section 2, the Norwegian Census of 1910 is introduced, with a

review of data collection, the historical context, and census publications. Sections 3–5 present the methods outlined in the three paragraphs above. A comparison of the approaches and a synthesis of the results is provided in Section 6. First, however, we review the literature on female labor force participation in the early twentieth century, with an emphasis on studies using individual-level data.

An early study using historical microdata on women to understand labor market mechanisms was performed by Fernandez *et al.* (2004), examining the role of World War II mobilization in women’s participation in the labor force over two generations. Using U.S. census data from 1940–1980, where the number of weeks working in the previous year is reported, they find that women in areas where more men were mobilized were more likely to work and that this increase also spilled over to the next generation. There are few historical censuses before 1940 that ask directly about labor force participation. Rather, this is inferred from occupation codes, which are derived from the occupation strings written down by census enumerators. For studies of the economic impact of fathers’ labor market characteristics, there has been limited interest in assessing labor force participation, as most men are typically assumed to have been working.

To the knowledge of this author, there are only two studies examining mother-child intergenerational mobility directly in samples going more than 100 years back in time. The first, Delabastita & Buyst (2021), is based on marriage registers in West Flanders between 1830 and 1900. In these registers, both husbands and wives recorded their occupations on marriage certificates. More than eighty per cent of wives reported occupation information, mostly in textiles. Delabastita & Buyst find that mother-daughter mobility was higher than father-son mobility in this period. The second, Dalman (2026), has data from Swedish censuses between 1880 and 1910 (as well as in the second half of the twentieth century) and does not find clear gender differences in mobility. For neither study is it possible to externally validate the precision of their measurement of mothers’ occupational positions.<sup>1</sup>

The present paper also contributes to the literature that attempts to revise and update estimates of the labor force participation of women. Goldin (1990), as part of a comprehensive project to describe the economic history of American women, re-estimated female labor force participation in the United States in 1890 using three principal corrections to the census. First, boardinghouse keepers were included (where income from boarders was high enough to merit a definition of the woman’s role as “gainfully employed”), second, the number of agricultural workers was adjusted by comparing 1890 to a more inclusive worker definition used in the 1910 Census, and third, the number of women in manufacturing was adjusted by comparing population census tabulations to tabulations from the census of manufacturing. In addition, an attempt was made to account for women’s market-oriented household production. This results in an increase of the estimated LFP of married women from 2.5 per cent to 12.5 per cent in 1890, predominantly driven by the inclusion

---

<sup>1</sup>There are also alternative approaches, such as using data on spouses to proxy for the women themselves. For example, Olivetti & Paserman (2015) study a woman’s social mobility by comparing her father’s characteristics to that of her husband.

of boardinghouse keepers and unpaid workers on family farms.

These adjustments, described by Goldin as a lower bound, are a key part of her main finding on female labor force participation in the United States: that the trend over time is shaped as a U, with a high level in the nineteenth century and a low point in the early 1900s, before an increase throughout the twentieth century. This U-shape is confirmed to be also found in other countries by Olivetti (2014), using a combination of historical and modern data for a large set of western countries.

There are several studies of female labor force participation based on smaller samples that all point in the direction of some under-measurement in official sources. Examples are Horrell & Humphries (1995) and Burnette (2004) for the United Kingdom and Goldin (1979, 1980) for the United States. Arguments against an undercount of historical labor force participation of women have also been made. Stanfors (2014) argues that historical female labor force participation in Sweden is often overstated, leading to the misunderstanding that Sweden historically was more “gender equal” than it actually was.<sup>2</sup>

Four recent studies use large data sets to improve the measurement of historical female LFP at the aggregate level. Van Nederveen Meerkerk & Paping (2014) compare enumeration practices across districts around Groningen in the Netherlands in the latter part of the nineteenth century and find that the measurement of female labor force participation is underreported in official sources. You (2020b) examines English censuses between 1881 and 1911 and finds that wives’ economic status, often enumerated as the wife of the husband’s occupation (e.g. “shoemaker’s wife”), should not necessarily be counted as in the labor force. Molinder (2022) performs a comprehensive re-estimate of female labor force participation in Sweden between 1870 and 1960, going through the economy sector by sector to revise the estimates. This results in a dramatic increase in female labor force participation, with estimates of the 15+ participation rate of more than 50 per cent in 1910, up from 20–40 per cent in earlier work. Finally, Ngai *et al.* (2024) re-estimate female (and male) labor force participation in the United States. Building on earlier work (Ruggles, 2015; Chiswick & Robinson, 2021), they re-classify family members on farms and households that run a business as working. This results in an increase in measured female labor force participation before 1920 from less than 25 per cent to more than 50 per cent.<sup>3</sup>

---

<sup>2</sup>Female labor force participation can also be studied as an outcome in its own right. You (2020a) study regional drivers of female labor force participation in England and Wales in 1881, and find that the degree of industrialization is the most important driver of such participation.

<sup>3</sup>Ngai *et al.* (2024) also estimate labor force participation on the extensive margin (hours of work) for men and women and provide information from a wide range of sources on time use and participation in paid work in the United States before 1940.



## 2 The 1910 Census of Norway

### 2.1 Historical context

The analysis in this paper is based on the 1910 Census of Norway. It was carried out in the first weeks of December 1910, with the reference date being 1 December (Statistics Norway, 1916, p. 150f). According to the instructions, census takers should be recruited among “the men or women most suited to this work”; it is a stated assumption that this would typically be teachers. The results were processed by Statistics Norway and published in seven volumes during the period 1912-1916.<sup>4</sup> The most relevant tables with respect to female labor force participation are found in the main reports of occupations. (Statistics Norway, 1913, 1914). Of interest is also a report linking the census with tax data (Statistics Norway, 1915), which we will return to below. An overview of the main results, with discussion, is given in Statistics Norway (1916).

The enumerators were instructed to record employment and social status, including that of housewives and children, and in such a way that both industry, trade, and occupational position could be discerned. For relatives (including adult children), employment was also to be stated if they engaged in work — including agricultural work and housework.<sup>5</sup> This corresponds closely to contemporary labor force definitions (ILO, 2013), even though a time definition (e.g., at least one hour a week) is not given.<sup>6</sup>

Little academic debate has been conducted on the measurement of female labor force participation in the Norwegian census, although the precision of its measurement has been frequently discussed in census publications. Lie & Roll-Hansen (2001) distinguish the Norwegian censuses before and after 1865, arguing that the 1865 Census was the first to base the information collection on work (industry and occupation) rather than social status. They also summarize the views of Anders Kiær, head of the Norwegian Census office (which became the statistical agency, Statistics Norway, from 1876). Kiær argued that the categorization of women at home in the 1865 Census, together with children, led to an under-appreciation of their economic contributions. For this reason, a different categorization was used in 1875, continuing onward to the 1910 census that we consider here.

The final report on the Census (Statistics Norway, 1916) includes some discussion on the classification of womens’ work. Participation in the labor force is conceptualized as a combination of

---

<sup>4</sup>All census publications can be downloaded from Statistics Norway at <https://www.ssb.no/historisk-statistikk/folketellinger/folketellingen-1910>

<sup>5</sup>The relevant enumerator instructions are given in Statistics Norway (1916) and are reproduced here in the Appendix, Section A.1. The data collection shares many similarities with other censuses in Western countries in the same period, such as those in the United States and United Kingdom. However, it should be noted that the Swedish censuses before 1930 followed a different practice, with parish priests extracting data from parish records rather than enumerators interviewing residents (Molinder, 2022; Dalman, 2026)

<sup>6</sup>Instructions also refer to unemployment, indicating that both employed and unemployed are to be recorded, as in modern definitions. The unemployment information does not appear to be used in published tables and it is not clear how comprehensively it was recorded.

two factors: *erhverv*, which refers to the source of income here translated as “employment”, and *livsstilling*, position in life, here translated as “social status”. The report establishes a typology where paid employment (A) is distinguished from work within the house benefiting the family (B), income from pensions, wealth or public support (C), and those supported by the family (D). Gunnar Jahn, who wrote the report, acknowledges that it could make sense to merge group A and B together for conceptual reasons but that this would make the census less useful, as it would hide characteristics that the census has as a purpose to uncover. In the analysis of occupation groups, housewives and family members helping at home are removed from the tabulations.

Although the discussion of participation in the various occupation groups is predominantly done without reference to gender, a couple of measurement challenges are explicitly brought up. First, that it can be hard to distinguish those helping at home from those privately supported, and to some extent also from servants. For example, widows or unmarried women who run the household for relatives were reported as housekeepers (employed) in 1910, but as family members in earlier censuses. Second, there are challenges in the agricultural sector. A crude measurement of the labor force participation of women age 15 and older is reported (p. 124), but only for “industry, trade and immaterial work”. 7 per cent of rural women and 28 per cent of urban women are found to be working in these sectors.

An account of the general development of women’s role in the labor force in Norway in the beginning of the twentieth century is beyond the scope of this paper. A comprehensive review is given by Blom *et al.* (2005, pp. 177-212, 306-323), highlighting the role of family work and the paid employment of women. Education gradually became more important, though the share of Norwegians with a high school degree was still very low. The first woman lawyer started practicing in 1904 and the first professor in 1912. In rural areas, traditional gender segregation still prevailed, for example in large parts of Northern Norway where fishing meant that men would be away from the house for months while women ran the farm (Balsvik, 1991).

In general, technology and the change in social norms contributed to the increasing role of women in the labor market and to the entry of women into new occupations. Two central developments highlighted by Statistics Norway (1916) are the increase in the number of women working as public servants and technological change. The entry into the public sector is described as covering most sectors, with the exception of “some that women have not yet attempted to enter (for example, railways)” and “some that they by nature are excluded from (defense)”. Technological development is described as working both for and against women’s employment: In textiles, an ongoing change from cottage production to industrialization reduced the number of independent female workers in 1910 compared to earlier years, while in the production of shoes, the number of (male) shoemakers decreased while there were many women employed in the increasing number of shoe factories.

## 2.2 Labor force participation according to the census

We begin with a presentation of the census results as given in the original reports. We focus here on two key overviews. First, the distribution of the population on employment status and second, on social status.

Population by employment NOS VI 77, p. 70		Population by social status NOS VI 20, p. 4-5	
1. Employed	269237	I. Agriculture, forestry, fishery	53214
		II. Mining and manufacturing	65186
		III. Trade and transportation	36478
		IV. Public work; private immaterial work	13638
2. Housewives	355366	V. Housework: work of indet. character	577334
3. Children at home			
<i>a. in agriculture</i>	15537		
<i>b. at housework</i>	105710		
5. Income from wealth	40094	VI. Income from wealth; retirees	40094
4. Privately supported			
<i>a. students</i>	5391		
<i>b. others age 15+</i>	12229		
6. Publicly supported persons	19349		
		VII. No employment	36969
7. Status not given	3184	VIII. Employment not stated	3184
Total female pop. 15+	826097		826097

Table 1: Labor market status according to original reports of the 1910 Census of Norway. Women age 15 years and older.

Table 1 combines two tables from the original census report. Each table categorizes the female and male population; only the numbers for women are shown here. The left column shows the results of the population table by employment (*erhverv*) (Statistics Norway, 1916, p.70).<sup>7</sup> 269,237 women, or 32.6 per cent, are listed as “employed” (category 1). In addition, 121,247 women are family members older than 15 years (children at home, category 3). We do not know the details of their contribution to the household economy, except that the census tabulators have grouped them into agriculture or housework. Although most of the men (not shown here) are in agriculture, a large majority of women are listed as in housework. If we were to include all of them in the labor force, we would get a female labor force participation of 47.3 per cent.

Finally, there were 355,366 housewives (category 2) in the 1910 Census. This is 43 per cent of the adult (15+) female population. If we add all of them to the labor force, in addition to categories 1 and 3, we would obtain a labor force participation rate of 90.3 per cent. While it is likely that some of these women would be classified as working according to modern statistical

<sup>7</sup>The display here is restricted to only individuals 15 years and older. The original table also has a category 4c: children under 15 years of age (423,693, of which 13,106 with own employment)

practices, it is unlikely that they would apply to all of them and as such it is not reasonable to pursue this high number any further. We return to the measurement of married women's positions in the labor market later in the paper. The final three panels in the left column of Table 1 are groups that are not reasonable to classify as in the labor force. 40,094 women (4.8 per cent) have income from wealth (category 5); 17,620 are privately supported as students or otherwise (category 4), and 19,349 are listed as publicly supported (category 6). Finally, there are 3,184 women for whom it is not possible to determine any employment status.

We can compare the classification of the population according to employment status with the population by social status (*livsstilling*) (Statistics Norway, 1914, p.4-5). This is shown in the right column of Table 1. The female labor force are classified here into five broad categories I-V. Categories I-IV correspond roughly to the modern partition of the labor market into primary, secondary, and tertiary sectors, with the tertiary sector being split into two groups. Each of these three sectors employs around six per cent of the female adult population.

Category V includes housework and work of indeterminate character. From more detailed tabulations in the same publication, we learn that housewives and children at home at housework, categories 2 and 3b in the left column, are placed here.<sup>8</sup> This gives a remainder of 116,258 in category V who are neither housewives nor daughters. Categories VI, VII and VIII in the right column match exactly the three bottom panels in the left column, showing that the categorization is consistent across the different tabulations produced from the census.

Occupational category (across) Social position (down)	I Agric.	II Industr.	III Trade	IV Publ./ immat.	V House- work/oth.
A. Private independent businesspersons	13,617 (5)	883 (2)	8,497 (6)	1,039 (3)	
B. Independent workers	967 (4)	34,114 (6)	1,771 (3)	478 (2)	355,409 (2)
C. Appointed civil servants				1 (1)	
D. Employed civil servants	8 (1)	50 (1)	2,607 (5)	5,850 (3)	117 (1)
E. Private managers and directors	28 (1)	219 (1)	434 (3)	76 (1)	86 (1)
F. Private clerks	1,072 (1)	989 (1)	15,678 (5)	2,454 (4)	23,865 (2)
G. Foremen	5 (1)	30 (1)	15 (1)	31 (1)	47 (1)
H. Servants	16,031 (3)		3,492 (2)	22 (1)	79,347 (2)
I. Workers	21,486 (6)	28,901 (17)	3,984 (3)	3,687 (2)	118,463 (4)
J. Income from wealth			(Category VI)	40,094 (9)	
K. No occupation			(Category VII)	36,969 (10)	
L. Occupation unknown			(Category VIII)	3,184 (1)	

Table 2: Employment as reported in 1910 Norwegian Census publication, re-tabulated to show occupational category and social position. Women aged 15 years and older.

In addition to the overview tables, the publication also provides detailed cross-tabulations of social status against a more detailed set of occupation categories, as well as age and marital status

<sup>8</sup>We also learn that the children at home in agriculture, 3a, are classified in category I of this table.

(Statistics Norway, 1914, p. 46f). A condensed version of this information is given in Table 2, which again categorizes all women 15 years and older. Rows show the contemporary grouping in nine occupational categories A-I, while columns show the five main categories I-V from the right column of table 1.<sup>9</sup> The number in parentheses shows the number of occupational subcategories (below the nine shown) reported in the original tables.

In the agricultural category, three groups stand out: the 13,617 private independent businesspersons, of which 12,139 are farmers; a majority of these again are widows. There are 16,031 servants in agriculture, mainly related to animal husbandry (cattle). 21,486 are listed as workers, and 15,537 of them are the previously mentioned “daughters and relatives at home employed in agriculture”. In both of the latter categories, the majority of employees are young and unmarried.

In category II, industrial work, there are 34,114 independent workers, of which a majority are seamstresses. The 28,901 workers are spread across 17 different categories, of which none have the plurality. The largest are seamstresses in small-scale industry and workers in clothing industry. In category III, trade, the groups are of more even size, but two categories stand out: category A, private independent businesspersons, and category F, private clerks. Of the 8,497 businesspersons the largest category is managers of hotels or eating establishments, while among the 15,678 private clerks, the largest category is commercial agents, clerks, assistants, and office clerks, mostly young and unmarried. The largest group in category IV, public and immaterial work, is civil servants; most of these work in teaching.

As discussed above, Category V, housework and work of indeterminate character, is by far the largest category of female employment. The 355,366 housewives are classified as “independent workers” in the household sector. In this group, there are an additional 43 individuals who are listed as employed as “earning a living from raising children”, presumably fostering children in public care. There are 23,865 private clerks in this category; no more details are given about this group except that 22,957 of them are in the housework category. Of the 79,347 servants, 77,878 are servants in private households, and the remainder are servants in hospitals and prisons. Servants are generally young and unmarried. Finally, there are 118,463 workers in Category V; of these, 105,710 are daughters and relatives employed at home.

As we can see, the 1910 census results that were published in the 1910s already have a detailed classification of womens’ work and a great deal about its distribution can be learned from the tabulations. This tells us that care was taken in the enumeration of women’s economic activity, but we still cannot know how precise these tabulations were. For example, we do not know whether being a housewife indicates that there was no economic activity carried out outside the home. As a first step to going into more detail on the original census records, we move to the digital micro-data versions, where all occupations are transcribed.

---

<sup>9</sup>Categories VI-VII from Table 1 have corresponding occupational categories J-L and are shown at the bottom of the table for completeness.

### 3 Digital versions of the 1910 Census

The availability of individual-level transcriptions of the 1910 Census data makes it possible to go into far more detail than one could do using the original census reports exclusively.

#### 3.1 Versions and coding

Individual-level records from the census were kept by Statistics Norway after the completion of the census publications in the late 1910s, and eventually transferred to the Norwegian National Archives. They were transcribed between 2001 and 2009, in preparation for publication in 2010, after the 100-year privacy protection had expired (Digitalarkivet, 2024). The transcription has been coded for research use by the Norwegian Historical Data Centre and published online as part of the comprehensive IPUMS database by the Minnesota Population Centre (Minnesota Population Center (2020); Thorvaldsen (2018, p. 262ff)).

The IPUMS version of the file includes variables coded in a harmonized way for age, marital status, household position, place of birth, and occupation. Occupations are specified using an extension of the HISCO standard (van Leeuwen *et al.*, 2002; Registreringssentralen for historiske data, 2014). There are a total of 534 distinct occupation codes in the published research file. Information on family relationships in the status variable is coded using a separate status variable. For example, a farmer’s son is coded as “farmer” in the occupation variable and “son” in the occupational relationship variable (van Leeuwen *et al.*, 2002, p. 28-29). Similarly, there is a status variable that registers whether the occupation string lists if someone is retired. As will be made clear below, these are not always consistently applied across genders.

For most purposes, the IPUMS files contain all the relevant information necessary for quantitative research. However, in the present case, it is also of interest to use the original occupation strings from the census file, which are not included in the IPUMS data file. For this reason, the most recent version of the original transcription file was obtained directly from the National Archives, and merged at the individual level, with the IPUMS file.<sup>10</sup>

The key variable of interest is the variable that indicates participation in the labor force. From the documentation on the IPUMS web page:

*LABFORCE is a dichotomous variable identifying whether a person participated in the labor force. Labor force participation generally means working or seeking work within a specified reference period.*

---

<sup>10</sup>As the order of observations, and in many cases the number of observations per municipality, is different across the files, an algorithm was developed to merge the files. This algorithm is described in Appendix A.3. There are 2,483,123 observations in the merged data. 14,300 observations from the National Archives file and 7,875 observations from the IPUMS file could not be merged.

No further detail is given, but it is clear that the variable is closely related to occupational codes. Individuals are coded as “in the labor force” if they have an occupation that is coded in HISCO. The variable is defined for individuals 15 years and older. A tabulation of this variable gives 315,707 working women out of a total of 819,640; 38.5 per cent. For men, the corresponding number is 624,964 out of 679,600 (92.0 per cent).

There are slightly more than 300,000 unique occupation strings for the population 15 years and older.<sup>11</sup> The most common strings in the transcription file for individuals 15 years and older, each accounting for two per cent or more of the population, are *hustru* (wife), *gaardbruker* (farmer), *tjenestepike* (servant girl), *gaardmandskone* (farmer’s wife), *husgjerning* (housework), and *fisker* (fisher[man]). These are all nearly completely gender stereotypical, the only one of these six occupation strings with less than 99 per cent of the majority gender is *gaardbruker* with 2.8 per cent women. *Gaardbruker*, *tjenestepike* and *fisker* are coded as in the labor force, while the other three are coded as not in the labor force. *Hustru* and *gaardmandskone* are coded as “wife” in the occupational relationship variable, with no HISCO code for occupation.

### 3.2 Comparing contemporary definitions to the modern data set

As nearly a century separates the original tabulation of the census results from the modern coding, it is of interest to see whether it is possible to recreate the tables in the census publications using micro-data. Table 3 shows the result of this exercise, for the grouping of the population by employment status as shown above in Table 1. Here we show both men and women to see if there are any systematic differences between the genders.

The total number of employed as tabulated from the micro-data file is similar in magnitude to that reported in the census publication, but the difference is larger for women (10 per cent more in the IPUMS data set) than for men (1 per cent more in the IPUMS data set).<sup>12</sup> The number of housewives is also similar across sources; in the tabulation of the IPUMS data set presented here, this category is constructed by including all married women not coded as employed. This suggests that the definitions imposed by IPUMS comes a long way in capturing the classification scheme intended by the original enumerators.

There is a larger discrepancy when it comes to the category “children at home”, both for men and women. Using the variable “occupational relationship” in the IPUMS data, and selecting a son or daughter, recovers 24,488 men and 32,769 women, much less than in the original sources. The 10% surplus of employed women in IPUMS is likely to cover some individuals who were reported as family members in the original census table, but this only accounts for less than half of the

<sup>11</sup>Including both men and women. Some minor adjustment for differing spelling conventions are applied to the file, for example *bruker* and *pike* can be spelled both with *k* and *g*, and *aa* can also be spelled as *å*.

<sup>12</sup>In the IPUMS data set, this is for all individuals coded as “in the labor force”, except those who fit any of the categories below.

	Census publication		IPUMS micro data	
	Men	Women	Men	Women
1. Employed	578,869	269,237	583,602	297,493
2. Housewives		355,366		349,181
3. Children at home, of which:			24,488	32,769
a. in agriculture	51,224	15,537		
b. at housework	5,927	105,710		
4. Privately supported, of which:				
a. students etc.	9,438	5,391	10,615	6,174
b. others above 15 y.o.	6,577	12,229		
b. under 15 years old	423,693	408,533	421,262	406,762
b.1 of which with own employment	13,106	11,684		
5. Persons supported by income from wealth, etc.	34,381	40,094	34,590	24,886
6. Publicly supported persons	10,586	19,349	9,292	21,201
7. Status not given	2,465	3,184	17,013	87,936

Table 3: Labor force participation according to Census publication and constructed from IPUMS micro data. All individuals, 1910 Norwegian census

difference. Across categories 4, 5 and 6 there are also some discrepancies; there are more privately supported individuals in the original census publication. Finally, there are many more with “status not given” in the IPUMS data set, simply because all the nuances in the original tabulations related to the complex interrelationship between household position, social status and employment are not possible to reproduce using the HISCO system supplemented with the status variables.

We recall from above that the IPUMS data set indicates a female labor force participation rate of 38.5 per cent. Table 3 indicates a slightly lower number, 36.3 per cent, when some groups are recoded according to family status or indicated dependency on private wealth or public support. However, it is not clear that all “children at home” should be coded as not working, so this lower number can probably be disregarded.

At this point, we can establish that IPUMS does not appear to undercount working women compared to original publications. Before we turn to an external verification using other sources from around 1910, we examine in more detail how working women are distributed across occupational types, according to the IPUMS data.

### 3.3 The distribution of working women across occupational groups

Moving to a description of the types of occupations held by working women in the 1910 census, we rely on the HISCO classification. Table 4 shows all women 15 years and older in the six major occupational categories, as well as marital status.<sup>13</sup> The largest occupational category is sales and

<sup>13</sup>The first digit indicates the HISCO major group, with 0/1 and 7/8/9 referring to the same category, giving a total of seven major groups. For the purpose of later analysis of married women, it is reasonable to consider groups



service workers, with a total of 132,610 women, 42 per cent of those employed. This is a category that was not as distinct in the original tabulation, as illustrated in Table 2 above. Most women in these occupations are single. Slightly less than half are servants (HISCO 54010), the other occupation groups with more than 10,000 workers are housekeepers (52020), maids (54020) and salespeople (45120).

HISCO major group	Marital status				Total
	Single	Married	Widowed	Unknown	
0/1 Professional, technical	9,239	1,567	671	144	11,621
2+3 Administrative, managerial, clerical	9,481	594	623	150	10,848
4+5 Sales, service	111,210	7,300	12,251	1,849	132,610
6 Agricultural, forestry, fishery	46,987	21,407	20,525	425	89,344
7/8/9 Production and related, laborers	42,446	6,196	6,959	847	56,448
99 Laborer/Not elsewhere classified	10,569	1,348	2,673	246	14,836
Total in labor force (sum of above)	229,932	38,412	43,702	3,661	315,707
Not in labor force	108,367	339,544	51,748	4,274	503,933
Total	338,299	377,956	95,450	7,935	819,640
<i>Share in labor force</i>	<i>68%</i>	<i>10%</i>	<i>46%</i>	<i>46%</i>	<i>39 %</i>

Table 4: Employed women 15 years and older, by occupational category and marital status. Note: Widowed includes a small number with status “separated”.

The second largest category is agriculture, with a total of 89,344 women. Of these, 50,012 are farm workers in dairy (HISCO 62500), 20,974 are farmers (61110) and 14,435 are farm workers of unknown specialization (62110). There is also a substantial share in production and related activity. Of the 56,448 in this category, 22,975 are seamstresses (HISCO 79520).

We also see from Table 4 that the labor force participation rate varies by marital status, with 68% of unmarried women coded as employed, but only 10% of married women. As was evident from the historical tabulation, where a large share of the population was coded as “housewives”, it is of particular interest to see if we can go into more detail on the economic participation of married women. As a step in this direction, we begin by examining the occupation strings — as recorded by the enumerators — in more detail.

### 3.4 Biases in the occupational coding process

Has any information on married women’s work been lost in the coding process? By revisiting the occupations the enumerators wrote down, we can try to address this issue. For each individual, we have both the transcribed occupation string (from the National Archives file) and the HISCO en-

2 (Administrative and managerial) and 3 (Clerical and related) together, as well as groups 4 (Sales workers) and 5 (Service workers). Moreover, since a substantial share of the workforce is laborers and “not elsewhere classified” (code 99), we count them separately from the other HISCO groups. This gives a total of six occupational categories, as shown in the table.

coding (from the IPUMS file). We restrict the sample to married women where we have information on both wife and husband (e.g. not those whose where there is no record for a spouse).

First, the occupation strings were harmonized to account for minor spelling differences. Then, all occupation strings assigned to more than 50 women were manually coded to categorize the type of information present. In this manual coding process, fourteen different groupings of labels were found to capture much of the relevant heterogeneity. These categories are shown in Table 5, cross-tabulated with the husbands' occupations (as coded in the IPUMS file), as well as whether the women are classified as in the labor market. Grouping outcomes by husbands' occupations gives some indication of which economic sectors the women's households are connected to, and helps distinguishing patterns across agricultural and non-agricultural households.

Wife's occupation field indicates...		Husbands' occupation (HISCO)			Total	Of which wives in the labor force
		0-5 Prof., adm, sales	6 Agric.	7-9 Prod., laborers		
Own occupation	(1) Farmer	31	348	38	417	417
	(2) Other agricultural	189	5,159	494	5,842	5,842
	(3) Other	392	409	820	1,621	1,621
Wife of...	(4) Farmer	915	82,724	1,644	85,283	164
	(5) Fisherman	7	4,313	17	4,337	0
	(6) Other occupation	1,602	2,237	6,659	10,498	0
(7) Housework		5,710	14,105	14,048	33,863	451
(8) Housewife		7,362	6,762	12,827	26,951	0
(9) Wife		20,296	13,392	39,878	73,566	0
(10) Pensioner/retired		37	6,589	36	6,662	75
(11) Rentier		4	1	2	7	0
(12) Pauper		1	4	6	11	0
(13) Nothing (explicitly)		390	102	323	815	0
(14) No information (empty string)		11,795	5,408	18,777	35,980	0
Total classified strings		48,731	141,553	95,569	285,853	8,570
Not classified ( $N < 50$ )		9,507	26,343	16,210	52,060	21,261
Total		58,238	167,896	111,779	337,913	29,831

Table 5: Married women, by information in own occupation string as well as husband's occupation category. Note: Women with husbands not in labor force (HISCO=99999) excluded from table.

Rows (1)-(3) summarize individuals for whom there is a clearly defined occupational description. All are coded as in the labor force in the IPUMS data set. 417 women are described as farmers; most of these are married to men who are also farmers. 5,842 married women, mostly married to farmers, have other agricultural descriptions. These are predominantly related to animal husbandry and are also coded as in the labor force. Finally, 1,621 married women have a non-agricultural occupation listed. Of these, 440 are seamstresses. It should be emphasized that this is not the entire married female working populations, as only strings with more than 50 individuals were considered, and the

share of working women is higher among women with less common occupation descriptions.

Rows (4)-(6) consider descriptions in which the woman is described as “wife of” an occupational position. 85,283 are farmers’ wives, 4,337 are fishermen’s wives and 10,498 are the wife of some other occupation or profession. Here, it is unclear from the description whether this describes participation in the main employment (such as running the farm) or whether it merely describes the woman’s role in the household. The classification in the IPUMS data has consistently coded these strings as “not in the labor force”.<sup>14</sup> It is challenging to map these 100,000 women (thirty per cent of all married women) to a labor force status, and it is likely that there is substantial variation within this group. Even if there is some disagreement as to how to map family production in farming to labor market activity (see e.g. Burnette, 2021, who argues this is a futile endeavor in non-industrialized societies) it is reasonable to assume some level of participation for this population. Regarding the population in fishing, in large parts of Norway there was a traditional division of labor between men (fishing, often away from home for a long time) and women (farm work) (Balsvik, 1991).

Around half of the occupations in the “other” category are simply “worker” (the women are listed as “worker’s wife” or similar), while many of the others refer to occupations clearly not taking place in the home (teacher, seaman, forest worker). It is also evident from the categorization across husbands’ occupational categories from the census data that many of these are outside agriculture, and for this reason it may not be reasonable to categorize this population as in the labor force. On the other hand, according to the instructions (see Appendix A.1), the text string is supposed to refer to a type of employment. In addition, there is an even higher share of women who are described as doing “housework” or simply being a “wife”, or with the occupational field left blank. Does this mean that the wives of farmers, fishermen and other occupations really should be counted as working? That would revise the labor force participation of married women in the classified subsample from 3 per cent to 38 per cent.

The study by You (2020b) on British census data between 1851 and 1911 argues that women enumerated as an “occupation’s wife” should not necessarily be counted as working in the British case. While the enumeration instructions would indicate some economic information from these types of text strings, You argues that the variation in such enumerations across occupation and farm sizes does not indicate that they should be considered as actively participating in their husband’s economic activity in a substantially different way from those enumerated otherwise. In the end, it is likely that some information on married women’s economic activity in Norway was intended to be captured by some enumerators, but we do not know if the enumeration was systematic enough that we can use this to systematically revise estimates of female LFP upwards.

Rows (7)-(9) of Table 5 refer to wives described as doing “housework” or being a “wife”, in some

---

<sup>14</sup>The exception is 164 farmers’ wives who also have some information about animal husbandry in their occupation string.

way. All are treated equally in the IPUMS data, as not working, with the exception of 451 women reported as housekeepers. However, there is substantial variation in the underlying text, and some text overlaps with what would have been reported for servants had the individual been in another household position. Again, it is likely that the harmonized coding masks heterogeneity within the group.

The fourth panel — rows (10)-(12) — of Table 5 shows married women that are retired, as well as rentiers and paupers. The retired includes individuals on *føderåd* (or *kår*), the Norwegian traditional system where farm owners would retire and hand the farm over to the next generation, based on a contract guaranteeing room and board on the farm. This presents an interesting example of differential occupation coding for men and women in comparable positions. Of the 6,662 with a common occupation string indicating retirement (*føderåd*, *kår*) only 75 are coded as in the labor force in IPUMS (those who do not have “kone”, wife, as part of the string). The remaining 6,587 are coded as not in the labor force, while their husbands are for the most part coded as in the labor force. This appears to be a clear inconsistency in the IPUMS coding.

The fifth panel lists the number of women for which it is either literally indicated “nothing” in the wife’s occupation string (row 13) or where it is empty (row 14). This sums up to a total of 285,853 married women with common (shared by at least 50 individuals) occupation strings.

To sum up the insights from reviewing the digital data file for the 1910 census, it is clear that it captures a lot of variation in the characteristics of women’s labor force participation. 68% of unmarried women are coded with an occupation and in the labor force. There is a relatively high degree of agreement with the contemporary census publication. However, there is some indication that women with recorded occupation information are discarded because it includes “wife”; moreover, retired farmers are clearly treated differently based on gender in the coded census file.

We now turn to reports of economic activity from other sources than the census to assess if these numbers are substantially different. This will help improve our understanding of how to interpret census numbers.

## 4 Triangulation between historical sources

This section compares results from the IPUMS data set and the original tabulation with other contemporary sources, to compare the number and share of women recorded in economic activity. We begin with two studies specifically looking at industrial employment and compare their coverage with that from the census. Then, we turn to a study based on tax records that also counts workers outside manufacturing.

## 4.1 Industrial statistics

The first source of employment statistics is the 1909 Census of industries. It was carried out by Statistics Norway, using a set of forms sent to manufacturing companies. The scope of the investigation was the companies covered by the Factory Inspection Authority (*Fabriktilsynet*); the publication notes that this does not perfectly correspond to the definitions used in the population census and the employment insurance statistics. The total responses consist of a total of 3,821 companies with a total employment of 105,042 individuals, corresponding to a response rate of 94 per cent.<sup>15</sup>

The results of the factory census were reported in a set of four publications between 1909 and 1915. Results of interest to the present paper are all in the first publication, listing the number of companies, workers, and owners (Statistics Norway, 1911). The discussion of the results refers to the gender composition of the workforce as one of several compositional aspects in addition to age, company size, and other characteristics. An overview of the number of men and women in industrial employment, separated into position (managers, submanagers, office workers, foremen, and workers), is given (p. XXIV). Of the 98,313 employed people in the surveyed firms, 20,099 (20.4 per cent) were women. Of these, the vast majority were workers, but there are also 1,118 office workers (27.4 per cent of all office workers) and a few foremen, submanagers, and managers. The text of the report does not indicate any concern about under- or overcount with respect to gender.

The data we use here are taken from the table of the gender composition of workers across industries on p. XXIV, listing the total number of workers, the number of women and the per cent women among workers in each industry. The highest female employment shares are found in textiles (6,664 women; 67.6 per cent of all workers in the industry) and clothing industry (1,691 women; 56.3 per cent of all workers in the industry). Other important industries are the food industry, employing 4,689 women (43.0 per cent) and the paper, leather, and rubber industry (1,927; 14.7 per cent).

Our second contemporary external source of employment data is the annual industrial statistics compiled by the national insurance board (*Riksforsikringsanstalten*). We base our comparison on the 1910 statistics, published in 1912 (Statistics Norway, 1912). The scope is limited by the coverage of the national employment accident insurance: manufacturing, mining, forestry, and some other related industries. It is explicitly stated that home production is not covered.

The basis of the analysis in the following is a segment of Table 1 reported on pages 164-165, listing, by industry, the number of companies, days of work, and salaries. Our interest lies in the male and female workers across type of work. In the employment insurance statistics report it is the number of days worked, not number of people employed, that is listed. For this reason, we have

---

<sup>15</sup>Information taken from the introductory pages (Statistics Norway, 1911, p. I-IV); listing 244 nonresponding companies in addition to the 3,821 responding.

to convert to workers; we will keep the reports’ convention of using 300 days per year, though this might give an undercount of female employment if women are more likely to work part time than men. The share of women among workers range from none (Forestry) to 64.8 per cent (Textiles).

## 4.2 Comparison of industrial employment

We now have four sources across which we can compare industrial employment. First, there is the original census tabulation, collected from household by census enumerators and tabulated by Statistics Norway in the 1910s. Second and third, the two comprehensive surveys conducted by specialized agencies on the employer side. Fourth, we have the IPUMS tabulations of the same census, coded in the 2000s based on the strings recorded by the census enumerators and tabulated electronically.

A reasonable expectation could be that the data collected from employers and workplaces would be more “gender neutral” than the data collected from the household side. After all, for the employer, a female employee is an employee just like a male employee, but for someone visiting the household, the woman of the household would be juggling several roles and the enumerator’s judgement could affect which one would be reported in the census form. However, as we shall see, there is no systematic evidence that this is the case.

We limit the analysis here to industry (manufacturing), where the coverage in the industry statistics is most comprehensive. In addition to aggregate numbers for industry, we also examine in more detail the four industries where the number of female employees is highest: Textiles, clothing, paper/leather/rubber, and food.<sup>16</sup>

Table 6 shows employment for men and women, as well as the share of women in total employment, according to each of the four data sources. The first panel reports numbers for all industries. The census reports more workers (both men and women) than the industrial statistics, showing that the coverage of the industrial statistics in terms of number of workers is not as comprehensive as that in the census. Even so, the share of women according to each of the two sources of industrial statistics (21 and 19 per cent) is very close to that reported in the census (18 per cent).

The number of women counted as working in the modern data set (IPUMS) is much higher than that in the original census tabulations. Based on the original tabulations, 18% of women would be counted as working, while the micro data set reports 30%. This suggests that the original processing of the census forms undercounted women’s labor force participation to some degree, although there could also be other explanations. For example, there could be differences in what is counted as “industry”. It is hard to separate manufacturing from cottage industry in the IPUMS data, and hard to discern any home production or putting-out systems in either data source. Nevertheless, we

---

<sup>16</sup>When using the IPUMS micro data, the corresponding occupation categories chosen are (at the 2-digit HISCO level, with exceptions at the 3-digit level given): Textiles: 75; clothing: 79 and 80; paper/leather/rubber: 73 (except 733), 76 and 90; Food: 77 (except 776) and 78.

		Men	Women	Share women
Total employment, all industries	<i>Population census (original report)</i>	133,237	28,901	18 %
	<i>Census of industries</i>	64,490	17,612	21 %
	<i>Industrial statistics (insurance board)</i>	84,211	19,137	19 %
	<i>IPUMS micro data from pop. census</i>	124,212	55,344	31 %
Textiles industry	<i>Population census (original report)</i>	3,196	6,190	66 %
	<i>Census of industries</i>	2,944	6,315	68 %
	<i>Industrial statistics (insurance board)</i>	3,020	6,179	67 %
	<i>IPUMS micro data from pop. census</i>	3,258	12,856	80 %
Clothing industry	<i>Population census (original report)</i>	9,489	10,472	52 %
	<i>Census of industries</i>	1,223	1,613	57 %
	<i>Industrial statistics (insurance board)</i>	1,688	2,293	58 %
	<i>IPUMS micro data from pop. census</i>	19,016	30,098	61 %
Paper, leather and rubber industry	<i>Population census (original report)</i>	12,035	2,115	15 %
	<i>Census of industries</i>	11,074	1,860	14 %
	<i>Industrial statistics (insurance board)</i>	11,656	2,179	16 %
	<i>IPUMS micro data from pop. census</i>	9,953	1,277	11 %
Food industry	<i>Population census (original report)</i>	9,706	4,435	31 %
	<i>Census of industries</i>	5,830	4,410	43 %
	<i>Industrial statistics (insurance board)</i>	7,991	5,275	40 %
	<i>IPUMS micro data from pop. census</i>	7,743	5,587	42 %

Table 6: Employed men and women, around 1910. Comparison of information from four sources: Original Census report (age 15 and older); Census of industries (age 16 and older); Public insurance statistics (age 16 and older); IPUMS micro data (age 15 and older).

can conclude that there is no evidence here that the IPUMS data is systematically under-counting womens' labor force participation compared to other sources. As the selectivity of what counts as industry likely affects the comparison, we now turn to the four specific industries, which should give a more precise comparison. These are reported in panels two to five of Table 6.

In the textile industry, the number of men is approximately the same in all four sources, while the number of women is similar in the three historical publications but not in the IPUMS data. This suggests that the definition used in IPUMS is more comprehensive than in census data, in a way that only affects women.<sup>17</sup> For clothing industry, both census data sets report many more workers, both men and women, than the industry statistics, but again the IPUMS overcount is higher for women. For paper, leather and rubber industry, it is the other way around: consistency across sources, but fewer women in the IPUMS data. Finally, in the food industry, the number of workers has some variation, but the number of women is relatively consistent.

For the four industries studied in detail, the share of women reported is relatively consistent, even if the total number of workers varies. The share varies by between thirteen percentage points (textile industry, where IPUMS reports the highest female share) and five percentage points (paper, leather, and rubber industry, where the insurance statistics report the highest female share). Substantial variation in total employment numbers for two of the industries indicates that industry definitions across sources are not perfectly aligned. However, the limited variation in the reported female shares suggests that there is no systematic discrepancy in the count of men versus women. One could expect that the census of industries and public insurance statistics would be less restrictive in counting women than the census, as information is collected from workplaces where it is clear who shows up to work. However, there is no indication of such a bias; for all the four industries, at least one of the two census sources report female shares equal to or higher than one of the sources reported from the employer side.

We now move to a more comprehensive external source of employment data, namely tax records, to see whether this changes our conclusion from the industrial statistics.

### 4.3 Census and tax records

At this point, we have compared the recording of labor force status from the household side (population census) as well as via employers (employment census and insurance statistics). We now turn to a third source, namely income tax statistics. Income taxation at the central government level was introduced in Norway in 1892 (Aaberge *et al.*, 2020); several municipalities also had income taxes before this. In principle, all individuals were subject to income tax, but the tax thresholds meant that low-income individuals did not pay the tax. Married couples were jointly taxed, recorded

---

<sup>17</sup>Examination of the occupation strings in the census enumeration suggests this is the case; there are many seamstresses coded in the relevant HISCO categories, but it is likely that a substantial number are not part of enterprises large enough to be counted as factories, or to be counted in the insurance statistics.



under the name of the husband.

As part of the presentation of the results of the 1910 Census, Statistics Norway performed an analysis of the “income and wealth conditions” by linking census records to tax records from the 1911 income taxes at the individual level. The results were based on a sample of one third of all rural municipalities, all small cities, and parts of the largest cities.<sup>18</sup> The results were presented in a 1915 report (Statistics Norway, 1915).

The report includes a detailed tabulation of how many men and women from the census were “found” in the tax records. Some of the omissions are due to mismatching, such as a person being present in a municipality different from the one they were taxed in. However, in general, when an individual is not found it is predominantly because they did not pay tax, and hence are not reported in the tax lists at all. For this reason, we can use the presence in tax records to assess if there is a large population of women that pay taxes but are not recorded as working in the census. The tax rules means that we have to limit our analysis here to unmarried women.

Table 7 compares the coverage in the tax registry for men and women, by marital status, based on the much more detailed tabulation in Statistics Norway (1915). The first panel, covering the entire population except married women, shows that around 58% of men 15 years or older are found in the tax records, while only 28% of women. Even for married men, the share found in the tax records is only 74%, even though their labor force participation rate, as recorded in the census, is close to 100%. This reflects that individuals with incomes lower than the tax threshold are not present in the tax records, even if they are working.

The next five panels group the population by employment status, in a similar manner to the overview in the left column of Table 1. Among those employed, the share found in the tax records is not dramatically different between unmarried men (41%) and unmarried women (36%). This indicates a high degree of correspondence in the definition of “working” for both genders in the census definition. For widows/widowers there is a much higher difference, though some of this is likely driven by the average age of widows being higher than of widowers (higher longevity among women). For children and relatives at home, the numbers found in the tax statistics indicate that more men than women are active in the labor market. For example, 39% of unmarried men who are relatives at home are found in the tax records, compared to 23% of unmarried women.

In Section 3 above, it was highlighted that retirees in farms, *kårfolk*, are treated differentially in modern microdata, with men counted as employed and women as not employed. The fourth panel of the table shows that there are some differences between male and female retirees with respect to taxable income, though not dramatically so; for widows and widowers 47% of men and 34% of women have taxable income. The fifth panel compares the tax coverage among urban retirees and

---

<sup>18</sup>One fourth of Kristiania (Oslo), the largest city; half of Bergen, Trondhjem and Stavanger, the next three; and full coverage of all the other cities. The sampling in cities is random (every x building is sampled), while the rural municipalities are selected to be representative also of those not included. In the presentation here, we scale the estimates to be representative of the entire population, in a similar way as the original report.

	Marital status	Census	Men		Census	Women	
			Tax	%Tax		Tax	%Tax
Entire population	Unmarried	283,551	111,892	39%	339,175	99,178	29%
	Married	355,788	262,823	74%	—	—	—
	Widow/er	41,427	21,656	52%	94,452	23,312	25%
	Total	680,766	396,371	58%	433,627	122,490	28%
Employed	Unmarried	218,693	90,220	41%	199,693	71,130	36%
	Married	321,083	250,104	78%	—	—	—
	Widow/er	27,867	16,212	58%	53,024	12,244	23%
	Total	567,642	356,536	63%	252,717	83,374	33%
Children/relative at home	Unmarried	50,723	19,775	39%	106,383	24,962	23%
	Married	2,971	1,242	42%	—	—	—
	Widow/er	488	122	25%	4,986	373	7%
	Total	54,182	21,139	39%	111,369	25,335	23%
Retirees (farms)	Unmarried	999	441	44%	1,128	291	26%
	Married	12,009	7,590	63%	—	—	—
	Widow/er	7,821	3,699	47%	17,616	5,907	34%
	Total	20,829	11,730	56%	18,744	6,198	33%
Retirees (urban); rentiers	Unmarried	2,629	1,361	52%	5,912	2,458	42%
	Married	6,541	3,656	56%	—	—	—
	Widow/er	3,764	1,605	43%	11,645	4,714	40%
	Total	12,934	6,622	51%	17,557	7,172	41%
Others not employed	Unmarried	10,507	95	1%	26,059	337	1%
	Married	13,184	231	2%	—	—	—
	Widow/er	1,487	18	1%	7,181	74	1%
	Total	25,179	344	1%	33,240	411	1%

Table 7: Individuals in Norway in 1910, by gender and marital status, as reported in census and matched to tax statistics. Persons 15 years and older.

rentiers. This is the most gender neutral category, with a difference in the tax coverage of around 10 percentage points. It is also worth noting that there are more female than male retirees/rentiers, both in the census records and the tax lists.

Finally, the bottom panel shows that among those counted as not employed, almost no-one is found in the tax lists. This is yet another indication that the census has not classified working women as “not active”, at least not if they are unmarried or widowed.

In total, the results so far indicate that census takers were not averse to classifying women as employed if they participated in economic activity. The triangulation across sources has so far failed to find any indication of substantial undercount in either the original census publications or the modern microdata files, to the extent that this can be assessed more than one hundred years later.

## 5 Enumerator triangulation: Separate enumerators comparing the same person

So far we have examined the precision of the 1910 Census and the corresponding micro data file by comparing original tabulations to modern reconstructions, as well as other reports from the same period. All these approaches carries some inherent problems, in that classifications differ, and, in the case of the industrial statistics, also the population covered. This section assesses the data in the original census forms from a different angle by comparing the occupational classification of individuals who were counted twice. This provides a unique opportunity to observe to what extent two independent enumerators choose to classify individuals differently and whether the differences differ systematically between men and women.

### 5.1 Combining two information records for the same individual

We can compare different enumerators because of a feature of the 1910 Census: The objective was to count both the “de jure” and “de facto” population. To achieve this, individuals who were not present at their regular place of residence on census day were counted both at their legal place of residence (as “absent resident”) and where they actually were (“present nonresident”). Although most of the people were counted at their place of legal residence (that is, as “present residents”), a substantial part of the population was not (Statistics Norway, 1916, p. 3). There does not appear to have been any coordination of the two separate counts of the individuals and they appear as two different records in the census data sets.<sup>19</sup> This double-enumeration can be used to assess the degree of consistency between the enumerators, once we are able to identify which two records

---

<sup>19</sup>A review of the use of both de facto and de jure censuses in Norway is given by Thorvaldsen (2018, p. 154ff).

refer to the same person. In this way, the accuracy for enumeration of womens' occupation can be compared to that of men.

In order to match the two records of each person, the one at home (where she was not present) and the one from the municipality where she was present, we have to resort to algorithmic linking. A method based on that used in Modalsli (2017) is used to link individuals across censuses.

We consider similarities across six variables: First name, year, month, day and place of birth, and the other municipality in which the individual is found (the "away" municipality for the home municipality and vice versa).<sup>20</sup> For each of the six variables, a similarity score is conducted, which is zero with perfect match. If a match is imperfect, a positive number is reported, increasing in the imprecision of the match.<sup>21</sup> Last names are not used in the match as there was no clear standard for reporting married women's last names at the time (as her own patronymic or her husband's), and the use of surnames could also introduce a bias with respect to married and unmarried women. The scores for each variable are weighted and aggregated to a final score for each match. All potential pairs within each birth cohort are compared and those with the best match are chosen, given that the score is below a set threshold and sufficiently different from the second-best score for each of the two candidate records.

It is not straightforward to calculate the share of individuals that have been successfully matched by the algorithm. There are a total of 25,584 female absent residents age 15 and older, but only 21,499 "visitors" in other municipalities. For men, the numbers are even more lopsided: 99,243 absent and 51,834 visitors. Based on the text strings in the variables denoting where the absent individuals are, it is likely that much of the difference is made up of individuals residing abroad, including those (mostly men) on ships. It could also be the case that some of the "visitors" were missed by the enumerators and therefore not counted in the census, even if they were counted at home.

The algorithm, which is conservative in that matches where there are multiple candidates are discarded, matches a total of 5,734 women and 17,105 men. This is 27 per cent of women (counting the visiting as the base population) and 33 per cent of men. The restrictiveness of the algorithm should limit the number of spurious matches, but there could be a small bias in the sense that individuals with less common names and originating in smaller municipalities are matched to a larger extent. However, we should keep in mind that the sample as such is already highly selected: this is individuals who are away from home at census day, indicating a type of activity different from that of the common Norwegian at the time.

---

<sup>20</sup>Birthplace locations are coded in the IPUMS data set, but not home/away municipalities; these are coded by assessing string similarities with birthplace municipalities and then adapting the codes that correspond to the string used.

<sup>21</sup>For first names, it is the Levenstein distance between the two text strings; for birthdays and birth months it is based on the absolute difference in numbers; birth years must match exactly.

## 5.2 High agreement between enumerators

We now turn to the comparison of the two separate census records for the same individuals. At an aggregate level, the shares of men and women classified as in the labor force are similar in home and away locations. Using the classification of the IPUMS data set, we see from Table 8 that 54% of the matched women enumerated at their home location are counted as in the labor force, while 55% of those enumerated at their away location are counted as in the labor force. The women in this sample have a higher participation in the labor force than the general population. This is expected as it is likely that at least some of them are traveling for work. We are therefore looking at a sample selected on having higher economic activity than the average.

	Men	Women
In labor force (reported at home location)	96%	54%
In labor force (reported at away location)	96%	55%
In labor force at away location, if reported as...		
not in labor force at home	65%	78%
in labor force at home	97%	79%
In labor force at home location, if reported as...		
not in labor force at away	64%	26%
in labor force at away	97%	77%
Sample size	17,105	5,734

Table 8: Comparison of labor force enumeration of individuals enumerated twice in the census

There does not appear to be a large systematic difference with respect to cross-classification. Among women reported as part of the labor force at home, 80% are reported as in the labor force at the away location; similarly, among women reported as in the labor force at the away location, 78% are reported as in the labor force at home. For men, the reported participation in the labor force is as high as in the aggregate statistics. However, also for men, there is some discrepancy between home and away location. This means that it is not necessarily the case that the reporting on women’s labor force participation is less systematic than that for men.

Looking at the occupation strings for the women, it is evident that many of the disagreeing classifications have a description such as “daughter”, or simply an empty occupation field, in one location, while there is an occupation such as seamstress or servant in the other field. An interesting case is that housework (*husgjerning*) has been coded as not working in the IPUMS file, while servant (*tjenestepike*) is coded as working. That these terms are used by different enumerators for the same person can indicate that the housework term in some cases was meant to also encompass individuals doing paid work.

We can examine in more detail the consistency in occupational status classification by comparing occupational codes between the doubly-enumerated individuals. Again, it is useful to look at both men and women, as some misenumeration is inevitable, and it is of interest how inconsistencies are

different across men and women.

Occupation group in home municipality	Men				Women			
	<i>N</i>	In lab. force	Same group	Same code	<i>N</i>	In lab. force	Same group	Same code
0/1 Professional, technical	603	96%	77%	58%	283	90%	96%	70%
2+3 Administrative, managerial, clerical	357	96%	69%	55%	76	78%	78%	61%
4+5 Sales, service	681	94%	61%	48%	822	75%	75%	56%
6 Agricultural, forestry, fishery	9,728	97%	81%	62%	630	65%	54%	46%
7/8/9 Production and related	4,594	98%	82%	60%	1,044	88%	90%	74%
99 Laborer/Not elsewhere classified	440	96%	24%	19%	235	74%	48%	42%

Table 9: Comparison of labor force enumeration of individuals enumerated twice in the census, by HISCO occupation category.

Table 9 shows how reporting consistency varies between occupation groups. Each line reports a major occupation category, as reported for an individual in their home municipality. In other words, all individuals on which this table is based are reported to be working in their home municipality. Then, for men and women separately, it is reported how well their status in the “away” municipality fits that in their home municipality.

Starting with men, our comparison group, most are reported as in the labor force in the away municipality; between 95 and 98 per cent depending on occupation category. The share reported in the same HISCO major group is between 61 and 82 per cent, with the exception of laborers. The low concordance for laborers reflects that those classified as laborers often are in this category because of missing information about the industry in which they work. The share of workers with the exact same HISCO code is lower again, but generally more than half.

For women, the consistency in whether a worker is classified as in the labor force is generally lower, as previously reported. This means that there is also more variation. Among professionals and production workers, there is a high probability that a woman reported as working in her home municipality will also be reported as working in her away municipality. The lowest consistency is found in agriculture.

Moving to the sample in which women are counted as working in both locations, there is very high precision in occupational major group classification for professionals (96 per cent) and low precision in agriculture. As for men, general laborers are also frequently classified differently, in a more specific group, in their away municipality. The share of women with the exact same occupation code, contingent on being counted as working both home and away, is between 43 and 75 per cent, depending on the major category. We note that this is around the same level as for men and that for most occupation groups, the consistency is higher for women than for men.

In summary, we see that while two independent enumerators assessed an individual at two different places in 1910, there was not a perfect consistency across the classifications. However,

there is no indication that less care was taken in classifying women than in classifying men.

## 6 Discussion

### 6.1 High consistency in the measurement of female labor force participation

The analysis in this paper has shown a high degree of consistency in the measurement of Norwegian female labor force participation in 1910. Based on the original census tabulations, we get a participation rate of 33 per cent if we exclude family members, and 47 per cent if all family members (“children at home”) are included (see Section 2.2). The original reports acknowledge the challenges in classifying labor force participation in a society where much of the economic activity takes place within households such as farms and do not explicitly report any participation rate. Women are found in occupations in all types of economic activity.

The commonly used micro data set, provided by IPUMS, gives a female participation rate of 39 per cent when the “in labor force” variable is tabulated (see Section 3.2). This is intermediate to the two above estimates and is based on a coding of the original occupational strings into the HISCO standard. If an individual is placed with an occupation, they are counted as employed. It is clear from further analysis of the occupation strings that a cautious approach has been taken; for example, farm wives are only included if they have recorded an explicit task such as animal husbandry in the occupation string, which is rarely the case. The average rate of 39 per cent masks large differences between married and unmarried women: among unmarried women, the participation rate is 63 per cent; for married women, 10 per cent.<sup>22</sup>

Comparison with data from industrial statistics does not reveal any systematic difference between sources. For some industries, data reported from the employer side (industrial statistics) gives higher female employment than data from the household side (censuses); for other industries, it is the other way around. It is likely that differences in industry classifications and coverage are larger sources of these discrepancies than disagreement on how to measure womens’ employment. The main source of uncertainty is likely to concern those for whom economic activity is also associated with the household in which they live. Likewise, evidence from tax data (that only covers unmarried women) does not indicate any systematic under-measurement in the census. A review of a subsample of people that are enumerated in two separate municipalities does not indicate any higher bias or mis-measurement for women than for men.

The tentative conclusion, therefore, is that the labor force participation of women is well measured in the 1910 Norwegian census; that the organizers of the census succeeded in their ambition

---

<sup>22</sup>An overview of participation rates when we change the age definition, look at present vs. resident population, or re-code labor force definitions, is given in the Appendix, Section A.2.

to get a measure of labor force participation that was as accurate as possible. Although some measurement error at the individual level will always remain and is likely to be higher in 115-year-old sources than in contemporary statistics, there is nothing that indicates any special caution in using this type of microdata compared to other sources from the same period. However, a closer examination of the recorded occupations, as well as studying the cases where the same person was enumerated in two different locations, does reveal some room for an upward revision of the participation rate, mainly for married women.

## 6.2 Revised estimates of female labor force participation

An examination of recorded occupation text strings reveal three sources of downward bias in the research data file. First, retired women on farms are counted as not in the labor force, while retired men are counted as working. There is no clear justification for such differential treatment and since it was common for those who had handed over their farm to the children to help in the farm work, it would make sense to count these women as working. Second, many women are reported as the wife of a farmer, fisherman or other occupation (“farm wife”, “fisherman’s wife” etc) are all single compound words in the Norwegian language). These have systematically been counted as not working. Although there might be a stronger justification for treating men and women differently here (men are never reported as the husband of a female occupation), it is still likely that these women would on average have a stronger labor market attachment, such as helping in the production of market goods, than women without such annotation. Hence, a case could be made to also count these women as working. Finally, there are important nuances in the way housework is reported. Although more than 100,000 women are reported to be a “wife”, there is also a large population of married women with reported occupational activity related to housework. Treating all of these similarly — as not employed — could mask important heterogeneity. Again, a simple adjustment would be to count “housework” as work.

These three adjustments should be seen as upper bounds; while the baseline estimate of 39 per cent of all women and 10 per cent of married women working is likely too low, a comprehensive upward adjustment might give a too high rate. However, for the purpose of comparison, it is of interest to see how an implementation of such corrections would affect the reported female labor force participation rate.

To study the impact of the adjustments outlined above on female labor force participation, we now implement them on the 1910 data set. Of interest is the resulting aggregate female labor force participation, but also in the subgroup of married women and women with children. After all, micro-data on occupations is often used to study marital matching (the relationship between a husband’s and wife’s economic characteristics) and intergenerational mobility (the relationship between a parent’s economic position and that of their child).



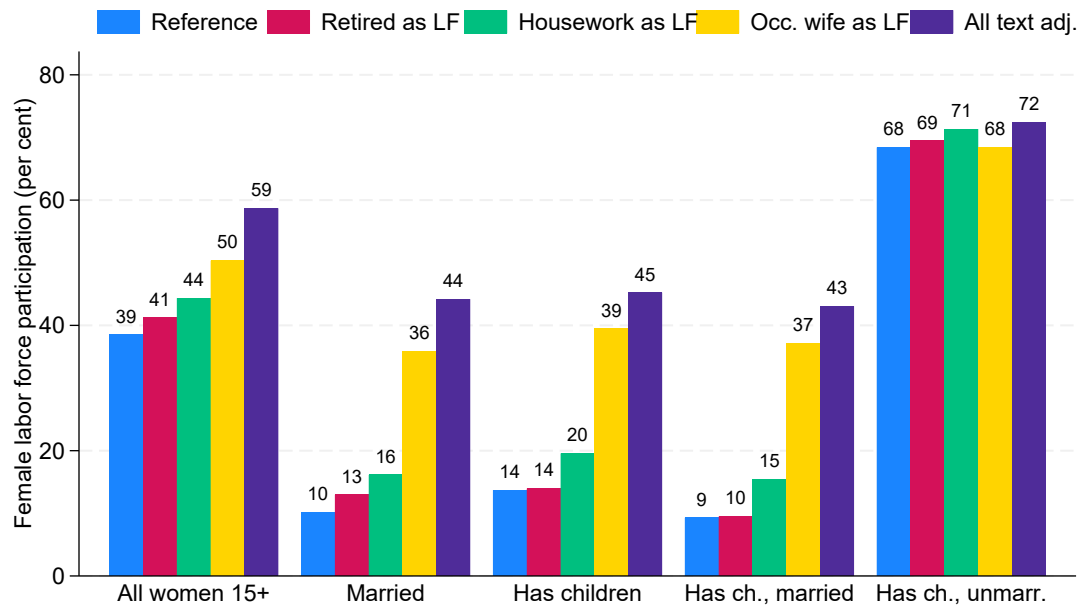


Figure 2: Share of women in the labor force, by demographic group and labor force definition. Note: “Reference” refers to the labor force participation variable as defined by IPUMS. “Retired as LF” re-codes retired women on farms as in the labor force (similar to how men are treated). “Housework as LF” re-codes occupation strings with some information on housework as in the labor force. “Farm wife as LF” re-codes occupation strings with information on husband’s occupation as in the labor force, if the husband is in an agricultural occupation (including fisheries). “All adj” applies all the three previous corrections at the same time.

Figure 2 shows estimates of female labor force participation in Norway in 1910, for different population subgroups and assumptions. The leftmost bar in each group is participation calculated directly from the IPUMS micro-data set – the share of women working. As already mentioned, this is 39 per cent for all women and 10 per cent for married women. If, rather than married women, we look at women who live together with at least one (own) child aged 15 or younger, the participation rate is 14 per cent. For parents, there is also a large discrepancy between the married and unmarried. 9 per cent of married mothers participate in the labor force, but 68 per cent of unmarried (single, divorced, or widowed) mothers do. Only seven per cent of mothers are unmarried.

We can now turn to an application of the three corrections to the data set and see how it affects participation. The second column in each group recodes retired women on farms (*kårfolk*) as participating, but this only has a limited effect of 0-3 percentage points for each of the subpopulations. Classifying those who state some kind of housework as working has a larger effect, five percentage points up from baseline for the reference population, and six percentage points for those with children. The largest effect is found when we reclassify the women with occupational information related to their husbands’ employment as working — even if we restrict this to women whose husbands work in agriculture or fisheries (HISCO major category 6). In this case, the aggregate participation rate is 50 per cent, while the participation rate of mothers is 39 per cent. The corrections for the most part apply to separate subgroups, so that applying all three of them at the same time (the rightmost column in each group) yields even higher participation rates: 59 per cent for the aggregate population 15 and older, 43 per cent for married mothers and 72 per cent for unmarried mothers.

What does this mean for historical labor market studies? The base rate of 39 per cent of all women is high enough that studies of occupational position or life-cycle mobility are clearly relevant. If one rather looks at intergenerational mobility, 9 per cent of mothers working is on the low side, but if there is some information in the text strings that could be used, even a modest increase to 15-30 per cent would make the data more informative. However, utilizing information on the husband’s occupation does have some limitations with respect to what can be analyzed. For example, knowing that a fisherman is more likely to be married to a “fisherman’s wife” does not add much to a study of occupational patterns in marital matching.

### 6.3 International comparisons

Is the estimate of a labor force participation of slightly less than forty per cent for women age 15 or older high, compared to other countries? This is challenging to answer, as there are few comprehensive studies to compare with. Some existing work takes a more extensive approach than this paper, reclassifying a large part of the population based on assumptions on household economic

activity. In this way, Ngai *et al.* (2024) obtain a participation rate for the United States in 1910 of around 54 per cent. A simplified application of their approach to the Norwegian data would yield a labor force participation rate of around sixty per cent (see Appendix, Section A.2).

Part of the data used in this paper is obtained from the IPUMS-International database (formerly the North Atlantic Population Project). Here, data is also available for other countries where censuses were conducted in a comparable way to this Norwegian census (e.g. by enumerator filling out census forms and reporting on all individuals). We can then calculate similar unadjusted participation rates for other countries and time periods. These are reported in Figure 3.<sup>23</sup>

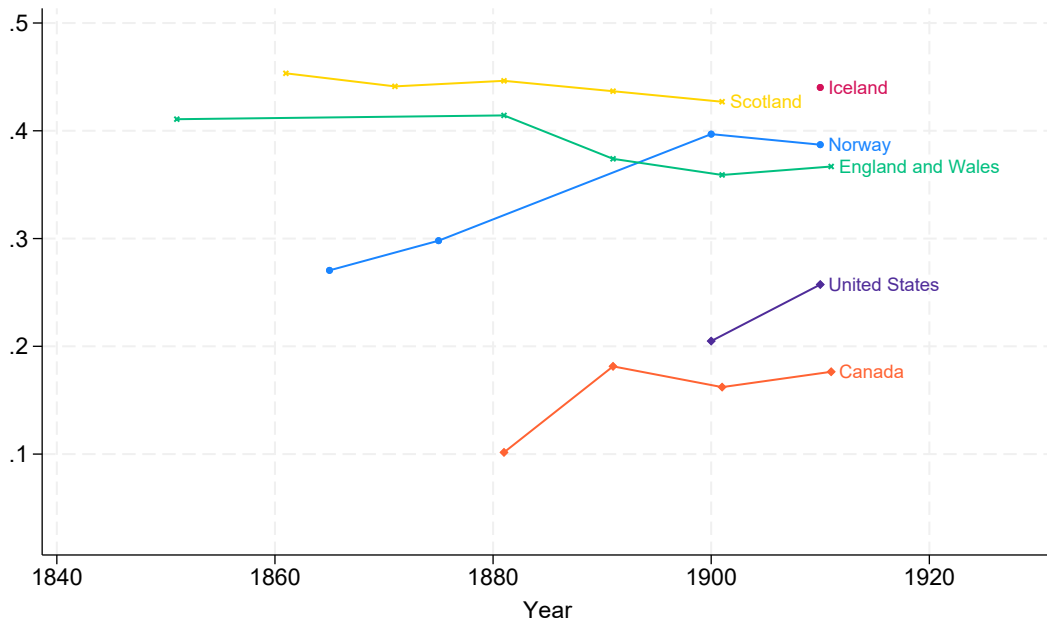


Figure 3: Share of women in the labor force (15 years and older), selected countries. Source: IPUMS micro data

For Norway, there are observations for 1865, 1875, 1900 and 1910. There is an increasing trend over the period, but a small decrease from 1900 to 1910. The comparison challenges between these years are acknowledged in the 1910 census publications. The United Kingdom (represented here by England/Wales and Scotland separately, as the censuses were conducted separately) have labor force participation rates comparable to Norway around the turn of the century, but not the same

<sup>23</sup>To ensure as high comparability as possible in the labor force definition across countries, only censuses after 1840, and only countries with a labor force participation of at least 80% for men 15 years and older are included in the figure. Samples with a high share of the population with an “unknown” value for labor force participation, or where there is no labor force participation variable at all, are also excluded.

increasing trends; these countries have high participation throughout. The observation for Iceland also gives a participation rate of more than forty per cent, while the United States and Canada have much lower participation rates.

Based on this rudimentary comparison, female labor force participation rates in Norway do not appear particularly high in a European context. With adjustments, the rate for Norway can be argued to be higher than 50%, but similar arguments could be made — and in some cases have been — for other countries. It is rather North America that appears as an outlier, with a labor force participation lower than the European countries. The United Kingdom had high industrialization rates; this could be one explanation for the high participation rates there. Although Norway had also entered more rapid industrialization in 1910, levels were still much lower than in other West European countries. Another candidate explanation for high participation rates could be emigration; Norway’s relative overseas emigration rates, mainly to the United States, were very high in a European context. For some cohorts, as many as 24 per cent of women and 36 per cent of men emigrated (Bævre *et al.*, 2001). Compared to the lower rates in the United States and Canada, countries with large immigration and large agricultural production in this period, this could suggest a yet unexplored relationship between female labor force participation and international migration in this period.

## 6.4 Concluding comments

The analysis presented here documents that the labor force participation in Norway in 1910 is well measured. The corresponding micro-data file, made available through IPUMS, is well suited for analysis using quantitative methods. There are some measurement issues stemming from inconsistent coding and/or unclear information in the original sources, and it is likely that the official rates are lower than what one would obtain using modern statistical measures. Future research should take into account the interplay between family relationship, household production, social status, and occupational activity that is present and often well documented in statistical sources from this period.

There is no inherent reason that future work on occupational trajectories, intergenerational mobility, matching in the labor market, and similar topics should be limited to the study of men, as has often been the case. Future research more explicitly including women in historical empirical labor economics will provide new insights both about the economy as a whole and about changes in gender roles over time.

## References

- Aaberge, Rolf, Atkinson, Anthony B., & Modalsli, Jørgen. 2020. Estimating long-run income inequality from mixed tabular data: Empirical evidence from Norway, 1875–2017. *Journal of Public Economics*, **187**(July), 104196.
- Abramitzky, Ran, Boustan, Leah, Eriksson, Katherine, Feigenbaum, James, & Pérez, Santiago. 2021. Automated Linking of Historical Data. *Journal of Economic Literature*, **59**(3), 865–918.
- Bævre, Kåre, Riis, Christian, & Thonstad, Tore. 2001. Norwegian cohort emigration. *Journal of Population Economics*, **14**(3), 473–489.
- Bailey, Martha J., Cole, Connor, Henderson, Morgan, & Massey, Catherine. 2020. How Well Do Automated Linking Methods Perform? Lessons from US Historical Data. *Journal of Economic Literature*, **58**(4), 997–1044.
- Balsvik, Randi Rønning. 1991. Kvinner i nordnorske kystsamfunn. *Historisk tidsskrift*, **70**(4), 636–655.
- Blom, Ida, Sogner, Sølvi, Hagemann, Gro, Melby, Kari, Sandvik, Hilde, & Øye, Ingvild. 2005. *Med kjønnsperspektiv på norsk historie*. Oslo: Cappelen Akademisk.
- Burnette, Joyce. 2004. The wages and employment of female day-labourers in English agriculture, 1740–1850. *The Economic History Review*, **57**(4), 664–690.
- Burnette, Joyce. 2021. Why we shouldn’t measure women’s labour force participation in pre-industrial countries. *Economic History of Developing Regions*, **36**(3), 422–427.
- Chiswick, Barry R., & Robinson, RaeAnn Halenda. 2021. Women at work in the United States since 1860: An analysis of unreported family workers. *Explorations in Economic History*, **82**(Oct.), 101406.
- Dalman, Elien. 2026. The impact of mothers. Intergenerational mobility in Sweden 1865–2015. *European Sociological Review*, **forthc.**
- Delabastita, Vincent, & Buyst, Erik. 2021. Intergenerational mobility of sons and daughters: evidence from nineteenth-century West Flanders. *European Review of Economic History*, **25**(2), 300–327.
- Digitalarkivet. 2024. *Folketellingen 1910*. <https://www.digitalarkivet.no/content/125/folketellingen-1910>.
- Fernandez, R., Fogli, A., & Olivetti, C. 2004. Mothers and Sons: Preference Formation and Female Labor Force Dynamics. *The Quarterly Journal of Economics*, **119**(4), 1249–1299.

- Goldin, Claudia. 1979. Household and market production of families in a late nineteenth century American city. *Explorations in Economic History*, **16**(2), 111–131.
- Goldin, Claudia. 1980. The Work and Wages of Single Women, 1870 to 1920. *The Journal of Economic History*, **40**(1), 81–88.
- Goldin, Claudia. 1990. *Understanding the Gender Gap: An Economic History of American Women*. New York Oxford: Oxford University Press.
- Horrell, Sara, & Humphries, Jane. 1995. Women’s Labour Force Participation and the Transition to the Male-Breadwinner Family, 1790-1865. *The Economic History Review*, **48**(1), 89–117.
- Humphries, Jane, & Sarasúa, Carmen. 2012. Off the Record: Reconstructing Women’s Labor Force Participation in the European Past. *Feminist Economics*, **18**(4), 39–67.
- ILO. 2013. *ILO estimates and projections of the economically active population: 1990-2030 : methodological description*. ILO.
- Lie, Einar, & Roll-Hansen, Hege. 2001. *Faktisk talt. Statistikkens historie i Norge*. Oslo: Universitetsforlaget.
- Maas, Ineke, & Van Leeuwen, Marco H. D. 2023. Modernization and Long-Term Trends in Social Mobility. In: Gangl, Markus, Platt, Lucinda, Polavieja, Javier G., & Van De Werfhorst, Herman G. (eds), *The Oxford Handbook of Social Stratification*, 1 edn. Oxford University Press.
- Minnesota Population Center. 2020. *Integrated Public Use Microdata Series, International: Norway, 1910*.
- Modalsli, Jørgen. 2017. Intergenerational Mobility in Norway, 1865–2011. *The Scandinavian Journal of Economics*, **119**(1), 34–71.
- Molinder, Jakob. 2022. Historical roots of the dual-earner model: Women’s labour force participation in Sweden, 1870–1960. *Lund Papers in Economic History*, Dec.
- Ngai, L Rachel, Olivetti, Claudia, & Petrongolo, Barbara. 2024. Gendered change: 150 years of transformation in US hours. *working paper*, Nov.
- Olivetti, Claudia. 2014. The Female Labor Force and Long-Run Development: The American Experience in Comparative Perspective. *NBER Chapters*, 161–197.
- Olivetti, Claudia, & Paserman, M. Daniele. 2015. In the Name of the Son (and the Daughter): Intergenerational Mobility in the United States, 1850-1940. *American Economic Review*, **105**(8), 2695–2724.

- Registreringsentralen for historiske data. 2014. *Kodedokumentasjon for folketellingene 1865 til 1910*. Tech. rept. Internal document. RHD.
- Ruggles, Steven. 2015. Patriarchy, Power, and Pay: The Transformation of American Families, 1800–2015. *Demography*, **52**(6), 1797–1823.
- Stanfors, Maria. 2014. Women in a changing economy: the misleading tale of participation rates in a historical perspective. *The History of the Family*, **19**(4), 513–536.
- Statistics Norway. 1911. *Fabriktællingen i Kongeriket Norge, 1909. Første hefte: Bedrifter, arbeidere og eiere*. Tech. rept. NOS V 147. Oslo.
- Statistics Norway. 1912. *Industristatistikk. Omfattende opgaver over ulykkesforsikringspligtige bedrifter og arbeidere for aaret 1910*. Tech. rept. NOS V 183. Oslo.
- Statistics Norway. 1913. *Folketællingen i Norge 1 December 1910. Fjerde hefte: Folkemængde fordelt efter livsstilling*. Tech. rept. NOS V 211. Oslo.
- Statistics Norway. 1914. *Folketællingen i Norge 1 december 1910. Sjette hefte. Folkemængde fordelt efter livsstilling med angivelse av alder og egteskabelig stilling*. Tech. rept. NOS VI 20. Oslo.
- Statistics Norway. 1915. *Indtægts- og formuesforhold efter skatteligningen 1911 i forbindelse med Folketællingen 1910*. Tech. rept. NOS VI 24. Oslo.
- Statistics Norway. 1916. *Folketællingen i Norge 1. december 1910. Hovedoversigt*. Tech. rept. NOS VI 77. Oslo.
- Thorvaldsen, Gunnar. 2018. *Censuses and census takers: a global history*. Routledge studies in modern history, no. 35. Abingdon: Routledge, an imprint of the Taylor & Francis Group.
- van Leeuwen, Marco, Maas, Ineke, & Miles, Andrew. 2002. *HISCO Historical International Standard Classification of Occupations*. Leuven University Press.
- Van Nederveen Meerkerk, Elise, & Paping, Richard. 2014. Beyond the census. Reconstructing Dutch women’s labour market participation in agriculture in the Netherlands, ca. 1830–1910. *The History of the Family*, **19**(4), 447–468.
- You, Xuesheng. 2020a. Women’s labour force participation in nineteenth-century England and Wales: evidence from the 1881 census enumerators’ books. *The Economic History Review*, **73**(1), 106–133.
- You, Xuesheng. 2020b. Working with Husband? “Occupation’s Wife” and Married Women’s Employment in the Censuses in England and Wales between 1851 and 1911. *Social Science History*, **44**(4), 585–613.

## A Appendix

### A.1 Enumerator instructions, 1910 Census

Instructions for field 9 (Occupation), author's translation, from Statistics Norway (1916) p. 155 (instructions for field 9a and 9b) and p. 156 (remark 7).

**Instructions for field 9a, Employment and social status.** Includes the employment of housewives and children. State clearly and specifically industry or trade that the person exercises or is employed in, and so that the individuals' position in the occupation can be seen (e.g. tenant farmer; shoemakers' apprentice; cellulose worker). If an individual has multiple occupations, list the primary occupation first. See also remark 7.

(9b.) If unemployed at the time of enumeration, write here the letter: l.

**Remark 7.** The type of industry or employment has to be clearly and specifically stated. For adult children at home, or other relatives and servants, it should be stated whether they are employed in housework, agricultural work, care of livestock or other types of work, and if so, what type of work. For widows and adult unmarried women it should be stated whether they live off savings or conduct any trade, such as sewing, small-scale retail, boardinghouses or similar.

For lodgers and visitors the employment should similarly be given.

For artisans and others in similar industries should be stated what kind of industry they are in; for example, it is not sufficient to write craftsman, factory owner, factory manager, etc; one must specify shoemaker master; brickworks owner; manager of sawmill, etc.

For clerks, office workers, supervisors, machinists, firemen, etc., specify the type of enterprise they are employed in.

For workers, tenants, and day laborers, specify the enterprise where they are employed at the time of the census or where they were regularly employed before, e.g., in agriculture, sawmills, wood pulp mills, dock work, etc. In every business, the position should be described in such a way that it is clear whether the person operates the business as an employer, is self-employed, or works in someone else's service as a manager, clerk, foreman, journeyman, apprentice, or worker.

As unemployed (l), those are counted who, at the time of the census, were without work (not due to illness, disability, or labor disputes) but who would otherwise usually be employed or in a similar position.



For all such positions that may be both private and public, the nature of the employment must be specified (e.g., civil servant, public employee in the service of the state or municipality, teacher at a private school, etc.).

If a person mainly lives off wealth, a pension, an annuity, or private or public support, this should be noted, but their employment should also be recorded if it is of any significance.

For former business owners, civil servants, etc., "fv" (formerly) should be written before their previous employment.

**Original text** (Norwegian, spelling as of 1910):

9 a Erhverv og livsstilling. Ogsaa husmors eller barns særlige erhverv. Angi tydelig og specielt næringsvei eller fag, som vedkommende person utøver eller arbeider i, og saaledes at vedkommendes stilling i erhvervet kan sees (f. eks. forpagter, skomakersvend, cellulosearbeider). Dersom nogen har flere erhverv, anføres disse, hovederhvervet først. (Se forøvrig bemerkning 7).

9 b Hvis arbeidsledig paa tællingstiden, sættes her bokstaven: I.

Bemerkning 7:

Næringsveiens eller erhvervets art maa tydelig og specielt betegnes.

For hjemmeværende voksne barn eller andre paarørende samt for tjenere oplyses, hvorvidt de er sysselsat med husgjerning, jordbruksarbeide, kreaturstel eller andet slags arbeide, og i tilfælde hvilket. For enker og voksne ugifte kvinder maa anføres, om de lever av sine midler eller driver nogen slags næring, saasom søm, smaahandel, pensionat o. l.

For losjerende eller besøkende maa likeledes næringsveien opgives.

For haandverkere og andre industridrivende m. v. maa anføres, hvad slags industri de driver; det er f. eks. ikke nok at sette haandverker, fabrikeier, fabrikkbestyrer o. s. v.; der maa sættes skomakermester, teglverkseier, sagbruksbestyrer o.s.v.

For fuldmægtiger, kontorister, opsynsmænd, maskinister, fyrbøtere o. s. v. maa anføres, ved hvilket slags bedrift de er ansat.

For arbeidere, inderster og dagarbeidere tilføies den bedrift, ved hvilken de ved optællingen har arbeide eller forut for denne jevnlig hadde sit arbeide, f. eks. ved jordbruk, sagbruk, træsliperi, bryggearbeide o.s.v. Ved enhver virksomhet maa stillingen betegnes saaledes, at det kan sees, om vedkommende driver virksomheten som arbeidsgiver, som selvstændig arbeidende for egen regning, eller om han arbeider i andres tjeneste som bestyrer, betjent, formand, svend, lærling eller arbeider.

Som arbeidsledig (1) regnes de, som paa tællingstiden var uten arbeide (uten at dette skyldes sygdom, arbeidsudygtighet eller arbeidskonflikt), men som ellers sedvanligvis er i arbeide eller i anden underordnet stilling.

Ved alle saadanne stillinger, som baade kan være private og offentlige, maa forholdets beskaffenhed angives (f. eks. embedsmand, bestillingsmand i Statens, kommunens tjeneste, lærer ved privat skole o. s. v.).

Lever man hovedsagelig av formue, pension, livrente, privat eller offentlig understøttelse, anføres dette, men tillike erhvervet, om det er av nogen betydning. Ved forhenværende næringsdrivende, embedsmænd o.s.v. sættes "fv" foran tidligere livsstillings navn.

## A.2 Supplementary tables

This section provides a summary of labor force participation rates calculated from individual-level micro data, for different subpopulations and population definitions.

Participation rates as discussed in the main text are, as is conventional, given for individuals 15 years and older. This means that we include individuals too young to have started in the labor force (in particular, if they come from affluent families) and individuals too old to work. For this reason, it is of interest to reexamine labor force participation rates for more narrow age definitions. Rows 2-6 of Table A1 show participation rates for more narrow age groups. Increasing the lower cut-off to 20 or 25 decreases the measured participation rate, indicating that occupational activity is higher than average in the youngest age brackets. However, exclusion of individuals above a certain age threshold increases the average participation rate. The highest rate is found when we limit the population to individuals between 15 and 54 years of age. Of these, 42.1 per cent work. Changing the population definition to resident rather than present population gives only a minor change in measured labor force participation (row 7).

As expected, participation rates for married women are much lower than for unmarried women (rows 8-9). Based on the micro-data, married women have a labor force participation rate of 10.2 per cent, while nearly two thirds — 62.8 per cent — of unmarried women work. Those living together with their own children (as indicated on census records) have an average labor force participation of 13.7 per cent, however there is a large discrepancy between married and unmarried women also when we compare those who have children (rows 10-12). Only about 19,000 women report being unmarried (including widowed) and living with at least one young child, but of these, 68.4 % participate in the labor force.

The next four rows (13-16) of Table A1 report the impact of making the adjustments described in Section 3. All of these increase the measured female participation in the labor force, but to varying degrees. Recoding retired women in the IPUMS data set to be considered in the same way as men — that is, as being in the labor force — increases the participation rate from 38.5% to 41.2%.

	Female labor force participation rate	Pop. size
(1) Reference (population 15 years and older)	38.5%	819,640
Adjustments to population definition:		
(2) Population 20+	35.5%	703,239
(3) Population 25+	31.8%	605,191
(4) Population 15-75	39.4%	784,453
(5) Population 15-66	40.3%	734,940
(6) Population 15-54	42.1%	628,738
(7) Resident population (instead of present population) 15+	38.7%	823,653
Population (15+) by marital status		
(8) Married	10.2%	377,956
(9) Not married	62.8%	331,684
Children in household (population 15+)		
(10) Living with own children younger than 15	13.7%	257,813
(11) Living with own children < 15, married	9.3%	238,697
(12) Living with own children < 15, unmarried	68.4%	19,116
Re-defining labor force participation rate (population 15+)		
(13) Recoding retired women into the labor force (similar to men)	41.2%	819,640
(14) Recoding "occupation's wife" into the labor force	50.4%	819,640
(15) Recoding housework (husgjerning) into the labor force	40.0%	819,640
(16) All of the above (13, 14 and 15)	58.7%	819,640
Extensive re-classification of household members (population 15+)		
(17) Counting as in labor force all household members of farmers and business owners	60.6%	819,640
(18) Labor force re-definition (16) as well as extensive re-classification (17)	67.5%	819,640

Table A1: Female labor force participation rate, with alternative definitions of population and participation rate. Norway, 1910, present population (except line 7 resident population). Note: "Not married" includes widowed and divorced. Source: IPUMS micro data. Adjustments (lower three lines) also calculated from Norwegian National Archives micro data.

Instead counting women described as the wife of some occupation (when the husband is a farmer or a fisherman — in HISCO major category 6) as working increases the overall participation rate to more than half, 50.4%, while recoding those reporting housework as working gives a participation rate of 44.3%. As discussed above, the latter two corrections are likely to yield slight overestimates. The aggregate increase of applying all three corrections gives a calculated female labor force participation rate of 58.7% for the population as a whole.

Finally, we can do a more extensive recode, as in Ngai *et al.* (2024). A simplified replication of their approach, recoding family members (including wives) on farms and those living with business owners (those with “mester” - meaning “master” - in their occupation title) gives a labor force participation rate of 60.6% (row 17). Combining this with the adjustments in the previous paragraph would give a total rate of 67.5%.

## A.3 Record matching

### A.3.1 Combining two micro data sets

This section details the merge of the data sets for the 1910 Norwegian census from the National Archives and IPUMS. The merge is necessary because not all relevant information exists in both data sets and the person records do not have 1:1 correspondence.

Both data sets were accessed in March 2023. The National Archives data set was transferred directly from the National Archives, while the IPUMS data set was downloaded from IPUMS’ data portal. The National Archives data set has 2,475,248 person records, while the IPUMS data set has 2,468,823 records. The difference likely reflects that some of the records in the National Archives data set contains very little information, while the IPUMS records appear to be restricted to those with at least some relevant information. The IPUMS data set originates from the National Archives, but was published around 15 years ago. For this reason, later adjustments and corrections to the file by the National Archives are not reflected in IPUMS, and IPUMS also processes the files they receive from their partners.

In order to match the records, the files were first processed in the following way:

- Adjust inconsistent municipality numbers in the IPUMS file (1017 Songdalen to 1018 Søgne; 1699 unknown municipality Sør-Trøndelag to 1665 Tydalen; 2099 unknown municipality Finnmark to 2026 Polmak).
- Construct a birth date variable that has the same format in both files (yyyy.mm.dd)
- Calculate an ID number within each municipality that is based on the person records’ order in that file (the first person is number 1, the second is number 2, and so on).

- For Kristiania (Oslo), containing 10 per cent of Norway’s population at the time, the ID variable was instead generated within each enumeration district. Enumeration district numbers and sizes match when IPUMS enumeration district numbers are divided by 100. Some enumeration districts were merged as sizes were inconsistent across the sources (first: 021 Johannes menighet and 121 Grønland menighet, second: 038 Trefoldighet menighet, 160 Fesningingen and some IPUMS records with inconsistent/missing district numbers (9999 or less than 10)).

The key matching units are hence 659 municipalities, of which one (Kristiania) is subdivided into 89 enumeration districts, for a total of 747 regional units. Although there are enumeration districts for all municipalities in both the IPUMS and National Archives files, the numbering and subdivision is generally not perfectly consistent across the files and it was not found cost-effective to manually align all 6,013 enumeration districts.

65 per cent of person records are unique when only the birth date and the regional unit is considered. Using these individuals as a basis, the rest of the individuals are then matched on the basis of their order in each of the original files. If an individual fills a gap in the list of records both when considering the ID number of the previous person, plus 1, as well as the ID of the next person, minus 1, they are considered matched (except in very few cases when the ID number is already matched elsewhere in the file). The key feature allowing this match is that while municipalities and enumeration districts are not necessarily similarly sorted across the files, individuals within districts mostly are.

This procedure gives a set of records that each contain information from one National Archives record and one IPUMS record. The reliability of the matches are then assessed by comparing characteristics that exist in both files (birth dates, sex, enumeration district, household position) and are found to be reliable.

A total of 2,460,948 records are matched and used as the analysis sample in this paper. This constitutes 99.42 per cent of the National Archive records and 99.68 per cent of the IPUMS records. The unmatched individuals are geographically spread across most municipalities, and relatively equally balanced with respect to sex.

### **A.3.2 Individual record matching**

This section explains how the “absent resident” and “present non-resident” records are merged, as the basis for the analysis in Section 5. Individuals who were away from home were enumerated twice, and the records tagged to belong to one of the two categories, but not with any attempt at denoting which records refer to the same individuals.

The approach follows the matching procedure used in Modalsli (2017) with some adjustments. The following information is used:

**First name.** First names are converted to lower case, and the letter “c” is converted to “k” (as these letters are not being used consistently in spelling of names in the data).

**Place of birth.** Place of birth is coded with municipality numbers in the IPUMS file. These codes are used.

**Time of birth.** Year, month, and date.

**Reference to home location.** Each of the records denoting a present non-resident has a field giving this person’s home location. This location is not in the IPUMS file and is hence not converted to a municipality code. Locations can refer to something other than a municipality and may not be standardized. In our case, we use the text strings from the place of birth field to map the strings to municipality codes.

**Reference to present location.** Similarly, each of the records denoting an absent resident has a field denoting where the person currently is, and this is converted to a municipality code using the mapping from place of birth.

Last names (family names) are not used as they are sometimes used inconsistently for both women and men, but even more so for married women (i.e. sometimes maiden names would be used and sometimes not).

A score is then constructed for each pair of records in the data set. To reduce the comparison space, matches are only considered between records that have (a) the same first letter initial (c and k grouped together) and (b) the same birth year and (c) the same sex. The weights given to each category (i.e. the scaling of the score) reflect a desire for similar weighting for each characteristic as well as a trade-off between too many and too few matches and are found by trial and error. The scores are computed as follows:

- First names: The Levensthein distance between the two strings. Only comparisons with a distance of 2 or less are considered.
- Birth month and birth date: The difference between the number for each of the two, if both are nonmissing. If one of the records have a missing date or month, the difference is set to zero. The difference is multiplied by 0.1 for month and 0.05 for day
- Birth municipality: Set to 0 if equal, 1 if different
- Comparison of “home municipality” to municipality of resident record: Set to 0 if equal municipality code, 0.5 if different
- Comparison of “away municipality” to municipality of present non-resident record: Set to 0 if equal municipality code, 0.5 if different

This gives a total of five numbers, which are then added together to obtain a difference between record  $i$  (of a present non-resident) and record  $j$  (of an absent resident). A lower score reflects a better match (more similar characteristics across the two sets of records).

A match  $(i, j)$  is accepted if it is better than the second best match for both records  $i$  and  $j$ . This includes both unique matches (where record  $i$  only matches record  $j$  and vice versa) and matches where there are other options for record  $i$  or  $j$  or both, but they are further away by the metric defined above.

Matches are then verified by observing the names and other characteristics that are present in both files.

## EHES Working Paper Series

### Recent EHES Working Papers

- EHES 281** Colonial Governance and Resource Allocation in the British West Indies (1838–1938)  
*Luisito Bertinelli, Fabio Gatti, Eric Strobl*
- EHES 280** Why railways fail: Colonial railways and economic development in Habsburg Bosnia-Herzegovina  
*Magnus Neubert, Stefan Nikolić*
- EHES 279** Text Analysis Methods for Historical Letters, The case of Michelangelo Buonarroti  
*Fabio Gatti, Joel Huesler*
- EHES 278** The Rákóczi family's estate wine production, wine trade, and wine prices at the end of the 17th century  
*Attila Ulrich*
- EHES 277** Transport and Biological Living Standards: Evidence from Late 19th-Century Catalonia  
*Ramon Ramon-Muñoz, Guillermo Esteban-Oliver, Josep-Maria Ramon-Muñoz*
- EHES 276** Climate adaptation in early Australian wheat farming (1860–1960)  
*Costanza Maria Fileccia, Eric Strobl*
- EHES 275** The Myth of Nordic Mobility: Social Mobility Rates in Modern Denmark and Sweden  
*Gregory Clark, Martin Hørlyk Kristensen*
- EHES 274** Teacher-directed scientific change: The case of the English Scientific Revolution  
*Julius Koschnick*
- EHES 273** The mortality impact of cholera in Germany  
*Kalle Kappner*

All papers may be downloaded free of charge from: <https://www.ehes.org/>

The European Historical Economics Society is concerned with advancing education in European economic history through study of European economies and economic history. The society is registered with the Charity Commissioners of England and Wales number: 1052680.