LIS
Working Paper Series

No. 773

Women in the Top of the Income Distribution – What Can We Learn From LIS-Data?

Roman Bobilev, Anne Boschini, Jesper Roine

September 2019

Women in the top of the income distribution –
What can we learn from LIS-data?[a]

Roman Bobilev[*]
Anne Boschini[♣]
Jesper Roine[♦]

July, 2019

Abstract

We explore the extent to which LIS-data can be used to shed light on the presence of women in the top of the income distribution. We show developments of the share of women in top groups (P90-100 and P99-100) of the labour income distribution for 28 countries and, when possible, compare to outcomes when including capital incomes. These turn out not to matter much for the share of women in top groups with some important exceptions. Relating our findings to the existing evidence on women in the top of the income distribution based on aggregate tax data, we find that LIS-data give a relatively accurate picture of the basic findings. However, we also note that once we divide the top1 group further, samples quickly become too small to allow further study. For countries where data allows such analysis, we find that having a partner and having children are positively associated with being in top income groups for men, but negatively associated for women. However, time interactions suggest that these differences have decreased over time. Also, top income men are more likely to have partners who are not in the top of the income distribution while this is not the case for top income women. All these results are surprisingly consistent across country groups.

Keywords: Income inequality, income distribution, top income women, gender inequality, top incomes, capital incomes,
JEL: D13; D31; H20; J16; J31

1 Introduction

In recent years the so called top-income literature has received a lot of attention. Following the seminal work by Piketty (2001, 2003), Piketty and Saez (2003), and Atkinson (2005) a large number of studies have shown how important different aspects of the very top of the income distribution are for fully understanding both the recent increase in inequality observed in many countries, as well as for its long-run evolution.[1] In particular, this literature has stressed the importance of looking carefully at developments *within the top*, and also the importance of including *all sources of income*. For example, the developments of the top1 group are often very different from the rest of the top decile, and a key factor in explaining this difference often turns out to be the role of non-labour incomes.

One very basic aspect of top incomes has, however, received little attention in this literature, namely the gender dimension. If we observe a growing share of total incomes going to top groups, and also that top incomes are different in terms of income composition, it seems natural to also ask questions about the gender composition of this group and to what extent important dimensions of top incomes are different for men and women. What share of the top ten or top one group is made up of women? How has this changed over time? Are top income women and men similar in terms of income composition and in terms of observable characteristics and has this changed over time?

In this paper we will explore the extent to which LIS-data can be used to shed light on questions like these. We proceed as follows: we first explain why the gender aspect has not received more attention in the standard top income literature and also relate to how the gender dimension in this literature differs from the extensive work on different aspects of gender inequality in related literatures. We discuss how we select countries and years in the LIS data given the limits that especially sample size puts on the questions when focusing on groups as small as the top one percent of the distribution.[2] We then present basic results on the share of women in the top ten and top one group, for labour income, and whenever possible, compare

---

[1] The collected volumes by Atkinson and Piketty (2007, 2010) contain much of the early work and Leigh (2009), Atkinson, Piketty and Saez (2011), Alvaredo, Atkinson, Piketty and Saez (2013), and Roine and Waldenström (2015) provide overviews of this literature. Data is available from the top income database at http://www.wid.world.

[2] LIS-data has obviously been used extensively to study long-run, cross-country aspects of gender equality in the population more broadly, recently in e.g. Kleven and Landais (2017). Our focus here is different in that we want to explicitly connect to the top income literature, which creates special challenges that we will discuss more below.

the outcomes to those for labour plus capital income (as well as the total income as defined in the LIS data), for as many countries and years as possible. When possible, we also compare these results to the few recent studies in the top income literature that use individual tax data to study the share of women in top incomes (in particular, Akinson, Casarico and Voitchovsky, 2018, and Boschini, Gunnarsson and Roine, 2017). Comparissons suggest that our LIS-based results, though not perfectly aligned, in general come very close to the shares of women in top groups observed in comprehensive income tax data, but we note also that there are some important caveats to this overall conclusion.

We then move on to using the richness of the LIS-data to study determinants of the probability that men and women respectively are found in the top of the income distribution. In particular, we study gender differences in some family variables such as children and the effect of having a partner, as well as some characteristics of that partner. We find that having children is associated with a lower likelihood of women being in the top of the income distribution, but with the opposite being the case for men (and with the gap generally increasing in the number of children). We also find that top income women are less likely to have a partner, but conditional on having a partner, it is likely that this partner is also a high-income earner. Top income men, on the other hand, are more likely to have a partner, and this partner is more likely not to be in the top of the income distribution. These associations are surprisingly similar across different country groups. Overall, the results underline the importance of understanding the combination of individual characteristics as well as family circumstances when thinking about what determines the relative presence of men and women in the top of the income distribution. This section, however, also serves as an illustration of when we quickly approach the limit of what can be meaningfully studied in the top of the distribution using LIS-data due to sample sizes becoming too small. We conclude with some remarks on what we find and suggestions for future research.

1.1 Why the delay in studying women in the top income literature?

Given the great interest in top incomes, one can wonder why the gender dimension has not received more attention earlier. The main reason lies in the fact that the unit of analysis in the top income literature has been determined by the availability of historical tax data, which for most countries has meant that married couples count as one unit making the division between

(married) men and women problematic.[3] While many countries have, at some point, switched from household based to individual based taxation, the treatment of married couples as one unit continuous to be the case in many countries still today. When it comes to answering the main question posed in this literature – What share of all income is earned by some top group? – household based taxation creates a problem (albeit with clear boundaries, see Atkinson 2007, p. 28-29, for an explanation of how to calculate these boundaries), but when it comes to studying the share of women in the top groups this in not possible without more detailed data about individual incomes.

This explains why the few recent papers, that have begun to answer questions about the presence of women in the top of the distribution, have focused on countries and time periods when men and women file taxes independently. Atkinson, Casarico, and Voitchovsky (2018) study the share of women in top income groups, as well as differences in income composition, in eight countries with independent taxation for men and women. They follow the methodology of the top income literature, in terms of defining the reference total for income and population, but then look separately at men and women in the different top groups. Piketty, Saez and Zucman (2016) report results for the share of women in top groups in the U.S. since the early 1960s. However, given joint tax filing for married couples, they are restricted to differences stemming from labour earnings. Similarily, Garbinti, Goupille-Lebret, and Piketty (2018) report the share of women in top fractiles of labour income in France starting in 1970. Boschini, Gunnarsson and Roine (2017) study gender aspects of top incomes in the case of Sweden starting in the early 1970s when Sweden switched to independent taxation of men and women. Having access to panel data they are able to also study aspects of gender differences in top income mobility, individual characteristics and family structure of top income men and women respectively. Ravaska (2018) studies similar questions in the case of Finland starting in 1995.

These papers show that there are several important gender dimensions to top incomes. A first insight is that the share of women in the top of the distribution (at least in the countries studied) has grown steadily since at least the 1970s. However, they also show that the top is still far from equal in terms of gender. The share of women in the top10 group has roughly doubled since the 1970s but only to reach around 30 percent, and the higher up in the distribution we move, the lower is generally the share of women. Another important insight is

---

[3] The key initial motivation for the top income literature was "a general dissatisfaction with existing income distribution databases", especially in terms of the lack of long run series and lack of decomposition of inequality into a labour and capital income component (e.g., Piketty, 2007, p. 1), hence the focus on historically available tax data.

that income compositions typically differ across genders. In several countries, other types of income than labour income, in particular capital income, make up a larger share of women's income as compared to men's. The time trend, however, is that the income compositions of men and women have mostly been converging. Using individual panel data, Boschini et al. (2017) also show that there are some differences in mobility (women are more likely to fall out of the top) and in some aspects of family characteristics. In particular, they show that the largest difference between top income men and women is not in terms of age or education,, but in terms of marital status and partner income. While as many women as men in the top1 group nowadays have a partner, this was not the case a couple of decades ago. In the 1970s, almost all Swedish top1 men were married, in contrast to less than half of the top1 women. Moreover, among those top income individuals having a partner, most of the married top income women in Sweden are married to a man who also has a high income (and virtually none are married to men with low income). For top income men the reverse is true: most married top income men have a wife who is not a top income earner.

Together these papers suggest, first, that there are indeed interesting gender related developments in the top of the income distribution and that these are related to aspects highlighted in the top income literature, such as within top group differences and income composition. Second, the findings in Boschini et al. (2017) suggest that some gender differences in top incomes become apparent only when looking at longitudinal data and individual as well as family characteristics of top earners.

1.2   What does it mean to study "top income women"?

The fact of there being little research on gender aspects in the top income literature does not, of course, mean that there has been a general lack of interest in gender differences in the top of the income distribution. On the contrary, some of the most well-known results in gender economics, such as gender differences in executive compensation (e.g., Bertrand, Goldin and Katz, 2010; Smith, Smith and Verner, 2013; Keloharju, Knupfer and Tåg, 2016) and the so-called "glass-ceiling" results (Albrecht, Björklund and Vroman, 2003; Arulampalam, Booth and Bryan, 2006; Albrecht, Skogman Thoursie, and Vroman, 2015) are explicitly about gender differences in the top. Recent work by Guvenen, Kaplan, and Song (2014) study gender dimensions of top wage earners in the US, 1981-2014, and in an overview Marianne Bertrand (2018) summarizes the state of current knowledge in her introduction: "Despite decades of

progress, women remain underrepresented in the upper part of the earnings distribution, a phenomenon often referred to as the 'glass ceiling'."

It is important to note that these studies are primarily concerned with the top of the *earnings* distribution. In general, the focus in this literature has been on labour market outcomes, hence, excluding capital incomes, which are known to be important especially in the top of the income distribution. Moreover, the population is also typically restricted to the working age population and comparisons of wage gaps are usually made conditional on working full time. Often, when the focus is on detecting potential discrimination it is also natural to control for individual characteristics and sector, etc. For many questions these restrictions are, of course, perfectly sound and even necessary, but for others we may instead want to know the actual total income (from all sources) regardless of the choices underlying the outcome (such as labour supply) and without restrictions on the population. And importantly, gender dimensions may not be the same across these different comparisons.

The bottom line is that these different research strands illustrate how studying the role of "women in top incomes" can clearly mean many different things depending on the questions asked, and how these may overlap when thinking about gender differences in total income. In this paper, as in the top income literature, we focus on women in the top of the individual income distribution.[4] As far as possible, we include income from all sources before taxes and transfers. The reference population is ideally the full adult population (18+). Since we only focus on the gender composition of different top groups, and not on the income share of these groups, the reference total for income will be of less importance but in principle we would like this to be all incomes.

2   The LIS data and its relation to top income data

Recent work by Yonzan, Milanovic, Morelli, and Gornick (2018) compares the coverage of top incomes in LIS to fiscal data (i.e. data from tax returns) used in the World Inequality Database (WID), for the entire tax population, which depending on the tax code of a country can be either households or individuals. Their study focuses on the U.S., but also contrasts results for Germany and France. They note that LIS allows them to match the (total) income concept used in WID and based on this they compare income shares of different top groups, as well as the

---

[4] To be precise, the unit of analysis in the top income literature is typically the tax unit, which in many cases is singles and married couples. For this reason, the top income studies concerned with women in the top have been restricted to countries where the individual is also the tax unit.

different components (basically pre-tax labour, capital and business income). Their preliminary conclusion is that LIS and WID seem to give very similar answers up to the top1 group, but beyond this LIS seems to underestimate the total income accruing to the top group. This is mainly due to missing non-labour income in the top1 group. The finding confirms what has been noted before in the overview of the top income literature by Atkinson, Piketty and Saez (2011) and studied in more detail by, e.g., Burkhauser, Feng, Jenkins and Larrimore (2012) for the U.S., and in Burkhauser, Hahn, and Wilkins (2016) for Australia, namely that survey data tends to underestimate the incomes in the very top.[5]

Our challenges are slightly different. We want to study the development of the share of women in top groups, with the top group being defined as the top of the individual distribution of total income from all sources (labour, business income, and capital) over the full adult population. As mentioned in the introduction this has previously been done in a few papers using the methodology and data typically used in the top income literature. Our aim is to see to what extent we can use LIS data to study this question and how our results relate to findings in these papers.

2.1 Restrictions used to select data in LIS

An obvious first consideration, well known from previous discussions about the pros and cons with survey data when studying top income share, has to do with requirements on the sample size to get reliable estimates, and related to this, the problem with top coding. LIS data are not top coded but the sample size varies making this a first limitation in terms of how many waves and countries can be included.[6] To illustrate the challenge of studying women in top groups using LIS data, consider the following back-of-the-envelope calculation: For a sample size of 50 000 individuals – a relatively large sample in LIS – the top1 percent consists of 500 individuals. Given what we know from previous work, our expectations on the share of women would range from below 10 percent to 20-30 percent at most. This translates into some 50-150 women in absolute numbers. Often samples in LIS are smaller. For a 20 000 sample, the absolute number of top income women would be 20-60 women. Clearly, any further study of

---

[5] This is, of course, part of a bigger discussion about what really has happened to inequality depending on what is included in the income concept and how one treats top observations in surveys (see, e.g. Burkhauser, Hérault, Jenkins and Wilkins (2018). An important point in this discussion is that the under-coverage has two components: one is the unreliability of observations that we actually have in the survey data; the other is that parts of the top may not be covered at all (either due to truncation or due to non-response). These problems require different types of adjustments.

[6] LIS does not apply bottom- or top-coding to the microdata sets themselves, but typically bottom-and top-coding is applied when calculating summary measures and the Key Figures. For some countries this can have large impacts (see, e.g., Thewissen, Nolan and Roser, 2016). We do not apply any top coding to our data.

characteristics of these women (education, employment, age profile, etc.) quickly brings us down to numbers where it is no longer meaningful to proceed.

Another concern is the coverage of all income sources. In principal, LIS data contains income from all sources (so constructing the equivalent to total income in the top income literature is possible), but all data is not available for all waves. This is especially the case for capital incomes. These are available for most countries in LIS only starting in 2007. Before that only Italy (from 1995), Germany (from in 2001), and the U.S. (from 1979) have individual level capital income data.[7] We will contrast the share of women in top group for the labour income distribution, the sum of labour income and capital income, and the total income distribution whenever data is available. Throughout we use the incomes from the personal-level data (the P-file).

Based on these considerations, we restrict the LIS data samples in the following way. First, we limit our study to the adult population, leaving only individuals over 18 years old in the data. Second, to be able to follow the development of the share of women in top income groups over time, we restrict our study to countries with 5 or more years in the LIS data. This implies a loss of 15 out of the 49 countries covered in LIS. Third, we impose a restriction that the sample size must either exceed 0.0005 times the country's population, or consist of more than 20 000 observations.[8] This leaves us with 223 country-years of a total of 329 in LIS.[9] Finally, for the part of our analysis concerning partner income, we restrict the sample even further, and require at least 5 partnered women in the top1 percent of the labour income distribution for a country-year. In this part of the analysis, we are able to include only 8 countries and 80 country-years.

In order to construct the shares of women in top income groups for a particular country, we merge personal-level data files, and bottom-code negative income to zero, whether labour, total, or labour plus capital. For each country and each year, we then weight the observation by inflated population weights and obtain cut-off points for the joint income distribution of both

---

[7] This availability is based on the 2011 template of LIS. In the just released 2019 template capital income is available only at the household level.

[8] We make one exception, Italy, where we include the three most recent years of data despite observation count falling just short of the latter threshold (20 000 observations) and the sample size being less than 0.0005 times the population of Italy.

[9] In Appendix A we show a table with the sample size for the countries and waves included in our study based on there being non-missing individual labour income for a sample size that meets our above conditions. We also display the absolute number of women in the group where this number is the smallest, the top1 group. In top10, the number is always more than 10 times this number (often 15-20 times), since the group is 10 times as large and also has a larger share of women.

men and women. We use these cut-off points to classify an individual as belonging to a particular percentile group of the income distribution focusing on the top10 and top1 groups.

For the part where we study of top income women and their partners, we rely on variables in LIS describing partnership, relation to the household head, and age. We classify an individual as having a partner if the LIS *partner* variable indicates him or her as having a partner (100), as living with a partner (110) or as not living with a partner (120).[10]

3 The development of the share of women in top income groups – What does LIS data show?

Given the minimum requirements on sample size we arrive at 28 countries for which we have relatively comparable observations of individual labour income (including self-employment income) since at least the early 1990s and in many cases since the late 1970s. For most countries we also have individual capital income starting in 2007, but only in three cases do we observe individual capital income before 2007, and only in one of these, the U.S., do we have both individual labour income and capital income starting in the late 1970s (the other two with individual capital income before 2007 are Germany and Italy).

We start by looking at the share of women in top10 and top1 groups of the labour income distribution. Below the different country developments have been divided into five country groupings: Anglo-Saxon countries, Continental European countries, Scandinavian countries, Eastern European countries, and "rest-of-the-world" countries (this group consisting of Israel, Taiwan, Paraguay, and Mexico). The groups are similar to those in the overview of top income developments by Roine and Waldenström (2015), which in turn are based on common patterns noted already by Atkinson and Piketty (2007).[11]

In addition, the grouping also makes sense from an economic gender inequality perspective. For instance, once men and women have equal rights in society, it is reasonable to, as a first step, group countries according to their extent of female labour force participation (FLFP). FLFP reflects not only how advanced a country is in terms of industrialization, but also parental leave and child care policies that enables mothers (traditionally the prime care takers)

---

[10] We do not use the marriage data also available in LIS, as the definition of marriage changes over time and across countries.

[11] The title of the first of two volumes (Atkinson and Piketty 2007, 2010) collecting much of the early top income work is "Top Incomes over the Twentieth Century: A Contrast between European and English-Speaking Countries". Clearly, the residual category "the rest of the world" does not have the same status as the others in terms of the grouping actually capturing common institutional features and historical legacies.

continuing to earn a wage after having children. The Scandinavian countries had already in the 1970s a high female labour force participation since the early expansion of high-quality child care enabled mothers to work while having small children. Another group of countries where mothers worked early was in Central and Eastern Europe, where these countries' joint Communist past have left a legacy of relatively high rate of FLFP. The Anglo-Saxon country group share a similar institutional set-up in terms of rather limited maternal leave benefits (and limited, if any, paternal leave provisions), and also characterized by high child care costs.[12] This institutional frame generally hampers the possibility for families with small children to be dual-earner households. Moving on to Continental Europe, most countries have had a relatively low FLFP rate in the past decades, except for France. The main difference between France and the rest of the continental group has been in France providing public day care solutions that allows parents to work full time. Finally, the four remaining countries in our sample (Israel, Mexico, Paraguay, and Taiwan) are not only geographically scattered but differ also from an economic gender inequality perspective. While Israel has a high rate of FLFP, Mexico is one of the countries in the OECD with the largest gender gap in employment. Therefore, we will refrain from making any group-specific analysis of this group. For further updated details on the status of economic gender inequality, see OECD (2017).
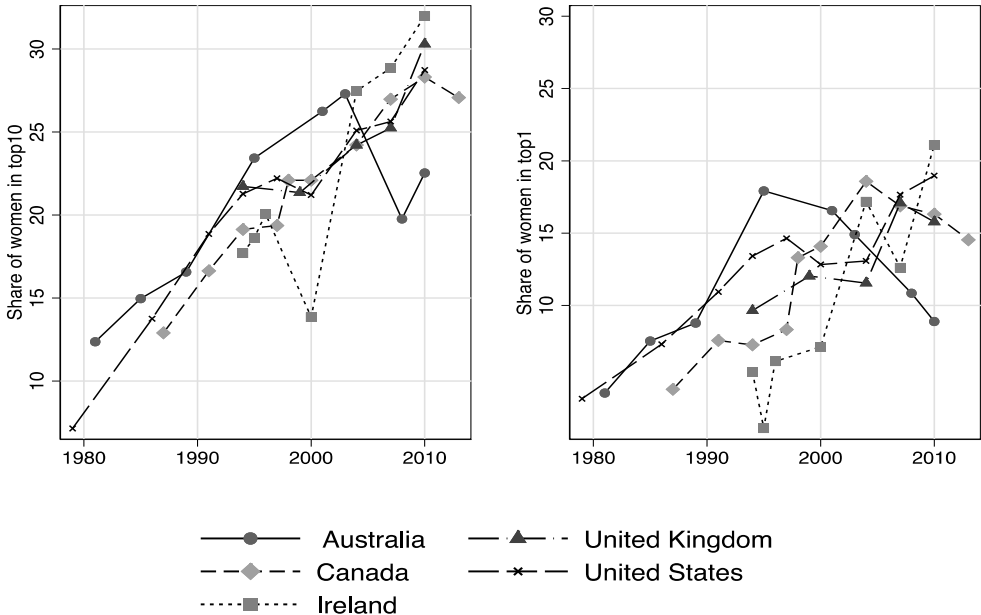

Figure 1. Share of women in top groups in the labour income distribution, Anglo-Saxon countries.

---

[12] See Adema et al. (2015) for more details.

Figure 1 shows the development of the share of women in the top10 group (left) and the top1 group (right) for the five English-speaking countries in our sample. The overall picture is clear. The share of women is far from equal to that of men, but it has at least doubled in the two top groups since the early 1980s, from low levels to around 25-30 percent women in P90-100 and to around 15-20 percent in the P99-100.

Figure 2 shows the same development for nine continental European countries. The overall trend is similar; growth of the share of women in the order of a doubling (or tripling) since the 1980s, and levels in the most recent waves around an average of 25-30 percent in the top10 and 15-20 percent in the top1 group. A noticeable, interesting difference here is that the spread is larger with some countries, like Spain, France and Greece, being at or above 30 percent women in the top10 group, while countries such as the Netherlands, Switzerland and Germany are around 20 or below.
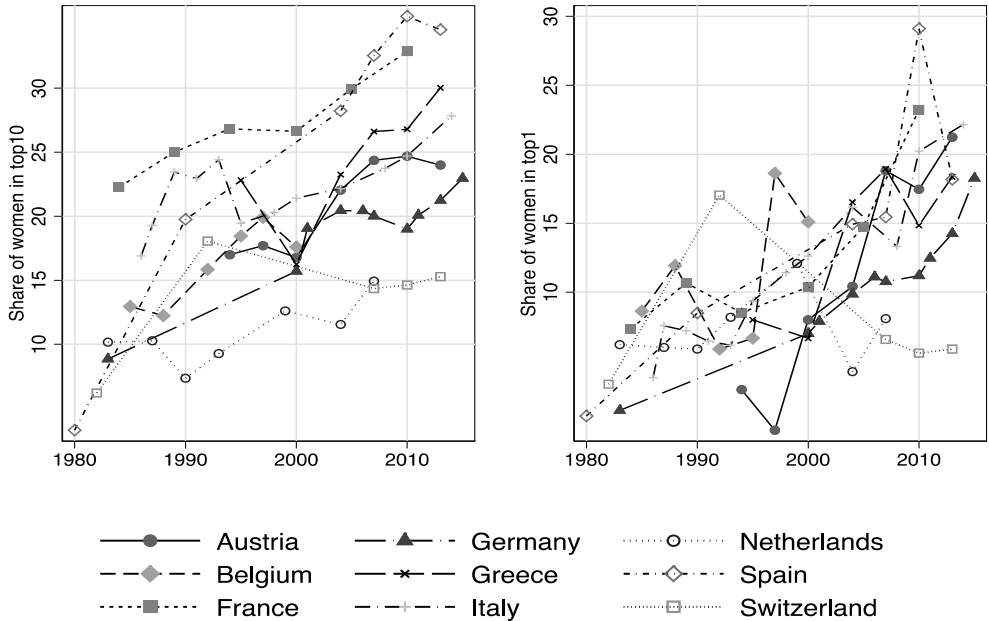

Figure 2. Share of women in top groups in the labour income distribution, Continental European countries.

Scandinavian countries also display similar trends and interestingly enough these countries, known to be comparatively gender equal, do not display higher shares of women in the top – see Figure 3.
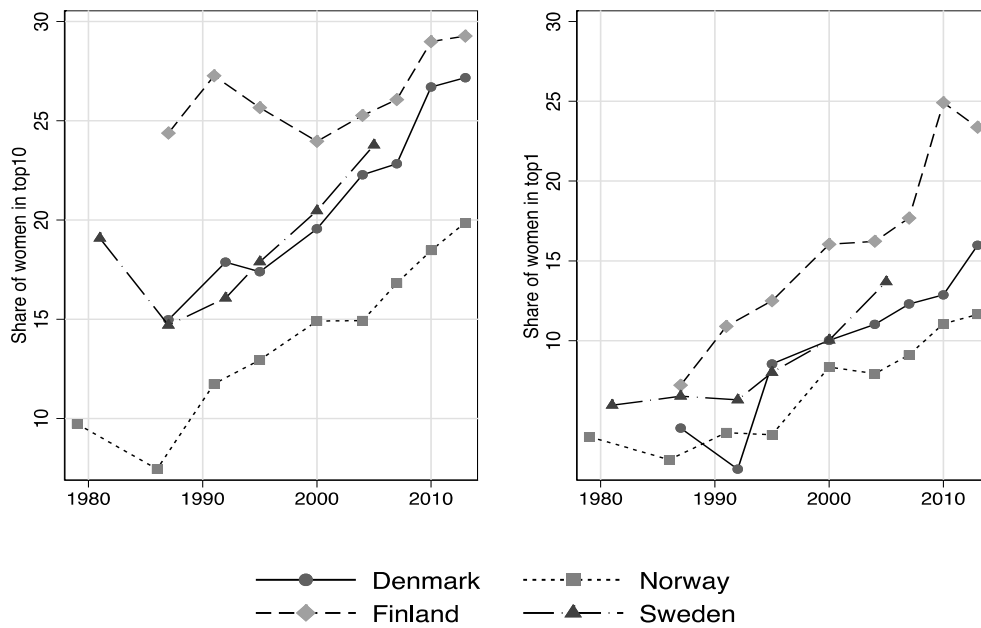
Figure 3. Share of women in top groups in the labour income distribution, Scandinavian countries.

If anything, the values at least for Denmark, Norway and Sweden are low compared to other countries (as also shown in Boschini, Gunnarsson and Roine, 2017, and discussed in Boschini and Gunnarsson, 2018).
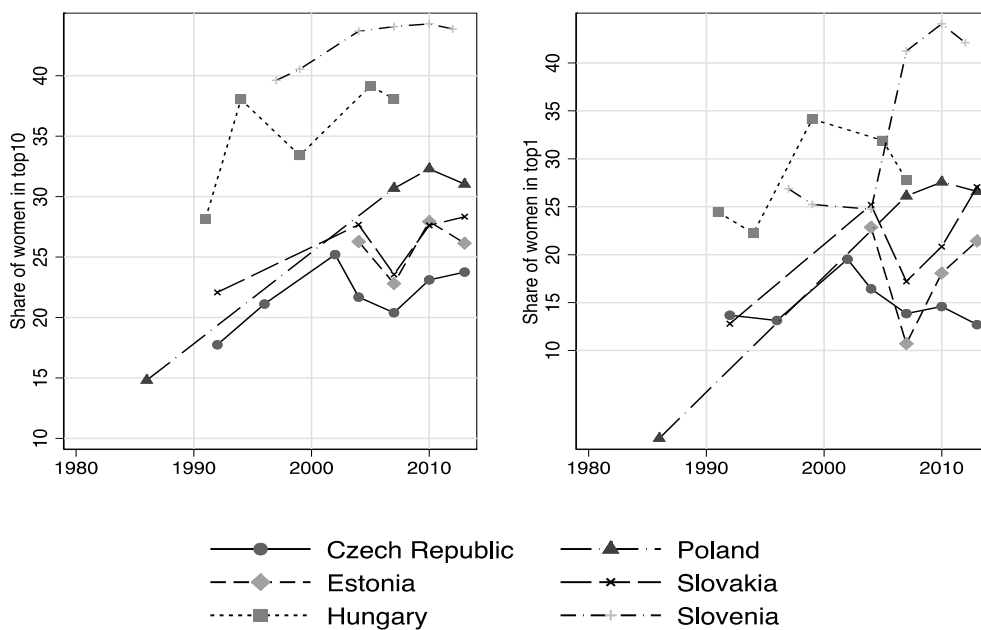

Figure 4. Share of women in top groups in the labour income distribution, Central and Eastern European countries.

Looking at the group of Central and East European countries in our sample – shown in Figure 4 - we note that they display on average higher levels of women in top income groups. This is well in line with the legacy of former communist countries being relatively gender equal, at least when it comes to labour market participation.

Figure 5 shows that the trend has been remarkably similar for as different countries as Israel, Mexico, Paraguay and Taiwan. The share of women in the top10 group has increased over time from low levels to around 30 percent in 2015, while the share of women in the top1 has experienced a less pronounced positive development so as to arrive at around 20 percent in 2015.
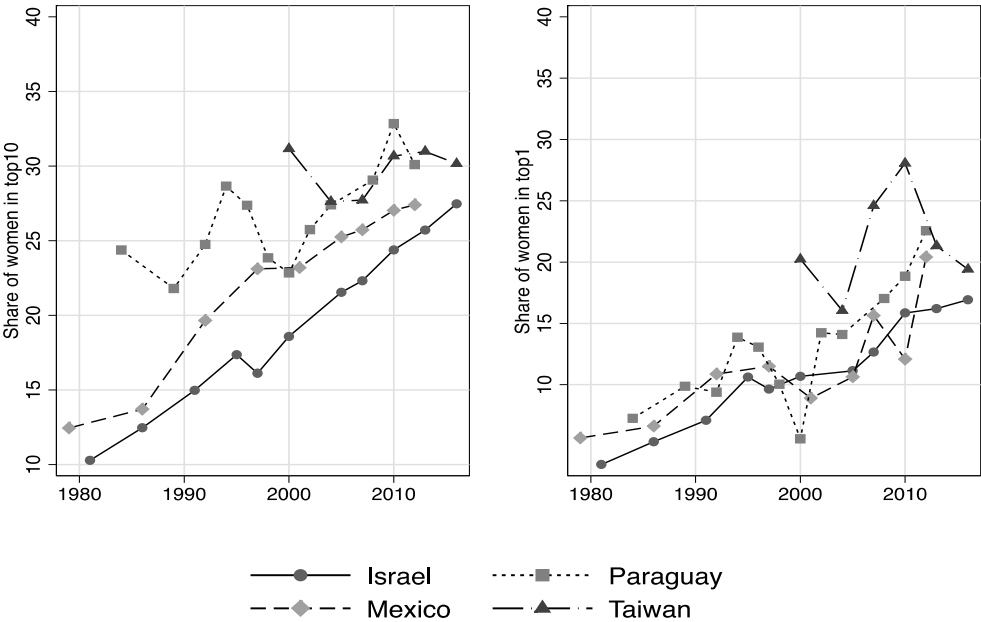

Figure 5. Share of women in top groups in the labour income distribution, for Israel, Mexico, Paraguay, and Taiwan.

Overall, there seems to be a relatively common trend across countries and the orders of magnitude are also relatively similar. In general, women's share of the top of the labour income distribution has increased a lot since the 1980s. There is also quite consistently a fanning out of the share of women, in the sense that there are consistently fewer women higher up in the distribution. But there are also some interesting differences across country groups, most notably the Central and Eastern European countries having the highest shares of women both in top10 and top1 (and also relatively similar in the top10 and top1 groups). The relatively low share of women in top1 in the Nordics is also notable.

The differences across country groups are in many ways in line with the results in the gender economics literature. Research in recent decades has shown that there is a complex relation between female labour force participation and the varying gender gaps across the earnings distribution, beyond the relevance of family policies - more about this in Section 4.2. below. On average, the mean gender gap has been shown to be negatively correlated with the rate of FLFP, which might at first appear surprising. But as shown in Olivetti and Petrongolo (2008), the European countries with a low gender wage gap tend to have strongly selected FLFP in that women in those countries either work out of necessity to support themselves or to pursue their careers (and have the means to arrange for private child care). As also emphasized in Arulampalam, Booth and Bryan (2006) the countries in Europe having a generous welfare state in terms of child care provision tend to have high glass-ceilings. The suggested mechanism goes through part-time work: Since mothers tend to retain the lion part of the responsibility for children, publicly provided child care enables them to be on the formal labour market, but only part-time – see e.g. Olivetti and Petrongolo (2017) and Johnsen and Løken (2015) for a more in-depth account of the mechanisms. Another way of capturing women's top earnings potential is to analyse the share of female managers, which vary considerably over country groups. These patterns match the extent of glass-ceilings, with Anglo-Saxon countries, and other countries with low public provision of child care having a higher fraction or at least the same fraction of female managers as the Scandinavian countries.

In relation to these patterns the high levels of women in top groups in Eastern Europe are interesting with their history of high labour force participation and a state socialist legacy in terms of less traditional gender-roles in the population (that remain today) with respect to work (see Campa and Serafinelli, 2018). The share of female managers in these countries is also comparatively high.[13]

---

[13] Out of 28 EU countries the top seven places in terms of women in management positions are Eastern European countries according to recent data from Eurostat.
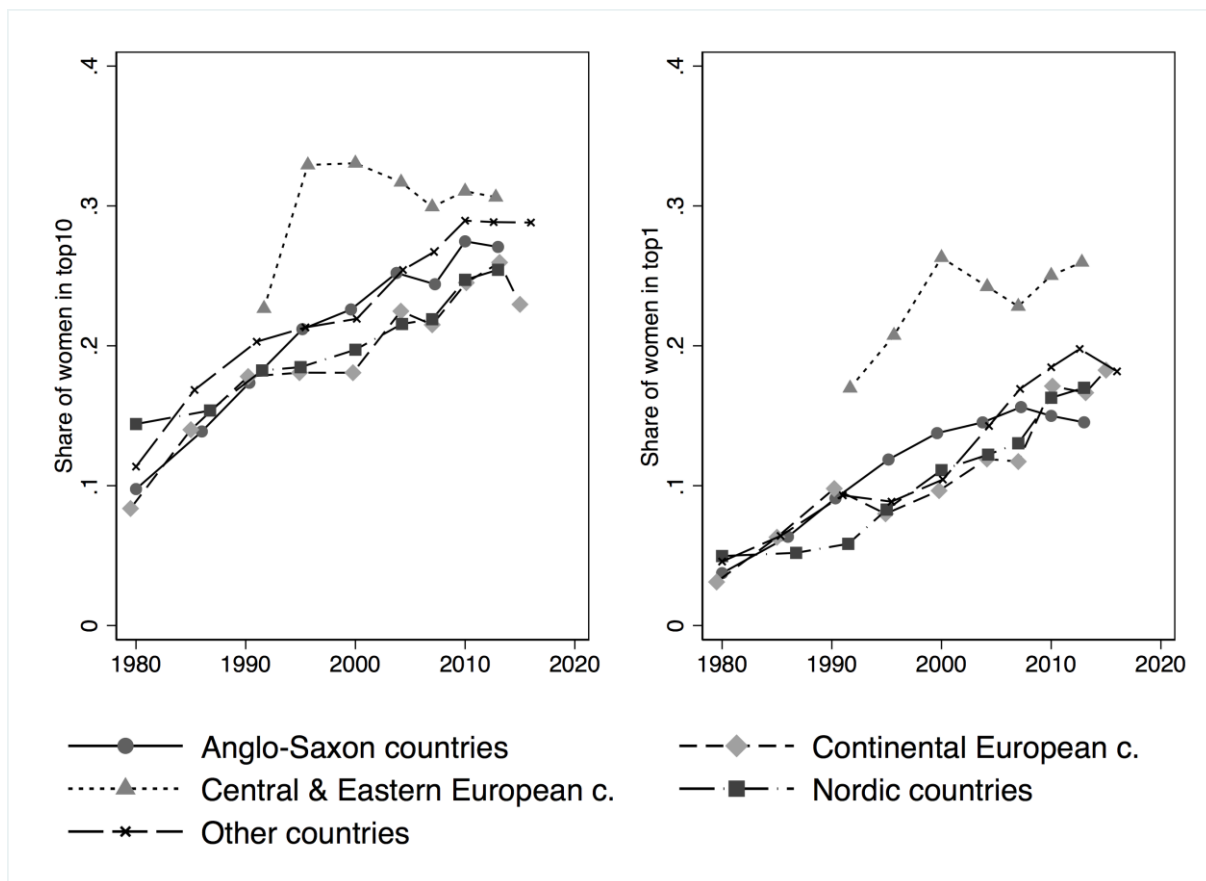
Figure 6. Average share of women in top groups in the labour income distribution for the five country groups.

Figure 6 shows the average share of women in the top10 and top1 in the labour income distribution across the five country groups. The average patterns are surprisingly similar with the Eastern European observations being the positive outliers, and with the Continental European and Nordic countries displaying the lowest shares.

3.1 Differences between labour income, labour plus capital income, and total income

As mentioned above, an important finding in the top income literature, especially historically, is that top income shares and their development depends a lot on capital incomes. Several papers studying long-run developments show that much of the decline of top shares in the first half of the century turn out to be driven by diminishing capital incomes in the top, while the top share of the wage bill in this period did not change much. In recent decades the picture is more mixed.

This, of course, raises the question to what extent the share of women in top groups vary when including capital. As also noted above, LIS data has some important limitations with respect to this. Only after 2007 do we observe individual capital income for most countries that

we study.[14] This means that for all observations after this we can compare the share of women in the top1 and top10 in the distribution of total income (now always including capital income) and corresponding shares in the labour income distribution. Figure 7 plots these share against one another.[15] The shares correspond surprisingly well to each other, and there is no clear pattern in over- or understating the share of women depending of which distribution is used. If anything, there appears to be a tendency of the top1 share of women being somewhat more sensitive to the income measure used. In 29 out of the 64 observations, the top1 share of women is larger in the labour income distribution, with the maximum difference being 0.05 percentage points. In the remainder 35 observations, the maximum understatement is 0.04 percentage points when using the labour income distribution instead of the total income distribution. Overall, it seems the top share of women in the labour income distribution serves as a good proxy for the top share of women in the total income distribution for this period and for this set of countries suggesting the income composition is not too different between men and women.

---
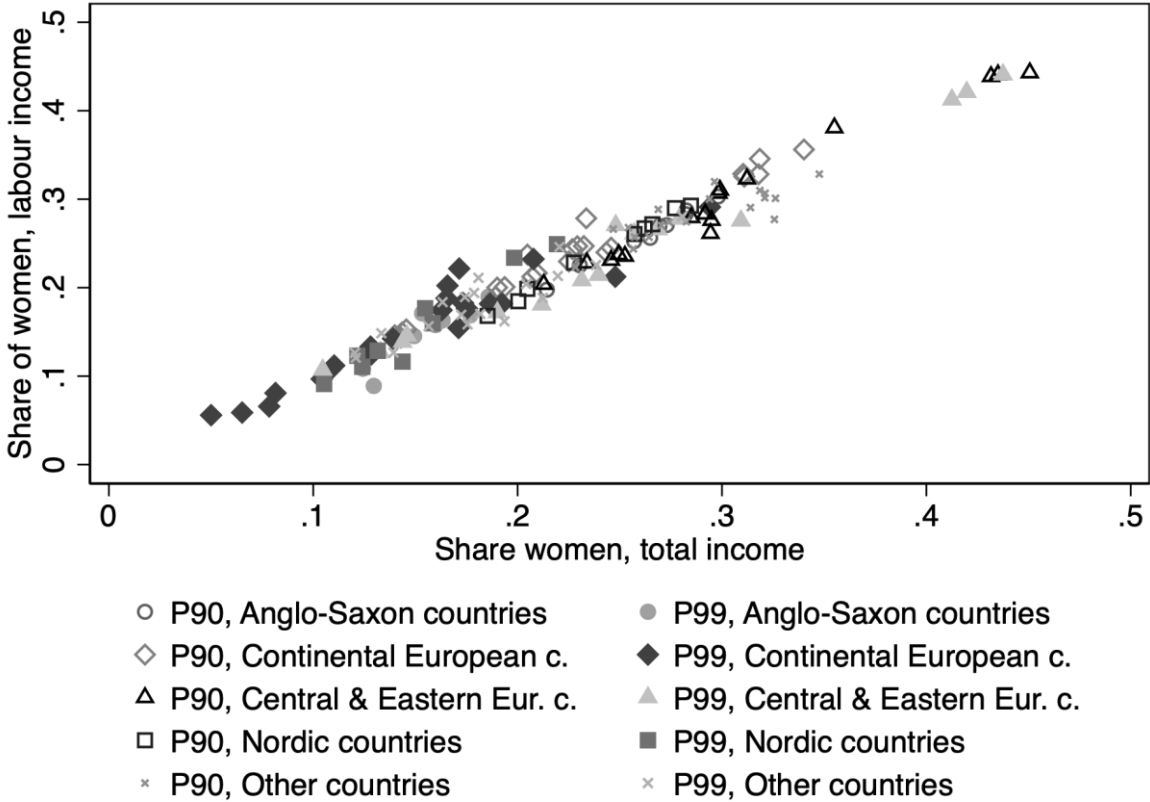
[14] This availability is based on the 2011 template of LIS. In the just released 2019 template capital income is available only at the household level.

[15] Only observations from 2007 and onwards are used.

Figure 7. The correlation of share of women in top1 and top10 in the labour income distribution and the total income distribution since 2007.

For the few countries where we have longer time series including capital incomes, however, we do see some differences when going back in time. Figure 8 illustrates this for the United States, for which the longest time series of capital income in LIS is available. It shows the share of women in top10 (left) and top1 (right) of the labour income distribution, of the total income distribution and in the distribution of the sum of capital and labour income. It suggests that the share of women in the top10 is consistently a few percentage points lower in the labour income distribution up until recent years when the difference is basically gone. The same is the case, but even more pronounced, for women in the top1 group. In line with the suggestive evidence in Edlund and Kopzuk (2010), the top1 share of women is higher in the 1980s in the US when using measures of total income (including capital) rather than labour income.
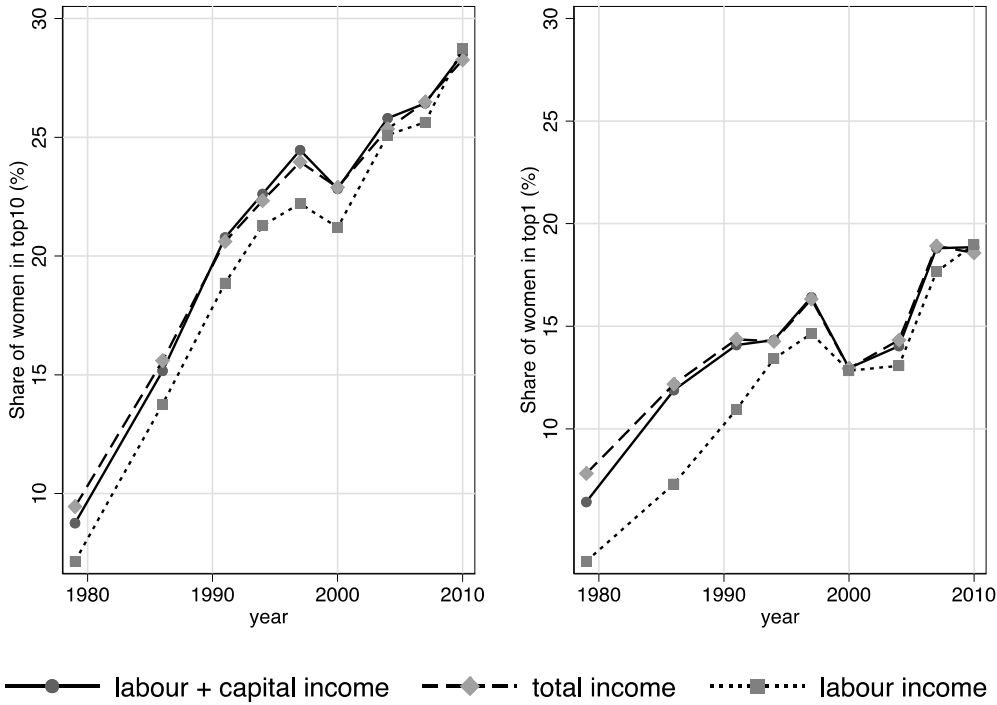

Figure 8. Share of women in top groups in the United States in different income distributions.

In Appendix B we present analogous graphs for all countries in our sample, and it appears as though the discrepancy between top1 share of women in the labour income distribution and in the total income distribution is rather limited in recent years.[16]

---

[16] Total income here is the total income concept defined in LIS. Given the difference in, for example, coverage of capital income these are not fully comparable over time.

3.2 Comparing LIS to top income results on women in top shares

Given the issues discussed above, it is interesting to see how the overall results compare to the previous studies done on countries where individual level income tax data is available for the full population (or in some cases larger samples). In Figure 8 we display our series (bold lines, solid for top10, dashed for top1) over the series for eight countries in Atkinson et al. (2018) and series for Sweden in Boschini et al. (2017). The overall trends and levels are very similar but there are also some important differences especially for individual years where the LIS-based series fluctuate. The clearest discrepancy is Australia where LIS data shows a decline in the share of women in the top1 group starting around 1995 to a level below 10 percent. This pattern is very different from what is found in Atkinson et al. (2018) where the trend is positive throughout and the level in the most recent year is above 20 percent. Canada also shows a similar discrepancy, though not as marked.

Our overall interpretation is that even the shares for top1 in LIS can be taken at least as suggestive for how large the share of women is in the top and also as giving a reasonably accurate picture of the long run trend. One should not, however, interpret individual year fluctuations, as these might just as well be a result of having a small sample. This is in line with what is found in the comparison between LIS and WID in Yonzan, Milanovic, Morelli, and Gornick (2018). But when it comes to studying finer details of gender differences within the top1 group this becomes too demanding given the sample nature for most LIS data.
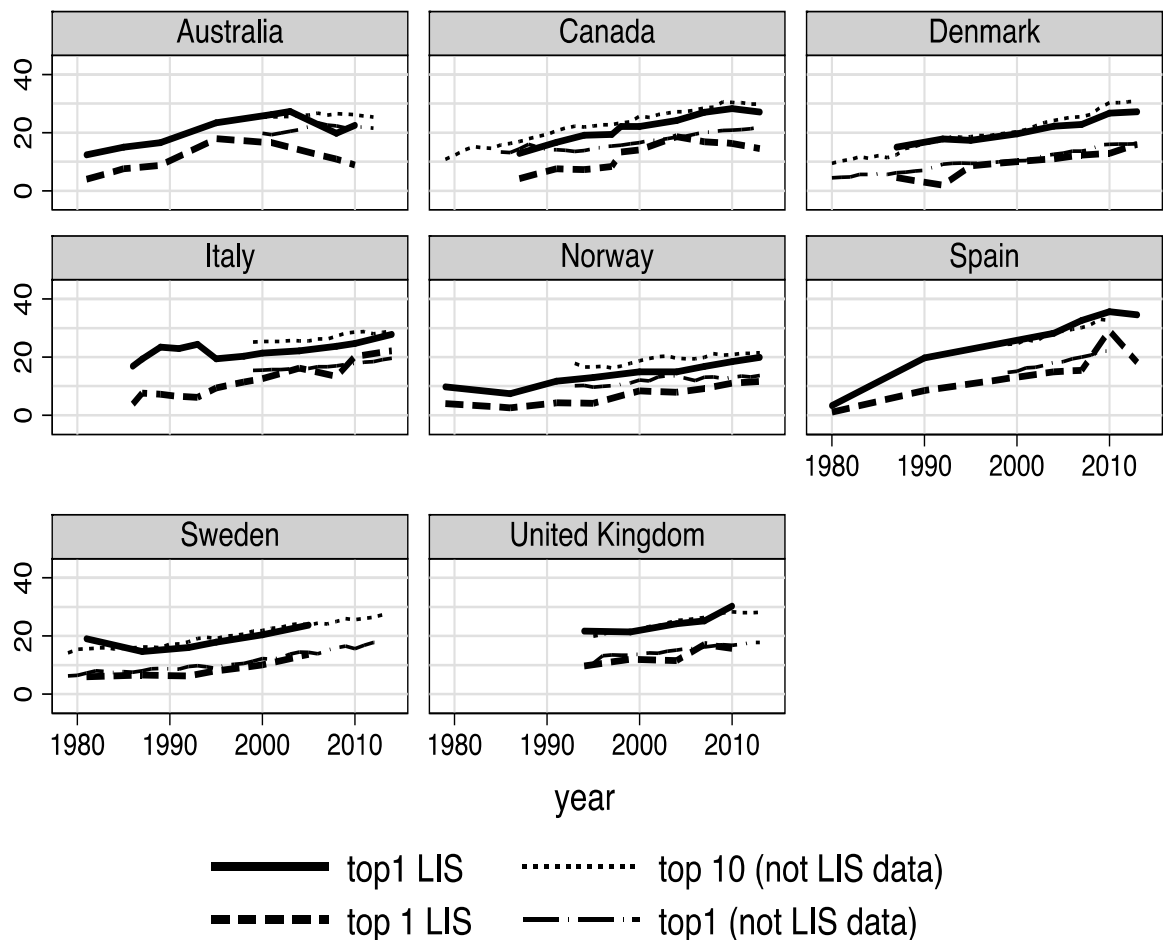
Figure 9. The shares of women in top1 and top10 in 8 countries with LIS data and tax register data.

4   What determines the share of women in top income groups?

A large literature has studied various mechanisms that could explain the lack of women in the top of the distribution, as well as the cross-country and time differences in this. Several papers, however, suggest that the women in the top of the labour income distribution are relatively similar to their male peers in terms of education, occupations and sector; if anything, women tend to be somewhat younger – Bertrand et al (2017), Guvenen et al (2014), Keloharju et al (2016). A potential explanation for the relative lack of top income women is that they often have been faced with having to choose between having a family and a career. Those women choosing the career path also tended to become childless to a greater degree –

see e.g. Goldin et al (2006), Bertrand et al (2010), Boschini et al (2011), Goldin (2014).[17] Another type of explanation has to do with the impact of children for those who have them. Angelov et al (2016), Kleven and Landais (2017) and Keloharju et al (2018) show for different Nordic samples, that the event of the first child severely hampers women's wages and future careers compared to those of their male partners. Kleven et al (2019) show that this extends to Germany, Austria, the UK and the US as well. In short, the nexus between having a career, a marriage and children is highly complex for women (and increasingly for men), since both partnership and fertility decisions are endogenous to succeeding in the labour market. Moreover, as argued already by e.g. Mincer and Polachek (1974) and Becker (1985), women's educational choice could also in the first place be endogenous to women's aspirations of being the prime caretaker later in life. In this paper we disregard these complexities and limit ourselves to explore the descriptive relation between income status and educational level, marriage and fertility (in terms of parity).

Guided by some of these previous papers we will in turn look at, first, a linear probability model of how the likelihood of men and women being in the respective top groups is associated with variables such as age, education, having a partner, having one or more children. Second, we look at the association between social norms concerning if women should stay at home when children are young and how this relates to the share of women in top groups. Third, we study asymmetries in how the likelihood of being in the top income group is related to having a partner, as well as – conditional on having a partner – where in the income distribution this partner is located.

4.1   Gender asymmetries in the probability of being a top earner

To explore some key associations, and in particular the difference between men and women, we run a linear probability model to study how the likelihood of being in the top10 and top1 groups, respectively, is related to a number of socio-economic factors, controlling for country and time fixed effects.[18]

---

[17] Nowadays this is a less binding trade-off, especially in cities with inflows of unskilled people that can help the high earning women with domestic chores at a low wage cost – see Cortes and Tessada (2011), Cortes and Pan (2018).

[18] We use, for ease of interpretation of the coefficients, a linear probability model (LPM) – even though LPMs usually generate biased and inconsistent estimates (Horrace and Oaxaca, 2006) of non-linear models. Reassuringly however, running a probit model yields similar qualitative results.

	Men P90	Women P90	Men P99	Women P99
Age in years	0.0174***	0.00639***	0.00246***	0.000658***
	(0.00013)	(0.00007)	(0.00005)	(0.00002)
Age squared	-0.000189***	-0.0000667***	-0.0000239***	-0.00000636***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Medium educ.	0.102***	0.0335***	0.0139***	0.00112***
	(0.00095)	(0.00050)	(0.00037)	(0.00015)
High educ.	0.315***	0.139***	0.0596***	0.00909***
	(0.00123)	(0.00086)	(0.00061)	(0.00028)
Living w. partner	0.0364***	-0.0141***	0.00300***	-0.000349
	(0.00106)	(0.00058)	(0.00040)	(0.00018)
One child	0.0464***	-0.00247***	0.00558***	0.0000109
	(0.00128)	(0.00067)	(0.00052)	(0.00021)
Two children	0.0917***	-0.00433***	0.0135***	0.000632**
	(0.00142)	(0.00074)	(0.00059)	(0.00023)
Three children	0.0842***	-0.0159***	0.0178***	-0.0000711
	(0.00183)	(0.00086)	(0.00083)	(0.00029)
Three+ children	0.0252***	-0.0319***	0.00659***	-0.00210***
	(0.00236)	(0.00097)	(0.00101)	(0.00031)
Constant	-0.237***	-0.117***	-0.0360***	-0.00140*
	(0.00531)	(0.00244)	(0.00201)	(0.00066)
Observations	3210985	3513825	3210985	3513825
R-squared	0.163	0.097	0.091	0.185

All regressions include fixed effects for LIS wave and country.

Standard errors in parentheses.

* $p<0.05$, ** $p<0.01$, *** $p<0.001$

Table 1. Linear probability regression of the likelihood of being in a top group for the pooled sample of all observations in the main country groups.

Table 1 shows the results for a pooled regression for all countries and years in our four main groups (Anglo-Saxon, Continental Europe, Central and Eastern Europe, the Nordic countries, and the residual group consisting of Israel, Mexico, Paraguay and Taiwan). As one would expect age is positively associated with the likelihood of being in the top group but decreasingly so (age squared being negative) for both men and women, education is also positively associated with being in the top groups, and the size of the coefficient is larger for higher education as compared to medium education (with low education being the base case). These effects are strongly statistically significant and larger in magnitude for men than for women. Having a partner is associated with an increased probability of appearing in the top income groups for men, while the reverse is true for women. Finally, the effects associated with children are also different for men and women. For men, having children is associated with a larger probability of being in the top, and more so for more children. For women,

however, the reverse is true. This is in line with the findings in previous literature – e.g. Goldin (1997, 2006) and Boschini et al (2011) – where until recently women in practice had to choose between pursuing a career and children, while male fertility increased with economic success.

Pooling the data like this, of course, runs the risk of potentially missing a lot of cross country (or country group) variation. However, when we run the same regression for the individual country groups the effects turn out to be surprisingly similar. Figure 10 illustrates the coefficients when running the above regression for the four country groups separately. To the left are point estimates for men, to the right those for women (P90 above, P99 below). With few exceptions the estimates are very similar (full regression tables are in Appendix C).
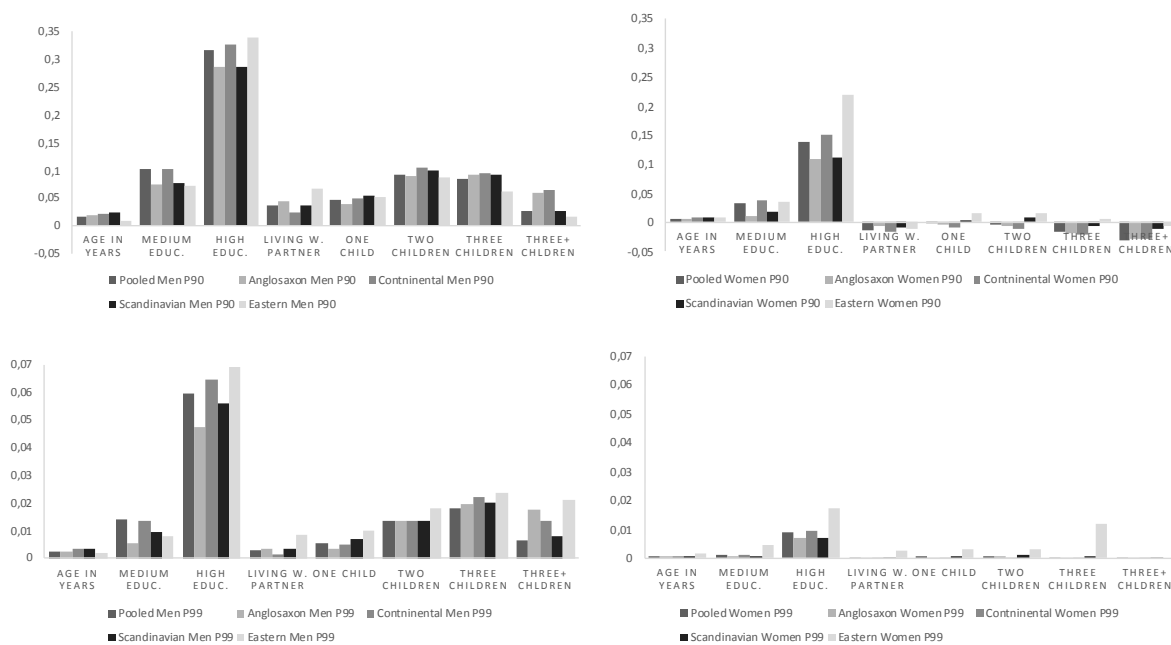

Figure 10. Regression coefficients for the pooled regression compared to the respective country groups.

While the above regressions include country and time fixed effects we have not explored the potential that some variables may have different over time. To do so, we interact the indicators for number of children and education with time dummies (one for each wave of LIS data) and repeat the analysis above separately for men and women. Coefficients remain qualitatively the same, but the added interaction effects suggest that the differences between men and women are somewhat decreasing over time. For men, the 'child premium', observed earlier in both P90 and P99 groups, tends to decrease over time for all number of children as the interaction terms are negative and increase in magnitude. Similarly, the 'child penalty'

observed for women is also decreasing over time, as the interaction terms remain positive and increase in magnitude. The effects for education categories are similar. For men in both income groups, the education premium they enjoyed compared to women seems to wane over time, especially so for men in P90. For women, the effect remains at a reasonably constant level throughout the period.

Overall, the results from the pooled regressions including time interactions are again surprisingly similar if we study groups of countries separately, with a notable exception in the Scandinavian countries, where the 'child penalty' for women was barely present to begin with (rather, the magnitude of premium for children was smaller for women, as compared to men). Of course, one should be careful interpreting these results, as the interaction terms do not represent the marginal effects and cannot be interpreted directly.[19] More importantly, these correlations are all purely descriptive and careful econometric modelling – beyond the scope of this paper – is required to shed light on any potential causal mechanisms.
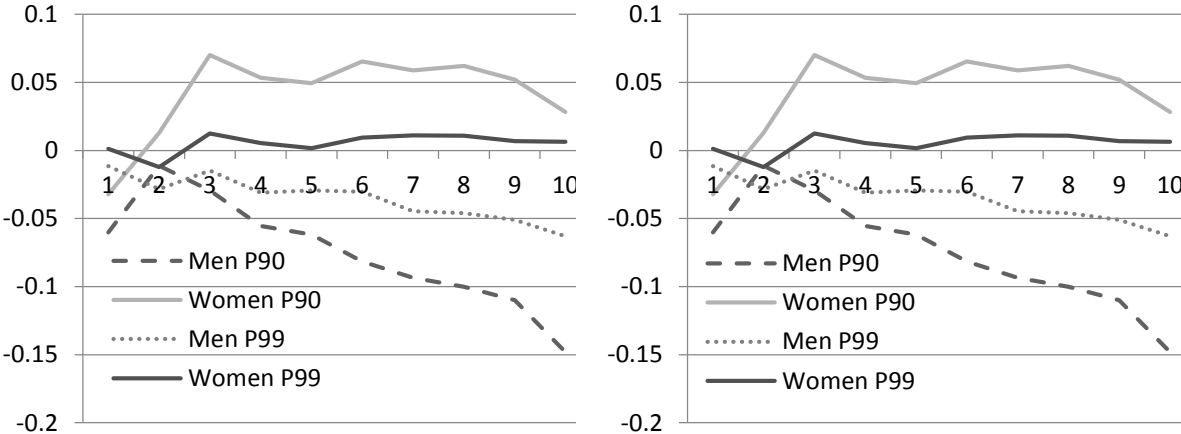

Figure 11. Interaction of medium (left) and high (right) education attainment and LIS wave indicators from the pooled OLS regressions.

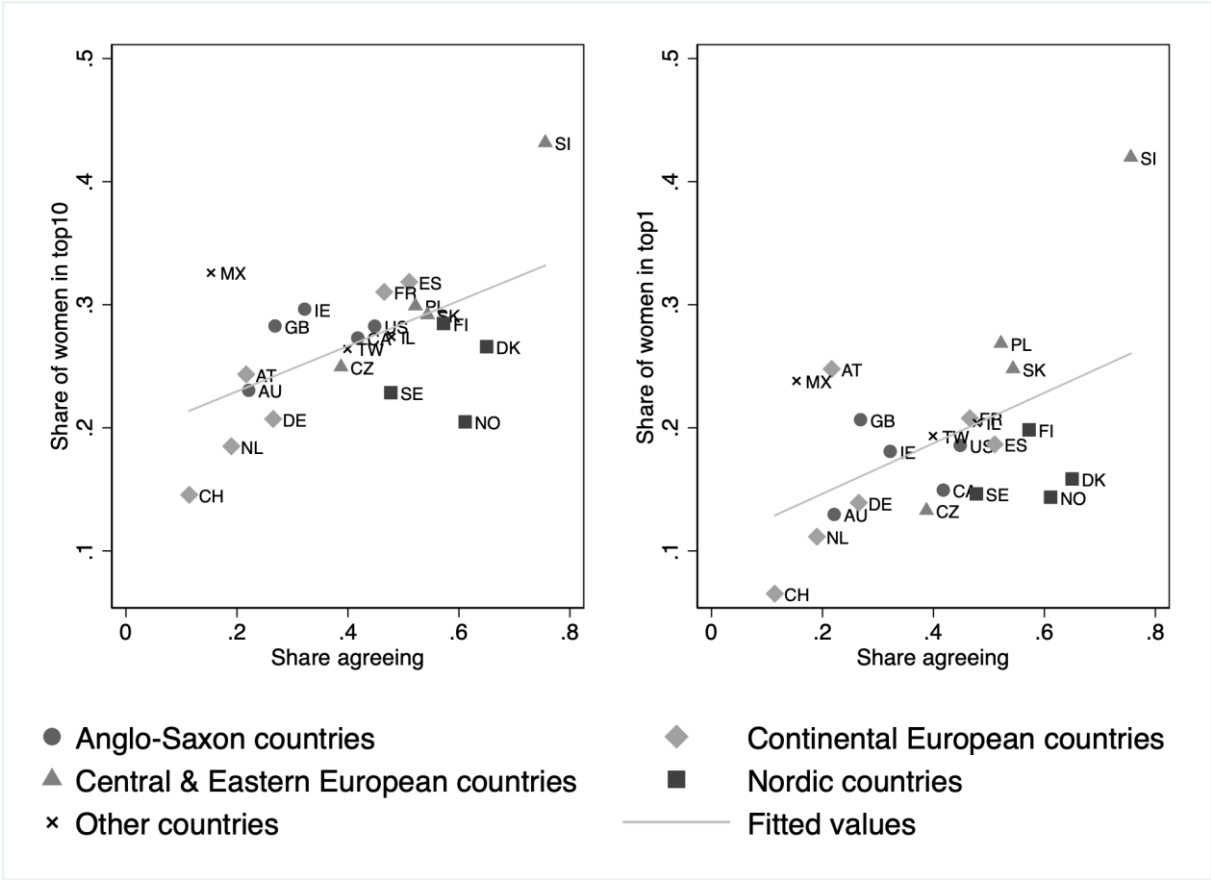## 4.2   Social norms and the share of women in top income groups

It has also been suggested that conservative social norms prescribing motherhood rather than career ambitions for women might be holding women back – see Bertrand (2011) and Ponthieux and Meurs (2015) for excellent overviews. Using questions in the World Value Survey on gender roles, Fortin (2005) indicates that there is a close relationship at the country level between female labour force participation and the share of the population agreeing with the statement "Being a housewife is just as fulfilling as working for pay". Moreover, in a

---

[19] Full results from these regressions are available upon request.

recent paper Kleven et al (2019) suggest that social norms might be even more important than changes in social policies for the wage penalties that mothers experience on the labour market. They measure social norms by the share of the population agreeing to that mothers should work full-time or part-time outside the home when having small children or school-age children using data from the International Social Survey Program (ISSP).[20]

To explore the extent to which social norms might influence the share of women in top income groups, we construct a similar variable as that used by Kleven et al (2019) from the 2012 wave of ISSP. More precisely, our social norm variable captures the share of the respondents that agree to the statement that "Women should work full-time outside the home when there is a child under school age". The underlying assumption is that to be a high-income earning woman (or man) you need to be working full-time most years (even if you have children under school age).

Figure 12 plots the correlation between this measure of social norms about working mothers with small children and the share of women in the top10 and the top1 groups of the total income distribution in 2012 (wave 9 in LIS).

Figure 12. Relationship between social norms encouraging women to work when children are young and the share of women in top groups.

4.3   Top income earners and their partners

As our regression results in Section 4.1 suggested, living with a partner is, on average, more likely for men than for women in our top groups. Given the increased female labour force participation and relatively lower gender wage gaps over the last decades, it becomes particularly interesting to explore how the share of the top men and top women with a partner has evolved. Moreover, given the increase in dual earner-households in general, could it be the case that more women in the top have partners that also in their own right belong to the top groups?

Studying partner choice and family composition in the top income groups with LIS data is challenging since the sample sizes are relatively small, and we easily end up with too few observations for it to be meaningful to conduct any analysis. Not only are there few women in the top group to start with, we also have the additional fact that not all of them have partners (and this has changed over time). In order to have longer time series, we focus on the top men and top women in the labour income distribution.

Figure 13 summarizes by country groups and waves, the share of top earners having a partner. (Appendix D contains the corresponding graphs by single countries.) First of all, we notice that the share of top men and women having a partner, on average, has converged over time in the top10 group of the labour income distribution. Especially in the Scandinavian countries in recent years, top10 men and women are almost as likely to have a partner. Back in the 1980s (for those countries where there is LIS data) only around half of the top women had a partner, compared to at least 90 per cent of the top men. Across top1 men and top1 women, there is a slight tendency towards convergence in that some more top1 women on average have a partner at the same time as somewhat less top1 men do not have a partner in recent waves. But the difference is still today at least 10 percentage points in share of top1 men and women that have a partner.
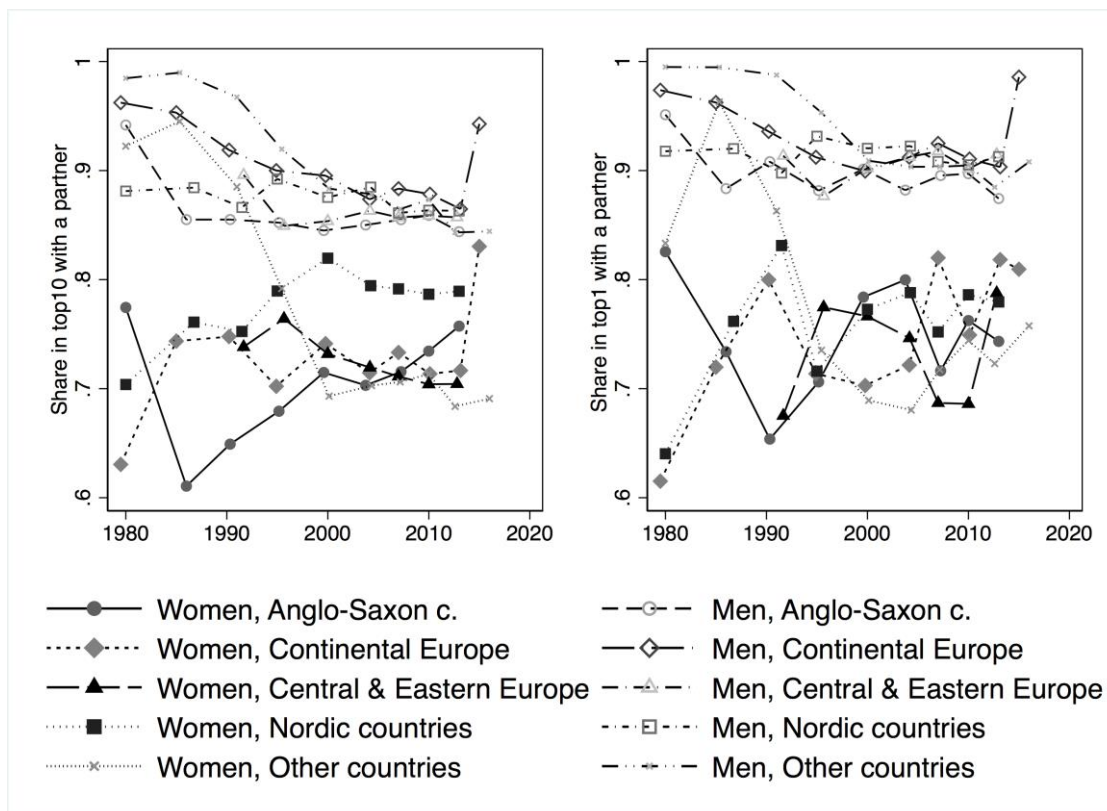
Figure 13. Share with partner in top10 and top1, women and men, for the respective country groups.

Yet another type of difference that has received relatively less attention until recently is the possibility that the choice of partner and asymmetries in relationships may matter a lot especially for top income women. There is, of course, a large literature on assortative mating (see Greenwood et al. (2017 for an overview) in general. But of more specific concern for the top, Bertrand, Kamenica, and Pan (2015) suggest that there is a norm prescribing women to earn less than their partners, and Folke and Rickne (2017) indicate that successful career women are likely to face a divorce as a consequence of their promotions. Our LIS sample can give an indication of whether top persons' partners also are top income earners or not.[21] However, the number of countries having a sufficient number of partnered women in the top during several waves is slim. Canada, Denmark, Finland, Norway and the US turn out to be the only countries having large enough samples for such an exercise to be meaningful. Figure 14 shows where in the labour income distribution the partners of the top men and top women are, when we pool the share of partners over these five countries, allowing for the partner to be either in P0-90, P90-99 or in P99-100. (Appendix E contains the corresponding country-

---

[21] In a related paper Aaberge et al. (2018) use LIS data to study assortative mating in relation to "perfect matching" and random matching, respectively.

specific graphs for Canada, Denmark, Finland, Norway and the US.) What stands out is that consistently a vast majority of top men, regardless of whether they are in top1 or top10 of the labour income distribution, have a partner that is not in the top10. For top women, that is not the pattern at all. To the extent that top women have a partner, they are more likely than men to have a high earning partner. These patterns are relatively stable over time. This is consistent with the patterns found for Sweden with register data by Boschini et al (2017).
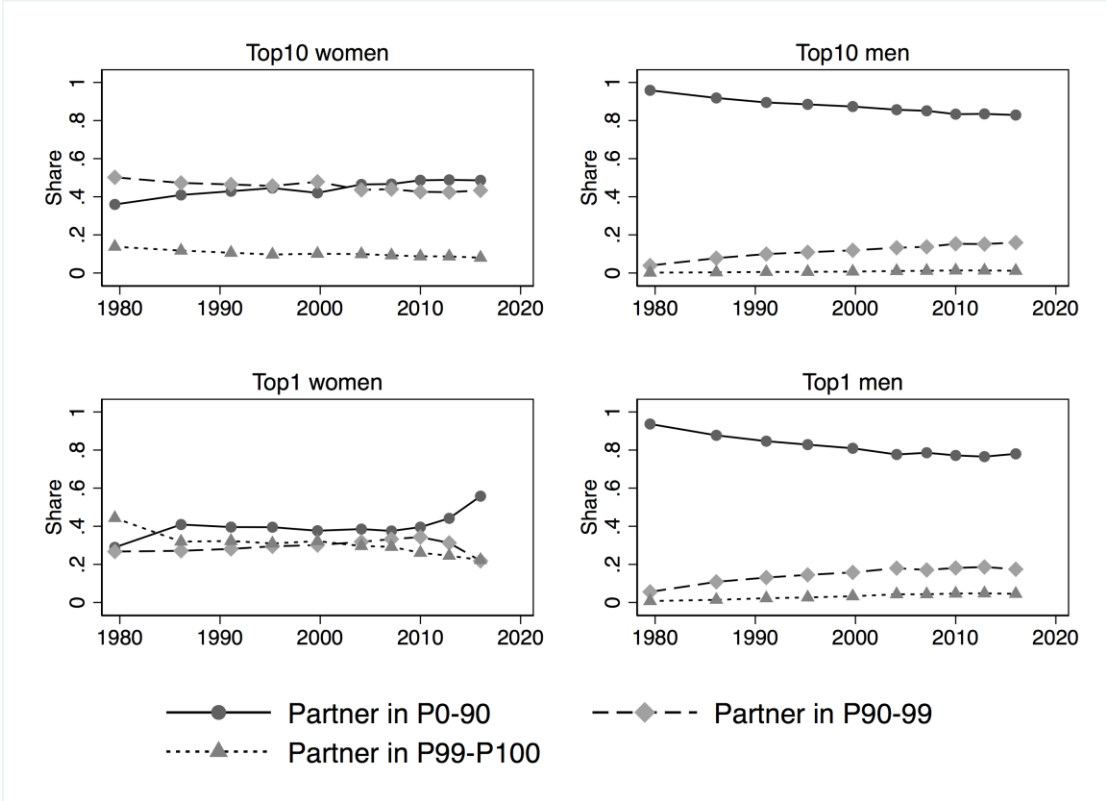

Figure 14. Income position of partner for men and women in top10 and top1 respectively. (Average for Canada, Denmark, Finland, Norway and the US.)

5 Summary and concluding remarks

This paper has explored what we can learn about gender differences in top incomes using LIS data. Overall, the main limitation lies in that the samples are rarely sufficiently large to allow careful study of women in the top group. This is, as shown above, especially true when one is interested in questions that would require a further division of characteristics of say the top1 group. Nevertheless, we think that our findings provide some novel insights.

First, we think that our overall trends for the share of women in the top10 and top1 groups give a reasonably accurate picture of both the level as well as the trend of this development since the 1980s. This is confirmed by comparing our results in this paper to previous findings that study a smaller number of countries (eight) but where the studies have

used data on the full population (or very large samples). Studying 28 countries we provide series that suggest that while women's share in top groups have increased (roughly doubled) over the past three, four decades, the representation of women in the top is still far from equal. The share of women in the top10 is around 25-30 percent in most countries, the max being above 40 percent in Slovenia, and the minimum being around 15 in Switzerland.

Second, for recent years when we can compare the shares of women in top groups across distributions of labour income and total income (including capital income) we find that the shares are not affected much (the maximum deviation being 4 percentage points, but most observations being very similar). Also, there is no clear pattern in over- or understating the share of women depending on which distribution is used. However, for the U.S. we can analyse this starting in 1979 and here we find a marked difference in that women's share in top1 was much higher when including capital income in the income concept, suggesting that even if recent observations show small differences depending on the income concept this is not necessarily representative for historical periods.

Third, regression results suggest that many associations between socio-demographic variables and the probability of being in the top income groups are different for men and women. Having higher education is positive for both, but more so for men. Interestingly, variables such as having a partner and having children have opposite signs for men and women. These family related variables are negatively associated with the probability of being in the income top for women while the opposite is the case for men. We also find that social norms, such as having a positive attitude to women working when children are young, are positively related to the share of top income women. Finally, when looking at income characteristics of the partners of top income earners, we find that top income men are much more likely to have a partner who is not in the top income groups, while this is not the case for women. Such asymmetries are likely to impact the ease of focusing on a top career differently between men and women.

Even though samples are small, the cases where it is feasible to study this show very clear consistent patterns to this effect across all countries in our data. Hopefully future studies on larger samples will be able to shed more light on this question.

The last point illustrates an important general gap to be filled in future studies of top income men and women. On the one hand, we need large samples (or preferably the full population) to observe sufficiently many top income individuals to study their characteristics, but we also need information on their family and household characteristics to fully understand who succeeds and who does not. Identical individuals in terms of observable individual characteristics may have

very different family situations, with important consequences for their individual success. Understanding these interactions seem like important avenues for future research.

References

Aaberge, R., Atkinson, A. B., Königs, S., 2018. "From classes to copulas: wages, capital, and top incomes". *Journal of Economic Inequality* 16(2): 295–320.

Aaberge, R., Thori Lind, J., Moene, K., 2018. "The inequality of equal mating", Working Paper, University of Oslo.

Adema, W., Clarke, C., Frey, V., 2015. " Paid Parental Leave: Lessons from OECD Countries and Selected U.S. States". OECD Social, Employment and Migration Working Papers 172.

Albrecht, J., Björklund, A., Vroman, S., 2003. "Is There a Glass Ceiling in Sweden?". *Journal of Labour Economics* 21: 145-177.

Albrecht J., Skogman Thoursie P., Vroman S., 2015. "Parental Leave and the Glass Ceiling in Sweden". *Research in Labour Economics* 41: 89-114.

Alvaredo, F., Atkinson, A. B., Piketty, T., Saez. E., 2013. "The top1 percent in international and historical perspective". Journal of Economic Perspectives 27: 3–21.

Angelov, N., Johansson, P., Lindahl, E., 2016. "Parenthood and the Gender Gap in Pay". *Journal of Labor Economics* 34(3): 545 – 579.

Arulampalam, V., Booth, A., Bryan, M., 2006." Is There a Glass Ceiling over Europe? Exploring the Gender Wage Gap Across the Wage Distribution". *Industrial and Labour Relations Review* 60(2): 163-186.

Atkinson, A. B., 2005. "Top Incomes in the UK over the 20th Century." *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 168(2):325–43.

Atkinson A B, Casarico A., and Voitchovsky S., 2018. "Top Incomes and the gender divide", *Journal of Economic Inequality*, 16(2):225-256.

Atkinson, A. B., Piketty, T. (Eds.), 2007. *Top Incomes over the Twentieth Century: A Contrast between European and English-Speaking Countries*. Oxford: Oxford University Press.

Atkinson, A. B., Piketty, T. (Eds.), 2010. "*Top Incomes: A Global Perspective, vol 2*. Oxford: Oxford University Press.

Atkinson, A. B., Piketty, T., Saez, E., 2011. "Top Incomes in the Long Run of History". *Journal of Economic Literature* 49: 3-71.

Becker, G. S., 1985. "Human Capital, Effort, and the Sexual Division of Labor". Journal of Labor Economics 3(1): S33-S58.

Bertrand, M, 2011. "New Perspectives on Gender". Handbook of Labor Economics 4B: 1543-1590.

Bertrand, M, 2018. "Coase Lecture – The Glass Ceiling". *Economica* 85: 205–231.

Bertrand, M., Goldin, C., Katz, L. F., 2010. "Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors". *American Economic Journal: Applied Economics* 2(3): 228-255.

Bertrand, M., Black, S., Lleras-Muney, A., Jensen, S., 2017. "Breaking the glass ceiling? The effect of board quotas on female labor market outcomes in Norway". University of Chicago Working Paper.

Boschini, A., Gunnarsson, K., 2018. "Gendered trends in income inequality". *Nordic Economic Policy Review* 2018: 100–127.

Boschini, A., Gunnarsson, K., Roine, J., 2017. "Women in Top Incomes: Evidence from Sweden 1974–2013". IZA working paper 10979.

Boschini, A., Håkanson, C., Rosén, Å., Sjögren, A., 2011 "Trading Off or Having it All? Completed Fertility and Mid-Career Earnings of Swedish Men and Women". Institute for Labour Market Evaluation Working paper 2011:15.

Burkhauser, R.V., Feng, S., Jenkins, S.P., Larrimore, J., 2012, "Recent trends in top income shares in the United States: reconciling estimates from March CPS and IRS tax return data". *Review of Economics and Statistics* 94 (2): 371-388.

Burkhauser, R. V., Hahn, M. H., Wilkins, R., 2016. "Top Incomes and Inequality in Australia: Reconciling Recent Estimates from Household Survey and Tax Return Data". Melbourne Institute Working Paper No. 19/16.

Burkhauser, R.V., Hérault, N., Jenkins, S. P., Wilkins, R., 2018. "Top incomes and inequality in the UK: reconciling estimates from household survey and tax return data". *Oxford Economic Papers* 70(2): 301–326.

Butikofer, A., Jensen, S., Salvanes, K.J., 2018. "The Role of Parenthood on the Gender Gap among Top Earners". *European Economic Review* 109: 103-123.

Campa, P., Serafinelli, M., 2018. "Politico-economic Regimes and Attitudes: Female Workers under State-socialism". *Review of Economics and Statistics (forthcoming)*.

Cortés, P., Tessada, J., 2011, "Low-Skilled Immigration and the Labor Supply of Highly Skilled Women". *American Economic Journal: Applied Economics,* 3(3): 88–123.

Cortés, P., Pan, J., 2018. "When Time Binds: Substitutes for Household Production, Returns to Working Long Hours, and the Skilled Gender Wage Gap". *forthcoming Journal of Labor Economics.*

Edlund, L., Kopczuk, W., 2009. "Women, Wealth and Mobility". *American Economic Review* 99(1): 146–178.

Folke, O., Rickne, J., 2016. "All the Single Ladies: Job Promotions and the Durability of Marriage". IFN working paper 1146.

Fortin, N., 2005. "Gender Role Attitudes and Women's Labour Market Outcomes across OECD Countries". *Oxford Review of Economic Policy* 21(3): 416–438.

Garbinti B., Goupille-Lebret, J., Piketty, T., 2018. "Income Inequality Dynamics in France 1900-2014: Evidence from Distributional National Accounts (DINA)", *Journal of Public Economics* 162: 63-77.

Goldin, C., 2014. "A Grand Gender Convergence: Its Last Chapter". *American Economic Review* 104(4): 1091–119.

Goldin, C., 2006, "The Quiet Revolution That Transformed Women's Employment, Education, and Family". *AEA Papers and Proceedings* 96(2):1-21.

Goldin, C., Katz, L. F., Kuziemko, I., 2006. The Homecoming of American College Women: The Reversal of the College Gender Gap". *Journal of Economic Perspectives* 20(4): 133–56.

Goldin, C., 1997. "Career and Family: College Women Look to the Past". In: Ehrenberg, R., Blau, F., *Gender and Family Issues in the Workplace*. New York: Russell Sage Foundation Press; 20-58.

Gornick, J., Milanovic, B., Morelli, S., Yonzan, N., 2018. "Comparing Top Incomes Between Survey and Tax Data: US Case Study". in Inequality Matters, LIS Newsletter 6: 10-11.

Greenwood, J., N. Guner, G. Vandenbroucke, 2017. "Family Economics Writ Large". *Journal of Economic Literature* 55 (4): 1346-1434.

Guvenen, F., Kaplan, G., Song, J., 2014. "The Glass Ceiling and The Paper Floor: Gender Differences among Top Earners 1981-2012". NBER Working Paper 20560.

Horrace, W., Oaxaca, R., 2006. "Results on the Bias and Inconsistency of Ordinary Least Squares for the Linear Probability Model". *Economics Letters* 90: 321-327.

Johnsen, J., Løken, K. V., 2015. "Nordic Family Policy and Maternal Employment". *Nordic Economic Policy Review* 2: 115-131.

Keloharju, M., Knupfer, S., Tag, J., 2016. "Equal Opportunity? Gender Gaps in CEO Appointments and Executive Pay". *Harvard Business School Research Paper* 16-092.

Keloharju, M., Knüpfer, S., Tåg, J., 2018. "What Prevents Female Executives from Reaching the Top?". "Working Paper Series 1111, Research Institute of Industrial Economics.

Kleven, H. J., Landais, C., 2017. "Gender Inequality and Economic Development: Fertility, Education, and Norms". *Economica* 84: 180-209.

Kleven, H. J., Landais, C., Søgaard, J. E., 2018. "Children and Gender Inequality: Evidence from Denmark". *American Economic Journal: Applied Economics, (forthcoming).*

Kleven, H., Landais, C., Posch, J., Steinhauer, A., Zweimüller, J. 2019. "Child Penalties across Countries: Evidence and Explanations". American Economic Association Papers and Proceedings 109(5): 122–126.

Leigh, A., (2009). "Top Incomes". In: Salverda, W., Nolan, B., Smeeding, T. (Eds.), *The Oxford Handbook of Economic Inequality*. Oxford: Oxford University Press.

Luxembourg Income Study (LIS) Database, http://www.lisdatacenter.org (multiple countries; May2018-July2019). Luxembourg: LIS.

Mincer, J., Polachek, S., 1974. "Family Investments in Human Capital: Earnings of Women". Journal of Political Economy 82(2): S76-108.

Olivetti, C., Petrongolo, B., 2008. "Unequal Pay or Unequal Employment? A Cross-Country Analysis of Gender Gaps". *Journal of Labor Economics* 26(4): 621-654.

Olivetti, C., Petrongolo, B., 2017. "The Economic Consequences of Family Policies: Lessons from a Century of Legislation in High-Income Countries". *Journal of Economic Perspectives* 31 (1): 205-30.

OECD, 2017. *The Pursuit of Gender Equality: An Uphill Battle*. OECD Publishing, Paris.

Piketty, T., 2001. *Les hauts revenus en France au 20ème siècle*. Paris: Grasset.

Piketty, T., 2003. "Income Inequality in France, 1901–1998". *Journal of Political Economy* 111(5): 1004–1042.

Piketty, T., Saez, E., 2003. "Income Inequality in the United States, 1913–1998". *Quarterly Journal of Economics* 118(1): 1–39.

Piketty, T., Saez, E., Zucman, G., 2016. "Distributional National Accounts: Methods and Estimates for the United States". NBER Working Paper No. 22945.

Piketty, T., Saez, E., Zucman, G., 2018. Distributional national accounts: methods and estimates for the United States, 1913–2013. *Quarterly Journal of Economics* 133(2): 553-609.

Ponthieux, S., Meurs, D., 2015. "Gender Inequality", in Atkinson, A. B., Bourguignon, F. (eds), *Handbook of Income Distribution 2*, Elsevier, 981-1146.

Ravaska, T., 2018. "Top Incomes and Income Dynamics from a Gender Perspective: Evidence from Finland 1995-2012". ECINEQ WP 2018-469.

Roine, J., Waldenström, D., 2012. "On the Role of Capital Gains in Swedish Income Inequality". *Review of Income and Wealth* 58(3): 569–587.

Roine, J., Waldenström, D., 2015. "Long-Run Trends in the Distribution of Income and Wealth". in Atkinson, A. B., Bourguignon, F. (Eds.), *Handbook of Income Distribution, Volume II*, Elsevier Science, North Holland, 469-593.

Smith, N., Smith, V., Verner, M., 2013. "Why are So Few Females Promoted into CEO and Vice President Positions? Danish Empirical Evidence, 1997–2007". *ILR Review* 66(2): 380-408.

Table A1. Number of non-missing observations of personal labour income included in the study.

COUNTRIES	HIST. DATA	WAVE I (~ 1980)	WAVE II (~ 1985)	WAVE III (~ 1990)	WAVE IV (~ 1995)	WAVE V (~ 2000)	WAVE VI (~ 2004)	WAVE VII (~ 2007)	WAVE VIII (~ 2010)	WAVE IX (~ 2013)	WAVE X (~ 2016)
Australia		42275	20445	39098	17986	16895	24674	22932	42531		
Austria					8950	6675	13042	13621	13933	12982	
					7962						
Belgium			18323	11139	7464	6399					
				10770	12260						
Canada				30739	53280	86614	72850	68542	64783	60362	
					97973	79433					
Czech Rep				43234	71836	18965	10333	26933	20629	18210	
Denmark			25621	25771	173177	175439	177269	179666	180266	183962	
Estonia							11934	12942	13331	14972	
Finland			34093	32381	25229	27841	29112	26481	23018	27142	
France		28779	33241	24335	25215	25803	25364		41285		
Germany	135088	128810	118367			28890	26824	26740	44134	41657	36949
						30261		24999	42531		
Greece					14300	11223	14897	16785	14913	20850	
Hungary				5848	5423	5469	5276	4298			
Ireland					11077	8638	15534	12545	10994		
					9692						
					8862						
Israel		8436	18616	19132	10980	19555	21046	20364	20225	28862	
Italy			25068	25150	24013	20901	20581	19907	19836	19366	
			25092	24930	23924	22268					
Luxembourg			6044	5498	4981	6226	9661	10146	14884	9977	
					6632						
Mexico			23985	57289	60353	48110	91738	118927	107781	33726	
				50862	64916	42535					
						72602					
Netherlands			13205	10861	13029	11661	23756	25448	25461	24494	
			10731								
Norway			25775	14271	24451	26305	34851	33989	468033	489750	
Paraguay						37435	34636	21046	20475	21185	37814
Poland			34201					111992	107967	102780	
Slovakia				47714			15379	16416	15334	15711	
Slovenia					8639	12658	11303	11094	11515	10805	
Spain		48655		72119			37491	35832	34495	31622	
Sweden	14857	24535	21589	28195	34208	33139	36918				
	29277										
Switzerland		16324		16784				16397	17602	15651	
Taiwan		73309	74441	68439	39273	49793	46386	46230	47900	50518	50569
					52491						
UK					62821	59010	64942	56926	57928	46166	
US	34244	181488	155407	155796	149642	218269	210648	206404	204983		
					131599						

Table A2. Number of women in top 1 percent of personal labour income included in the study.

COUNTRIES	HIST. DATA	WAVE I (~ 1980)	WAVE II (~ 1985)	WAVE III (~ 1990)	WAVE IV (~ 1995)	WAVE V (~ 2000)	WAVE VI (~ 2004)	WAVE VII (~ 2007)	WAVE VIII (~ 2010)	WAVE IX (~ 2013)	WAVE X (~ 2016)
Australia		16	15	32	31	24	31	22	35		
Austria					5	3	14	19	21	28	
					1						
Belgium			12	13	4	7					
				5	19						
Canada				15	41	72	67	75	65	59	
					41	50					
Czech Rep				43	67	35	12	24	21	17	
Denmark			8	6	122	143	157	170	186	231	
Estonia							16	9	18	23	
Finland			40	54	33	64	61	52	66	72	
France		23	18	19	13	28	33		48		
Germany	13	20	18			13	38	39	42	47	61
						36		32	47		
Greece					13	4	19	23	14	24	
Hungary				11	10	15	14	12			
Ireland					4	4	23	13	17		
					1						
					3						
Israel		6	9	14	13	13	15	23	18	39	
Italy			9	19	12	15	33	21	28	30	
			27	19	16	22					
Luxembourg			4	5	6	1	9	10	24	15	
					1						
Mexico			9	33	30	26	95	153	138	41	
				22	38	13					
						42					
Netherlands			8	8	10	10	10	19	32	31	
			4								
Norway			31	3	5	7	39	24	378	532	707
Paraguay						53	31	28	39	29	42
Poland			3					213	244	213	
Slovakia				50			34	23	28	31	
Slovenia					16	26	27	33	42	34	
Spain		3		67			50	45	85	47	
Sweden	4	30	10	15	25	26	49				
	26										
Switzerland		6		15				12	9	9	
Taiwan		16	26	32	26	48	62	56	73	84	75
					42						
UK					40	51	46	61	66	69	
US	4	93	133	157	150	195	211	286	274		
					151						

NOTE. To gauge the number of women observed in the top10 group, recall that with the same share of women in top10 and top1 it would be ten times as large, but given that the share of women in the top 10 group is typically 1,5-2 times the share in top1, it is more 15-20 times the above number.

. Comparing share of women in top10 and top1 in the capital + labour income distribution, in the total income distribution, and in the labour distribution

Figure B1. Share women in different top groups, Anglo-Saxon countries.
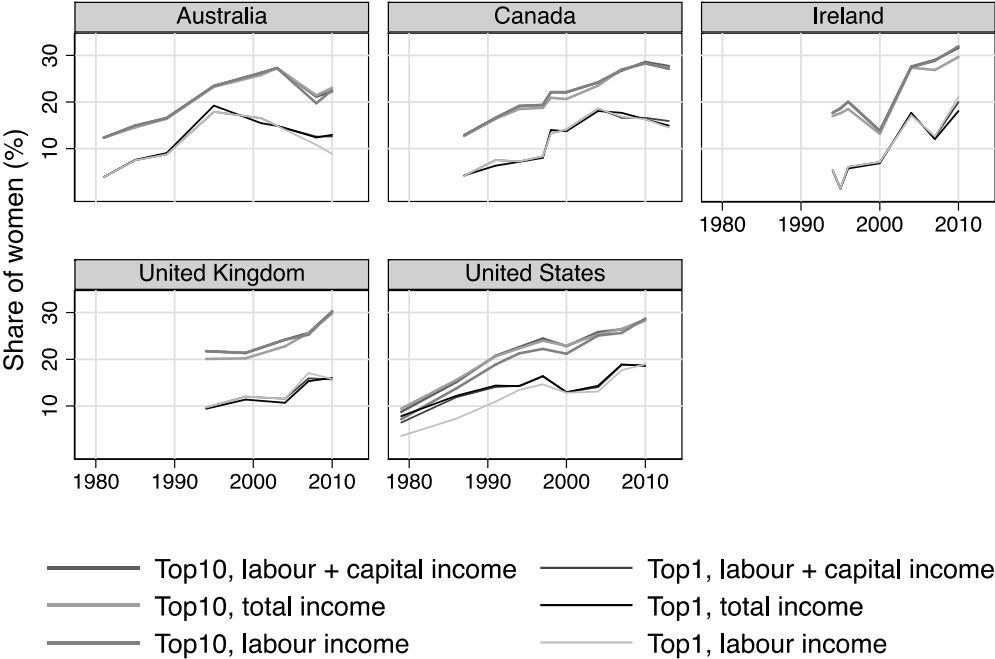
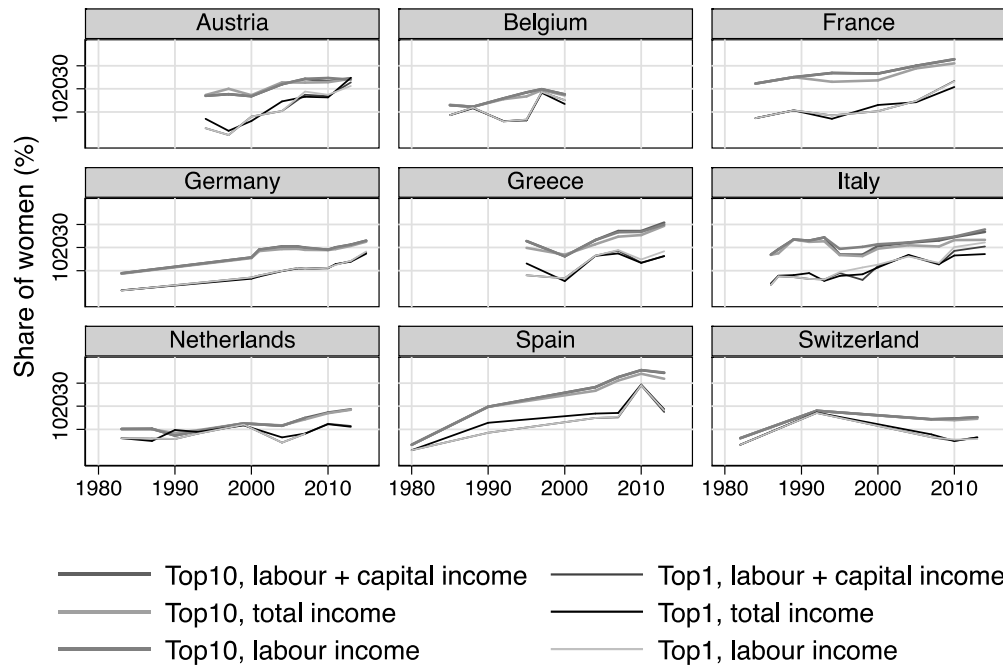*Figure B2. Share women in different top groups, Continental Europe.*

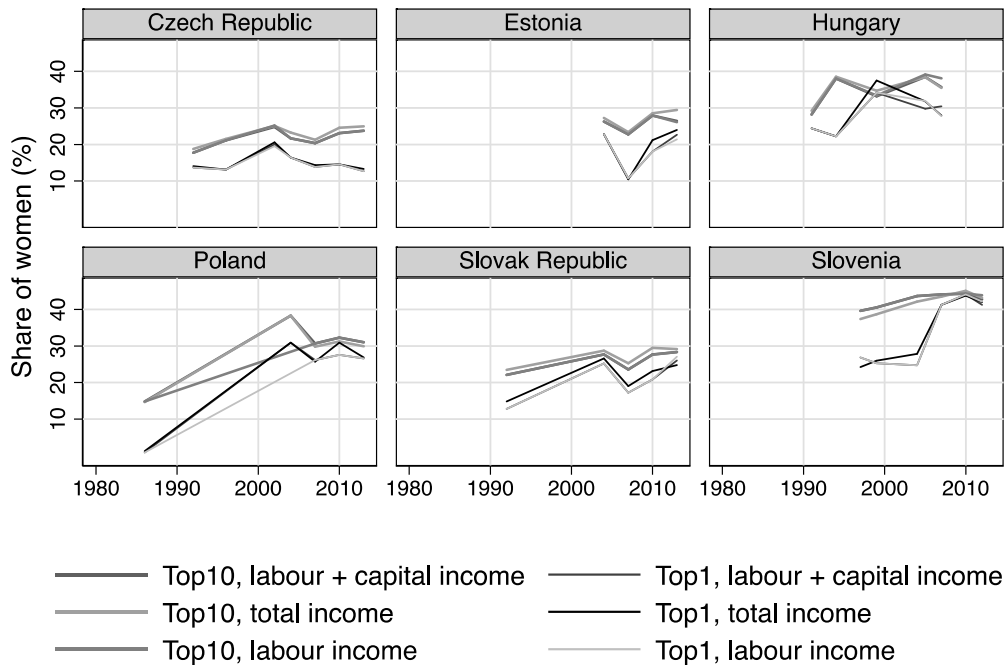*Figure B3. Share women in different top groups, Eastern Europe.*

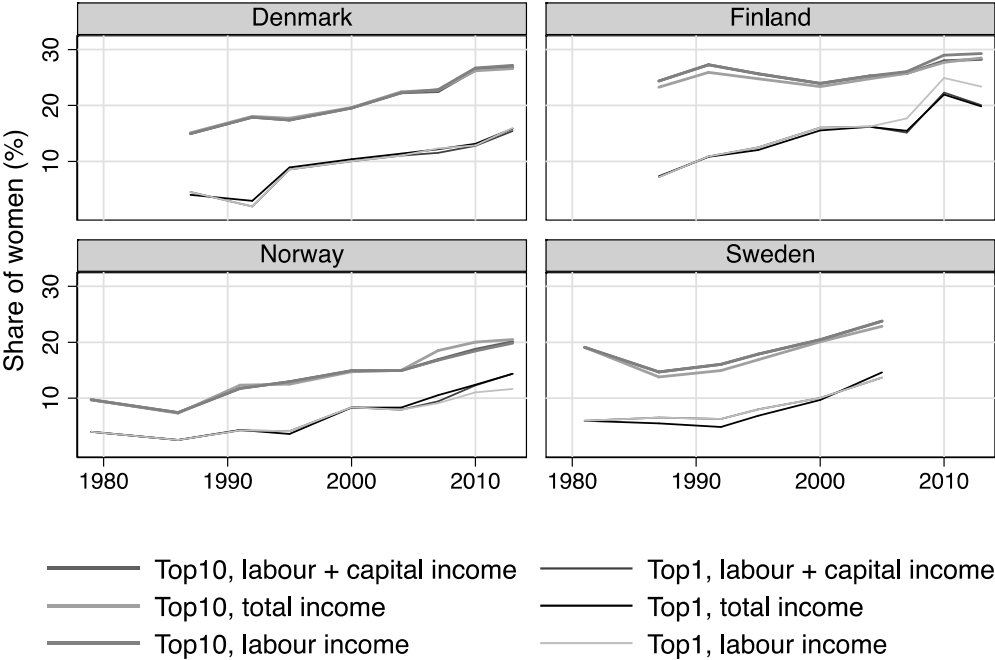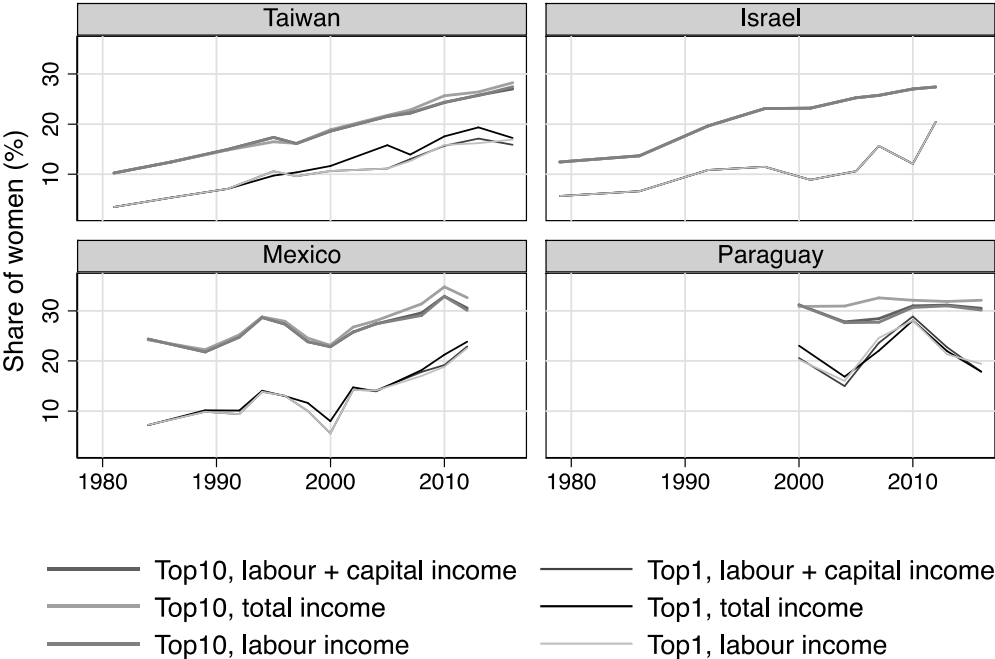*Figure B4. Share women in different top groups, Nordic countries.*

Figure B5. Share women in different top groups, Taiwan, Israel, Mexico and Paraguay.

Appendix C. Linear probability regression results for country groups separately.

Table C1. Linear probability regression of the likelihood of being in a top group for the pooled sample of Anglo-Saxon countries.

	Men P90	Women P90	Men P99	Women P99
Age in years	0.0194***	0.00656***	0.00231***	0.000429***
	(0,00014)	(0,00007)	(0,00004)	(0,00002)
Age squared	-0.000202***	-0.0000678***	-0.0000222***	-0.00000433***
	(0,00000)	(0,00000)	(0,00000)	(0,00000)
Medium educ.	0.0730***	0.00957***	0.00548***	0.000318***
	(0,00095)	(0,00031)	(0,00027)	(0,00008)
High educ.	0.285***	0.108***	0.0471***	0.00698***
	(0,00132)	(0,00076)	(0,00058)	(0,00021)
Living w. partner	0.0435***	-0.00647***	0.00317***	-0.000326*
	(0,00115)	(0,00059)	(0,00041)	(0,00015)
One child	0.0381***	-0.00443***	0.00327***	-0,000269
	(0,00149)	(0,00070)	(0,00057)	(0,00018)
Two children	0.0903***	-0.00645***	0.0133***	0.000570*
	(0,00172)	(0,00081)	(0,00071)	(0,00023)
Three children	0.0909***	-0.0197***	0.0194***	-0,000312
	(0,00249)	(0,00097)	(0,00111)	(0,00028)
Three+ chldren	0.0600***	-0.0278***	0.0173***	-0,000566
	(0,00349)	(0,00112)	(0,00163)	(0,00038)
Constant	-0.301***	-0.138***	-0.0434***	-0.0110***
	(0,00585)	(0,00254)	(0,00215)	(0,00061)
Observations	884399	992741	884399	992741
R-squared	0,169	0,068	0,031	0,005

All regressions include fixed effects for LIS wave and country.

Countries included: Australia, Canada, Ireland, United Kingdom, United States

Standard errors in parentheses.

* p<0.05, ** p<0.01, *** p<0.001

Table C2. Linear probability regression of the likelihood of being in a top group for the pooled sample of Continental European countries.

	Men P90	Women P90	Men P99	Women P99
Age in years	0.0200***	0.00707***	0.00312***	0.000671***
	(0,00025)	(0,00013)	(0,00009)	(0,00004)
	-	-	-	-
Age squared	0.000217***	0.0000720***	0.0000298***	0.00000605***
	(0,00000)	(0,00000)	(0,00000)	(0,00000)
Medium educ.	0.102***	0.0387***	0.0136***	0.00110***
	(0,00184)	(0,00093)	(0,00071)	(0,00029)
High educ.	0.327***	0.151***	0.0643***	0.00945***
	(0,00238)	(0,00185)	(0,00118)	(0,00065)
Living w. partner	0.0239***	-0.0173***	0,0014	-0,00039
	(0,00214)	(0,00110)	(0,00077)	(0,00034)
One child	0.0496***	-0.00890***	0.00505***	-0.00193***
	(0,00228)	(0,00116)	(0,00090)	(0,00034)
Two children	0.105***	-0.0103***	0.0137***	-0,00079
	(0,00260)	(0,00132)	(0,00104)	(0,00043)
Three children	0.0940***	-0.0200***	0.0221***	-0.00225***
	(0,00378)	(0,00179)	(0,00178)	(0,00063)
Three+ chldren	0.0636***	-0.0297***	0.0134***	-0.00338***
	(0,00607)	(0,00230)	(0,00293)	(0,00093)
Constant	-0.140***	0.0840***	0.131***	0.200***
	(0,00747)	(0,00415)	(0,00330)	(0,00208)
Observatons	547199	592880	547199	592880
R-squared	0,182	0,129	0,141	0,276

All regressions include fixed effects for LIS wave and country.

Countries included: Austria, Belgium, France Germany, Italy, Netherlands, Spain, Switzerland

Standard errors in parentheses.

* p<0.05, ** p<0.01, *** p<0.001

Table C3. Linear probability regression of the likelihood of being in a top group for the pooled sample of Eastern European countries.

	Men P90	Women P90	Men P99	Women P99
Age in years	0.00912***	0.00785***	0.00173***	0.00175***
	(0,00030)	(0,00020)	(0,00019)	(0,00014)
Age squared	-0.000120***	-0.0000809***	-0.0000179***	-0.0000160***
	(0,00000)	(0,00000)	(0,00000)	(0,00000)
Medium educ.	0.0710***	0.0352***	0.00766***	0.00463***
	(0,00205)	(0,00140)	(0,00130)	(0,00109)
High educ.	0.339***	0.218***	0.0692***	0.0172***
	(0,00349)	(0,00299)	(0,00235)	(0,00178)
Living w. partner	0.0662***	-0.0103***	0.00828***	0.00258*
	(0,00272)	(0,00170)	(0,00166)	(0,00109)
One child	0.0523***	0.0154***	0.00975***	0.00286**
	(0,00267)	(0,00173)	(0,00157)	(0,00111)
Two children	0.0872***	0.0159***	0.0178***	0.00320*
	(0,00315)	(0,00211)	(0,00179)	(0,00129)
Three children	0.0621***	0,0049	0.0237***	0.0120***
	(0,00463)	(0,00280)	(0,00254)	(0,00189)
Three+ chldren	0.0172**	-0,00692	0.0211***	0.0133***
	(0,00635)	(0,00362)	(0,00350)	(0,00271)
Constant	-0.148***	-0.231***	-0.122***	-0.136***
	(0,00730)	(0,00480)	(0,00419)	(0,00327)
Observatons	328589	380269	328589	380269
R-squared	0,294	0,405	0,52	0,611

All regressions include fixed effects for LIS wave and country.

Countries included: Czech Republic, Estonia, Hungary, Poland, Slovakia, Slovenia

Standard errors in parentheses.

* p<0.05, ** p<0.01, *** p<0.001

Table C4. Linear probability regression of the likelihood of being in a top group for the pooled sample of Nordic countries.

	Men P90	Women P90	Men P99	Women P99
Age in years	0.0227***	0.00712***	0.00313***	0.000482***
	(0,00023)	(0,00011)	(0,00007)	(0,00003)
Age squared	-0.000237***	-0.0000702***	-0.0000297***	-0.00000453***
	(0,00000)	(0,00000)	(0,00000)	(0,00000)
Medium educ.	0.0756***	0.0183***	0.00953***	0.000806***
	(0,00161)	(0,00075)	(0,00048)	(0,00016)
High educ.	0.285***	0.111***	0.0557***	0.00690***
	(0,00226)	(0,00140)	(0,00099)	(0,00035)
Living w. partner	0.0371***	-0.00990***	0.00348***	-0.00109***
	(0,00178)	(0,00096)	(0,00058)	(0,00024)
One child	0.0544***	0.00435***	0.00680***	0.000752**
	(0,00229)	(0,00115)	(0,00085)	(0,00028)
Two children	0.100***	0.00803***	0.0137***	0.000894**
	(0,00249)	(0,00125)	(0,00094)	(0,00029)
Three children	0.0910***	-0.00596***	0.0202***	0,00051
	(0,00364)	(0,00167)	(0,00153)	(0,00042)
Three+ chldren	0.0265***	-0.0123***	0.00808***	(0,00036)
	(0,00631)	(0,00281)	(0,00238)	(0,00051)
Constant	-0.406***	-0.135***	-0.0680***	-0.0110***
	(0,00580)	(0,00271)	(0,00187)	(0,00059)
Observatons	990203	1045498	990203	1045498
R-squared	0,16	0,065	0,034	0,005

All regressions include fixed effects for LIS wave and country.

Countries included: Denmark, Finland, Norway, Sweden

Standard errors in parentheses.

* p<0.05, ** p<0.01, *** p<0.001

Appendix D. Share of top men and women with a partner for individual countries.

Figure D1. Share with a partner in top10 and top1, women and men, Anglo-Saxon countries.
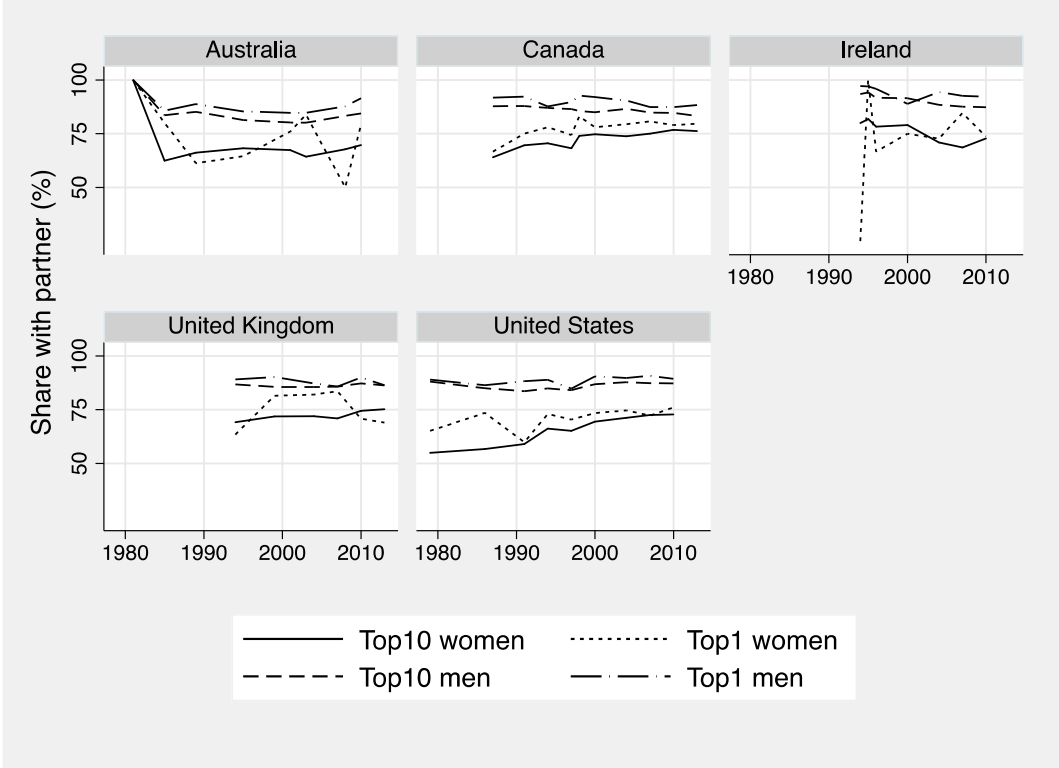
Figure D2. Share with partner in top10 and top1, women and men, Continental European countries.
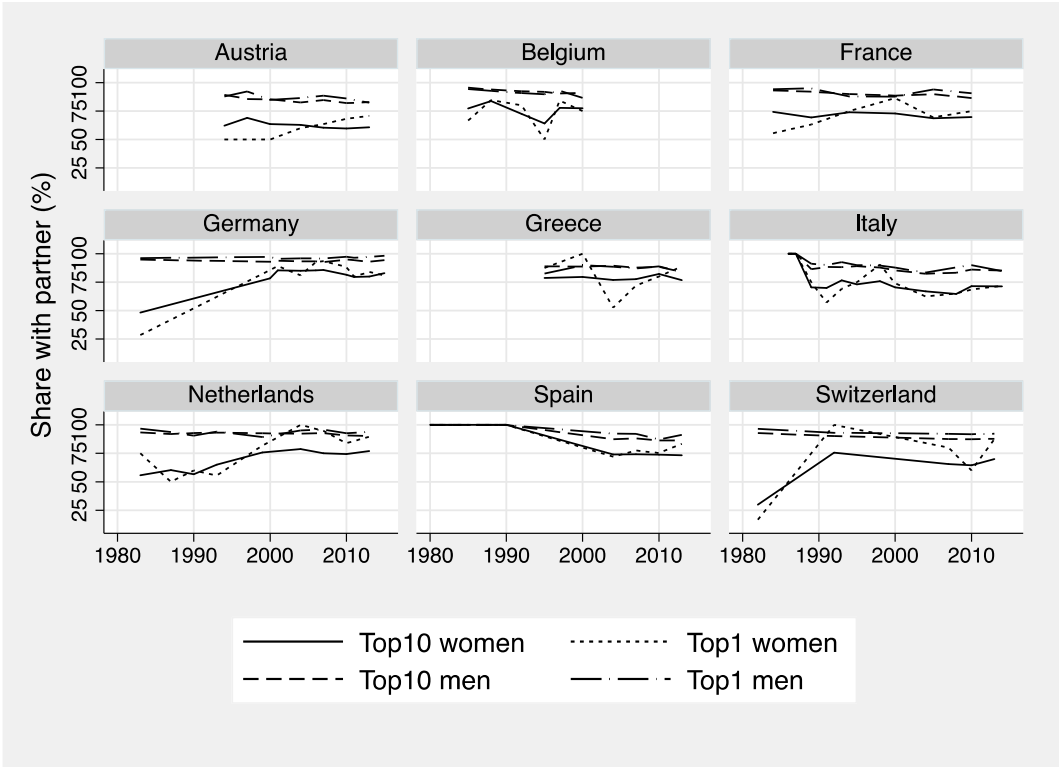
Figure D3. Share with partner in top10 and top1, women and men, Central European countries.
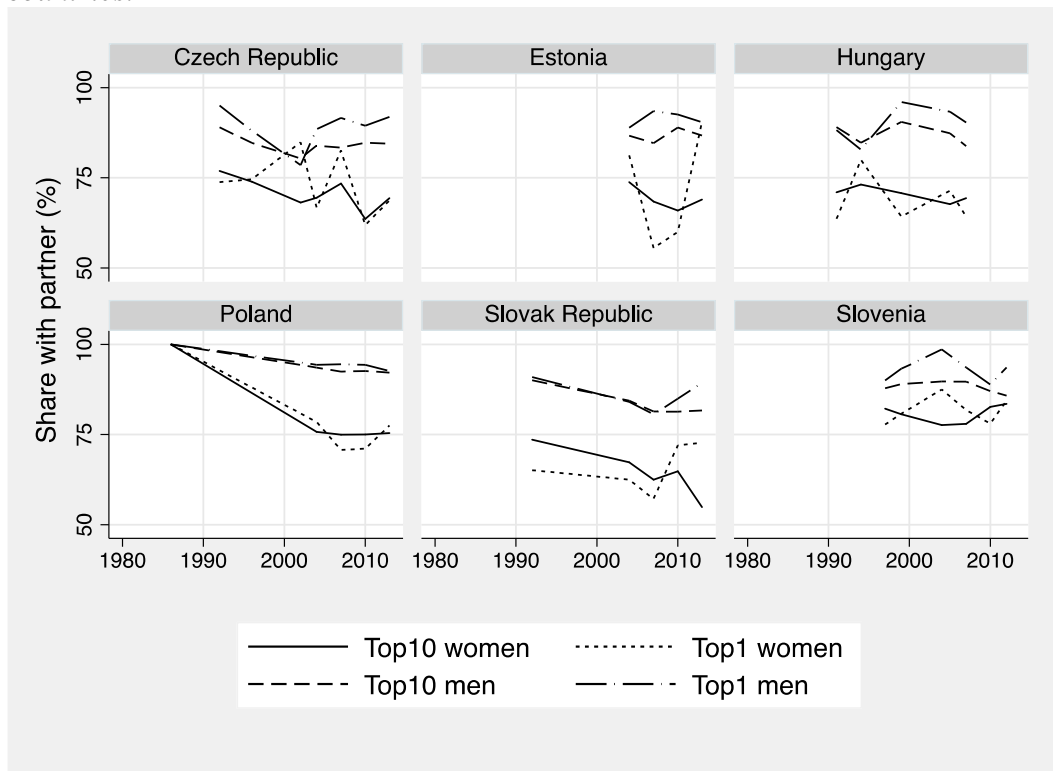
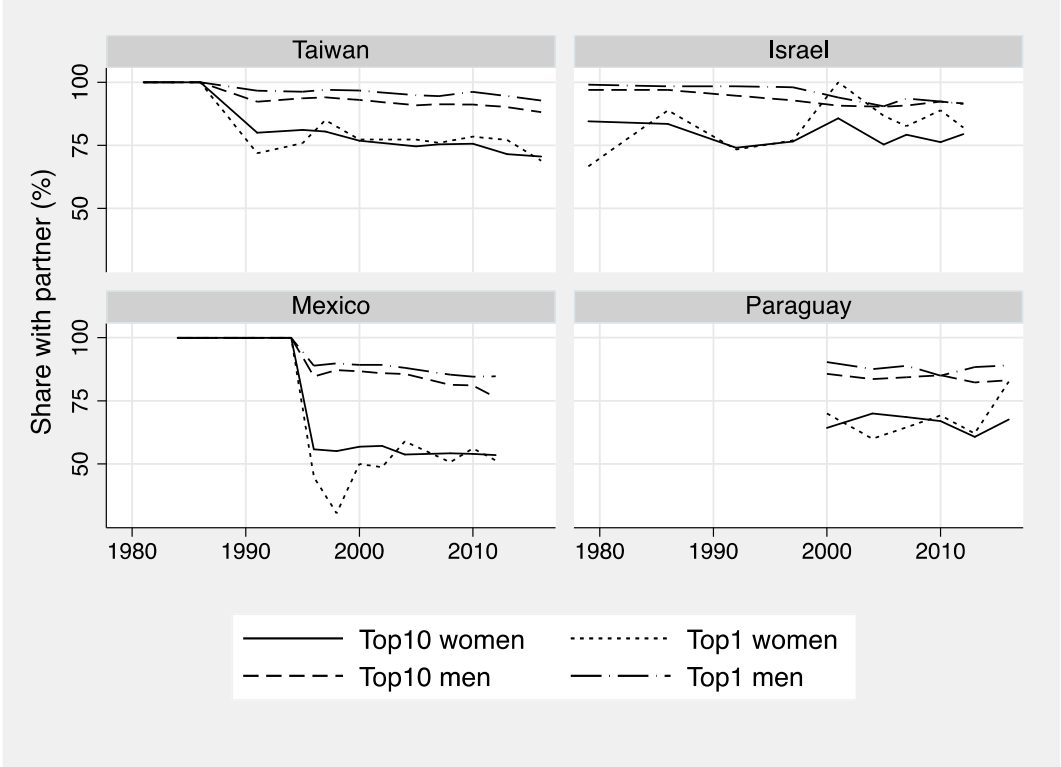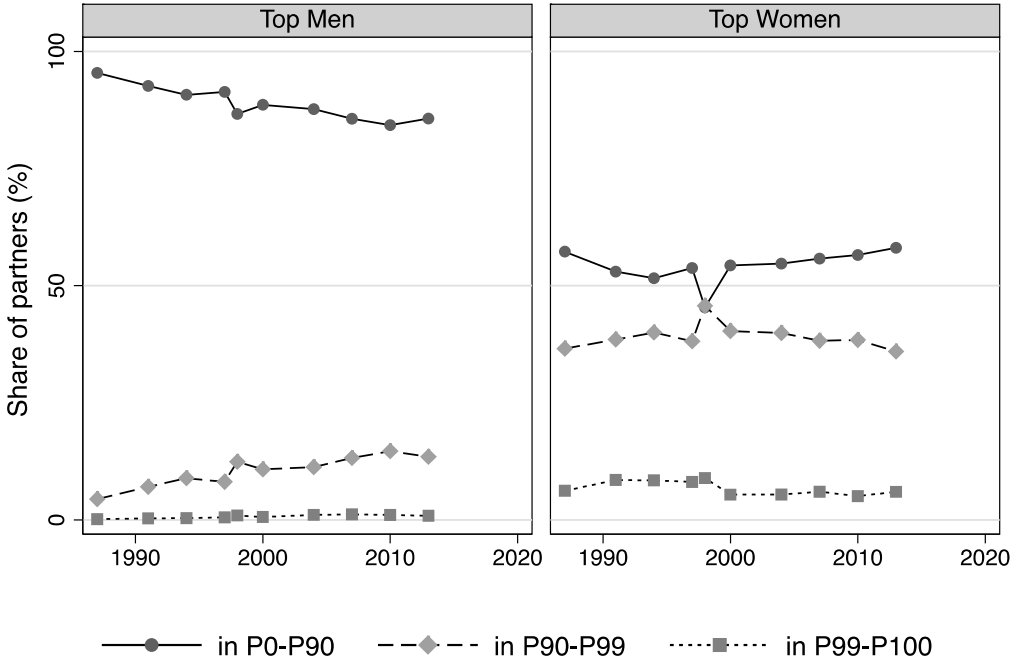*Figure D4. Share with partner in top10 and top1, women and men, Nordic countries.*
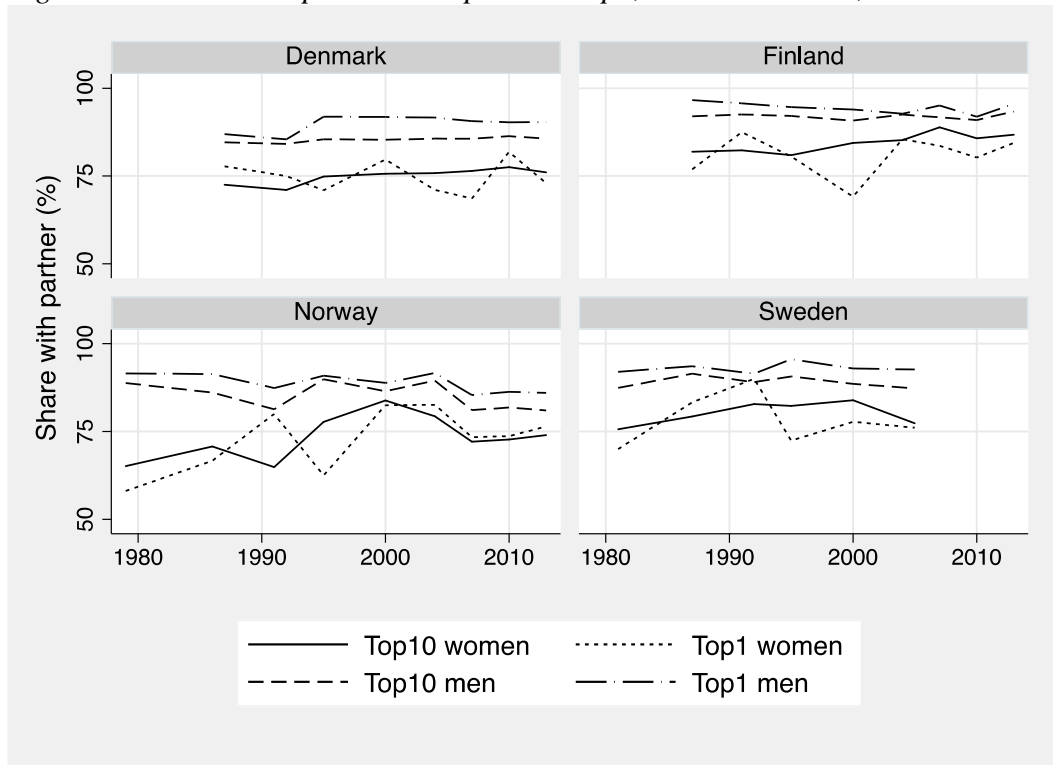
Figure D5. Share with partner in top10 and top1, women and men, Taiwan, Israel, Mexico and Paraguay.

Canada

Denmark
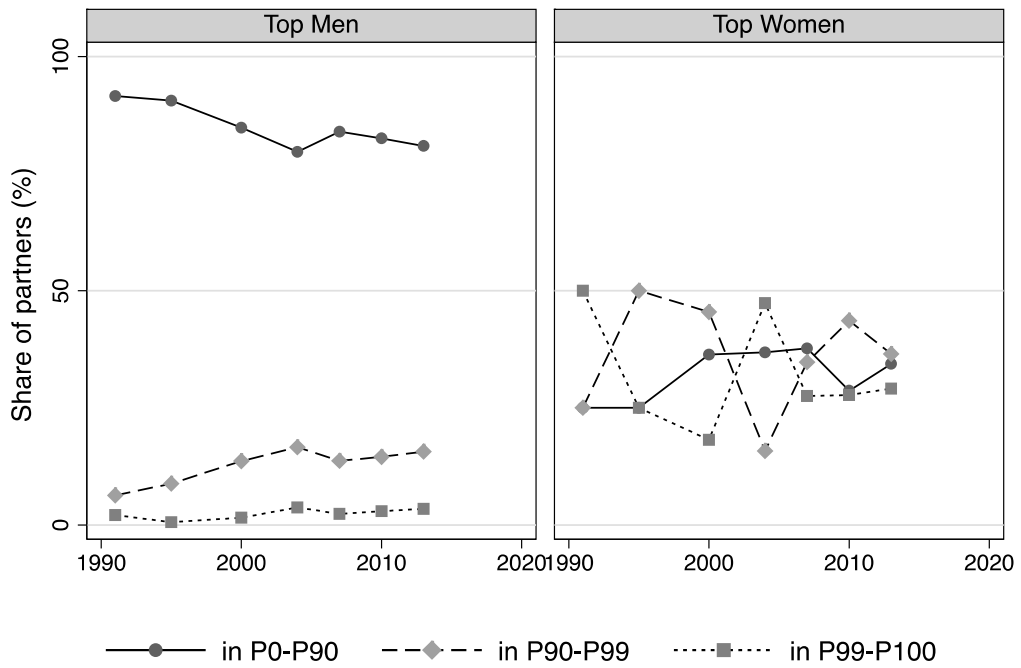

Graphs by women
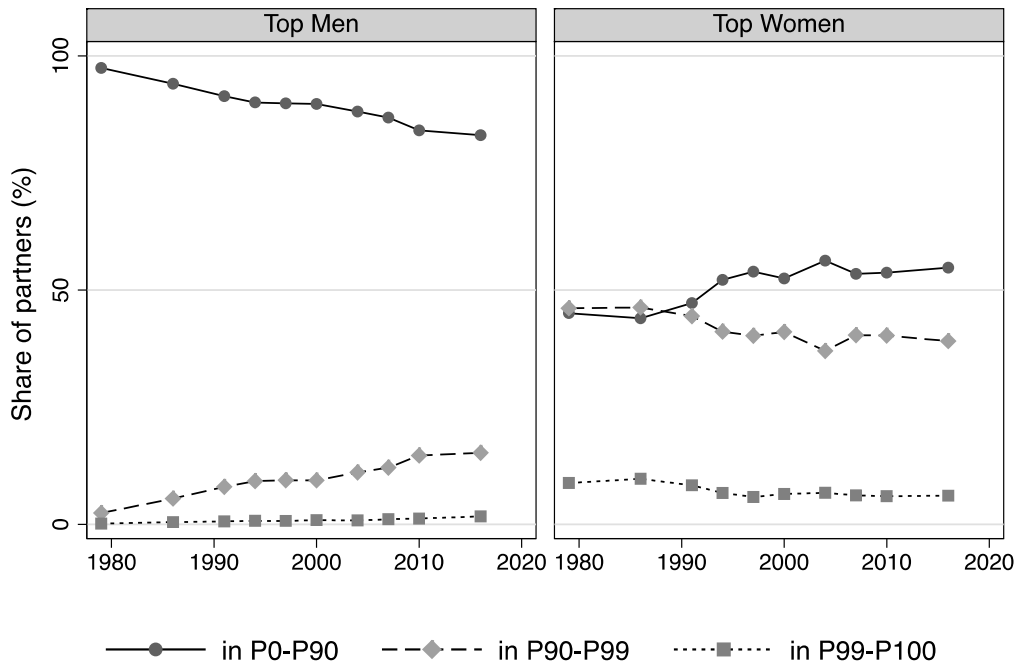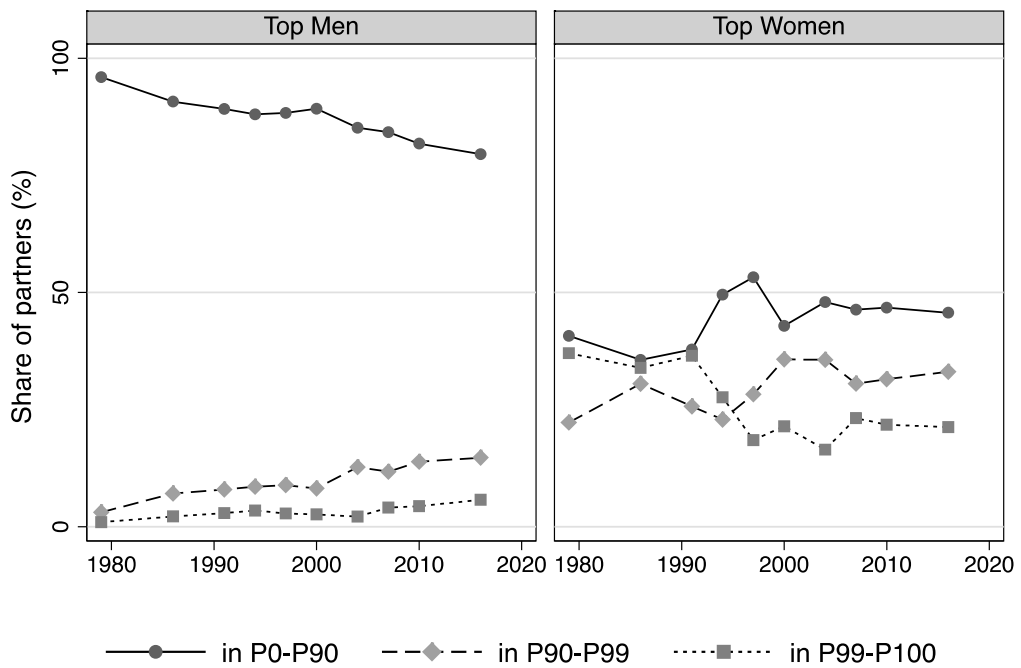

Graphs by women

47
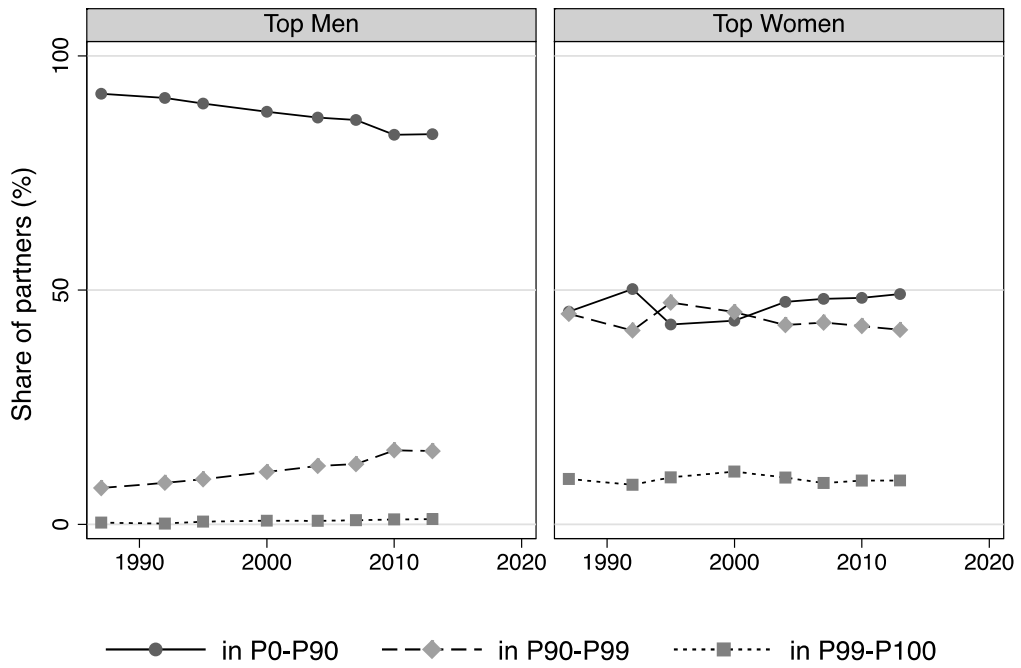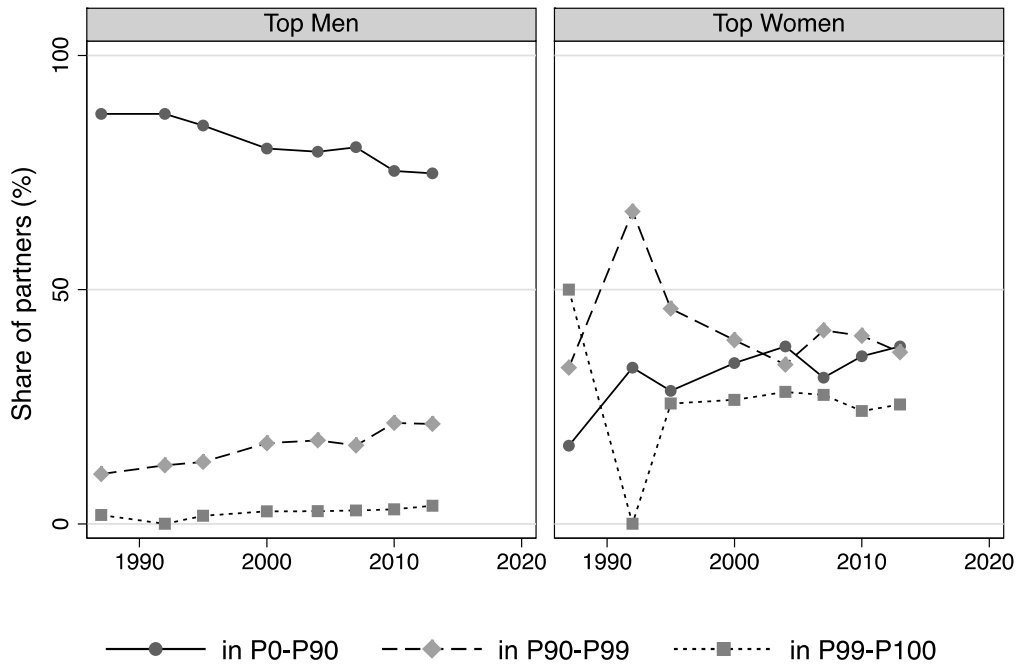
Finland
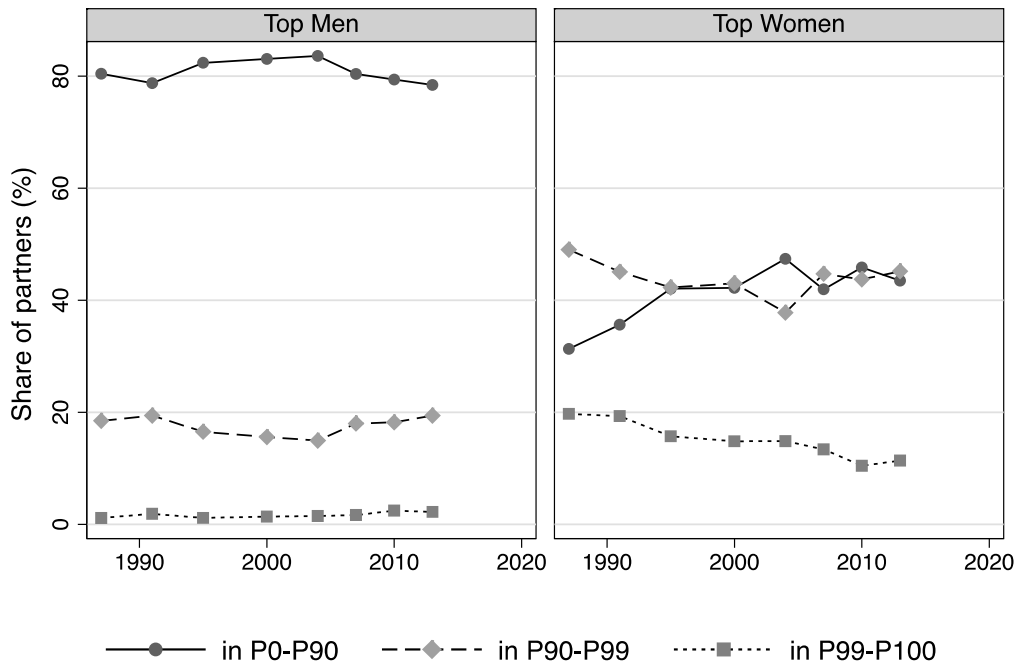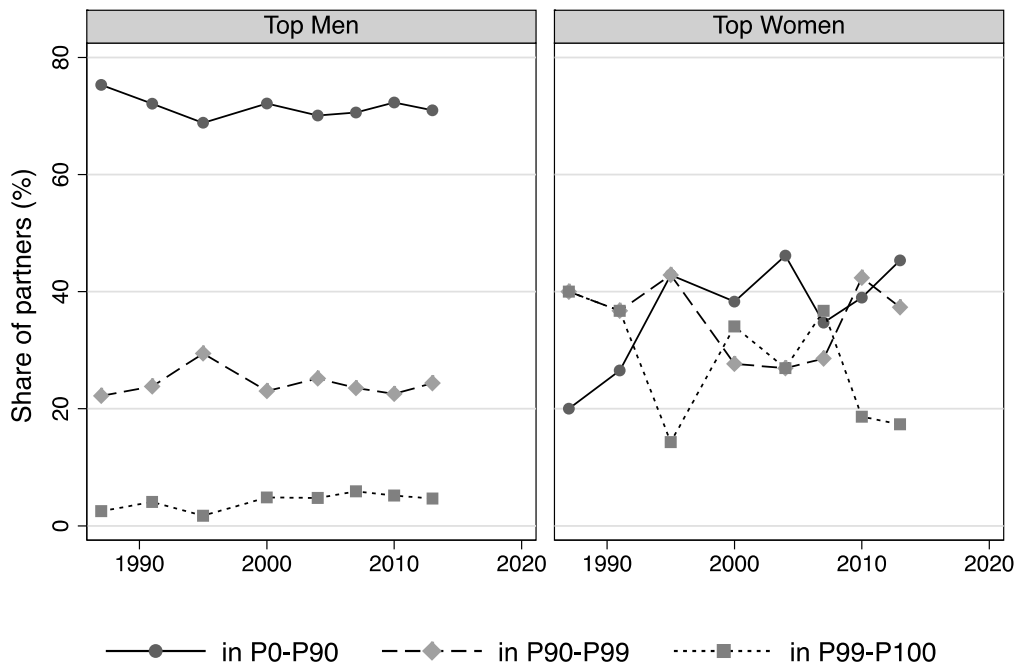
Norway
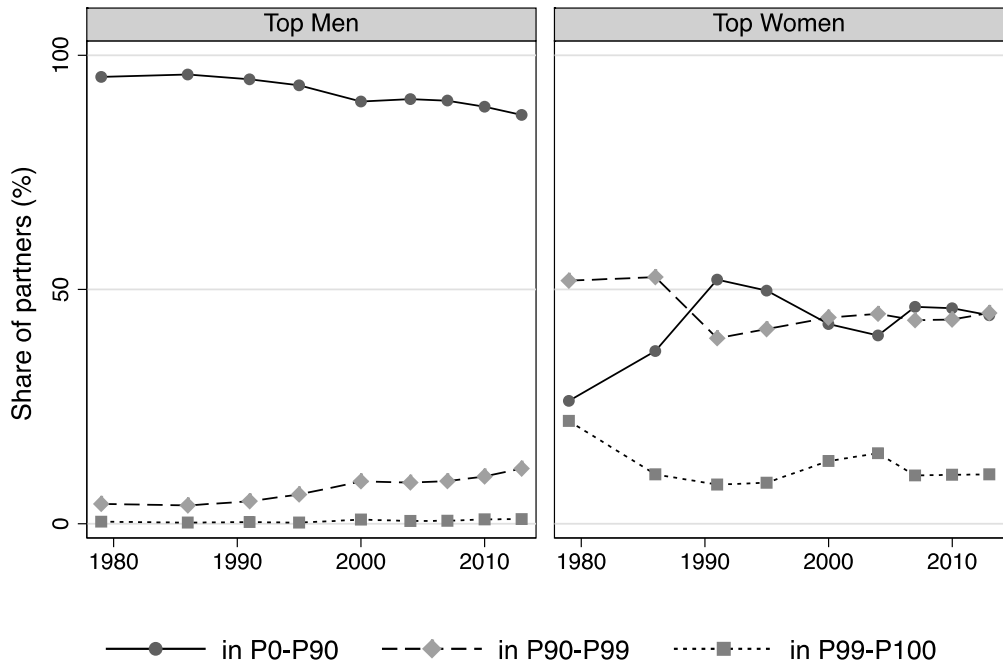
Graphs by women

Graphs by women

USA

Graphs by women

Graphs by women

50