

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

Roman Bobilev*

Anne Boschini*

Jesper Roine♦

October 26, 2018

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 find that, while sample sizes and lack of coverage of all sources of income (in particular, lack of capital incomes) for most years and most countries limits possibilities to arrive at robust findings, especially for the very top groups, LIS-data does give important suggestive evidence of patterns worth looking into using more detailed country-specific data sources. We show developments of the share of women in top groups (P90-100, and P99-100) of the labour income distribution for 28 countries. When possible, we compare to outcomes when including capital incomes and find that they do not matter much for the share of women with some important exceptions and caveats. We compare our findings to the existing evidence on women in top groups based on aggregate tax data and find that LIS data seems to give a relatively accurate picture of the basic findings, but we also note that once we would like to divide the top1 group further samples quickly become too small to allow further study. For a limited set of countries where samples are somewhat larger we also look at partner composition of top income earners. We find a strong asymmetry in that top income men everywhere typically have a partner who is not in the top of the distribution while top income women much more often have a partner who is also a top earner. This illustrates the importance of combining data on individual characteristics with household characteristics to understand determinants of gender differences in top incomes.

Keywords: Income inequality, income distribution, top income women, gender inequality, top incomes, capital incomes,

JEL: D13; D31; H20; J16; J31

* Department of Economics and SITE, Stockholm School of Economics, roman.bobilev@hhs.se

♦ SOFI, Stockholm University, anne.boschini@sofi.su.se

♦ SITE, Stockholm School of Economics, nejr@hhs.se

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) and Piketty and Saez (2003) 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, namely the gender dimension. Given that we observe a growing share of total incomes going to the top group, 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 these questions. 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. In the following section we then 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,

¹ The collected volumes by Atkinson and Piketty (2007, 2010) contain much of this 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 gender equality in the population more broadly, recently in 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.

compare 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, Atkinson, Casarico and Voitchovsky, 2018). Comparisons suggest that our LIS-based results are not perfectly aligned, but come very close to the shares of women in top groups observed in income tax data. We then move on to use the fact that the individual LIS-data has information about the household. In particular, we use the information on partner income to study gender differences in partner composition of top income men and women respectively. We find suggestive evidence that top income men are much more likely to have a partner who is not in the top of the income distribution, while top income women much more often have partners who are also in the top of the distribution. This section, however, also serves as an illustration of when we quickly approach the limit of what can be meaningfully studied using LIS-data. 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 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 continues 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? – this fact creates a problem 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 is 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 periods when men and women file taxes independently. Atkinson, Casarico, and Voitchovsky (2018) study the share

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).

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 but given joint tax filing for married couples they are restricted to differences stemming from labour earnings. Similarly, 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 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 that income compositions typically differ. In several countries other types of 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 have mostly been converging. Using individual panel data, Boschini et al. (2017) also show that there are 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 or number of children, but in terms of partner income. 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 above points about there being little research on gender aspects in the top income literature do 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; Arulamplam, Booth and Bryan, 2006; Albrecht, Skogman Thoursie, and Vroman, 2015) are explicitly about gender differences in the top of the distribution. Recent work by Guvenen, Kaplan, and Song (2014) study gender dimensions of top wage earners in the US, 1981-2014. In a recent 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’.”

The exact phrasing is important here; it is the top of the *earnings* distribution that is in focus. In general, the focus in this literature has been on labour market outcomes, hence, excluding capital income, which is known to be important especially in the top of the distribution. Moreover, the population is also typically restricted to the working age population and comparisons of wage gaps are typically made conditional on full time work. 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.

If one instead were to think about top income women through the lens of the standard income inequality literature, the analysis could be very different. Much of inequality research is concerned with disposable incomes adjusting for household size and composition using some

version of an equivalence scale. In such a setting the standard assumption is to treat the entire household as the income earning unit and then divide incomes (gender) equally within the household. This, of course, would mean that many women (and men) would end up in the top of the distribution of disposable incomes even if they earn nothing themselves. Even if this, in its starkest form, is unlikely to cause conceptual problems, there are situations where one may have to think more carefully about the connections between, on the one hand, gender differences between “atomistic” income earning units and, on the other hand, men and women who form households and have individual earnings from labour and potentially also from capital and other sources, but also joint income flows (e.g. from capital gains from jointly owned assets).

Taken together the different research strands illustrates 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. As far as possible we want to include income from all sources before taxes and transfers. The reference population is ideally the full adult population. 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 Gornick, Milanovic, Morelli, and Yonzan (2018) compare the coverage of top incomes in LIS to fiscal data for the entire tax population (the source underlying data in the WID data base used in the top income literature) focusing on the U.S., 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 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 going 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.⁴

Our challenges are slightly different. We want to primarily 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. Here we exclude realized capital gains, though we note that these may be important in general (see, Roine and Waldenström, 2012) and also for the representation of women in top groups (see, Boschini, Gunnarsson, and Roine, 2017). To the extent that whatever it is that we are missing in the top is not systematically different between men and women the under reporting need not bias the results for the *share* of women in the top group. However, if capital, for example, is more (or less) important for women and also to a larger extent missing in the top this will lead us to underestimate (or overestimate) the presence of women in the group. We will use the available data, as well as comparisons to previous estimates using tax data, to estimate the extent of this problem.

In terms of the unit of analysis and the reference population we make use of the LIS person-level data files to create a distribution of income for the full adult population. We then look at the top10 group and the top1 group, respectively, and report the share of women in this over time. As the types of income included in LIS change over time, we study labour income (as defined in LIS, that is including self-employment income) and labour plus capital income, and total income (as defined in LIS) separately. As we will explain below our main series end up being for the share of women in the top of the labour income distribution, but for the one country where we have sufficient data as well as coverage going back in time, the U.S. starting in 1979, we note that disregarding capital incomes when looking at the share of women in the top going back in time is not without problems. For the more recent years (after 2007), when we observe capital income as well, for most countries, the problem seems to be smaller and less systematic.

⁴ 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, Héroult, 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.

2.1 Restrictions used to select data in LIS

To illustrate the challenge with respect to studying women in top groups using LIS data, consider the following back-of-the-envelope calculation: For a sample size of 50000 individuals (relatively large for 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 characteristics of these women (education, employment, age profile, etc.) quickly brings us down to sample sizes where it is no longer meaningful to proceed.

Another issue concerns the coverage of all income sources. In principal, LIS provides series for 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 (starting 1995), Germany (starting in 2001), and the U.S. (starting in 1979) have individual level capital income data. For all countries and years we end up using, we can of course compare labour income to total income, but the component we have reason to believe can make a substantial difference, capital, is limited in coverage. We will contrast the share of women using the different concepts but it is important to note that total income in LIS for the most part before 2007 does not include capital incomes.

Our approach to all this is to be as transparent as possible about what we do and let the reader judge how much the results can be trusted. We will in the end be cautious about interpreting many of our results but, obviously, we have made decisions where we think that what we arrive at is informative about the developments. Also, for the countries where it is possible, we can compare our results to studies based on observations of the entire tax population (or much larger samples).

With this in mind, we restrict the LIS data samples as follows. First, to be able to trace development of the share of women in top income groups in time, we restrict our study to countries with 5 or more years in the LIS data that costs us 15 countries out of 49 covered in

LIS. Second, we impose a restriction that a dataset must either exceed 0.0005 as a share of the country's population, or exceed 20000 observations. We make one exception, Italy, where we include the 3 most recent years of data despite observation count falling just short of the latter threshold (20000 observations). This leaves us with 223 country-years of LIS total of 329.⁵ Third, for a part of our analysis concerning positions in income distribution within couples, we restrict the sample even further, requiring at least 50 observations for women in the top1 percent of labour income distribution for a country-year, leaving us with only 8 countries and 80 country-years. Finally, we limit our study to the adult population, leaving only individuals over 18 years old in the data.

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 men and women. These cut-off points we use further to classify an individual as belonging to a particular percentile of the income distribution.

For the more detailed study of top income women and their partners, we rely on variables in LIS describing partnership, relation to the household head, and age. The part of the study dealing with share of men and women at the top having or not having a partner relies on *partner* variable in LIS. We classify an individual as having a partner if the variable describes him or her as having a partner (100), living with a partner (110) or not living with a partner (120). Finally, for the study of income distribution in couples, we define a couple based on two variables in personal-level data. To be considered a couple, the *partnership* variable that must describe an individual as having a partner or living with a partner, and *relation* variable must describe an individual as a household head, a spouse of a household head, or a cohabiting partner. We have refrained from using the marriage data also present in LIS, as the definition of marriage changes over time and across countries and thus did not provide sufficient certainty of the point in question.

⁵ 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.

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

Given the above 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). For all years we can of course also use the variable called “individual total income” but the components included in this change across countries and over time. However, as pointed out above, such changes need not impact the share of women, unless the different components of total income have very different gender profiles in the different top groups. As we will show the choice of variable across these choices does not change the share of women in top groups much with some important exceptions.

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 (somewhat ad hoc) 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 panels to the left show the share of women in the top10 group over time and to the right the share of women in the top1.

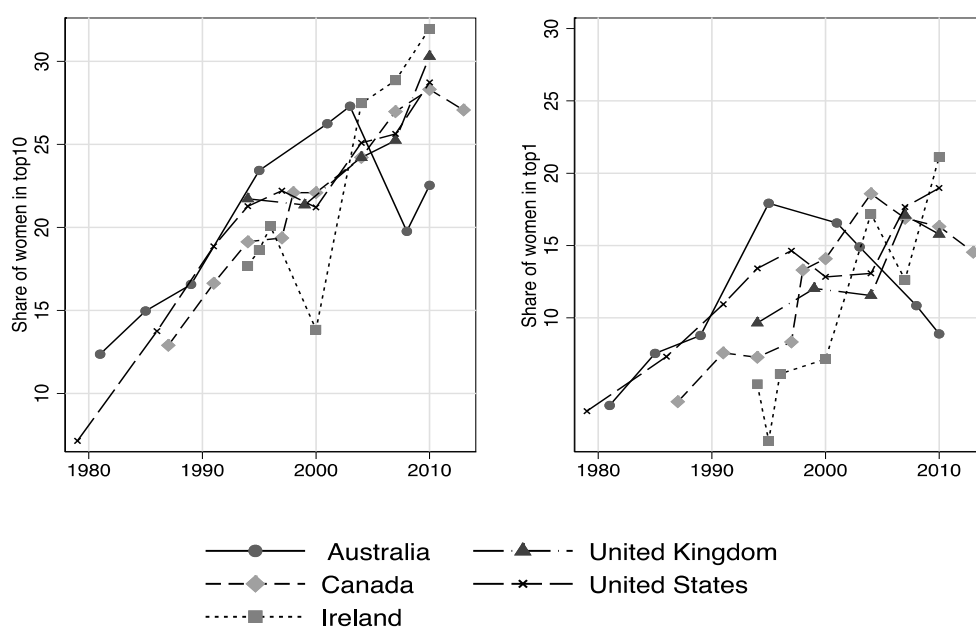


Figure 1. Share of women in top groups in the labour income distribution, Anglo-Saxon countries.

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.

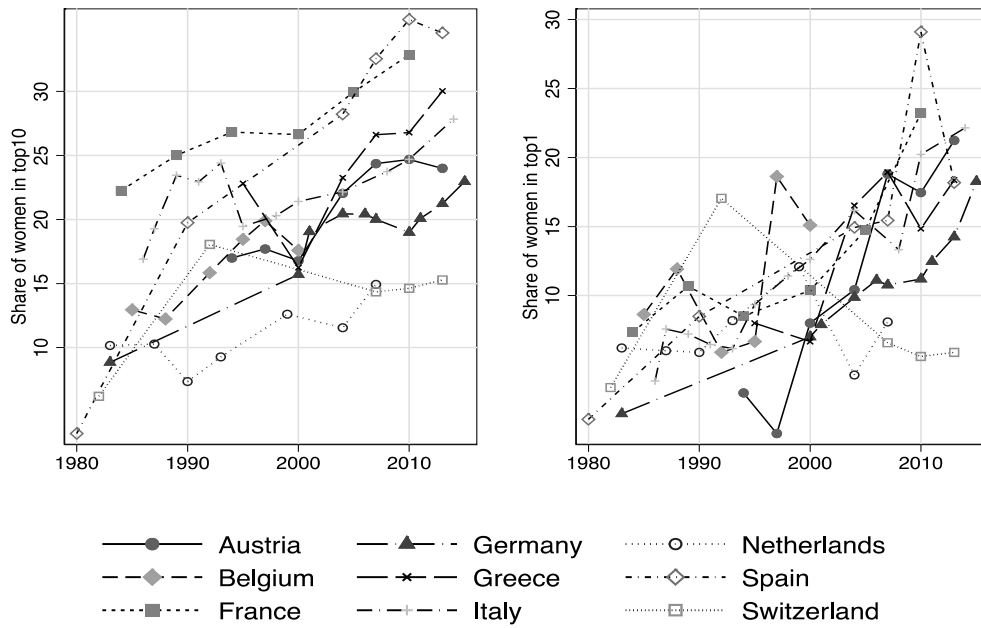


Figure 2. Share of women in top groups in the labour income distribution, Continental European countries.

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.

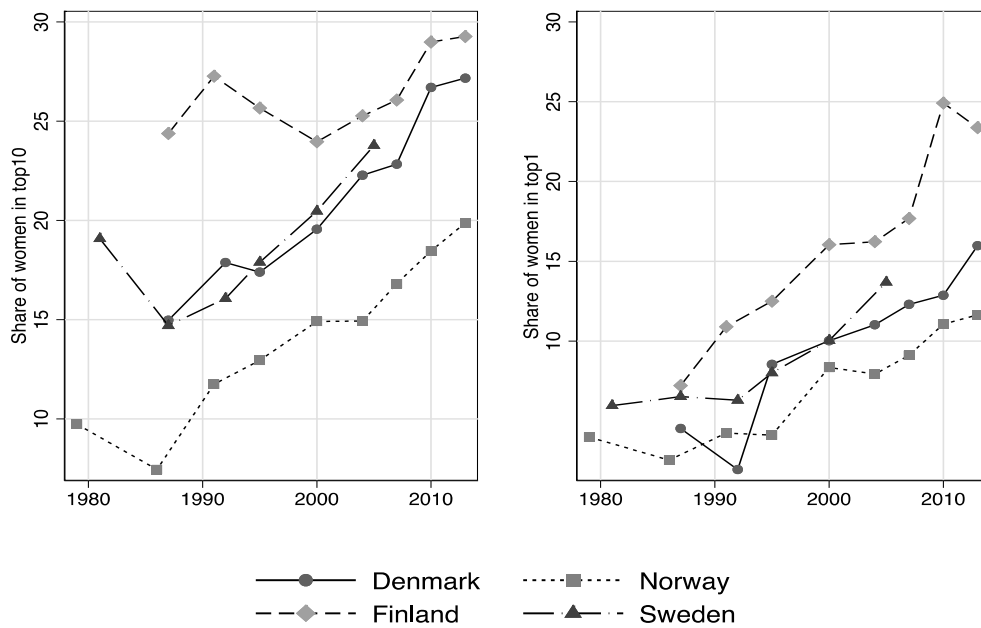


Figure 3. Share of women in top groups in the labour income distribution, Scandinavian countries.

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, 2007, and discussed in Boschini and Gunnarsson, 2018).

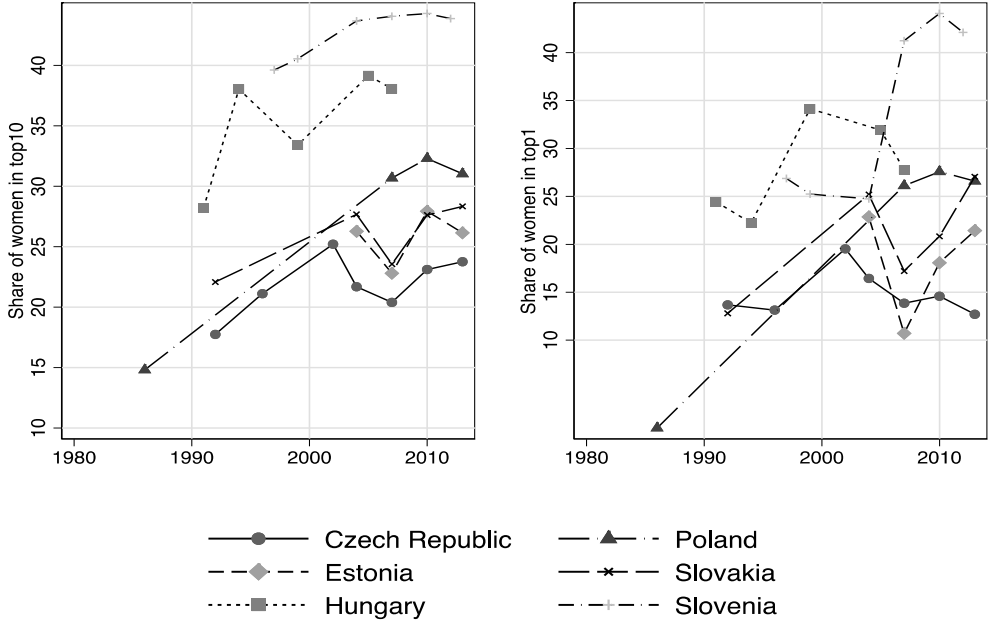


Figure 4. Share of women in top groups in the labour income distribution, Central and Eastern European countries.

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 in some dimensions, in particular when it comes to labour market participation. Interestingly, this seems to remain today and is also in contrast to the Nordic countries where overall labour force participation has been very high but the share of women in the top groups is markedly lower.

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.

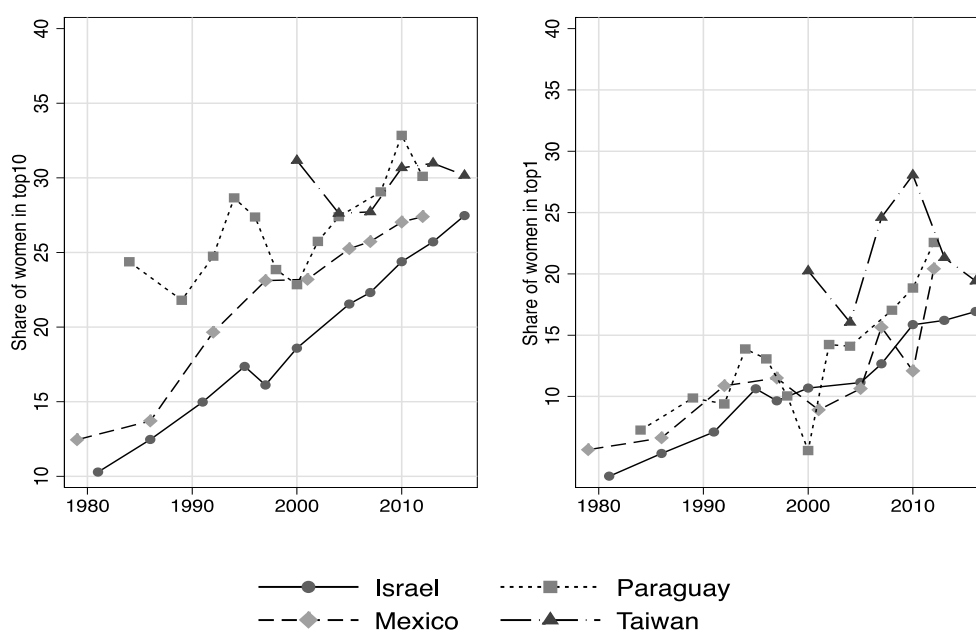


Figure 5. Share of women in top groups in the labour income distribution, for Israel, Mexico, Paraguay, and Taiwan.

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, especially in P90. There is also quite consistently a fanning out of the share of women as we reach the very top of the labour income distribution. But there are also some interesting differences across country groups, and most notable the Central and Eastern European countries similar share of women both in top10 and top1 and the low share of women in top1 in the Nordics. As the precision of the estimates is relatively low – as discussed in Section 2.1 - one should be careful when interpreting these results and further corroborating research of these trends is needed.

3.1 Differences between labour income and total income including capital income

An important finding in the top income literature, especially historically, is that top income shares and their development depend a lot on capital incomes. In several papers that study the development over the 20th century, 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. Figure 6 illustrates this difference for the United States, for which the longest time series of capital income is available in LIS, between the share of women in top1 of the labour income distribution, the top1 of the total income distribution and in the distribution of the sum of capital and labour income. As it turns out, the share of women in top groups is only marginally affected by the total income measure used. However, the long time series suggest that the share of women in the top1 group is affected by the inclusion of capital income in times when the income composition across genders in top groups differs. In recent decades, available evidence shows that top1 women are becoming more similar top top1 men, but in the 1970s and 1980s top women needed to have a larger capital share to compensate for their relative low earnings compared to top men. In line with the suggestive evidence in Edlund and Kopczuk (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. 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. However, this is a preliminary finding needing corroboration.

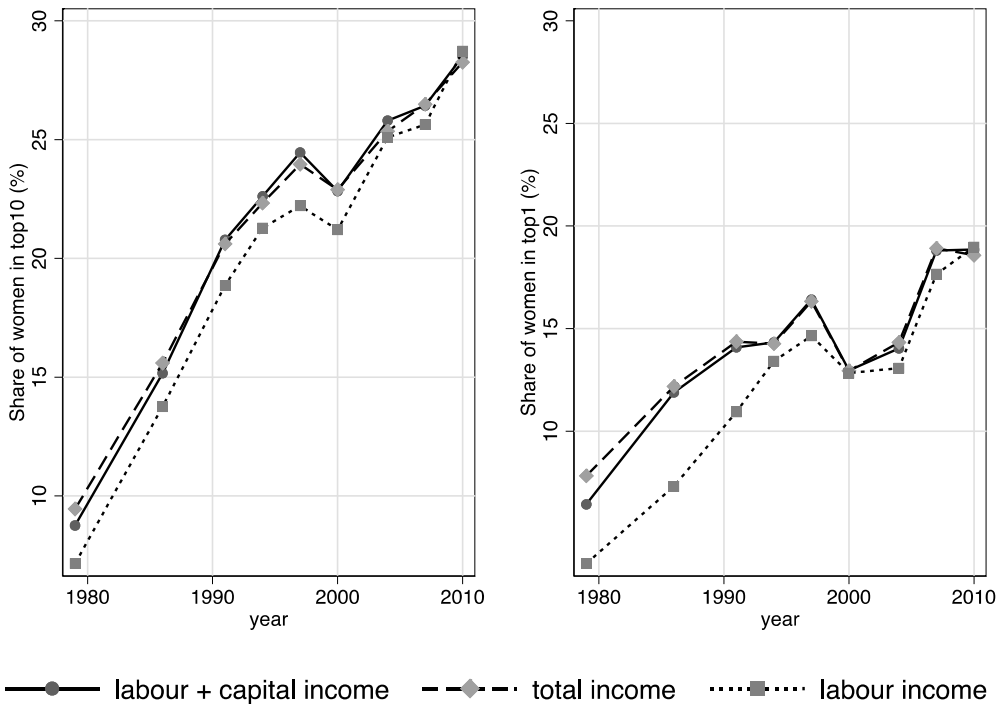


Figure 6. Share of women in top groups in the United States in different income distributions.

After 2007 we observe capital income for most countries that we study. 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.⁶ 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, the top share of women in the labour income distribution serves as a good proxy for the 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. However, as the case of the U.S. above illustrates, this does not imply that we can rely on this to be true for years going back in time

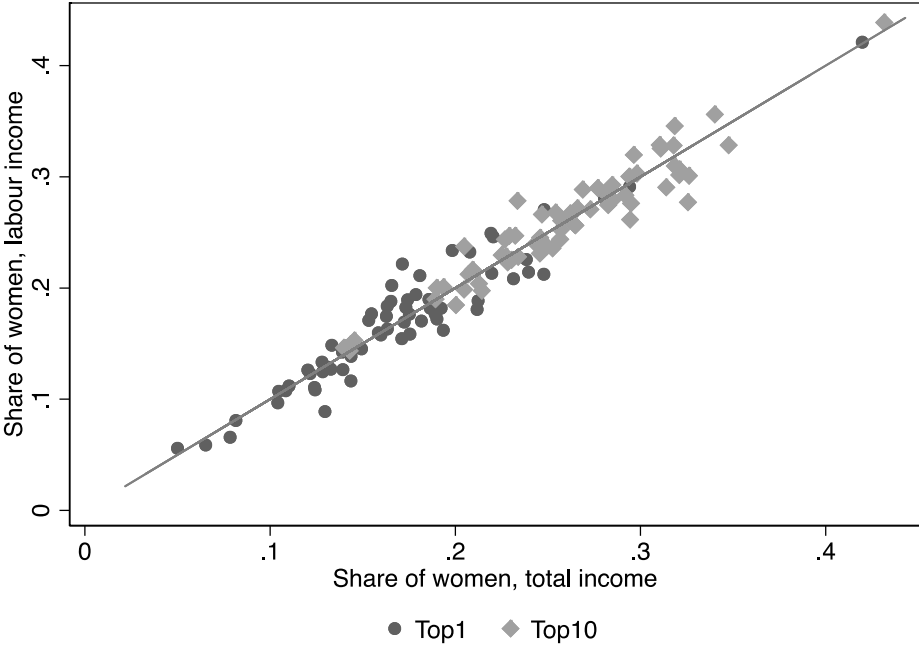
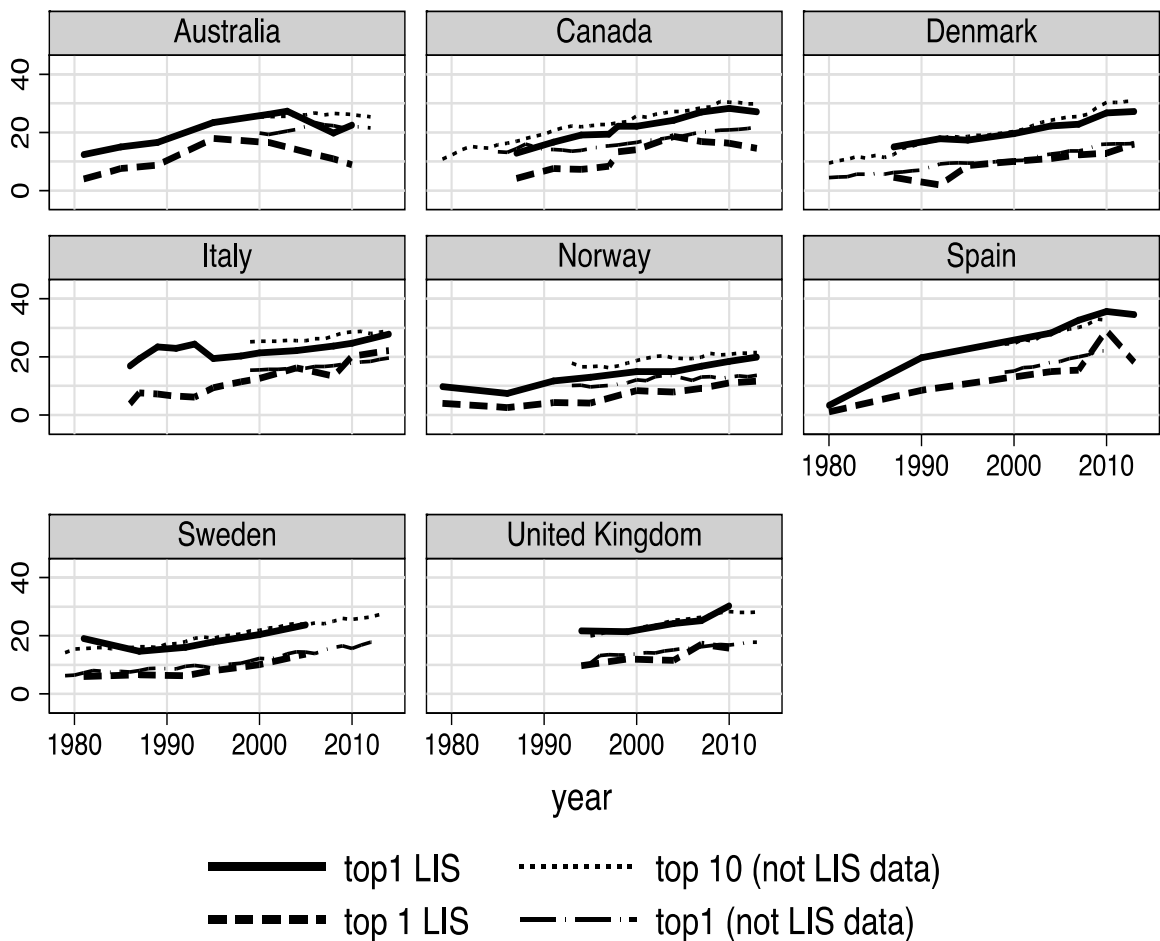


Figure 7. The correlation of share of women in top1 and top10 in the labour income distribution and the total income distribution since 2007.

⁶ Only observations from 2007 and onwards are used.

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. Our interpretation is that even the shares for top1 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.



Graphs by country

Figure 8. The shares of women in Top1 and Top10 in 8 countries with LIS data and tax register data.

4 Top income women and their partners

When thinking about what might explain differences in the presence of women in the top of the income distribution it is of course natural to first consider differences in various observable characteristics such as education, educational field, differences in occupations and sectors, etc. Several papers suggest however 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). Another potential explanation is that women often had to choose between having a family and a career, which in turn led to fewer women than men going for having a career. Those women choosing the career path tended to become childless – see e.g. Goldin et al (2006), Bertrand et al (2010), Boschini et al (2011), Goldin (2014). 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). Another type of explanation has to do with life-cycle events, and, in particular, children. 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 partner.

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). The figures in Appendix B show the share of men and women in top10 and in top1 that have a partner. Not surprisingly in light of the findings in Boschini et al (2017), top women are generally less likely to have a partner, albeit to a decreasing extent as we move towards the present day. 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.

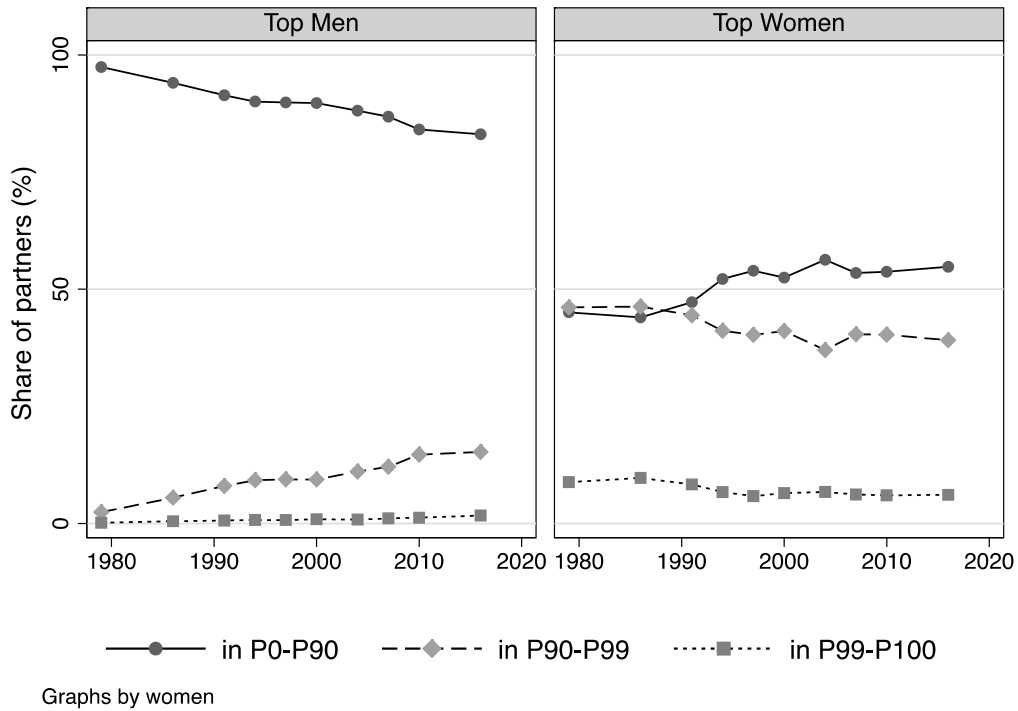


Figure 9. The partner distribution of men and women in top10 of the labour income distribution in the United States.

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.⁷ Figures 9 and 10 show where in the income distribution the partners of the top men and top women in the United States are. (Appendix D contains the corresponding graphs for Canada, Denmark, Finland and Norway – the only countries having enough women in the top and large enough samples for such an exercise to be meaningful). 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

⁷ In a related paper Aaberge et al. (2018) use LIS data to study assortative mating in relation to “perfect matching” and random matching, respectively.

to have a high earning partner. This is consistent with the patterns found for Sweden with register data by Boschini et al (2017).

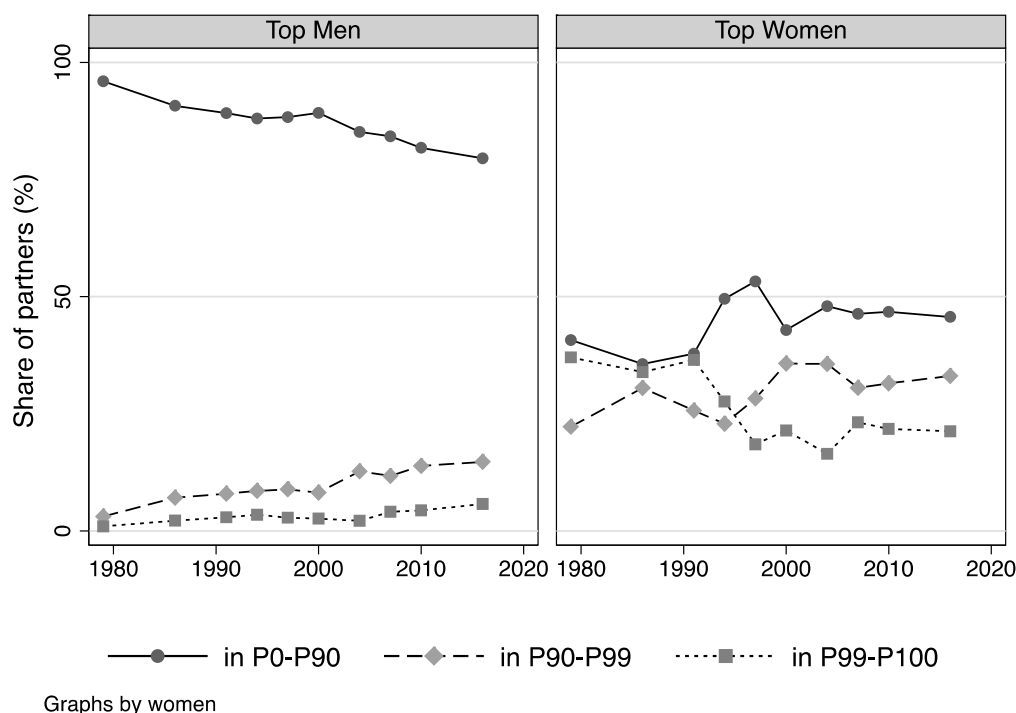


Figure 10. The partner distribution of men and women in top1 of the labour income distribution in the United States.

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 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 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, in the order of magnitude 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. The share of women in the top1 is lower in all countries we have studied and the average is around 15 to 20 percent.

Second, for recent years when we can compare the share's 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, we find suggestive evidence that top income men to a large extent have partners who are not themselves top income earners, while this is not the case for women, who much more often have partners also in the top income group. 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.
- 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 top 1 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, 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.
- 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., Héroult, 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.
- 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.
- 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., 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.
- 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*.

- 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.
- Keloharju, M., Knüpfer, 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., Søgaaard, J. E. 2018. "Children and Gender Inequality: Evidence from Denmark". *American Economic Journal: Applied Economics*, (forthcoming).
- Leigh, A., (2009). "Top Incomes". In: Salverda, W., Nolan, B., Smeeding, T. (Eds.), *The Oxford Handbook of Economic Inequality*. Oxford: Oxford University Press.
- 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.
- 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.

Appendix A. Sample sizes and minimum absolute number of women in top 1 group for countries and LIS waves included in the study

Table 1. Number of non-missing observations of personal labor income included in the study

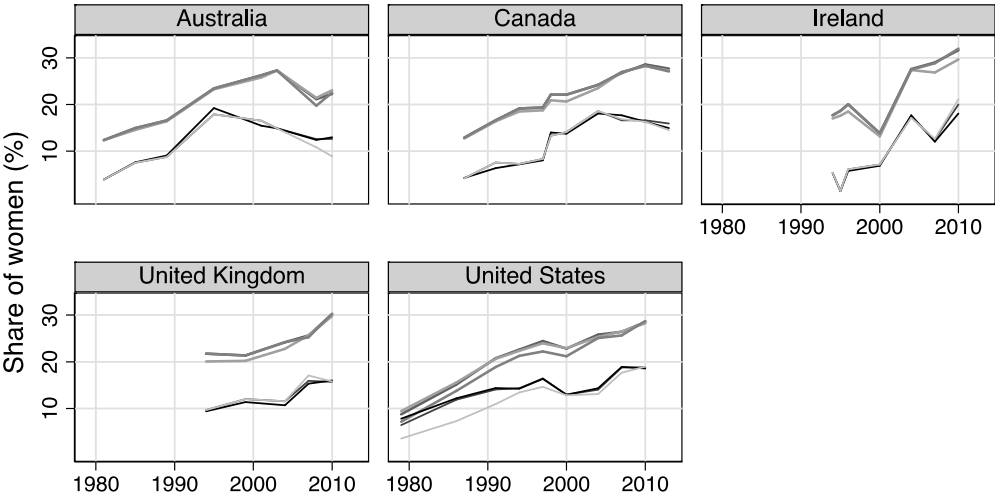
| COUNTRIES | HIST. DATA | WAVE I (~ 1980) | WAVE II (~ 1985) | WAVE III (~ 1990) | WAVE IV (~ 1995) | WAVE V (~ 2000) | WAVE VI (~ 2004) | WAVE VII (~ 2007) | WAVE VIII (~ 2010) | WAVE IX (~ 2013) | WAVE X (~ 2016) |
|-------------|----------------|--------------------|---------------------|----------------------|-----------------------|-------------------------|---------------------|----------------------|-----------------------|---------------------|--------------------|
| Australia | | 42275 | 20445 | 39098 | 17986 | 16895 | 24674 | 22932 | 42531 | | |
| Austria | | | | | 8950 7962 | 6675 | 13042 | 13621 | 13933 | 12982 | |
| Belgium | | | 18323 | 11139 10770 | 7464 12260 | 6399 | | | | | |
| Canada | | | | 30739 | 53280 97973 | 86614 79433 | 72850 | 68542 | 64783 | 60362 | |
| 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 30261 | 26824 | 26740 24999 | 44134 42531 | 41657 | 36949 |
| Greece | | | | | 14300 | 11223 | 14897 | 16785 | 14913 | 20850 | |
| Hungary | | | | 5848 | 5423 | 5469 | 5276 | 4298 | | | |
| Ireland | | | | | 11077 9692 8862 | 8638 | 15534 | 12545 | 10994 | | |
| Israel | | 8436 | 18616 | 19132 | 10980 | 19555 | 21046 | 20364 | 20225 | 28862 | |
| Italy | | | 25068 25092 | 25150 24930 | 24013 23924 | 20901 22268 | 20581 | 19907 | 19836 | 19366 | |
| Luxembourg | | | 6044 | 5498 | 4981 6632 | 6226 | 9661 | 10146 | 14884 | 9977 | |
| Mexico | | | 23985 | 57289 50862 | 60353 64916 | 48110 42535 72602 | 91738 | 118927 | 107781 | 33726 | |
| Netherlands | | | 13205 10731 | 10861 | 13029 | 11661 | 23756 | 25448 | 25461 | 24494 | |
| 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 29277 | 24535 | 21589 | 28195 | 34208 | 33139 | 36918 | | | | |
| Switzerland | | 16324 | | 16784 | | | | 16397 | 17602 | 15651 | |
| Taiwan | | 73309 | 74441 | 68439 | 39273 52491 | 49793 | 46386 | 46230 | 47900 | 50518 | 50569 |
| UK | | | | | 62821 | 59010 | 64942 | 56926 | 57928 | 46166 | |
| US | 34244 | 181488 | 155407 | 155796 | 149642 131599 | 218269 | 210648 | 206404 | 204983 | | |

Table 2. Number of women in top 1 percent of personal labor income included in the study.

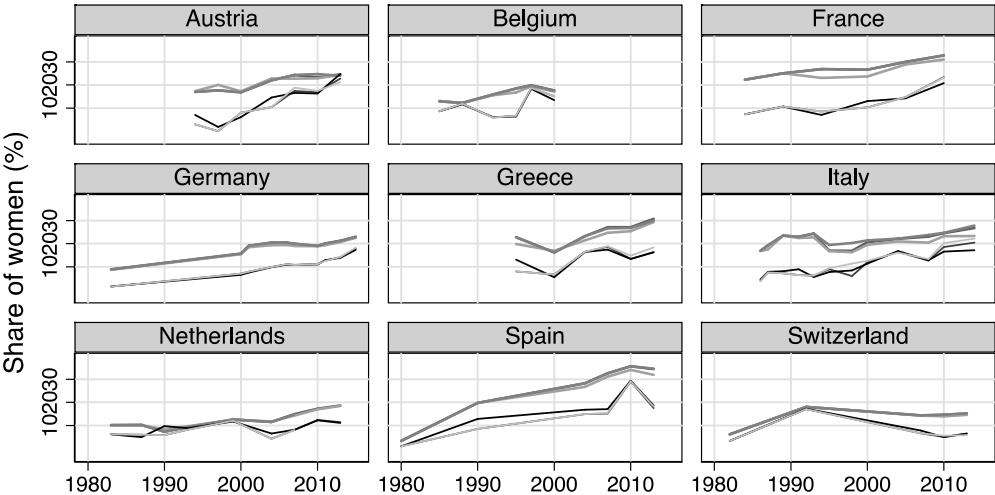
| COUNTRIES | HIST. DATA | WAVE I (~ 1980) | WAVE II (~ 1985) | WAVE III (~ 1990) | WAVE IV (~ 1995) | WAVE V (~ 2000) | WAVE VI (~ 2004) | WAVE VII (~ 2007) | WAVE VIII (~ 2010) | WAVE IX (~ 2013) | WAVE X (~ 2016) |
|-------------|------------|-----------------|------------------|-------------------|------------------|-----------------|------------------|-------------------|--------------------|------------------|-----------------|
| Australia | | 16 | 15 | 32 | 31 | 24 | 31 | 22 | 35 | | |
| Austria | | | | | 5 1 | 3 | 14 | 19 | 21 | 28 | |
| Belgium | | | 12 | 13 5 | 4 19 | 7 | | | | | |
| Canada | | | | 15 | 41 41 | 72 50 | 67 | 75 | 65 | 59 | |
| 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 36 | 38 | 39 32 | 42 47 | 47 | 61 |
| Greece | | | | | 13 | 4 | 19 | 23 | 14 | 24 | |
| Hungary | | | | 11 | 10 | 15 | 14 | 12 | | | |
| Ireland | | | | | 4 1 3 | 4 | 23 | 13 | 17 | | |
| Israel | | 6 | 9 | 14 | 13 | 13 | 15 | 23 | 18 | 39 | |
| Italy | | | 9 27 | 19 19 | 12 16 | 15 22 | 33 | 21 | 28 | 30 | |
| Luxembourg | | | 4 | 5 | 6 1 | 1 | 9 | 10 | 24 | 15 | |
| Mexico | | | 9 | 33 22 | 30 38 | 26 13 42 | 95 | 153 | 138 | 41 | |
| Netherlands | | | 8 4 | 8 | 10 | 10 | 10 | 19 | 32 | 31 | |
| 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 26 | 30 | 10 | 15 | 25 | 26 | 49 | | | | |
| Switzerland | | 6 | | 15 | | | | 12 | 9 | 9 | |
| Taiwan | | 16 | 26 | 32 | 26 42 | 48 | 62 | 56 | 73 | 84 | 75 |
| UK | | | | | 40 | 51 | 46 | 61 | 66 | 69 | |
| US | 4 | 93 | 133 | 157 | 150 151 | 195 | 211 | 286 | 274 | | |

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.

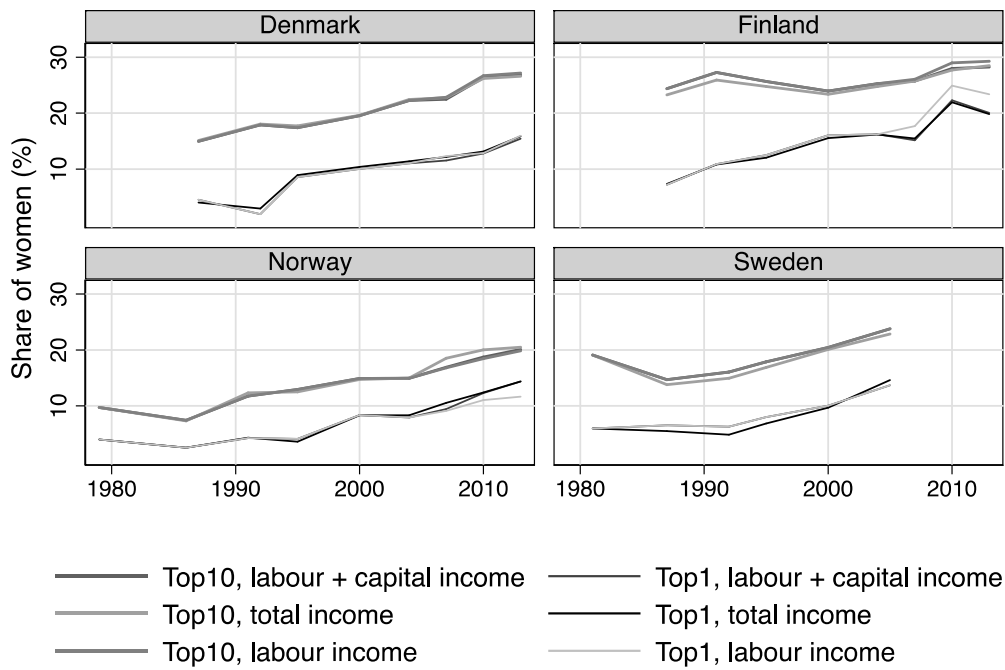
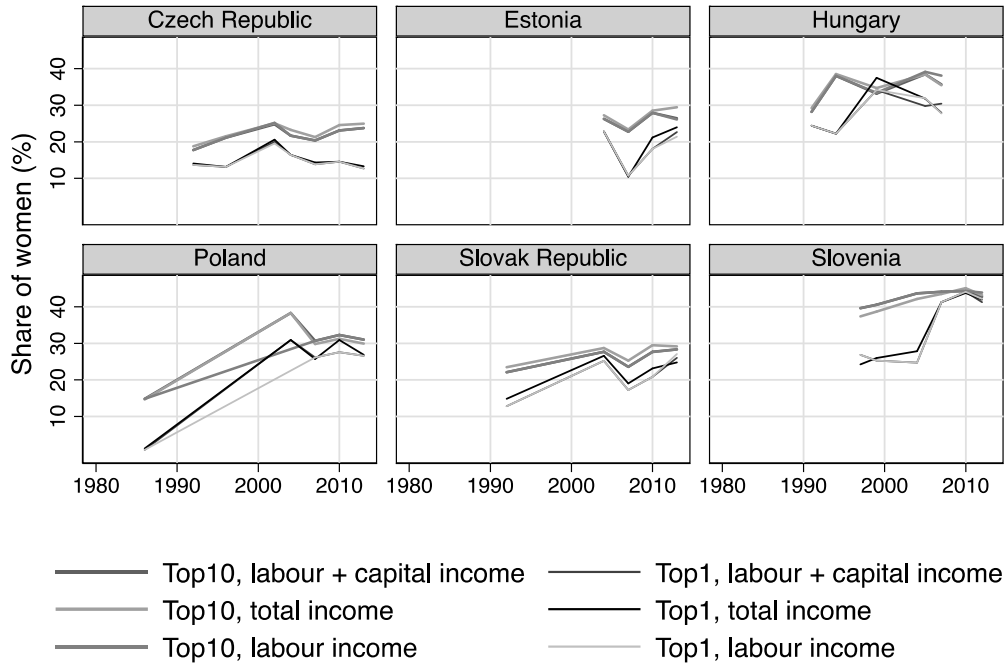
Appendix B. Comparing share of women in top10 and top1 in the capital + labour income distribution, in the total income distribution, and in the labour distribution

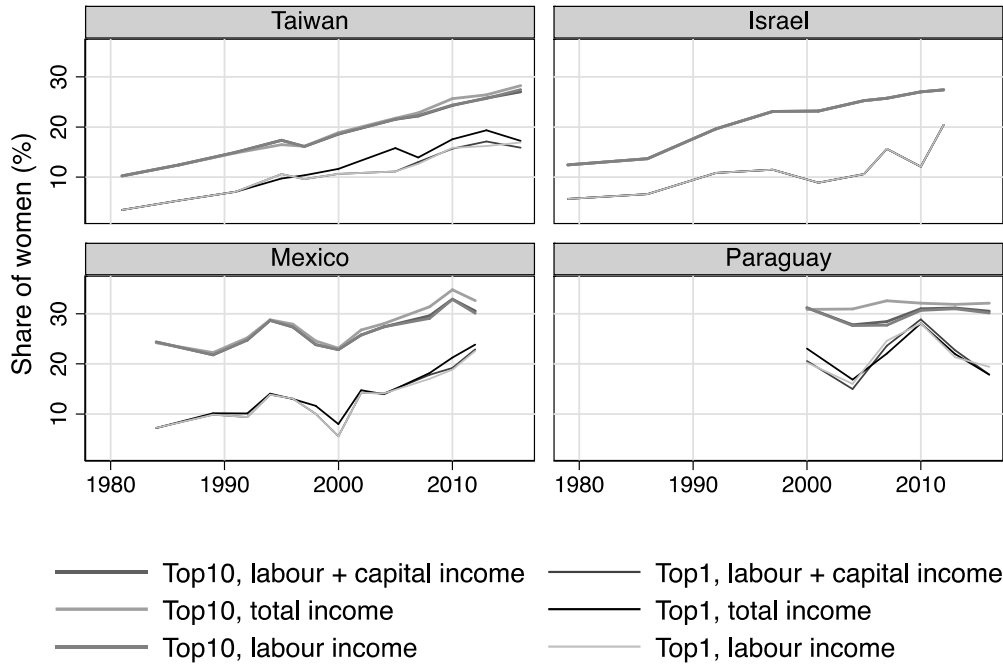


— Top10, labour + capital income — Top1, labour + capital income
 — Top10, total income — Top1, total income
 — Top10, labour income — Top1, labour income

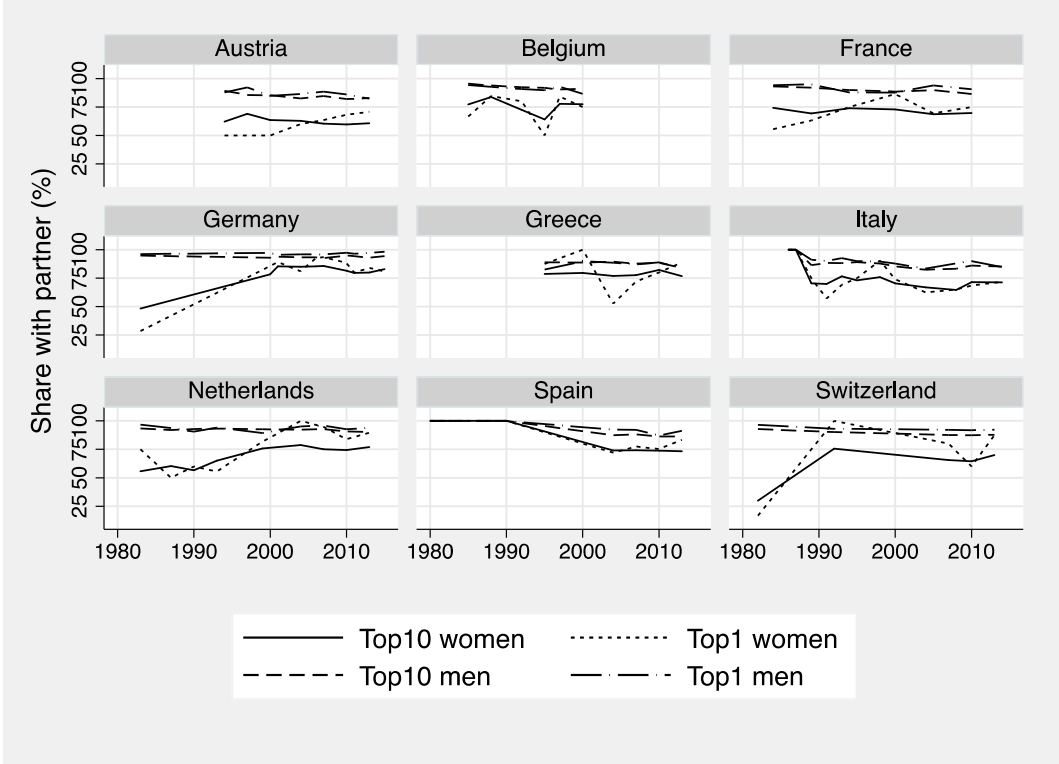
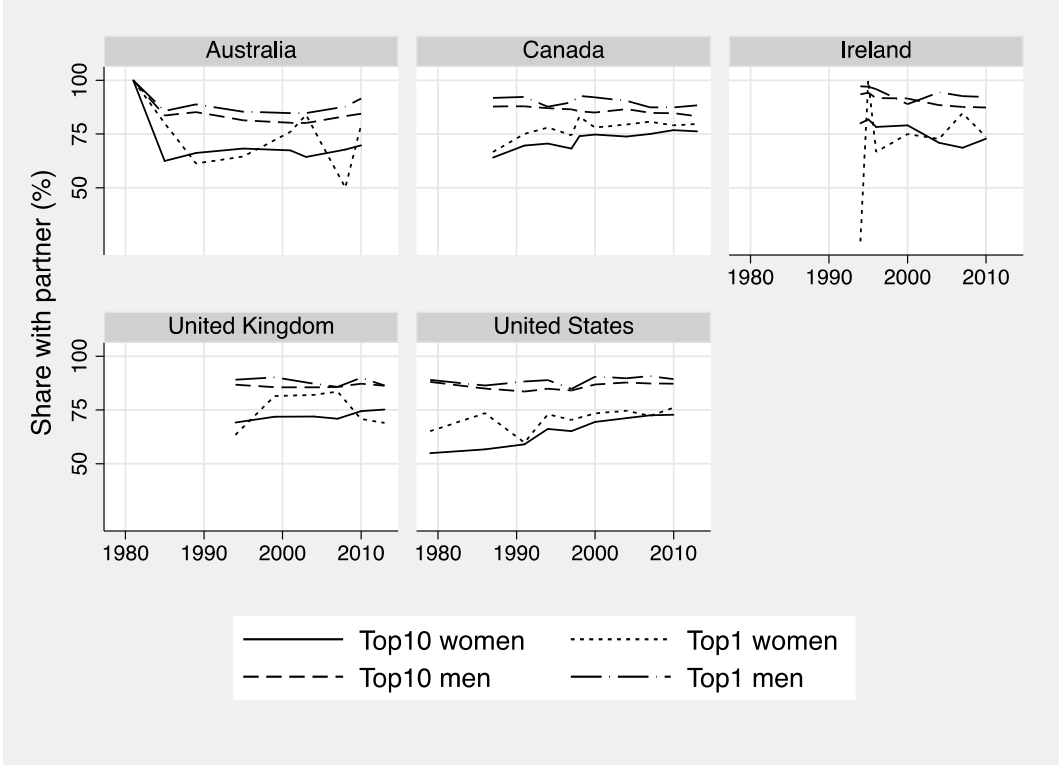


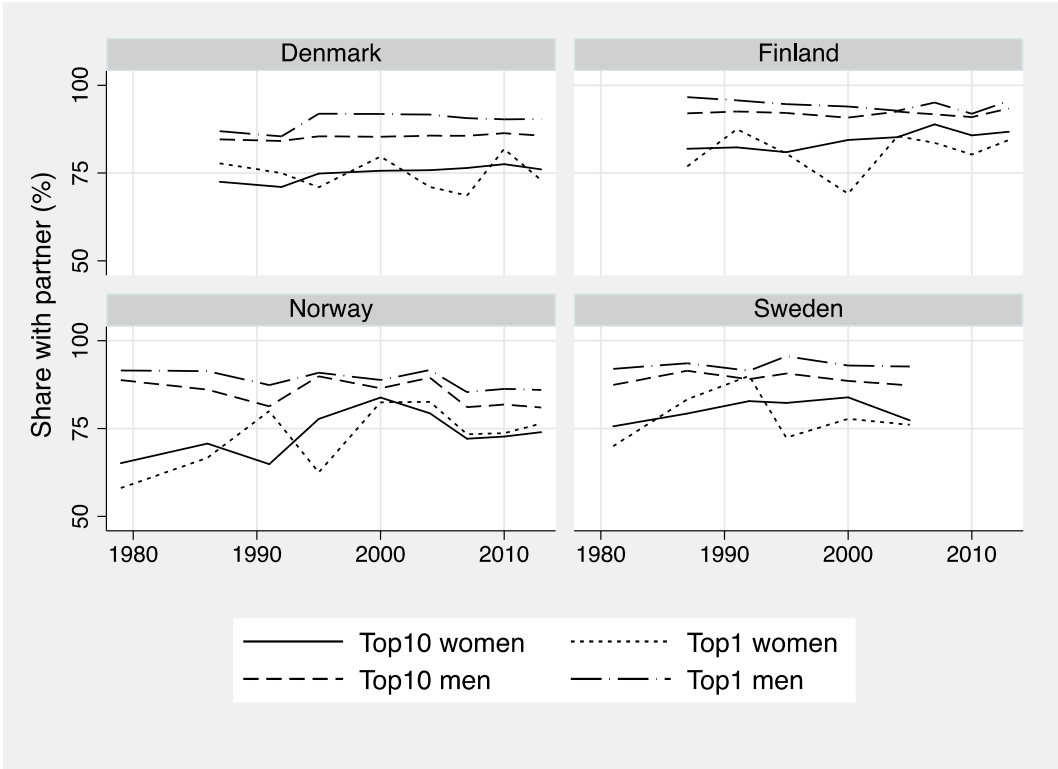
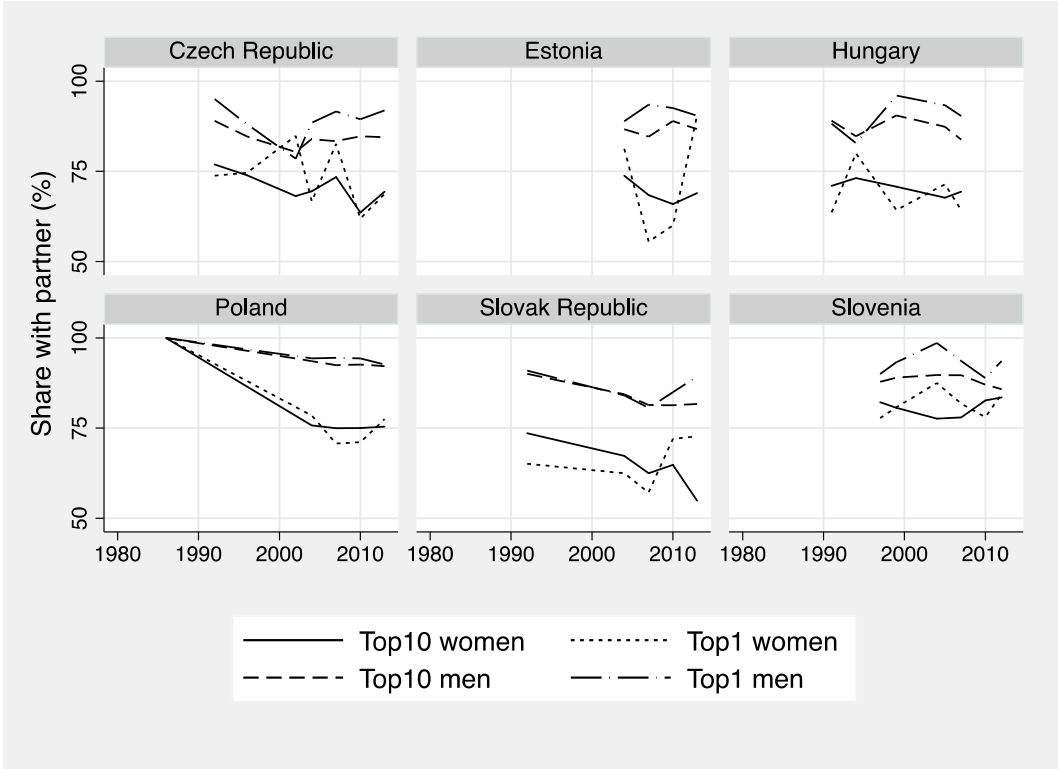
— Top10, labour + capital income — Top1, labour + capital income
 — Top10, total income — Top1, total income
 — Top10, labour income — Top1, labour income

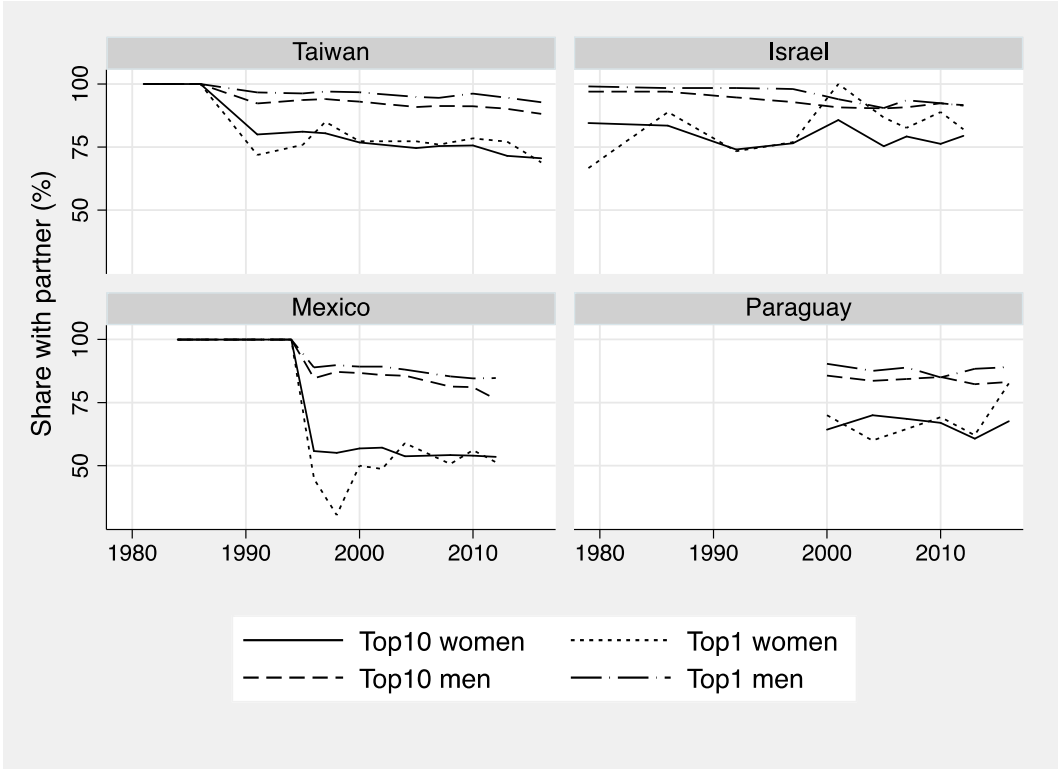




Appendix C. Share of top men and women with a partner.

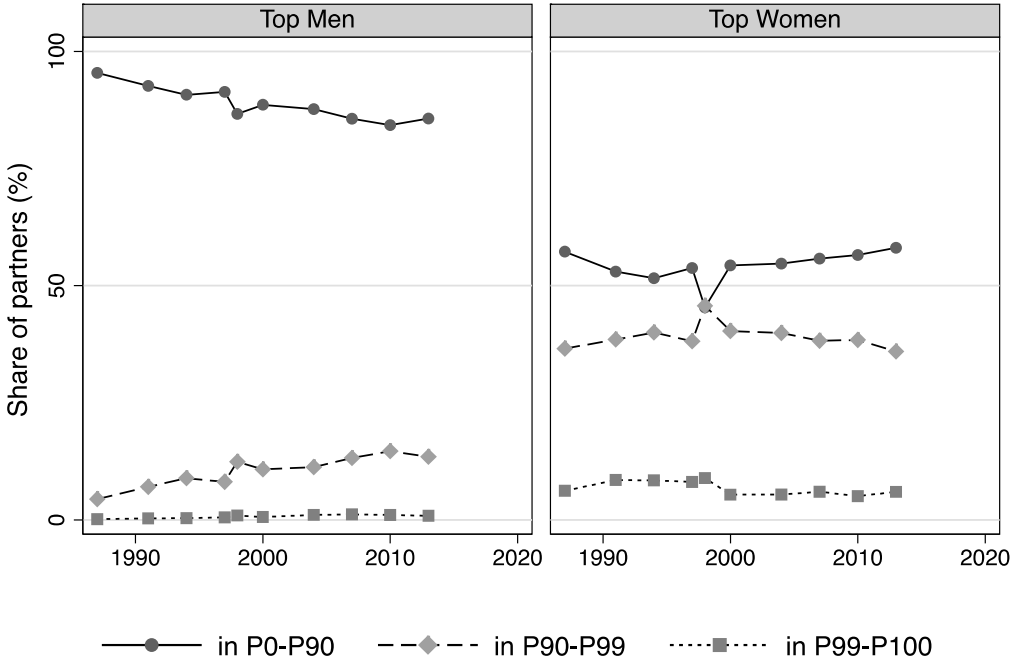




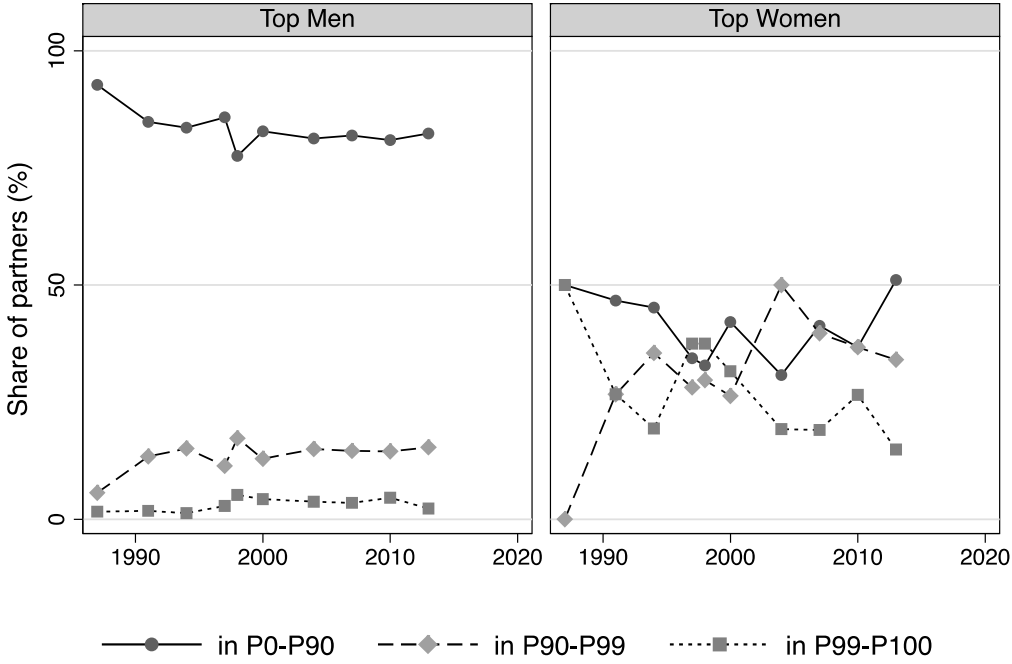


Appendix D. Partner distribution in four countries for top men and top women in the top10 and top1 groups of the labour income distribution.

Canada

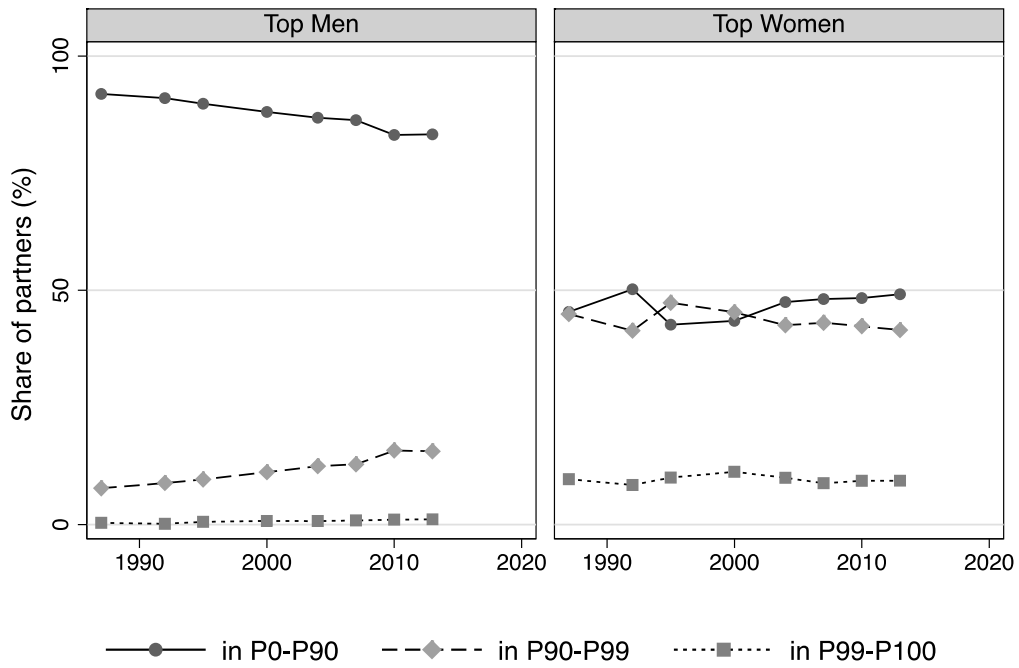


Graphs by women

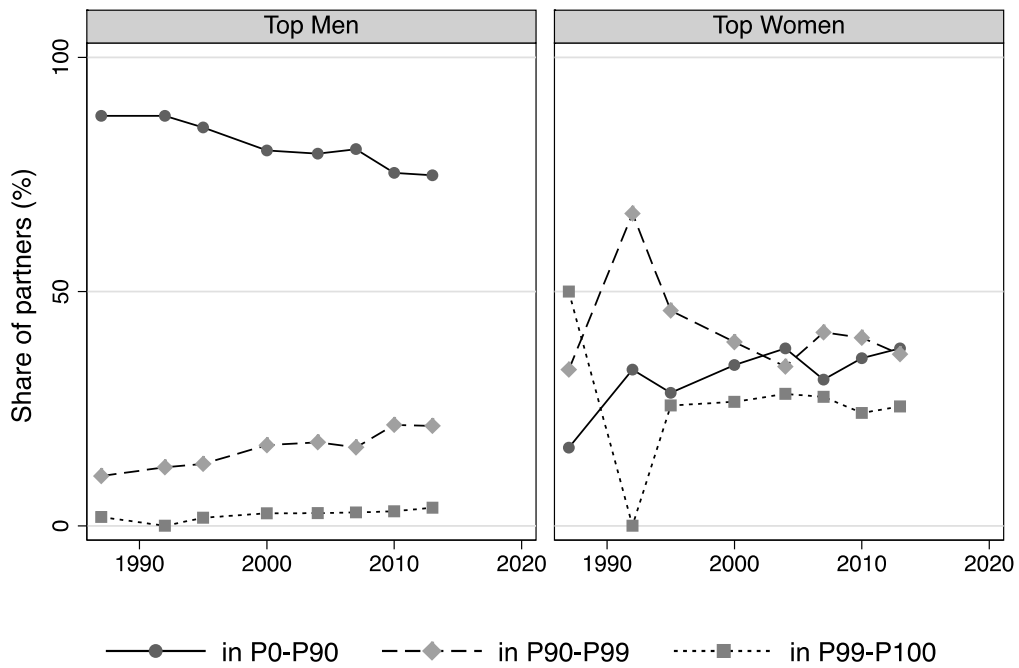


Graphs by women

Denmark

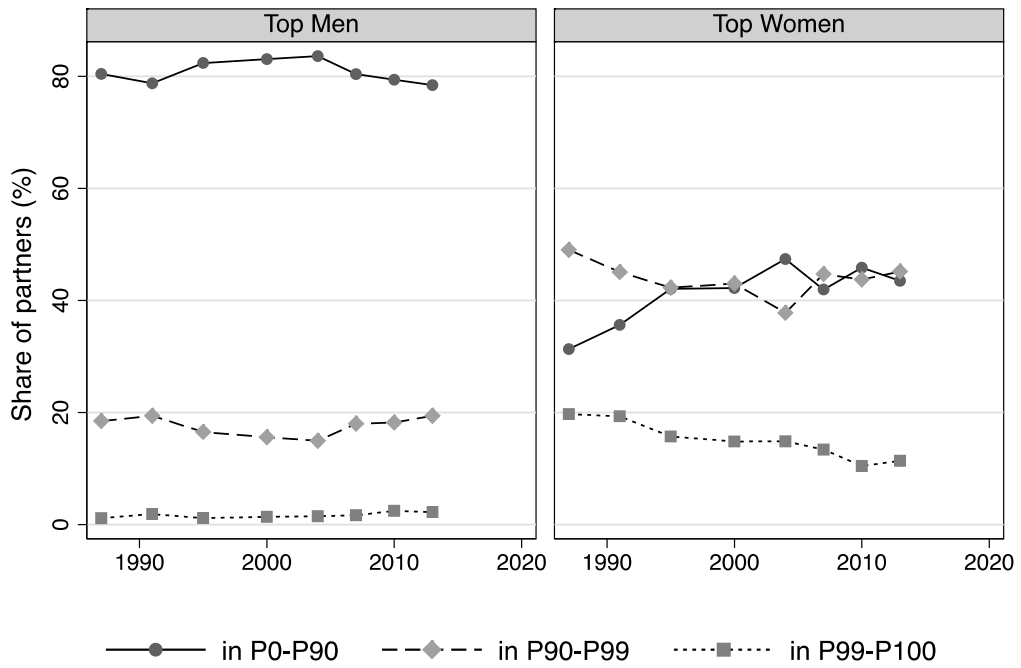


Graphs by women

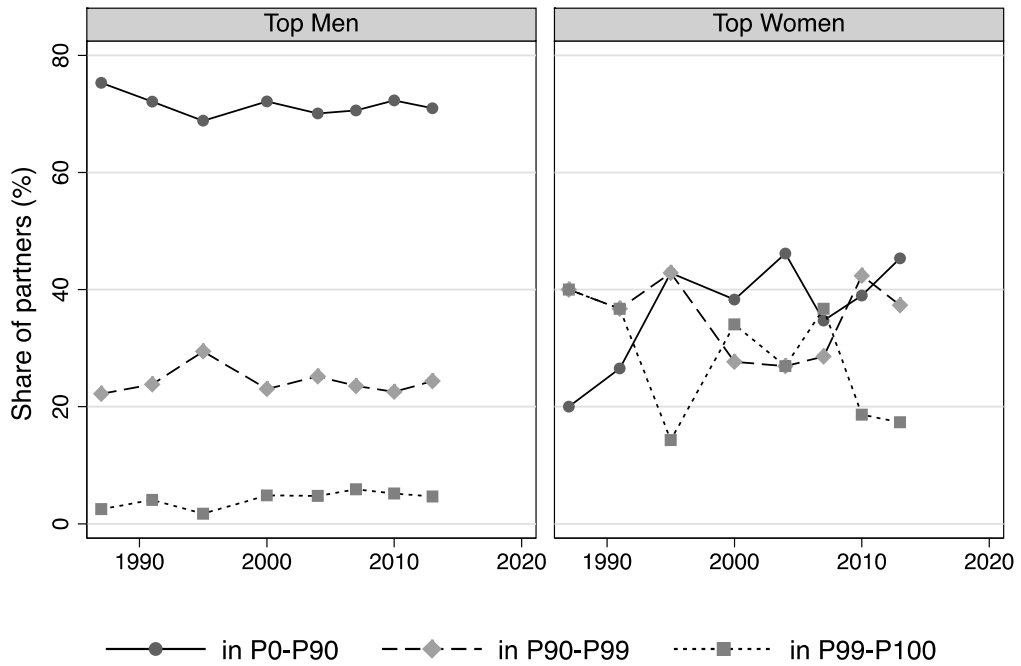


Graphs by women

Finland

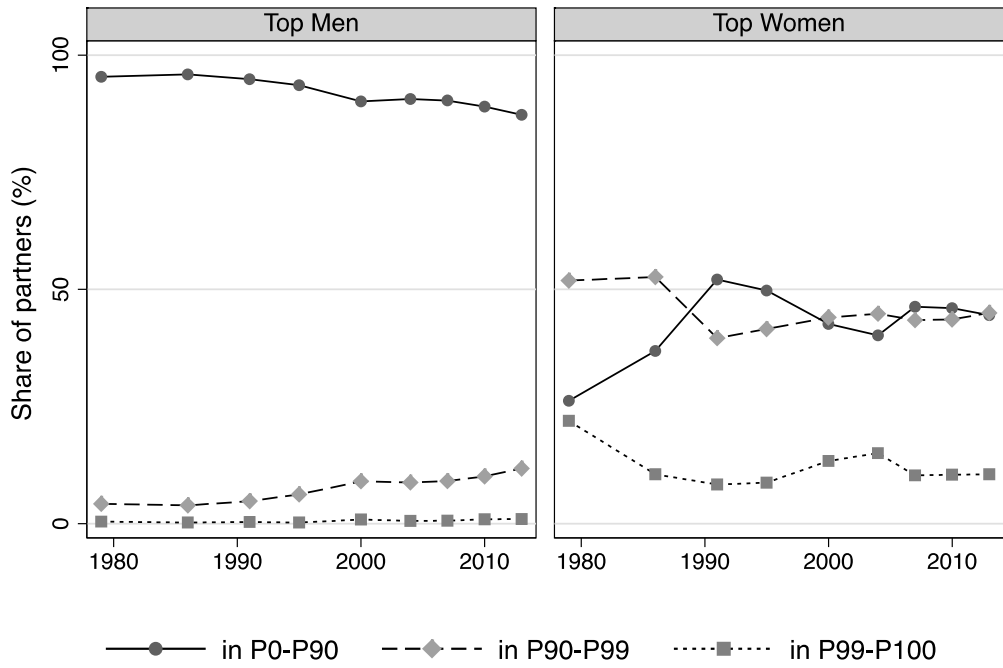


Graphs by women

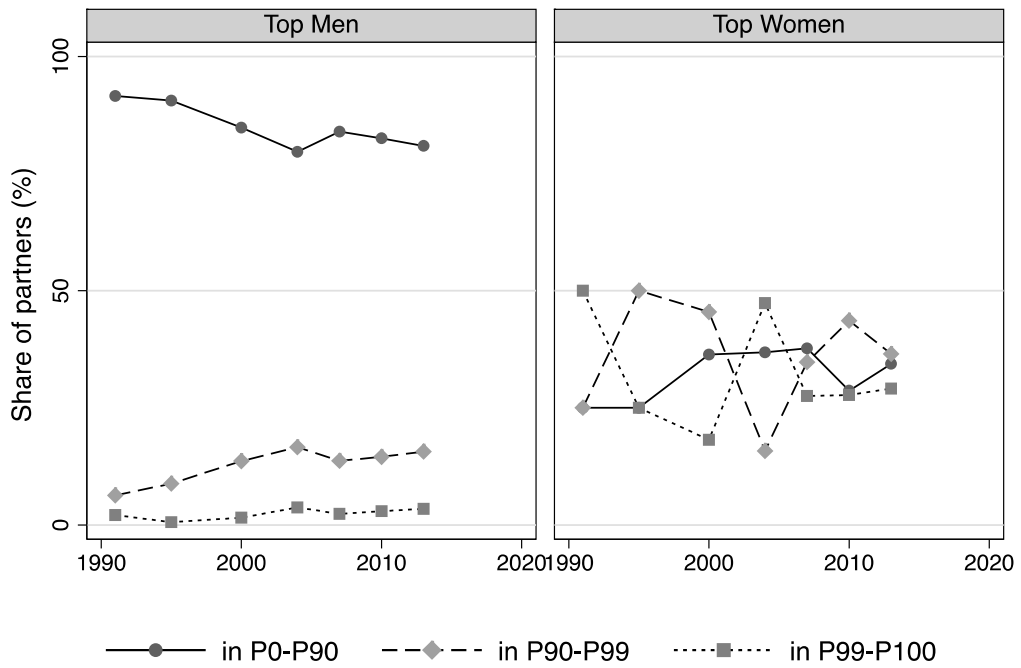


Graphs by women

Norway



Graphs by women



Graphs by women