// NO.20-001 | 01/2020

DISCUSSION PAPER

// ANJA ROTH AND MICHAELA SLOTWINSKI

Gender Norms and Income Misreporting within Households

Leibniz Association

ZEW

Gender norms and income misreporting within households

Anja Roth[†] and Michaela Slotwinski[‡]

January 13, 2020

Abstract

We show that the discontinuity in the distribution of surveyed female income shares at the margin where a woman would outearn her partner is primarily driven by norm induced misreporting in surveys. We draw on unique Swiss data combining survey and administrative information for the same individual and their partner. We demonstrate that individuals misreport incomes in surveys to comply with the male breadwinner norm. The male breadwinner norm does, however, not affect real labor market decisions around this margin. The resulting survey bias leads to a considerable overestimation of policy relevant measures like the gender wage gap.

Keywords: gender norms, female income shares, combination survey and administrative data, income misreporting, gender earnings gap

JEL classifications: D10, J01, J16

[†]University of Basel, Faculty of Business and Economics, Peter Merian-Weg 6, 4002 Basel, Switzerland. Phone: +41 (0)61 207 33 14, email: anja.roth@unibas.ch.

[‡]University of Basel, Faculty of Business and Economics, Peter Merian-Weg 6, 4002 Basel, Switzerland. ZEW – Centre for European Economic Research, L7,1, 68161 Mannheim, Germany. Phone: +41 (0)61 207 33 75, email: michaela.slotwinski@unibas.ch.

Michaela Slotwinski gratefully acknowledges financial support from the *NCCR on the move* financed by the Swiss National Science Foundation. We are grateful to the FSO, and particularly to Jonas Deplazes, for supporting our research and enabling us to connect the survey and the administrative information. We are further grateful to Felipe Carozzi, Marius Faber, Olle Folke, Patricia Funk, Matthias Krapf, Armando Meier, Georg Nöldeke, Reto Odermatt, Marcus Roller, Kurt Schmidheiny, Sebastian Siegloch, Sarah Smith, Alois Stutzer, Daniel Waldenström, and Conny Wunsch and the many seminar and conference participants for helpful comments.

1 Introduction

Women's lives have become more emancipated and empowered over the last century (Goldin, 2006). At the same time, there is still ample evidence that differences in the ascribed gender roles are deeply grounded and strikingly persistent, and that these differences play an important role in shaping individuals' outcomes (see, e.g., Alesina et al., 2013; Teso, 2019; Giuliano, 2017). In recent years, economists have been striving to understand the extent to which these gender norms impact economic agents' behavior (see, e.g., Alesina et al., 2013; Fernández et al., 2004; Fernández and Fogli, 2009; Fortin, 2005; Teso, 2019).

In a prominent contribution, Bertrand et al. (2015) document a striking discontinuity in the distribution of female income shares, measured as female incomes as a share of couple income. This discontinuity is located right at the point at which a woman would outearn her partner. There is sharp bunching in female income shares just below 50 percent and missing mass above in US data. The observed statistical pattern is attributed to the *male breadwinner norm*, which states that men are supposed to be the main earners in a couple and which leads couples to sort to below the threshold to comply with this norm. This study has motivated a whole literature trying to explain the discontinuity and to find the mechanisms behind it, mostly focusing on real behavior. In addition to couple formation and women actively adapting their labor market outcomes in order to not outearn their partner, various drivers unrelated to gender norms have been proposed, such as the tax schedule or collective wage agreements (see, e.g., Wieber and Holst, 2015; Lippmann et al., 2019; Binder and Lam, 2018; Eriksson and Stenberg, 2015; Zinovyeva and Tverdostup, 2018). However, there is still dissent on the exact mechanisms and on whether gender norms are indeed involved.

In this paper, we present a different behavioral channel which has thus far not been considered: We show that misreporting of incomes in surveys can account for the largest part of the discontinuity in female income shares in Switzerland.[1] Merging individual level Swiss administrative and survey data, we are able to precisely capture the differences between the surveyed and administrative earnings. Our results reinforce gender norms as the primary driver behind the discontinuity in the distribution of female income shares based on survey data, as there are no incentives other than norms to misreport incomes in a survey. Our findings suggest that these gender norms might, however, not affect real behavior as strongly as previously thought.

First, we show that the distribution of female income shares in survey data features a sharp drop of about 75 percent at the threshold where a woman would outearn her partner,

[1] Gil and Mora (2011), for example, document that social norms play a role in the misreporting of individuals' weight. Martinelli and Parker (2009) relatedly find that misreporting in self-reported program eligibility surveys is not only driven by underreporting due to material incentives, but also by overreporting of goods with 'status' value. Hariri and Lassen (2017) document that there is social desirability bias in income reporting of high income groups in Denmark. Furthermore, Funk (2016) documents for Switzerland that socially acceptable norms are an important driver of responses to post-vote surveys.

which is in line with prior studies for other countries. Second, using administrative data of all married couples in Switzerland, we document that the discontinuity prevails but is about 75 percent smaller (about 20 percent). This eliminates the possibility that the observed discontinuity is purely driven by institutional factors or real responses in the Swiss case and substantiates the suspicion that the largest part is a result of survey responses.

In order to examine misreporting of incomes in surveys as the mechanism behind this differential, we retained the administrative income of survey respondents and their partners for a sub-sample of the survey population. Based on this data, we show that the excess mass below the threshold in survey data is composed of individuals whose administrative income share lies above the threshold, who thus misreport incomes to comply with traditional gender norms. Individuals misreport both their own as well as their partner's income in order to adhere to traditional gender norms and place the couple below the threshold where the woman would outearn her partner. We further observe that individuals just below the threshold preemptively react to the norm and start to adapt their income reporting such that they do not violate the male breadwinner norm. This speaks against an equality norm argument, which would state that a couple's goal is to state equal earnings in surveys.

Descriptive statistics are in line with traditional gender norms being a major driver of this behavior. Focusing on couples where the woman outearns the partner based on administrative information, we find that the probability that the respondent reports a surveyed female income share below or equal to 50 percent is higher if the man is more or equally educated. Consistently, misreporting is more likely if the woman works the same or fewer hours but still outearns her partner. This is in line with the reasoning that situations which might pose a threat to the male identity lead to misreporting. Furthermore, misreporting is more prevalent for individuals from gender unequal countries, for German speaking individuals (compared to non-German speaking individuals), and in couples with a higher within-couple age difference; all measures frequently related to more traditional gender norms.

Our finding that individuals' responses in surveys are strongly affected by gender norms reveals that a large part of the discontinuity in female income shares based on survey data is due to systematic income misreporting. However, it also has broader implications. Survey data are widely used in economics and social sciences, e.g., for studies on the gender wage gap (GWG), unemployment rates, and other statistics used by economists and policy makers. If, as for example in our case, women's incomes are systematically underreported and those of men systematically overreported due to gender norm considerations, this creates an upward bias in estimates of the GWG based on survey information. As survey participants tend to also systematically misreport partner incomes, the problem is amplified if the survey sample is enriched by proxy interviews, which is frequently the case to reduce survey costs (Reynolds and Wenger, 2012; Lee and Lee, 2012). For instance, proxy interviews constitute almost 50 percent of the Current

Population Survey, the most widely used data source of the US Department of Labor. In our data, the use of surveyed incomes results in an 9.4 percent overestimation and the additional use of proxy information to a 13.5 percent overestimation of the true GWG as measured in the administrative data. This underscores the importance of learning more about social norms and potential biases in survey data when studying sensitive questions.

Empirical evidence on the distribution of female income shares is in line with the result that misreporting in surveys plays an important role: Studies using survey data find more distinct discontinuities than studies using administrative data. Wieber and Holst (2015) and Sprengholz et al. (2019) find a strong discontinuity in the distribution of surveyed female income shares in Germany of roughly 60 percent, i.e., the mass just above the threshold is 60 percent lower than the mass just below. Bertrand et al. (2015) themselves use a mix of survey and administrative data for their analysis. They clearly note that there is a much larger mass of couples earning the exact same income in survey data than in administrative data and distribute the excess mass at 50 percent in surveyed female income shares to the neighbouring bins.[2] Using the same data base, Binder and Lam (2018) show that there is a discontinuity of about 12.4 percent in US administrative data. Other studies relying on administrative data also find that the discontinuity is less distinct (Eriksson and Stenberg (2015), 22.5 percent in Sweden and Zinovyeva and Tverdostup (2018) 11.3 percent in Finland). Using administrative data Eriksson and Stenberg (2015), and Zinovyeva and Tverdostup (2018) show that the spike at the point where spouses earn the exact same income consists mainly of couples working in the same sector or for the same employer, and argue that the discontinuity at the point where the woman outearns her partner is therefore not related to traditional gender norms. Rather, the spike at the point where spouses earn the exact same income might reflect a country's institutional framework, e.g., collective agreements, minimum wages, or a progressive income tax schedule with individual taxation, which amplify incentives for couples to bunch at exactly 50 percent. This would be reflected in a large spike at this point of the distribution, which would produce a discontinuity in both survey and administrative data and which is unrelated to traditional gender norms.

Descriptive evidence for the US is in line with the hypothesis that misreporting might play a role. Murray-Close and Heggeness (2018) find that the deviation between surveyed and administrative incomes is higher in couples where the woman earns more than her husband than in couples where the woman earns less than her husband. While their findings are interesting, they remain descriptive and cannot contribute to explaining the bunching below the point where the woman outearns her partner, as they compare the average of

[2]More specifically, they attribute the difference in mass at 50 percent between the distribution based on administrative and on survey data partly to top-coding and imputing of incomes and drop these couples from the analysis. They ascribe the remaining differential between survey and administrative data to rounding and misreporting and randomly distribute the remaining mass differential to the neighboring bins.

all couples below and above the threshold.[3] Bursztyn et al. (2017) similarly document that single women tend to underreport their career ambitions in situations where social norms become salient, i.e., if classmates are more likely to observe the response and larger marriage market consequences can be expected.

We document that misreporting in surveys to conform with the male breadwinner norm causes the largest part of the discontinuity found in survey data and explains why it is not fully reflected in the distribution of female income shares based on administrative data. This explanation reconciles the results of prior studies and draws attention to a systematic bias in surveyed incomes. The question whether systematic misreporting of incomes is equally prevalent in other countries is up for future research. However, our evidence clearly reveals that potential survey biases have to be considered when working with survey items that are prone to be distorted by social norms. Such biases can have far reaching consequences for policy relevant measures like the GWG. Furthermore, we demonstrate that the combination of survey and administrative data may be a novel approach to measuring norms.

The remainder of the paper is structured as follows. Section 2 describes the data for our main analysis before Section 3.1 presents the results for the overall distribution of female income shares, once based on survey and once based on administrative data. Section 3.2 presents a detailed analysis of individuals' misreporting of their own and their partner's incomes for couples where the woman earns more based on administrative information. Section 3.3 presents some descriptive evidence on characteristics correlated with misreporting. Section 4 discusses and quantifies broader implications of the resulting survey bias for a frequently used policy measure, the GWG. Section 5 offers concluding remarks.

2 Data

Our main analysis draws on data of the largest Swiss labor market survey (Schweizerische Arbeitskräftererhebung, SAKE) from survey years 2012 and 2015. In these years, the special questionnaire 'Social Security' including questions about partner income was administered in addition to the basic questionnaire eliciting respondents' incomes. The survey is based on telephone interviews where the respondent within a household is randomly chosen.[4] The information from survey interviews allows us to calculate the surveyed female income share. Furthermore, we were able to retrieve the respondent's as well as his or her partner's administrative income based on the Swiss Central Compensation Office

[3] They use the full range of female income shares for their analyses and their results seem to be driven by individuals in the top and bottom quantile of the earnings' distribution. Also, the selection of women who outearn their husbands show very different characteristics in terms of their administrative income, education level, age, and number of children, to name a few. It thus remains unclear whether the differences found by Murray-Close and Heggeness (2018) are related to traditional gender norms or whether they simply reflect the different selection of individuals.

[4] More information on the exact interview process can be found in Appendix A.5.

(ZAS) for the same years.[5,6] Observing administrative and surveyed incomes for both the survey respondent and the partner gives us the opportunity to study misreporting of ones own and the partner's incomes, as well as the difference between the administrative and the survey based female income share.

Our sample consists of respondents with Swiss citizenship or permanent residence permit, and where both partners are in paid employment.[7] Survey income can be stated as either hourly, monthly, or yearly gross or net income.[8] The category most commonly chosen is monthly gross income, picked by 36 percent of individuals. In order to avoid biases from approximations, we focus on individuals who report both their own as well as their partner's income in the same mode (monthly net, monthly gross, or yearly gross), which is true for about 74 percent of our couples.[9] We use the survey information of those individuals who we believe are most likely to be able to report their own and their partner's income correctly. We therefore restrict the sample to individuals employed in the twelve months prior to the interview. We exclude all individuals who work shift since part of their income can vary from month to month. We exclude couples where any one partner is self-employed as administrative income may be distributed between partners such that taxes are minimized. We further exclude all same sex couples and interviews where the randomly chosen respondent within the household was not available, as well as couples where one or both partners are above the retirement age of 65. Finally, we exclude all couples where we observe a deviation between administrative and surveyed incomes of more than 100 percent, which is true for 44 respondents when it comes to their partner's income and 29 respondents when it comes to their own income.[10]

[5]Linking survey respondents' partners to registers has only technically been feasible since 2010. Survey information on the partner's income is only surveyed as part of a special questionnaire, which, since 2010, has only been administered in 2012 and 2015. The years 2012 and 2015 are therefore the only years for which we have both surveyed and administrative income information for both the respondent and his or her partner.

[6]The respective incomes in the social security register are third party reported and there is thus no scope for misreporting.

[7]Foreigners with temporary residence permit are exposed to a special tax scheme, which among other things entails taxation at the source: In contrast to foreigners with a permanent residence permit and Swiss citizens, taxes are withheld. They might thus report their income differently, as their wage statement entails systematically different positions (see, e.g., Schmidheiny and Slotwinski, 2018).

[8]In Switzerland, the difference between gross and net income amounts to approximately 11 percent and consists of contributions to social insurances and pension payments, which are directly deducted by the employer. Except for foreigners with a temporary residence permit, taxes are paid by the individual directly. Net income therefore refers to income after social security and pension payments, but before taxes. Consequently, we observe a baseline underreporting of incomes of about 11 percent in our data (see, e.g., the constant terms Table A.3).

[9]We convert yearly gross into monthly gross incomes by simply dividing yearly incomes by 12.

[10]We do this to exclude extreme cases where the individual reports completely unrealistic numbers in the survey.

The main variables of interest are:

- *Survey income*: Stated income of the survey respondent and of their partner, stated by the surveyed individual.[11]
- *Survey income share (RI^{survey})*: Income share of the woman in the couple based on surveyed incomes. We define it as
  $[Survey\ inc.\ woman/(Survey\ inc.\ man\ +\ Survey\ inc.\ woman)] \times 100$.
- *Administrative income*: Survey respondent's and partner's actual earnings as recorded in social insurance registers. The variable reports total monthly gross income from employment in the month of the interview.
- *Admin income share (RI^{admin})*: Income share of the woman in the couple based on administrative incomes of both partners. We define it as
  $[Admin\ inc.\ woman/(Admin\ inc.\ man\ +\ Admin\ inc.\ woman)] \times 100$.
- *Income deviation*: Deviation between survey income and administrative income for the surveyed individual and the partner. It is defined as
  $[(\ Survey\ inc. - Administrative\ inc.\ )/Administrative\ inc.] \times 100$.[12]

3 Results

Our analysis is presented in three steps: First, we test for a discontinuity in the distribution of survey income shares based on five waves of the SAKE data, which are all available waves in which partner incomes were levied. Subsequently, we test for the presence of the same discontinuity in income shares based on administrative information of the whole population of married couples. Second, we draw on our main sample for which we observe the respondent's and the partner's surveyed and administrative incomes. This allows us to calculate the surveyed and the administrative female income share for the same couple and assess the extent to which the observed discontinuity can be explained by systematic misreporting of one's own and the partner's income around the threshold. These data further allow us to analyze the exact misreporting of couples where the woman outearns her partner based on administrative information, but where the respondent places the couple below the threshold or at the point where they both earn the same based on surveyed incomes. We then explore descriptive statistics of misreporting couples. Third, we discuss broader implications of our finding that there is systematic income misreporting due to gender norms, demonstrating the consequences of the resulting survey bias for estimates of the gender wage gap (GWG).

[11] The original questions are displayed in Appendix A.5.

[12] Only survey income reported as gross can directly be compared to administrative data. We control for the fact that the income deviation between survey and administrative incomes should be larger when income is reported as net by construction by adding an indicator variable for income reported as net in all our estimations.

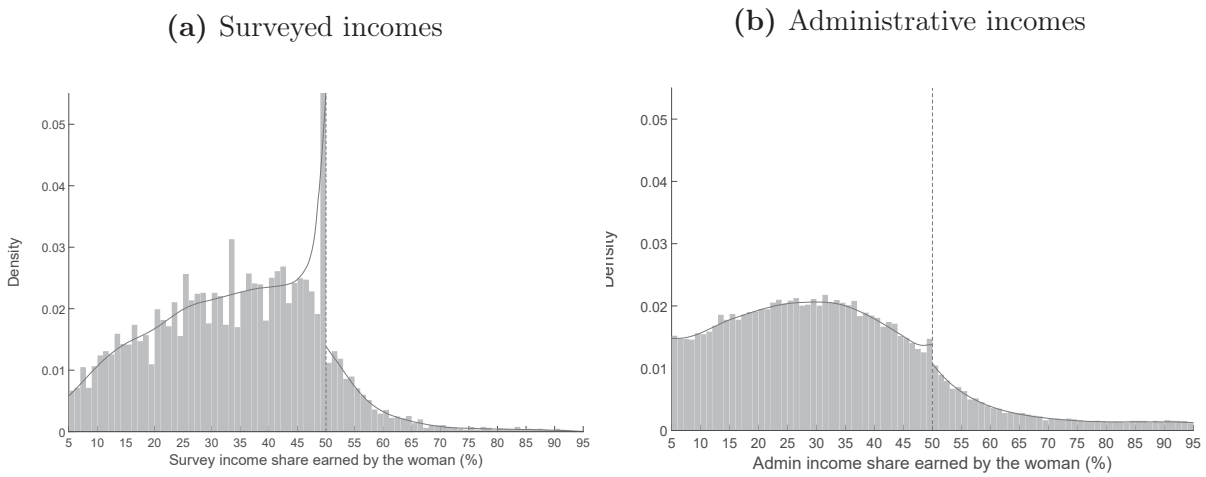

Figure 1: Overall distribution of female income shares and LLD fit of the income share earned by the woman in the couple. The solid line represents the LLD fit on both sides of the threshold using a bandwidth of 7 percent. The shaded area represents the histogram of the underlying data in 1 percent bins. The figure on the left visualizes the distribution observed in survey data (based on SAKE survey years 2002, 2005, 2008, 2012, and 2015). The figure on the right shows the same distribution based on administrative income data for married couples (this data are described in detail in Section A.2 in the appendix). The corresponding density discontinuity estimates can be found in row (1) and row (4) of Table A.1 in Appendix A.4.

3.1 Distribution of female income shares

To test for a discontinuity just above 50 percent in the distribution of female income shares based on survey data, we apply the empirical likelihood-based test by Otsu et al. (2013). This approach has several advantages over the previously proposed approach by McCrary (2008).[13] In a nutshell, it estimates the discontinuity in separate local (linear) likelihood density estimates (LLD) to both sides of the threshold. Consistently, for graphical evidence, we plot two LLD estimates to both sides of the threshold of an income share of women of just above 50 percent.[14]

For the analysis of pure survey responses, we are able to draw on a larger sample of survey data, i.e., all SAKE waves where the question about partner incomes was asked (years 2002, 2005, 2008, 2012, and 2015).[15] However, the link with the administrative information required to analyze misreporting is only possible for the last two waves, leaving us with data from 2012 and 2015 for the detailed analysis of misreporting.

The graph on the left of Figure 1 presents the overall distribution based on surveyed incomes pooling male and female respondents. The distribution visually features a clear

[13]The approach by Otsu et al. (2013) shares the good boundary properties of the local linear estimate. Additionally, the estimator is non-negative by construction, while the McCrary (2008) estimator can produce negative density estimates. See Otsu et al. (2013) for more details about the approach.

[14]We use the local likelihood implementation (locfit) in the Chronux software package for Matlab (Bokil et al., 2010) to fit LLDs to our data. As threshold we use the first observed value above 50 percent, indicated in the respective tables.

[15]In the years 2012 and 2015, the special questionnaire "Social Security" was only administered to two third of individuals questioned in the first wave of the year, which corresponds to about one third of the total sample.

spike just below the margin where the woman outearns her partner and a clear discontinuity. This suggests the presence of a discontinuity in the distribution of female income shares at the point where the woman outearns her partner, just as in other countries investigated in prior studies. We observe rather similar and systematic discontinuities no matter if women or men are surveyed, as shown in Figure A.1 and Table A.1 in Appendix A.4.[16]

The overall discontinuity in the distribution based on survey data amounts to about 4 percentage points. The point estimate just below the threshold is about 4 times as high as the estimate just above the threshold; or to put it differently, the mass drops by about 75 percent at the threshold.[17] The observed drop is very similar to the one found in other studies using survey data (Wieber and Holst, 2015). As this bunching is a local phenomenon, the relative size becomes even larger, about 9 or 88 percent, if we use half the bandwidth, i.e., 3.5 percent (see Table A.2 in Appendix A.4).[18]

The graph on the right of Figure 1 presents the overall distribution based on pooled administrative data for the whole Swiss population of married couples for the years 2010 to 2014. These data are based on the social security register for the whole population and are described in more detail in Section A.2 in the appendix. The graph shows that there is no pronounced spike just below the threshold in administrative data of the whole population. The discontinuity estimates presented in row (4) of Table A.1 are much smaller but remain significant. However, it must be noted that the sample size for this calculation is huge. The drop only amounts to about 20 percent compared to the 75 percent found in survey data. This finding is in no way conclusive, since the population differs between the two data sources and the distribution based on administrative data can therefore not directly be compared to the distribution based on surveyed incomes. It still shows that any discontinuity that might be driven by real responses or institutional factors is magnitudes smaller than survey data would suggest. Bunching driven by institutional factors or real responses would also appear in the administrative distribution.

Our main sample described above allows us to compare the surveyed income share earned by the woman to the administrative female income share for the very same couple and is composed of 3,081 observations.

In Figure 2, we compare the resulting distribution of surveyed female income shares and administrative female income shares for the exact same sample of couples. As above, we observe bunching of mass below the point where a women would outearn her partner in the survey data. When plotting the distribution for the very same couples based on administrative data, we do not observe any bunching below the 50 percent threshold.

[16] We choose the bandwidth to be 7 percent. The optimal bandwidth following McCrary (2008) would be 12.89 and that proposed by Cattaneo et al. (2018) would be 5 percent. Table A.2 in the Appendix repeats the density estimates for half the bandwidth, i.e., 3.5 percent.

[17] The conventional McCrary type approach is less precise would, however, lead to the same conclusions. See Figure A.2 in the Appendix.

[18] This is simply because the incentives to bunch are the strongest around the threshold and we see an exceptionally strong spike just below the threshold. Thus, if we reduce the bandwidth, these observations, i.e., the spike, get more weight in the estimation and the estimated discontinuity increases.

(a) Surveyed incomes **(b)** Administrative incomes

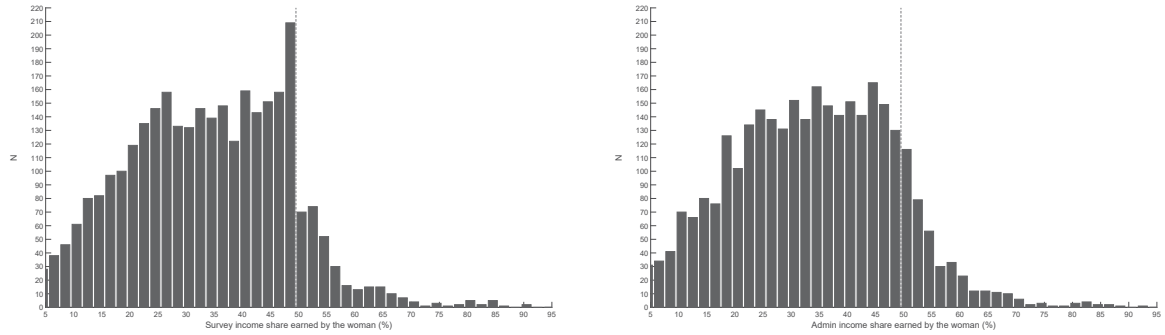

Figure 2: Pure survey and administrative distribution of female income shares for the very same couples. The distributions are presented as raw histograms in 2 percent bins. The corresponding density discontinuity estimates can be found in row (5) and row (6) of Table A.1 in Appendix A.4

This conclusion also holds for the discontinuity estimates: While the survey distribution features a discontinuity just above 50 percent the administrative one does not (see row (5) and row (6) of Table A.1).

The divergence between the surveyed and administrative distributions suggests that the discontinuity in survey data is a survey artifact. The finding that bunching only appears in the surveyed distribution further means that it is unrelated to any real responses to traditional gender norms around the threshold. In the following sections, we analyze whether systematic income misreporting can explain this divergence between the distribution of female income shares in surveyed and administrative incomes.

3.2 Misreporting of own and partner incomes

A violation of the male breadwinner norm entails a cost by contradicting individuals' self-perception. While real labor market adaptations are also costly, misreporting of actual incomes allows survey respondents to comply with the male breadwinner norm and, at the same time, maintain their self-image at a low cost (Akerlof and Kranton, 2000). It therefore allows individuals to resolve the cognitive dissonance of violating traditional gender norms without bearing the costs they would incur if they adapted their labor market behavior.

Such systematic misreporting would be reflected in women who earn more than their partner underreporting their own and/or overreporting their partner's income. If a woman earning slightly more than her partner underreports her income such that it falls below or equals her partner's, this will result in two changes: First, the percentage difference between surveyed and administrative income will increase in absolute terms. Second, since the woman now reports she earns less than or the same as her partner, the couple will find themselves below 50 percent in the distribution of surveyed female income shares. A man whose partner earns just more and who wants to conform with the social norm that

a woman should not earn more would, respectively, overreport his own income and/or underreport his partner's income. Such behavior would result in a strong selection of individuals around the threshold in the distribution of female income shares based on survey data. Those conforming with the norm by misreporting their own or their partner's income are placed below the threshold and those not conforming with the norm and who therefore state that the woman earns more above the 50 percent margin in the distribution of surveyed female income shares. This selection would be a consistent explanation for the divergence between the administrative and the survey based distribution of female income shares.

As we observe the position of couples in the distribution of female income shares based on administrative information, we are able to identify those individuals who cross the threshold with their survey response, i.e., couples where the woman outearns her partner based on their administrative incomes but earns less or the same as her partner according to the survey response (termed norm-complying misreporters in the following).

Figure 3 visualizes the correlation between the administrative female income share (on the x-axis) and the surveyed female income share (on the y-axis). Overall, there is a strong positive correlation between the surveyed and administrative female income share.[19] Couples where the woman outearns her partner based on administrative information but who report that the woman earns less or the same as her partner are shown in red, the opposite being true for couples in blue. We see that couples crossing the threshold from above (red) are placed between 51 and 70 percent in the distribution of female income shares based on administrative data and place themselves between the 20^{th} and the 50^{th} percent bin when surveyed, with the majority repositioning themselves from the range of 5 percent above 50 percent to the the five percent just below.

The number of couples crossing the threshold is visualized in Figure 4. The gray bars in graph (a) show the distribution of surveyed female income shares. The red bars show the number of couples in the survey distribution who cross the threshold from above and in blue the individuals who cross the threshold from below. We see that there are many couples we termed norm-complying misreporters. In particular, about 40 percent of couples in the bar just below the threshold are placed above in the administrative distribution. This constitutes direct evidence that a considerable part of the excess mass just below the threshold is driven by couples where the woman outearns her partner based on administrative information. There are considerably fewer individuals who place the couple above the threshold in the survey distribution and are placed below in the administrative distribution, i.e., whose deviation between survey and administrative incomes leads them to violate the norm (norm non-complying misreporters).

[19]This correlation is also visible when looking at individuals' surveyed and administrative incomes, as shown in Figure A.3 in the appendix. We see that individuals, on average, know what they and their partner earn. This is true for female and male respondents alike. The correlation between administrative and surveyed incomes is reasonably high at 0.93 for female respondents and their own income and 0.87 for female respondents and their partner's incomes. The same numbers for male respondents are 0.92 for their own and 0.92 for their partner's income, respectively.

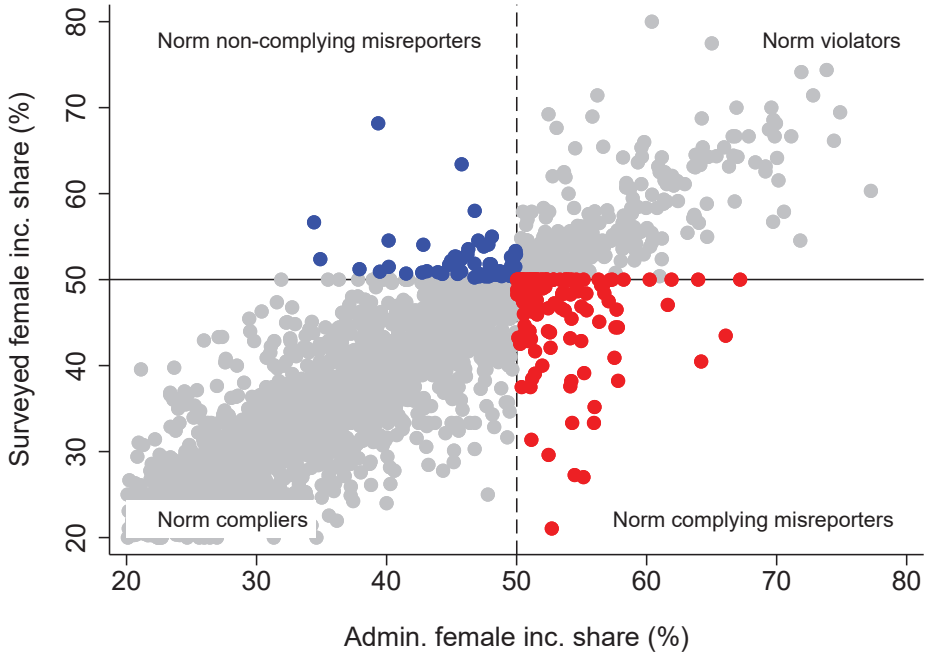

Figure 3: Surveyed vs. admin. based female income shares. This figure shows the scatter plot between the female income share resulting from survey data and from administrative data respectively, and pooling male and female respondents. The red dots mark individuals who cross the threshold from above (whose administrative female income share is above 50 percent but the surveyed one lies below or equals 50 percent) and the blue dots mark those individuals who crossed the threshold from below (whose administrative female income share is below or equal to 50 percent but the surveyed one lies above).

It is important to note that the pattern cannot simply be explained by random deviations in income reporting of one's own or the partner's income. Theoretically, if what we observe were caused by random deviations, we would observe more norm non-complying misreporters (in blue) than norm complying misreporters (in red), as the actual mass below the threshold is considerably higher than the mass above. Figure 4 (b) visualizes a simulated distribution of surveyed female income shares. It demonstrates what a distribution of female income shares based on survey data would look like if deviations between administrative and surveyed incomes were random.[20] As expected, if misreporting were random, there would be more norm non-complying misreporters in blue than norm complying misreporters in red. This is confirmed when we repeat this random assignment of deviations in surveyed incomes 1,000 times. We observe that 202 couples cross the threshold, of which 61 couples (30.2%) cross from below and 141 couples (69.8%) from above. In

[20]More precisely, in order to determine the average deviation between surveyed and administrative incomes for the simulation, we regress the deviation of surveyed and administrative incomes on administrative incomes and an indicator whether incomes are reported as gross or net (since this affects the deviation). We then use each individual's administrative incomes and add a random error to simulate random reporting deviations. This error is drawn from a normal distribution with mean and standard deviation determined by mean misreporting of an individual for each income, determined by the regression described before. We do this separately for men and women, as well as for their own and their partner's income.

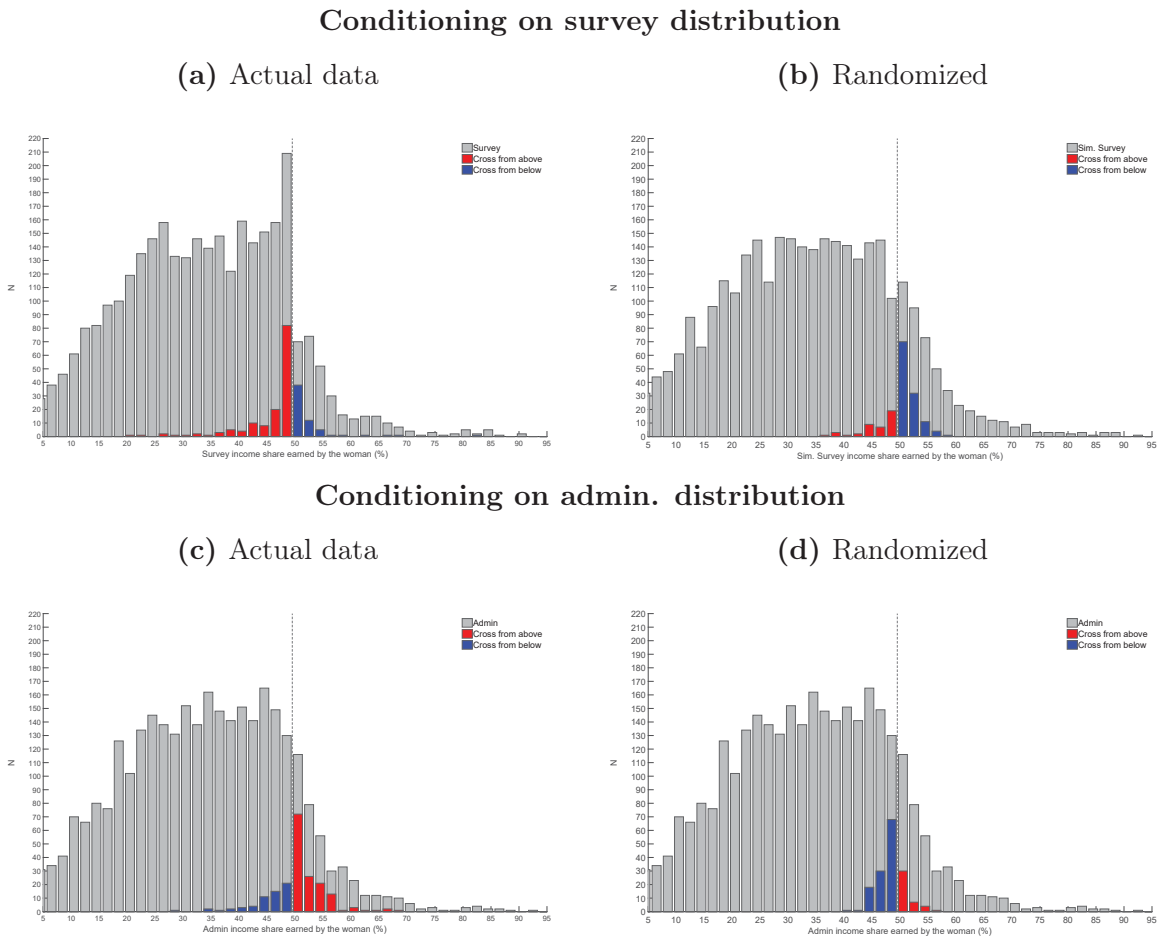

Figure 4: Histograms visualizing the number of couples finding themselves on a different side of the threshold when using survey vs. admin. data. In all four graphs, the red bars show the number of couples where the woman outearns her partner based on administrative information but earns less or the same based on survey information (norm complying misreporters). The blue bars show the number of couples where the woman earns less or the same based on administrative data and outearns the partner based on survey data (norm non-complying misreporters). The four graphs differ regarding the underlying distribution. The distributions are presented as raw histograms in 2 percent bins. The histogram in the first row and to the left shows the distribution of surveyed female income shares and the histogram to the right visualizes the distribution of income shares based on administrative data with random misreporting (in gray). The histograms in the second row show the distribution of administrative income shares.

the simulation, an average of 175.342 couples cross the threshold, of which 125.736 couples (72%) cross from below and only 49.606 couples (28%) cross from above. Consequently, in our sample, the number of individuals who place themselves below instead of above the threshold is disproportionately higher, alleviating any concerns that the pattern we observe could be driven by random reporting errors. This again leaves systematic repositioning of couples where the woman outearns the man to below the threshold as the most likely explanation.

Figure 4 (c) shows the same point, this time conditioning on the administrative female income share. Conditioning on the administrative female income share allows us to see

where norm complying misreporters and norm non-complying misreporters originate. The gray bars show the distribution of the administrative income share. The red bars show the number of couples for which the respondent reports a female income share below the threshold in the survey and their position in the administrative distribution (norm complying misreporters in red). As shown above, their administrative female income shares are mainly between the 51 and the 60 percent bin. There are much fewer couples for which the respondent reports a female income share above 50 percent and whose administrative female income share lies below (blue). As before Figure 4 (d) visualizes the equivalent if there were a random reporting error. Again, we see that if the reporting error were random we would observe only few couples who report to have an income share below 50 percent and whose administrative female income share is above, while we would observe more norm non-complying misreporters in blue. This clearly shows that there is a systematic repositioning of couples from above to below the threshold.

In what follows, we study the shift from above 50 percent in the administrative income share to below or equal to 50 percent in the surveyed distribution of income shares earned by the woman in more detail. This analysis is based on the following equation:

$$\begin{aligned} y_i = \ & \beta_1 \times \mathbb{1}[RI_i^{admin} \leq 40] + \textstyle\sum_{k=21}^{29} \beta_k \times \mathbb{1}[(2k-2) < RI_i^{admin} \leq 2k] \\ & + \beta_{30} \times \mathbb{1}[58 < RI_i^{admin}] + \rho \times mode_i + u_i, \end{aligned} \tag{1}$$

where y_i describes the outcome variable, which is either the reporting difference in the female income share (*surveyed female inc. share - admin. female inc. share*) or the income deviation (Δy =*(survey inc.- admin. inc.)/admin inc.*). The estimates of β_k therefore estimate the average in outcome y for each bin of the distribution of the administrative female income share (RI^{admin}), controlling for the mode income is reported in (*mode*) which can either be monthly gross, monthly net, or yearly gross. Two percent bins are included for the range between 40 and 58 percent, while the observations below 40 and those above 58 percent are grouped in wider bins.

We further interact the bins above 50 percent with an indicator taking value 1 when couples have an administrative relative income share above 50 percent and a surveyed relative income share below or equal to 50 percent (*norm compl. mirep.*), which allows us to distinguish between the behavior of norm complying misreporters and norm violators. We do this using the following equation:

$$\begin{aligned} y_i = \ & \beta_1 \times \mathbb{1}[RI_i^{admin} \leq 40] + \textstyle\sum_{k=21}^{29} \beta_k \times \mathbb{1}[(2k-2) < RI_i^{admin} \leq 2k] \\ & + \beta_{30} \times \mathbb{1}[58 < RI_i^{admin}] \\ & + \left\{ \textstyle\sum_{k=26}^{29} \gamma_k \times \mathbb{1}[(2k-2) < RI_i^{admin} \leq 2k] + \beta_{30} \times \mathbb{1}[58 < RI_i^{admin}] \right\} \\ & \times \mathbb{1}[norm\ compl.\ mirep._i] + \rho \times mode_i + u_i \end{aligned} \tag{2}$$

The coefficients γ_k therefore indicate the average differential between norm complying misreporters and norm violators for the respective bin.

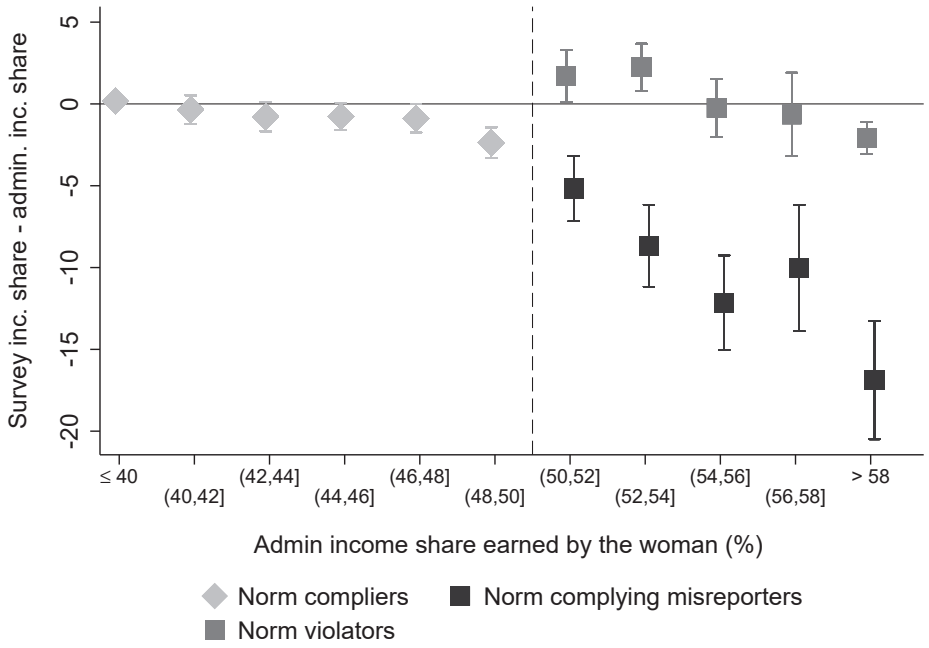

Figure 5: Deviation between the surveyed and administrative income share for bins of the distribution of administrative income shares earned by the woman. The estimates are based on regressing the deviation between the administrative and survey income share on bins of the administrative distribution and an indicator for the income response mode. For the bins above 50 percent of the administrative distribution, we additionally interact the indicator for the bin with an indicator for norm complying misreporters, i.e., for couples above 50 percent in the distribution based on administrative data but with a female income share below or equal to 50 percent in the distribution based on survey data (see, Eq. 2).

Figure 5 visualizes individuals' deviations between the surveyed and the administrative female income share conditioning on the position in the administrative distribution. Up to a female income share of 40 percent, there is on average no deviation between the surveyed and the administrative female income share of a couple. After that, we find a visible decrease in the surveyed income share compared to the administrative income share, indicating that individuals start to underreport female incomes and/or overreport male incomes in surveys. This deviation increases for couples just below the threshold and continues to increase for couples where the woman outearns her partner based on administrative data. More specifically, norm complying misreporters, i.e., individuals who cross the threshold from above by their survey response, systematically underreport their female income share compared to other individuals in the same bin.

The observation that the deviation is already negative in the bins just below the 50 percent threshold indicates that couples where the woman earns less or the same as her partner but who are close to the 50 percent threshold already start to underreport their income share in surveys, which speaks against an equality norm argument. Such an equality norm would state that a couple's goal is to state equal earnings in surveys.[21]

[21] Even if we focus only on norm complying misreporters, only about 35% percent report equal earnings. Equality is thus not the main driver behind our findings.

This observation indicates that individuals just below the threshold preventively react to the norm and start to adapt their income reporting such that they do not violate the male breadwinner norm. For norm violators, i.e., couples where the woman earns more and who report this as such, we see two slightly positive estimates for the first two bins above the threshold. This indicates a slight overreporting of the female income share. While it is orders of magnitudes smaller than the underreporting of those who cross the threshold from above, it still might indicate that individuals who are willing to violate traditional gender norms are more conscious about the female income share and want to prevent the situation from being misrepresented.

In a next step, we want to shed light on how specifically respondents misreport and achieve this systematic repositioning from above to below the threshold. In principle, there are three possible ways: One can either misreport one's own or one's partner's income or both.

In order to see which strategy is applied when it comes to income reporting, we analyze the average deviation between surveyed and administrative incomes of couples above 50 percent of the administrative distribution of female income shares. For bins spanning 2 percent, we calculate the average deviation of norm violators (couples where the woman earns more and who report this as such) and norm complying misreporters (couples where the woman earns more who, however, report that she earns less) controlling for the income response mode (see Equation 2 for the full specification).

Figure 6 visualizes the excess deviation of norm complying misreporters (the excess income deviation compared to other individuals in the same bin, estimated by $\hat{\gamma_k}$). The first row shows misreporting conditioning on administrative income shares. Panel a) shows that female respondents who cross the threshold from above underreport their own income (negative values, black dots) and simultaneously overreport the income of their partner (positive values, gray diamonds) compared to the average individual in the bin. The size of misreporting increases in absolute terms in the position in the distribution of administrative income shares: the farther above, the higher the required misreporting to conform with the norm.

We see a consistent picture for male respondents in panel b). Male norm complying misreporters overreport their own income and underreport the income of their partner. Consequently, both own and partner incomes are misreported in order to comply with traditional gender norms. For both male and female respondents, overreporting of male incomes seems to be clearer and more systematic.

The second row in Figure 6 shows another perspective on misreporting of norm-complying misreporters, this time conditioning on the reported position in the surveyed female income share and thus using norm compliers as the reference group. We see that female respondents who place themselves just below the threshold (in the range between 48 and 50 percent), underreport their own and overreport their partner's income systematically, but only to a small extent. The further below respondents place themselves in the surveyed distribution, the larger misreporting becomes. The picture for male respon-

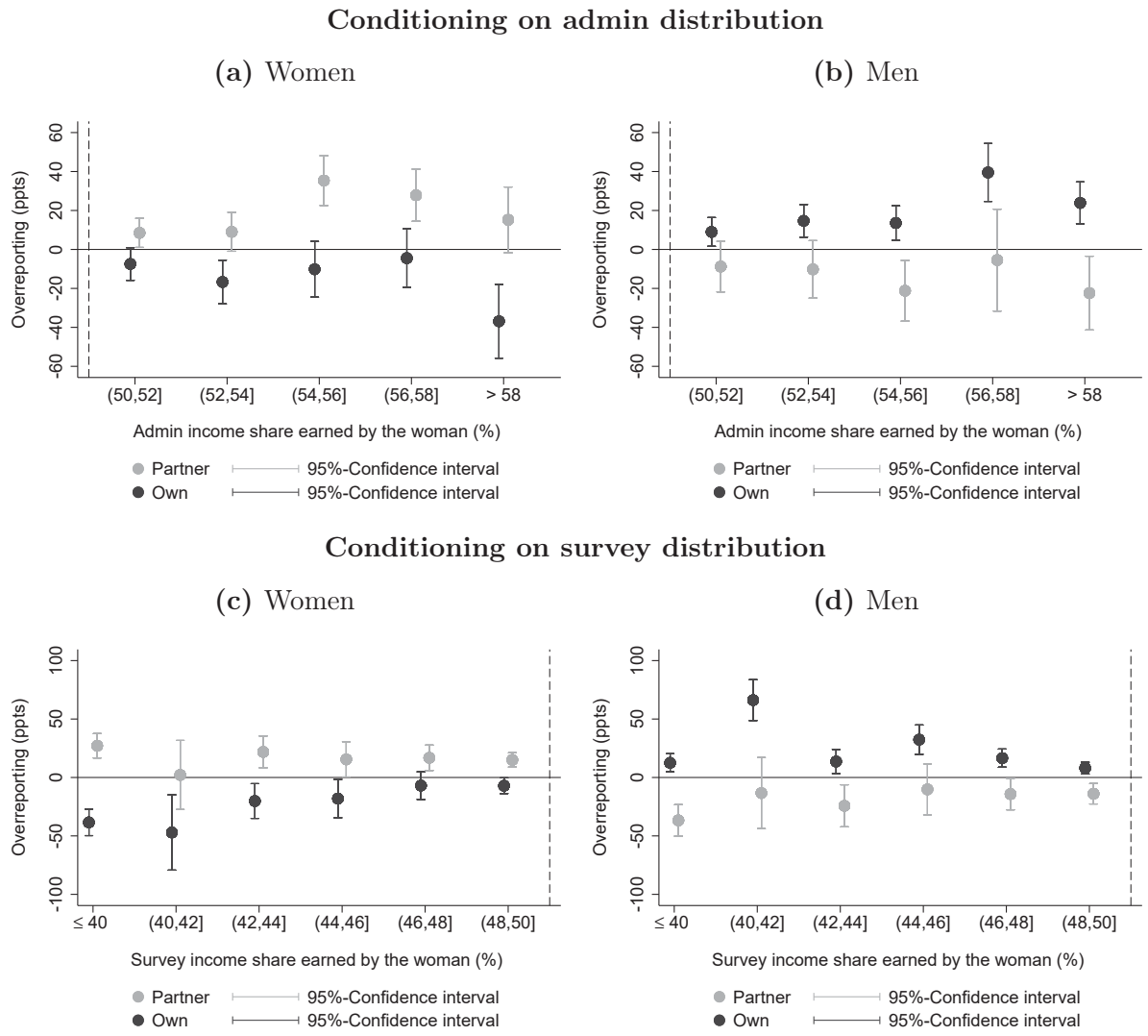

Figure 6: The upper panel shows excess misreporting of norm complying misreporters compared to other individuals in the same bin, conditional on the administrative distribution for women and men. It shows for each bin whether individuals misreport their own or their partner's income to fall below the 50 percent threshold. The lower panel shows misreporting conditional on the survey distribution for women and men. It shows for each bin whether individuals sorting themselves to the specific bin do this by misreporting their own or their partner's income.

dents in panel d) shows a analogous picture. They underreport their partner's income and overreport their own incomes. Many individuals choose to place themselves just below the threshold, contributing to the narrow confidence bands in the bin just below the threshold. While Figure 6 presents the estimated differentials, Figures A.4 and A.5 in the appendix show the respective level estimates.

Before, we have seen that couples where the women earns a similar amount or slightly less than her partner already start misreporting their income share. Figure A.4 in the Appendix shows the level estimates of their income misreporting of own and partner incomes. This slight precautionary misreporting seems to mostly be driven by overreporting (female respondents) or underreporting (male respondents) their partner's incomes. This

implies that there is already a slight bias in respondents' surveyed incomes for couples which fall below the 50 percent margin in administrative income shares and further, that the norm seems to even be salient for individuals who, in actual terms. conform. Additionally, this precautionary misreporting confirms that what we see is not driven by a pursuit of income equality within the couple but that it is very one-dimensional, raising male incomes and depressing female incomes.

We document that in our data, the entirety of the discontinuity at the point where the woman would outearn her partner observed in survey data is explained by misreporting of own and partner incomes. This, however, might not necessarily be fully applicable to other countries. In Switzerland, there is, for instance, no particular tradition of collective wage agreements applying to whole sectors and there is no general minimum wage. This makes it less likely for couples to earn the exact same income, even if working in the same sector. Additionally, married couples are taxed jointly and there is no tax incentive to equalize earnings. There are thus good reasons to expect that some spike due to responses to institutional incentives would persist in other countries.

Nevertheless, our results stress that studies investigating this very phenomenon in survey data capture systematic income misreporting to a large extent. This misreporting is indeed driven by gender norms. It does, however, not seem to affect real labor market decisions.

3.3 Descriptive evidence on norm-complying misreporters

In this section we present descriptively the characteristics correlated with being a norm-complying misreporter. Figure 7 shows how, on average, norm-complying misreporters differ from norm violators. We focus on couples where based on administrative data, the woman earns more than her partner. The distinction between norm complying misreporters and norm violators shows what characteristics are correlated with respondents reporting incomes in surveys which lead to a surveyed female income share below the threshold compared to couples reporting incomes such that they remain above the threshold. Panel a) presents the comparison of socio-demographics. Norm complying misreporters do not differ from norm violators when it comes to the average age of the man, the woman, nor the average household income. While there is no difference in the probability that the man in the couple has tertiary education, women in couples that place themselves below the 50 percent threshold in surveys are less likely to hold tertiary education. We find that norm complying miserporters are neither more nor less likely to have children. Except for the observation that couples where the woman is highly educated are less likely to misreport, norm complying misreporters and norm violators seem not to differ systematically.

Panel b) shows a comparison of observable characteristics known to be correlated with gender norms to proxy groups we expect to be more *traditional* and test whether these groups are indeed more likely to be norm complying misreporters. In a first step, we

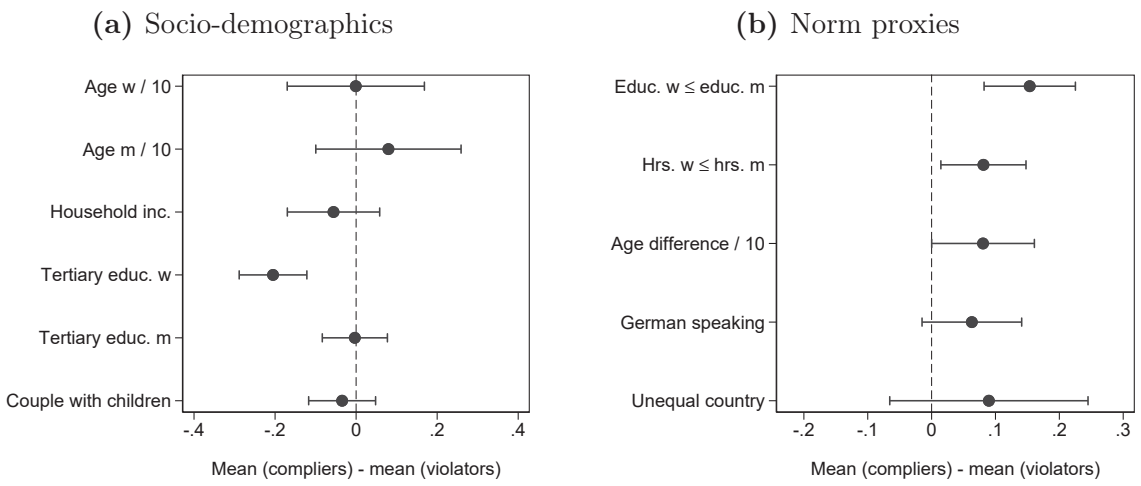

Figure 7: Average comparison between norm complying misreporters and norm violators. This Figure shows the difference in the group averages between norm violators (admin. female income share $>$ 50 percent, surveyed female income share $>$ 50 percent) and norm complying misreporters (compliers, admin. female income share $>$ 50 percent, surveyed female income share $\leq$ 50 percent) for different characteristics. The estimates only include individuals with an administrative female income share above 50 percent. Panel a) presents the comparison of sociodemographic characteristics and panel b) the comparison of observables known to be correlated with traditional gender norms. We pool the data for male and female respondents and report 90 percent confidence bounds. The corresponding estimates can be found in Tables A.5 and A.6. *Hrs.* stands for weekly work hours. *Age diff.* describes the within-couple age difference defined as $(age_m - age_w)$. *German speaking* describes an indicator set to one for German speaking individuals (in relation to French, Italian, or Romansh speaking). *Unequal country* describes an indicator set to one for individuals with origins in a country with more traditional gender norms.

try to capture situations which would jeopardize the status of a man holding traditional gender norms in a relationship. This approach is based on the findings of Fisman et al. (2006) and Bursztyn et al. (2017), who document that men consider women less attractive if the woman's ambitions exceed their own.[22] The first measure we employ is the relative education within a couple. The idea is that it might be perceived as fair if the women earns more if she is also more educated than the man. A situation where the woman outearns her partner and is equally or less educated might, however, provoke discomfort and be a threat to the male identity (Akerlof and Kranton, 2000). We test for a difference in the probability that a norm complying misreporter is part of a couple where the man is equally or more educated and is still outearned by the woman, and find that these couples are more likely to be norm complying misreporters. Or to put it differently, the share of couples where the woman is less or equally educated is 15 percentage points higher in the group of norm complying misreporters when compared to norm violators. The difference amounts to about 20 percent of the sample average (76 percent). This evidence is in line with the idea that situations which might be a threat to the male breadwinner identity provoke misreporting. Another situation, which could produce similar unease

[22]It is further in line with social structure theory (Eagly and Wood, 1999).

is when the woman works the same or fewer hours per week as her partner but earns more. Considering this measure, we find that couples where the woman works fewer or the same hours are systematically more likely to misreport (the difference amounts to about 10 percent of the sample average). These findings suggest that situations which jeopardize the male position within the relationship (or the male identity) seem to be a good predictor for norm-complying misreporting.

In a next step, we analyze direct proxies of traditional gender norms known in the literature. First, we exploit a proxy which we observe for all couples in the sample. Following Folke and Rickne (2018), we use the within-couple age difference as a proxy of an individual's gender norms when entering the relationship. Couples where the man is older than the woman are assumed to be, on average, more compliant with traditional gender norms.[23] Consistently we observe that the within-couple age difference is on average higher in misreporting couples (the difference amounts to 35 percent of the sample average).

Second, we use the language the survey interview was conducted in as an approximation of gender norms. Prior studies have shown that individuals in German speaking areas of Switzerland hold more traditional gender norms than individuals in the other language regions (Italian, French, and Romansh) (see, e.g., Steinhauer, 2013). We would thus expect that norm-complying misreporters are more likely to speak German than one of the other languages. This is what we find. Misreporters seem to be on average more likely to be German speaking. While the difference is not significant at any of the conventional levels, it amounts to about 9 percent of the sample average.

As a final norm proxy, we use cultural norms in the countries of origin of migrants to approximate individuals who we expect to hold more traditional gernder norms. We exploit the fact that Switzerland has a comparatively high share of immigrants and apply the epidemiological approach suggested in Fernández and Fogli (2009). We approximate an individual's norms by gender norms in the country of their ancestry. The basic idea is that individuals take part of the culture (through socialization) with them when emigrating. These norms are to some extent also transmitted intergenerationally. As these individuals live and partly grow up in the same country and institutional setting, differences in their behavior should emerge through these transmitted gender norms. Traditional gender norms are proxied by average agreement of all employed women in a country with the statement "When jobs are scarce, men should have more right to a job than women", as measured in the WVS. Splitting countries at the sample median, we define two types of origin countries: countries where average agreement is lower, which we would expect to hold less traditional gender norms, and countries where agreement is higher, which we expect to hold more traditional gender norms.[24] The difference for the likelihood that a norm complying misreporter originates from a more gender traditional country is not statistically significant at conventional levels. However, the sample is of course quite small

[23] In our sample, men are at the median two years older than their partner.

[24] Detailed information on how we determine an individual's ancestry and the definition of equal and unequal countries can be found in Section A.3 of Appendix 2.

(about 25 percent of the original sample). Still the difference is considerable. Given a baseline probability of 30 percent to be from a country with traditional gender norms, norm complyiers are about 9 percentage points more likely to originate from more traditional countries, which amounts to about 30 percent of the sample mean.[25] Summing up, measures of traditional gender norms seem to be good predictors for individuals' misreporting behavior, strengthening the interpretation that individuals self-conceptions and norms are the drivers of such behavior.

4 Discussion

Our evidence strikingly demonstrates that gender norms play an important role in individuals' responses to surveys and can lead to systematic misreporting. In this section, we demonstrate that the survey bias might also translate into a systematic bias in policy relevant measures frequently studied by economists and social scientists.

Due to a lack of high quality administrative data, social scientists have long been confined to survey data for their analyses. Furthermore, survey data often include information on topics such as voting behavior, division of family responsibilities, or employment in the informal sector, which cannot be elicited from administrative sources. It is all the more important for researchers to be aware of potential biases in survey information. This is relevant in individuals' responses regarding their own characteristics, but might be even more pronounced when proxy responses are used. Proxy interviews are frequently used to reduce survey costs and might lead to biased information, as pointed out by Reynolds and Wenger (2012) and Lee and Lee (2012). In 2009, for example, almost half of the labor force data in the Current Population Survey (CPS) are provided by proxy respondents (Reynolds and Wenger, 2012).[26] Particularly for information on sensitive topics or questions prone to the influence of norms, survey data may be considerably biased. According to our evidence this is a problem that concerns both information from interviews with the target person directly and proxy interviews (individuals' responses about their partner).

Our data offers the unique opportunity to assess how strongly the use of survey data and proxy interviews could bias, for example, estimates of the gender wage gap (GWG). In particular, we assess the bias in a sample for which we observe equally many women and men and in which the selection into response by gender does not play a role as the reporting person within the household is randomly chosen.

We start by calculating the *true* GWG based on administrative information for the respondents. In particular, we regress log hourly incomes on gender, controlling for the

[25]The fact that in only 30 percent of couples where the woman outearns her partner based on administrative data, the respondent is from an unequal country serves as additional indication that a couple's relative income share is related to gender norms. Based on the fact that we use the within sample median to split countries into equal and unequal, this number would be 50 percent if couples were equally distributed across relative income shares.

[26]The CPS is one of the most frequently used US surveys in economics. Among other things, many studies on the GWG (see, e.g., Blau and Kahn, 2000, 1997; Katz and Murphy, 1992; Macpherson and Hirsch, 1995) rely on CPS data.

Table 1: Gender wage gaps

	GWG admin.	GWG survey	GWG admin. including partner	GWG survey with proxies
	(1)	(2)	(3)	(4)
Female	-0.149***	-0.163***	-0.171***	-0.194***
	(0.014)	(0.012)	(0.010)	(0.009)
Net	Yes	Yes	Yes	Yes
Age	Yes	Yes	Yes	Yes
Education	Yes	Yes	Yes	Yes
Constant	Yes	Yes	Yes	Yes
N	3,078	3,078	6,071	6,071
R-squared	0.306	0.394	0.276	0.325

Notes: OLS estimates of the gender wage gap regressing log hourly income on indicator variables for female, education, and if income is reported as net, a continuous age variable and a constant. The four specifications distinguish between the data source (survey or admin.) and whether proxy incomes are used for the estimation. We loose some observations due to missing values in education of the respondent or the partner. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

level of education, individuals' age as well as the mode of the income response. This results in a GWG of 14.9 percent (see Table 1 column (1)). In a next step, we asses the GWG estimate based on surveyed incomes for the very same sample of individuals. As seen before, there is some misreporting just below the threshold and systematic and pronounced misreporting of norm complying misreporters and their own incomes above the threshold. Based on surveyed incomes, the GWG amounts to 16.3 percent (see Table 1 column (2)) and is thus overestimated by about 9.4 percent.

In a next step, we mimic the data structure of surveys with proxy information (like the CPS) and include administrative information on the respondent and the partner as independent observations. This naturally almost doubles the number of observations.[27] Based on this sample, we find a true (administrative) GWG of 17.1 percent.[28] In column (4), proxy income information is included, i.e., for each individual, we use both their own reported income, as well as the income they report for their partner as our data base; just like it is done in the CPS data. Based on the proxy information, the GWG now amounts to 19.4 percent (see Table 1 (4)), which overestimates the *true* GWG in the administrative data by about 13.5 percent and thus introduces a considerable bias. Both comparisons of the administrative and survey GWG differ significantly at the 10 percent level.

This additional analysis strikingly shows that the misreporting of incomes in surveys due to gender norms can lead to considerable biases in estimates of gender differences. This bias becomes more substantial if proxy information is used. The use of proxy interviews

[27] We loose some observations if there are missing values for the characteristics of the partner.

[28] In the survey, information on the partner's weekly work hours is only given in classes of five hours, ranging from 1 (0-5 hours) to 9 (more than 40 hours). However, we have the true weekly work hours for the respondent. Based on the this information we impute the respective average hours per class.

does not seem to be rare in datasets used by economists, which underscores the broader relevance of our findings.

5 Conclusion

A growing literature in economics strives to understand whether gender norms and roles drive economic agents' behavior. Due to the lack of high quality administrative data, survey data has been the main data basis available to learn about these questions for a long time. If, however, gender norms strongly impact individuals' response behavior, survey information might be less informative about agents' real choices than we thought.

Comparing surveyed and administrative outcomes, we demonstrate that individuals' survey responses are strikingly prone to the influence of social norms. Drawing on prior work by Bertrand et al. (2015), we introduce misreporting in survey data as an additional explanation behind the debated mechanism driving the discontinuity in the distribution of female income shares at the point where the woman outearns her partner.

In particular, we draw on a unique dataset available for Switzerland, which combines surveyed and administrative information for the same individual. We document that the distinct discontinuity in the distribution of female income shares in survey data is indeed driven by gender norms, which materialize in systematic income misreporting of couples where the woman earns more than her partner. There is, however, no clear indication of real labor market responses around the same margin. We substantiate our findings by demonstrating that the GWG, a policy relevant measure frequently retrieved from survey data, is vastly overestimated in survey data and even stronger if we were to use proxy information, a standard practice in many population surveys.

Data Statement

The confidential individual-level data (SESAM) were obtained under contract Nr. 180262 from the Swiss Statistical Office. The data from the Swiss Structural Survey (Strukturerhebung) were obtained under contract Nr. 170334. Access is granted for scientific research projects after review of a detailed application.[29]

[29] See https://www.bfs.admin.ch/bfs/de/home/statistiken/arbeit-erwerb/erhebungen/sesam.html

References

Akerlof, G. A. and R. E. Kranton (2000). Economics and identity. *Quarterly Journal of Economics 115*(3), 715–753.

Alesina, A., P. Giuliano, and N. Nunn (2013). On the Origins of gender roles: Women and the plough. *The Quarterly Journal of Economics 128*(2), 469–530.

Bertrand, M., E. Kamenica, and J. Pan (2015). Gender identity and relative income within households. *Quarterly Journal of Economics 130*(2), 571–614.

Binder, A. J. and D. Lam (2018). Is there a male breadwinner norm? The hazards of inferring preferences from marriage market outcomes. NBER Working Paper 24907.

Blau, F. D. and L. M. Kahn (1997). Swimming upstream: Trends in the gender wage differential in the 1980s. *Journal of Labor Economics 15*(1, Part 1), 1–42.

Blau, F. D. and L. M. Kahn (2000). Gender differences in pay. *Journal of Economic Perspectives 14*(4), 75–99.

Bokil, H., P. Andrews, J. Kulkarni, S. Mehta, and P. P. Mitra (2010). Chronux: A platform for analyzing neural signals. *Journal of Neuroscience Methods 192*(1), 146–151.

Bursztyn, L., T. Fujiwara, and A. Pallais (2017). 'Acting wife': Marriage market incentives and labor market investments. *American Economic Review 107*(11), 3288–3319.

Cattaneo, M. D., M. Jansson, and X. Ma (2018). Manipulation testing based on density discontinuity. *Stata Journal 18*(1), 2034–261.

Eagly, A. H. and W. Wood (1999). The origins of sex differences in human behavior: Evolved dispositions versus social roles. *American Psychologist 54*(6), 408.

Eriksson, K. H. and A. Stenberg (2015). Gender identity and relative income within households: Evidence from Sweden. IZA DP No. 9533, IZA.

Fernández, R. and A. Fogli (2009). Culture: An empirical investigation of beliefs, work, and fertility. *American Economic Journal: Macroeconomics 1*(1), 146–77.

Fernández, R., A. Fogli, and C. Olivetti (2004). Mothers and sons: Preference formation and female labor force dynamics. *Quarterly Journal of Economics 119*(4), 1249–1299.

Fisman, R., S. S. Iyengar, E. Kamenica, and I. Simonson (2006). Gender differences in mate selection: Evidence from a speed dating experiment. *The Quarterly Journal of Economics 121*(2), 673–697.

Folke, O. and J. Rickne (2018). All the single ladies: Job promotions and the durability of marriage. *forthcoming in American Economic Journal: Applied Economics*.

Fortin, N. M. (2005). Gender role attitudes and the labour-market outcomes of women across OECD countries. *Oxford Review of Economic Policy 21*(3), 416–438.

Funk, P. (2016). How accurate are surveyed preferences for public policies? Evidence from a unique institutional setup. *Review of Economics and Statistics 98*(3), 442–454.

Gil, J. and T. Mora (2011). The determinants of misreporting weight and height: The role of social norms. *Economics & Human Biology 9*(1), 78 – 91.

Giuliano, P. (2017). Gender: A historical perspective. In S. L. Averett, L. M. Argys, and S. D. Hoffman (Eds.), *The Oxford Handbook of Women and the Economy*, pp. 1–32. New York, NY: Oxford University Press.

Goldin, C. (2006). The quiet revolution that transformed women's employment, education, and family. *American Economic Review 96*(2), 1–21.

Hariri, J. G. and D. D. Lassen (2017). Income and outcomes: Social desirability bias distorts measurements of the relationship between income and political behavior. *Public Opinion Quarterly 81*(2), 564–576.

Katz, L. F. and K. M. Murphy (1992). Changes in relative wages, 1963–1987: Supply and demand factors. *The Quarterly Journal of Economics 107*(1), 35–78.

Lee, J. and S. Lee (2012). Does it matter who responded to the survey? Trends in the us gender earnings gap revisited. *ILR Review 65*(1), 148–160.

Lippmann, Q., A. Georgieff, and C. Senik (2019). Undoing gender with institutions: Lessons from the German division and reunification. *forthcoming in The Economic Journal*.

Macpherson, D. A. and B. T. Hirsch (1995). Wages and gender composition: Why do women's jobs pay less? *Journal of Labor Economics 13*(3), 426–471.

Martinelli, C. and S. W. Parker (2009). Deception and misreporting in a social program. *Journal of the European Economic Association 7*(4), 886–908.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics 142*(2), 698–714.

Murray-Close, M. and M. L. Heggeness (2018). Manning up and womaning down: How husbands and wives report their earnings when she earns more. SESHD Working Paper No. 2018-20.

Otsu, T., K.-l. Xu, and Y. Matsushita (2013). Estimation and inference of discontinuity in density. *Journal of Business & Economic Statistics 31*(4), 507–524.

Reynolds, J. and J. B. Wenger (2012). He said, she said: The gender wage gap according to self and proxy reports in the current population survey. *Social Science Research 41*(2), 392–411.

Schmidheiny, K. and M. Slotwinski (2018). Tax-induced mobility: Evidence from a foreigners' tax scheme in switzerland. *Journal of Public Economics*.

Sprengholz, M., A. Wieber, and E. Holst (2019). Gender identity and wives' labor market outcomes in West and East Germany between 1984 and 2016. DIW Berlin Discussion Paper 1799.

Steiner, I. and P. Wanner (2015). Towards a new data set for the analysis of migration and integration in Switzerland. nccr – on the move, Working Paper No. 1, nccr – on the move.

Steinhauer, A. (2013). Identity, working moms, and childlessness: Evidence from Switzerland. Mimeo, University of Zurich.

Teso, E. (2019). The long-lerm effect of demographic shocks on the evolution of gender roles: Evidence from the transatlantic slave trade. *Journal of the European Economic Association 17*(2), 497–534.

Wieber, A. and E. Holst (2015). Gender identity and women's supply of labor and non-market work: Panel data evidence from Germany. IZA DP No. 9471.

Zinovyeva, N. and M. Tverdostup (2018). Gender identity, co-working spouses and relative income within households. IZA DP No. 11757.

A Appendix

A.1 Data

A.2 Administrative data on the total population

In one additional sub-analysis, we assess whether the overall administrative distribution of female income shares also exhibits a discontinuity at the 50 percent margin as well as a spike at exactly 50 percent for the universe of Swiss couples. To this end, we draw on administrative information for all households in Switzerland for the years 2010 to 2014.[30] The data do not include an identifier for the spouse. However, the fact that we observe a household identifier, the marital status, and the date the marital status last changed allows us to get a close approximation of couples. We define individuals to be a married couple if they live in the same household (same address) and share the same date of change of marital status.[31] This should render a good approximation as it is unlikely that two married individuals live in the same household and share the same wedding date, but are not married to each other. The resulting within-couple age difference is similar in the survey and the administrative data. Men are on average about two years older than their wife. This supports our approach of matching spouses. We restrict the sample to Swiss citizens and individuals holding a permanent residence permit and who are aged between 18 and below 65 in order to match the sample used for the main analysis. We further exclude all self-employed. The income measure we use is the raw sum of gross incomes of an individual across all occupations in one year. This allows us to compute the administrative income share of a married woman within a couple for the whole population.

A.3 Definition of an individual's ancestry

In order to gain a comprehensive picture of individuals' migration background, we exploit information on the respondent's nationality and their parents' country of birth. Information on the individuals' nationality or their second nationality if they hold dual citizenship (Swiss and any foreign country) have been part of the regular SAKE survey since 2003, which means this information is available for all years we use for the analysis. We have a sample of 664 individuals with information on migration background. The information on an individual's background is supplemented with data on norms persisting in their country of ancestry, as reflected in the World Value Surveys.[32] We use average approval rates of full-time or part-time employed women with the statement "When jobs are scarce, men should have more right to a job than women" as a measure of gender norms. Norms

[30] These data combine the population register (Statpop) with income data from the social security register (ZAS by the Swiss Central Compensation Office). It was composed by Steiner and Wanner (2015).

[31] The date of marital status in our data source is available for all individuals who have changed their marital status within the last 40 years. We might thus lose observations of older cohorts. However, this information was not centrally registered before. Furthermore, we restrict the data to households with no more than 20 individuals to reduce the possible false positive rate.

[32] Data link: http://www.worldvaluessurvey.org/WVSDocumentationWV6.jsp.

for the countries Bosnia and Herzegovina, Kosovo, Serbia and Montenegro and Central Serbia are proxied by values for Serbia.[33] Based on this information, we generate a binary variable indicating whether an individual's country of origin shows average approval rates above or equal to the sample median (gender equal), or below the sample median (gender unequal). This results in two types of origin countries:

- **Equal countries:** Australia, Brazil, Canada, Finland, Germany, Italy, Montenegro, Netherlands, Peru, Slovenia, Spain, Sweden, United States
- **Unequal countries:** Albania, Algeria, Argentina, Belarus, Bosnia and Herzegovina, China, Colombia, Croatia, Ecuador, France, Hungary, India, Japan, Kosovo, Kyrgyzstan, Latvia, Lithuania, Macedonia, Pakistan, Poland, Romania, Russia, Serbia, Slovakia, South Korea, Thailand, Turkey, Ukraine, Venezuela, Vietnam

The within sample median is set by Germany (represents 28 percent of the sample) where 12 percent of surveyed women who work either full- or part-time agree with the statement that "When jobs are scarce, men should have more right to a job than women." The largest fraction of individuals with a background in a gender *equal* country is from Italy (31 percent) and the largest fraction of individuals from a gender *unequal* country is from France (11 percent).

[33] Country names defined by the BFS:
www.bfs.admin.ch/bfs/de/home/grundlagen/stgb.assetdetail.6166613.html.

A.4 Tables & Figures

Discontinuity in density estimates

In the following two tables with density discontinuity estimates following the empirical likelihood-based test by Otsu et al. (2013), c refers to the threshold used (the lowest value of female income shares exceeding 50 percent in order to only fit the density between realized values), h refers to the bandwidth, $\hat{f}_l$ reports the fit of the density coming from the left and $\hat{f}_r$ coming from the right respectively, $\hat{\theta}$ is the estimate of the discontinuity, $\hat{lr}$ is the value of the local likelihood ratio statistic under the null[34], and $\hat{f}_l/\hat{f}_r$ measures the relative size of the discontinuity.

Table A.1: Density discontinuity estimates

	c	h	$\hat{f}_l$	$\hat{f}_r$	$\hat{\theta}$	$\hat{lr}$	$\hat{f}_l/\hat{f}_r$	p-value	N
Large sample:									
Survey overall									
(1)	50.0905	7	0.055 5	0.013 8	−0.041 7	332.475 6	4.033 5	0.000 0	13,068
Survey female									
(2)	50.0905	7	0.063 9	0.016 2	−0.047 7	177.635 9	3.945 1	0.000 0	6,037
Survey male									
(3)	50.0952	7	0.048 2	0.011 7	−0.036 5	153.526 5	4.116 9	0.000 0	7,031
Administrative overall									
(4)	50.0007	7	0.013 6	0.010 6	−0.003 0	32.296 0	1.281 8	0.000 0	70,231
Small sample:									
Survey overall									
(5)	50.2203	7	0.048 1	0.015 6	−0.032 5	50.856 1	3.074 0	0.000 0	3,081
Administrative overall									
(6)	50.0225	7	0.021 4	0.023 7	0.002 2	0.437 3	0.905 3	0.508 4	3,081

Notes: Rows (1) to (3) present local likelihood ratio results for the discontinuity in the distribution of females' surveyed income shares (based on SAKE survey years 2002, 2005, 2008, 2012, and 2015). Row (4) presents the discontinuity estimate in the distribution of administrative female income shares in the full population of married couples which is described in Section A.2 of the appendix. Row (5) presents the local likelihood ratio results for the survey information and row (6) for the administrative information in our main sample including administrative and survey information for the very same couples (2012 and 2015). N stands for the number of observations with regard to the observations available for estimating the whole density in the respective sample.

[34]The null hypothesis $H_0 : \theta_0 = \theta$ for some θ can be tested by $lr(\theta)$ using $\chi^2(1)$ critical values. We test against $H_0 : \theta_0 = 0$.

Table A.2: Density discontinuity estimates, half of bandwidth

	c	h	$\hat{f}_l$	$\hat{f}_r$	$\hat{\theta}$	$\hat{lr}$	$\hat{f}_l/\hat{f}_r$	p-value	N
Large sample:									
Survey overall									
(1)	50.0905	3.5	0.116 1	0.012 3	−0.103 8	493.275 2	9.448 3	0.000 0	13,068
Survey female									
(2)	50.0905	3.5	0.123 4	0.014 5	−0.108 9	244.763 8	8.526 5	0.000 0	6,037
Survey male									
(3)	50.0952	3.5	0.111 3	0.010 5	−0.100 8	247.149 5	10.593 8	0.000 0	7,031
Administrative overall									
(4)	50.0007	3.5	0.016 3	0.011 1	−0.005 2	41.225 7	1.473 0	0.000 0	70,231
Small sample:									
Survey overall									
(5)	50.2203	3.5	0.095 3	0.014 8	−0.080 5	78.057 2	6.421 8	0.000 0	3,081
Administrative overall									
(6)	50.0225	3.5	0.020 7	0.023 9	0.003 2	0.485 5	0.865 7	0.485 9	3,081

Notes: Rows (1) to (3) present local likelihood ratio results for the discontinuity in the distribution of females' surveyed income shares (based on SAKE survey years 2002, 2005, 2008, 2012, and 2015). Row (4) presents the discontinuity estimate in the distribution of administrative female income shares in the full population of married couples which is described in Section A.2 of the appendix. Row (5) presents the local likelihood ratio results for the survey information and row (6) for the administrative information in our main sample including administrative and survey information for the very same couples (2012 and 2015). N stands for the number of observations with regard to the observations available for estimating the whole density in the respective sample.

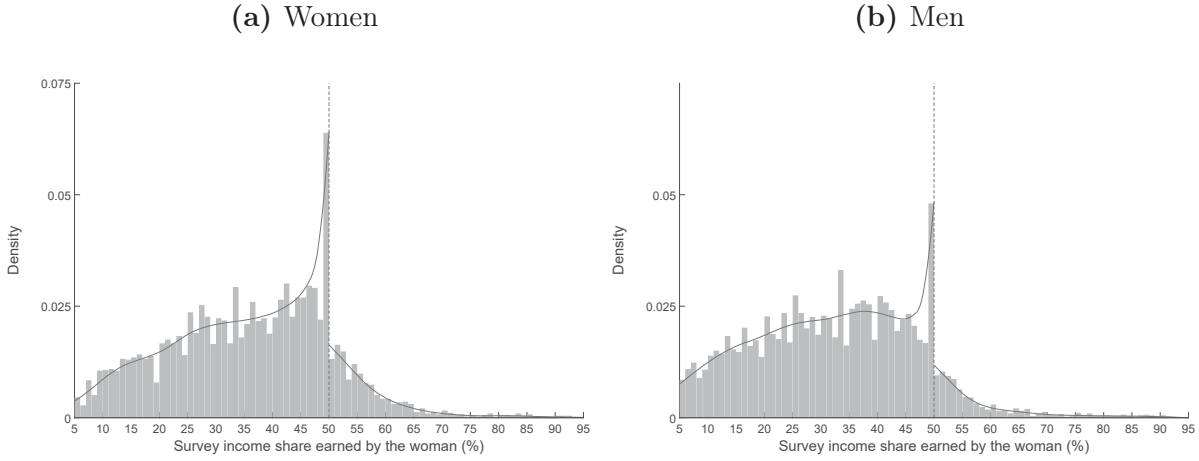

Figure A.1: Overall LLD fit of the income share earned by the woman in a couple separately for female and male respondents. The solid line represents the LLD fit on both sides of the threshold using a bandwidth of 7 percent. The shaded area represents the histogram of the underlying data in 1 percent bins. The corresponding density discontinuity estimates can be found in rows (2) and (3) of Table A.1.

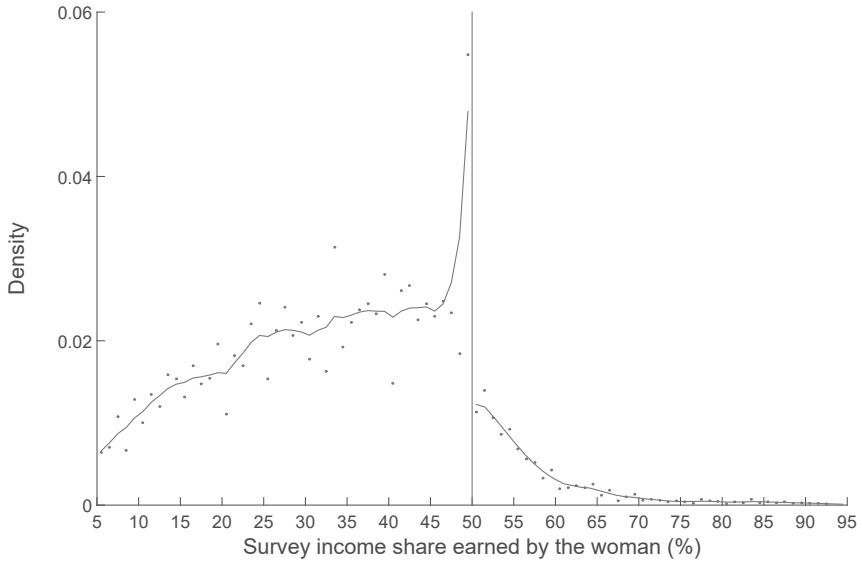

Figure A.2: Overall distribution of surveyed female income shares estimated using smoothed binned counts as proposed in McCrary (2008). The smooth is a local linear smooth using a triangular kernel and a bandwidth of 3.5%. The bin size is 1%. The automated procedure in McCrary (2008) proposes a bin size of 0.262 and a bandwidth of 13.11. It renders a point estimate of log difference in height of 0.556 with a p-value of 0.049. The manipulation test provided in the rddensity package in STATA (Cattaneo et al., 2018) proposes an optimal bandwidth of about 4 percent and also indicates a clear discontinuity with p-values<0.01.

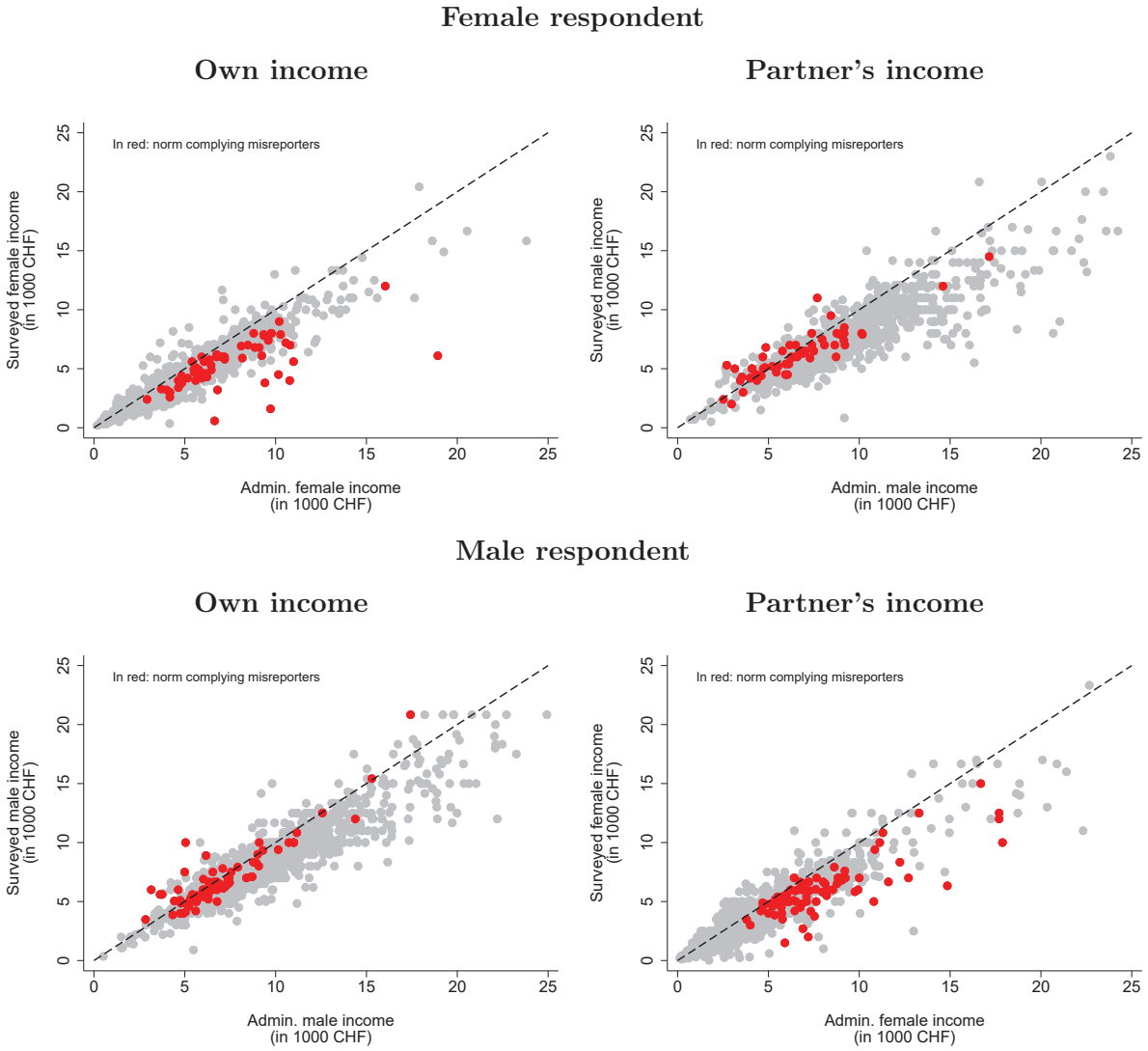

Figure A.3: Correlation between surveyed and administrative incomes for female and male respondents and for their own and their partner's income in the main sample. The dashed line shows the 45° line where surveyed and administrative incomes coincide. The red dots mark individuals who cross the threshold of just above 50 percent from above to conform with the norm that the women should not earn more than her partner (norm complying misreporters). For this correlation, we exclude 7 observations with monthly incomes of more than CHF 25,000 as these would distort this picture.

Misreporting estimates

Table A.3: Misreporting estimates, conditional on admin. distribution

Dependent variable	Diff. female inc. share	Overreporting (*Survey inc. – Admin. inc.*)			
		Female respondent		Male respondent	
		Own inc.	Partner inc.	Own inc.	Partner inc.
	(1)	(2)	(3)	(4)	(5)
P(0,40]	0.960**	1.816	-2.594	-3.821***	1.819
	(0.433)	(1.861)	(1.662)	(1.455)	(2.563)
P(40,42]	0.422	-1.485	-3.040	-2.752	-1.928
	(0.603)	(2.619)	(2.339)	(1.996)	(3.515)
P(42,44]	-0.008	1.903	0.581	-0.765	-1.070
	(0.613)	(2.672)	(2.386)	(2.023)	(3.562)
P(44,46]	Ref.	Ref.	Ref.	Ref.	Ref.
P(46,48]	-0.106	-2.502	1.088	-1.593	0.196
	(0.604)	(2.620)	(2.340)	(2.002)	(3.526)
P(48,50]	-1.597**	-2.496	5.603**	-2.263	-4.514
	(0.627)	(2.611)	(2.332)	(2.173)	(3.827)
P(50,52]	2.480***	2.723	-0.977	-6.616*	5.494
	(0.907)	(3.523)	(3.146)	(3.523)	(6.203)
P(52,54]	3.005***	1.874	-2.759	-7.735**	5.182
	(0.843)	(3.358)	(2.999)	(3.132)	(5.514)
P(54,56]	0.511	-1.145	-0.846	-0.221	1.059
	(0.994)	(4.251)	(3.796)	(3.344)	(5.887)
P(56,58]	0.131	-8.507	2.372	-13.944***	1.130
	(1.362)	(5.267)	(4.704)	(5.295)	(9.324)
P(58,60]	-1.297**	-4.523*	6.415***	-0.309	-2.443
	(0.644)	(2.556)	(2.283)	(2.449)	(4.311)
P(50,52] × norm complying misrep.	-5.169***	-7.455*	8.514**	8.986**	-8.796
	(1.022)	(4.274)	(3.817)	(3.763)	(6.626)
P(52,54] × norm complying misrep.	-8.669***	-16.711***	9.047*	14.682***	-10.216
	(1.279)	(5.714)	(5.103)	(4.271)	(7.521)
P(54,56] × norm complying misrep.	-12.145***	-10.173	35.350***	13.565***	-21.210***
	(1.475)	(7.303)	(6.523)	(4.516)	(7.952)
P(56,58] × norm complying misrep.	-10.027***	-4.443	27.851***	39.509***	-5.461
	(1.970)	(7.639)	(6.822)	(7.577)	(13.342)
P(58,60] × norm complying misrep.	-16.874***	-36.888***	15.183*	23.868***	-22.407**
	(1.848)	(9.644)	(8.614)	(5.492)	(9.670)
Constant	-0.770*	-10.289***	-12.735***	-7.556***	-14.538***
	(0.425)	(1.814)	(1.621)	(1.430)	(2.519)
Inc. mode	Yes	Yes	Yes	Yes	Yes
N	3,081	1,415	1,415	1,666	1,666
R-squared	0.115	0.150	0.144	0.182	0.055

Notes: Misreporting of female income shares and incomes by bin of the admin. female income share. We distinguish between misreporting of norm compliers (admin female inc. share $\leq$ 50 percent), norm violators (admin female inc. share $>$ 50 percent, surveyed female inc. share $>$ 50 percent), and norm complying misreporters (admin female inc. share $>$ 50 percent, surveyed female inc. share $\leq$ 50 percent). The interaction term between the bins above 50 percent and the indicator for norm complying misreporters shows the additional deviation for this group. In column (1), the dependent variable is the deviation in the female income share between survey and admin. data (*surveyed female inc. share - admin female inc. share*). For columns (2) to (5), the dependent variable is misreporting, as defined above. We use the specification described in Eq. 2 and standard errors are reported in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.4: Misreporting estimates, conditional on survey distribution

Dependent variable	Diff. female inc. share	Overreporting (*Survey inc. – Admin. inc.*)			
		Female respondent		Male respondent	
		Own inc.	Partner inc.	Own inc.	Partner inc.
	(1)	(2)	(3)	(4)	(5)
P(0,40]	-2.499***	-1.158	0.989	1.732	-11.782***
	(0.456)	(2.027)	(1.858)	(1.553)	(2.698)
P(40,42]	-0.756	1.177	-1.320	0.941	-7.193*
	(0.635)	(2.851)	(2.613)	(2.137)	(3.713)
P(42,44]	-0.381	2.779	1.845	1.859	-0.407
	(0.625)	(2.788)	(2.556)	(2.112)	(3.669)
P(44,46]	Ref.	Ref.	Ref.	Ref.	Ref.
P(46,48]	0.193	3.037	0.729	-1.061	-4.136
	(0.621)	(2.638)	(2.419)	(2.235)	(3.882)
P(48,50]	1.511**	3.824	-2.845	-0.786	1.510
	(0.641)	(2.678)	(2.455)	(2.375)	(4.127)
P(50,52]	-0.074	2.239	0.360	-1.354	-6.602
	(0.747)	(3.178)	(2.913)	(2.670)	(4.639)
P(52,54]	-1.216	-2.083	0.667	2.433	-6.512
	(0.761)	(3.199)	(2.932)	(2.762)	(4.798)
P(54,56]	-1.652**	-2.274	6.876**	1.163	-1.472
	(0.832)	(3.401)	(3.117)	(3.165)	(5.499)
P(56,58]	-0.926	1.163	2.403	4.555	-8.792
	(1.059)	(4.052)	(3.714)	(4.653)	(8.083)
P(58,60]	1.259*	0.441	-0.997	-3.917	-6.097
	(0.679)	(2.767)	(2.537)	(2.641)	(4.589)
P(0,40] × norm complying misrep.	-19.841***	-38.398***	27.148***	12.524***	-36.730***
	(1.229)	(5.794)	(5.312)	(3.956)	(6.873)
P(40,42] × norm complying misrep.	-17.456***	-47.065***	2.148	66.176***	-13.343
	(3.030)	(16.442)	(15.073)	(8.926)	(15.507)
P(42,44] × norm complying misrep.	-11.140***	-20.157***	21.889***	13.629**	-24.245***
	(1.627)	(7.569)	(6.938)	(5.294)	(9.197)
P(44,46] × norm complying misrep.	-10.193***	-18.059**	15.467**	32.336***	-10.206
	(1.887)	(8.381)	(7.684)	(6.405)	(11.128)
P(46,48] × norm complying misrep.	-7.592***	-6.950	16.815***	16.541***	-14.160**
	(1.241)	(6.038)	(5.535)	(3.958)	(6.877)
P(48,50] × norm complying misrep.	-6.258***	-7.066**	15.026***	8.049***	-13.894***
	(0.742)	(3.405)	(3.121)	(2.567)	(4.460)
Constant	1.419***	-9.816***	-14.162***	-12.113***	-4.171
	(0.448)	(1.979)	(1.814)	(1.529)	(2.656)
Inc. mode	Yes	Yes	Yes	Yes	Yes
N	3,081	1,415	1,415	1,666	1,666
R-squared	0.165	0.160	0.108	0.191	0.090

Notes: Misreporting of female income shares and incomes by bin of the surveyed female income share. We distinguish between misreporting of norm compliers (admin female inc. share $\leq$ 50 percent), norm violators (admin female inc. share $>$ 50 percent, surveyed female inc. share $>$ 50 percent), and norm complying misreporters (admin female inc. share $>$ 50 percent, surveyed female inc. share $\leq$ 50 percent). The interaction term between the bins below 50 percent and the indicator for norm complying misreporters shows the additional deviation for this group. In column (1), the dependent variable is the deviation in the female income share between survey and admin. data (*surveyed female inc. share - admin female inc. share*). For columns (2) to (5), the dependent variable is misreporting, as defined above. We use the specification described in Eq. 2 and standard errors are reported in parentheses. Significance levels: * p $<$ 0.10, ** p $<$ 0.05, *** p $<$ 0.01.

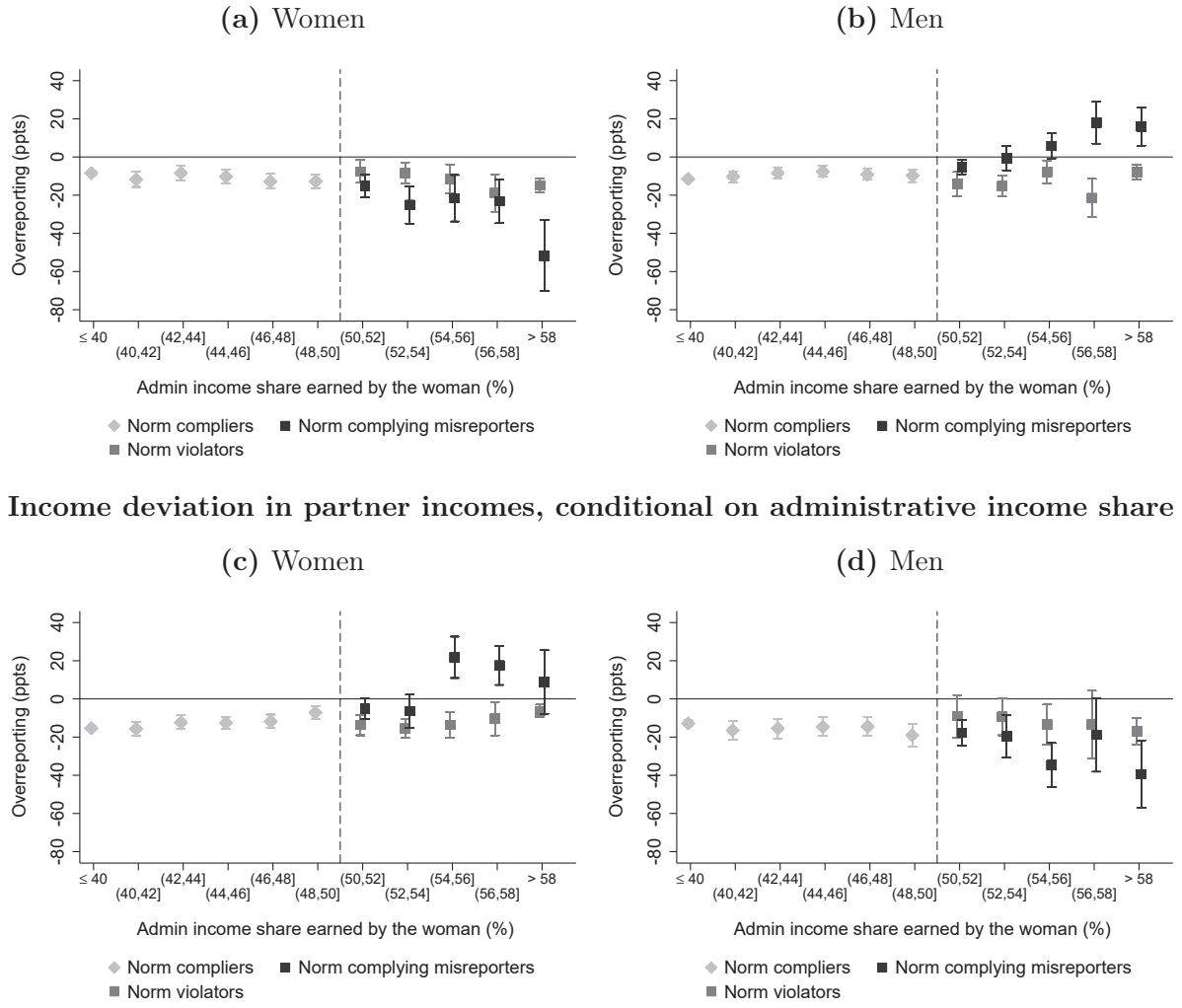

Figure A.4: These graphs show the level estimates of the deviation between surveyed and administrative incomes for norm compliers, norm violators, and norm complying misreporters conditioning on the position in the administrative distribution of female income shares. The upper panel shows the average deviations of own incomes for female (a) and male (b) respondents. The lower panel shows the average deviations of partner incomes for female (c) and male (d) respondents. 95 percent confidence bounds are displayed. We use the specification described in Eq. 2.

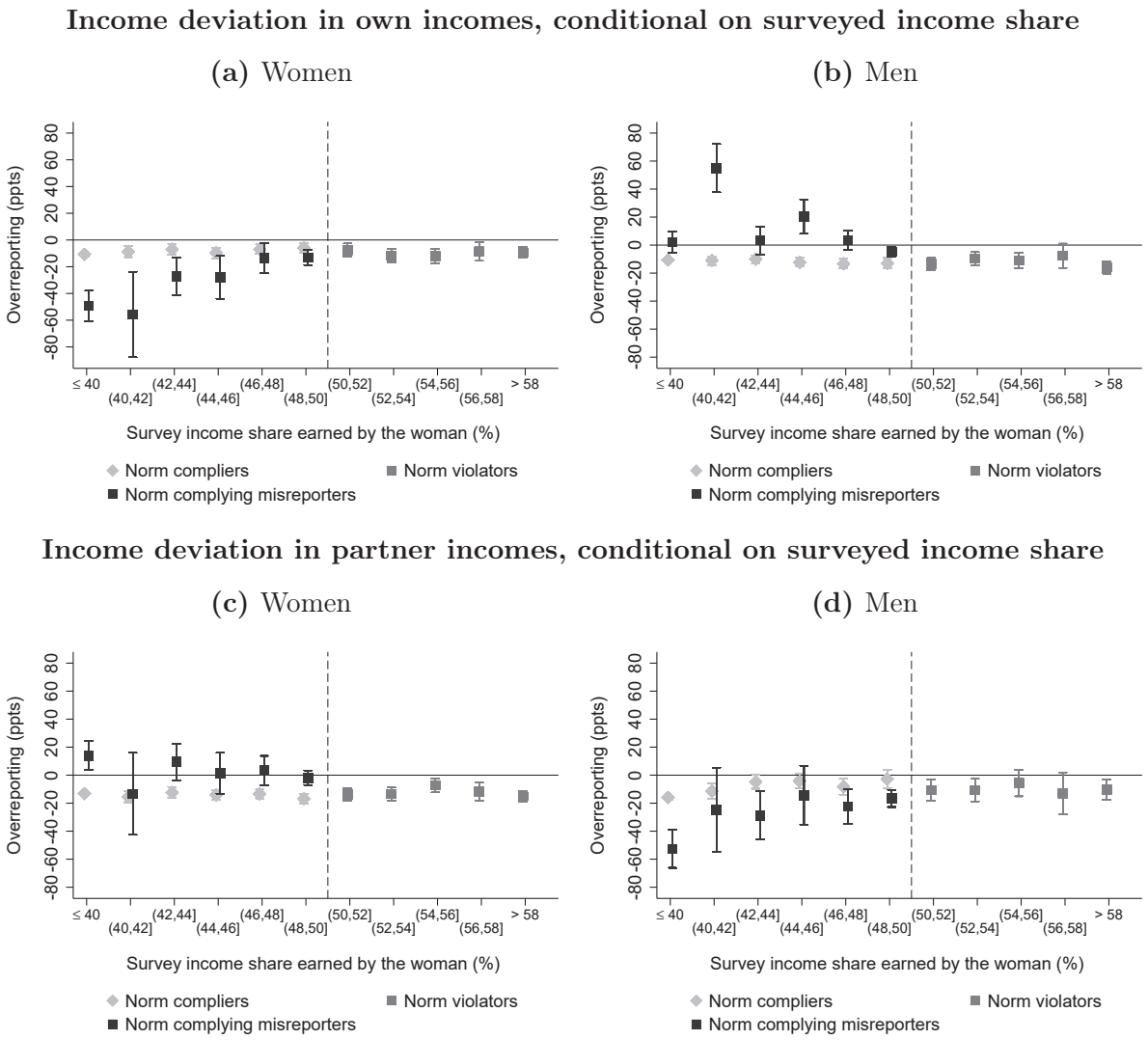

Figure A.5: These graphs show the level estimates of the deviation between surveyed and administrative incomes for norm compliers, norm violators, and norm complying misreporters conditioning on the position in the survey distribution of female income shares. The upper panel shows the average deviations of own incomes for female (a) and male (b) respondents. The lower panel shows the average deviations of partner incomes for female (c) and male (d) respondents. 95 percent confidence bounds are displayed. We use the specification described in Eq. 2.

Table A.5: Socio-demographic characteristics of couples where the woman outearns her partner

	Mean	Mean violators	Mean compliers	Diff.	SE diff.	t-value	p-value	N
Age w / 10	3.97	3.97	3.97	-0.001	0.103	0.01	0.994	408
Age m / 10	4.21	4.18	4.26	0.080	0.109	-0.73	0.466	408
Inc. admin. hh	1.49	1.51	1.45	-0.056	0.069	0.80	0.423	408
Tert. educ. w	0.43	0.50	0.29	-0.205	0.051	4.03	0.000	405
Tert. educ. m	0.32	0.32	0.32	-0.003	0.049	0.06	0.950	405
Couple w/ children	0.37	0.38	0.35	-0.035	0.050	0.69	0.494	408

Notes: Descriptive statistics for couples where the woman outearns her partner in administrative terms, distinguishing between norm violators (admin. female income share $>$ 50 percent, surveyed female income share $>$ 50 percent) and norm complying misreporters (compliers, admin. female income share $>$ 50 percent, surveyed female income share $\leq$ 50 percent).

Table A.6: Norm proxies of couples where the woman outearns her partner

	Mean	Mean violators	Mean compliers	Diff.	SE diff.	t-value	p-value	N
Educ. w $\leq$ educ. m	0.76	0.71	0.87	0.154	0.044	-3.52	0.000	408
Hrs. w $\leq$ hrs. m	0.81	0.78	0.87	0.081	0.041	-2.00	0.046	405
Age diff. / 10	0.23	0.21	0.29	0.081	0.049	-1.64	0.101	408
German speaking	0.70	0.68	0.74	0.063	0.048	-1.33	0.186	408
Unequal country	0.30	0.27	0.36	0.090	0.095	-0.95	0.346	106

Notes: Descriptive statistics of norm proxies for couples where the woman outearns her partner in administrative terms, distinguishing between norm violators (admin. female income share $>$ 50 percent, surveyed female income share $>$ 50 percent) and norm complying misreporters (compliers, admin. female income share $>$ 50 percent, surveyed female income share $\leq$ 50 percent). *Hrs.* stand for weekly work hours. *Age diff.* describes the within-couple age difference defined as ($age_m - age_w$). *German speaking* describes an indicator set to one for German speaking individuals (in relation to French, Italian, or Romansh speaking). *Unequal country* describes an indicator set to one for individuals with origins in a country with more traditional gender norms.

A.5 Survey procedure and questions

The income components are surveyed as follows: Close to the end of the standard SAKE questionnaire, the target person states her or his own income. Right after the standard questionnaire concludes, the special questionnaire starts. Consecutively, the respondents are asked whether they have a partner living in the same household. If yes, the follow up questions are whether the partner is employed and what is her or his income. The question about the partner's income is asked in exactly the same way and with the same options as the question about the target person's income. Thus, it is always the respondent's own income that is surveyed first before the focus changes to the partner's income. There are only a few questions in between the one on one's own and on the partner's income. It can thus be assumed that the respondent still has the response to the question about their own income present when asked about the partner's income.

```
77000 01
77000 02  Könnten Sie mir Ihren MONATSLOHN angeben ?
77000 03  Wenn's Ihnen leichter fällt, können Sie auch den Jahres- oder
77000 04  Stundenlohn angeben.
77000 05  ------------------------------------------------------------------------
77000 06
77000 07
77000 08        o LOHN ..............................<*******>
77000 09        - Arbeitet ohne Entlöhnung .........<0>
77000 10        ------------------------------------------
77000 11        - Weiss nicht .......................<X>
77000 12        - Keine Antwort .....................<Y>
77000 13
77000 14
77000 15                         *******
```

Figure A.6: Survey question on personal income. Survey question 77000 (variable IW04) is translated as: "Could you tell me your monthly salary? If it is easier for you, you may also tell me your yearly or hourly salary." There are four response options: 1. Salary (numeric) 2. Works without compensation 3. Don't know 4. No answer

.

```
77100 01
77100 02  ==> INT: Sind die angegebenen Fr. X ...
77100 03
77100 04  o BRUTTO (VOR Abzug der Sozialbeiträge):    - pro Monat .....<1>
77100 05                                              - pro Jahr ......<2>
77100 06                                              - pro Stunde ....<3>
77100 07  o NETTO (NACH Abzug der Sozialbeiträge):    - pro Monat .....<4>
77100 08                                              - pro Jahr ......<5>
77100 09                                              - pro Stunde ....<6>
77100 10
77100 11  o weiss nicht ..............................................<8>
77100 12  o keine Antwort ..............................................<9>
77100 13  o der angegebene Betrag von Fr. X ist falsch .......<0>
77100 14
77100 15
```

Figure A.7: Survey question on mode of income declaration. Survey question 77004 asks whether the reported income is net or gross; hourly, monthly, or yearly. The exact question is "Is the declared amount CHF X: 1. Gross (per month / per year / per hour), 2. Net (per month / per year / per hour)". Additional answer options are "3. "I don't know" 4. No answer 5. "The above amount of CHF X is wrong."

```
77900 01 Sie haben mir vorher gesagt, dass Ihre Ehe-/Lebenspartnerin
77900 02 erwerbstätig ist. Könnten Sie mir den MONATSLOHN von Ihrer Ehe-/
77900 03 Lebenspartnerin angeben?
77900 04 Wenn's Ihnen leichter fällt, können Sie auch den Jahres- oder
77900 05 Stundenlohn angeben.
77900 06 ------------------------------------------------------------------------
77900 07
77900 08
77900 09
77900 10       o MONATSLOHN ........................ <********>
77900 11       o Arbeitet ohne Entlöhnung .......... <0>
77900 12       ---------------------------------------------
77900 13       - Weiss nicht ....................... <X>
77900 14       - Keine Antwort ..................... <Y>
77900 15                        ********
```

Figure A.8: Survey question that defines the partner's income. Survey question 77900 (variable IW20) is translated as: "You have told me before that your spouse / partner is employed. Could you tell me the monthly salary of your spouse / partner? If it is easier for you, you may also tell me his / her yearly or hourly salary." There are four response options: 1. Salary (numeric) 2. Works without compensation 3. Don't know 4. No answer

.

```
77950  01
77950  02 Sind die angegebenen#bFr. X.-#e...
77950  03
77950  04 o BRUTTO (VOR Abzug der Sozialbeiträge):    - pro Monat .....<1>
77950  05                                              - pro Jahr ......<2>
77950  06                                              - pro Stunde ....<3>
77950  07 o NETTO (NACH Abzug der Sozialbeiträge):    - pro Monat .....<4>
77950  08                                              - pro Jahr ......<5>
77950  09                                              - pro Stunde ....<6>
77950  10
77950  11 o weiss nicht ..........................................<8>
77950  12 o keine Antwort .........................................<9>
77950  13 o der angegebene Betrag von Fr. X ist falsch ...........<0>
77950  14
77950  15                                                          FORMAT !
```

Figure A.9: Survey question that specifies the mode of the income declaration for the partner's income. Survey question 77950 asks whether the reported partner income is net or gross; hourly, monthly, or yearly. The exact question is "Is the declared amount CHF X: 1. Gross (per month / per year / per hour), 2. Net (per month / per year / per hour)". Additional answer options are "3. "I don't know" 4. No answer 5. "The above amount of CHF X is wrong."

Download ZEW Discussion Papers from our ftp server:

http://ftp.zew.de/pub/zew-docs/dp/

or see:

https://www.ssrn.com/link/ZEW-Ctr-Euro-Econ-Research.html
https://ideas.repec.org/s/zbw/zewdip.html

IMPRINT

ZEW – Leibniz-Zentrum für Europäische Wirtschaftsforschung GmbH Mannheim
ZEW – Leibniz Centre for European Economic Research

L 7,1 · 68161 Mannheim · Germany
Phone +49 621 1235-01
info@zew.de · zew.de