Measuring Gender Norms for Men and Women, and Their Implications*

Joshua Dean
Chicago Booth

Christine Exley
University of Michigan

Muriel Niederle
Stanford

Heather Sarsons
UBC & NBER & CEPR

August 2024

PRELIMINARY. PLEASE DO NOT CIRCULATE.

Abstract

A large literature documents the extent to which gender norms contribute to gender inequality in the labor market. Norms are typically measured as attitudes about women's proper roles in the household and labor market, taking attitudes toward men as given. Within a couple, however, the norms that men face may matter when deciding how to split childcare, housework, and earning responsibilities. This paper uses standard survey measures of gender norms and adjusts them to elicit beliefs about men's roles in the household and labor market. We find that respondents typically have less progressive views about men than about women, believing that women should be free to choose whether to work or stay home whereas men should work. We construct measures of gender norms facing both men and women and correlate them with respondents' labor market and marriage decisions, as well as state-level labor force participation and voting patterns. We find that both male and female norms matter for women's (and men's) decisions. In addition, male, but not female, norms are correlated with lower female labor force participation. Taken together, our results suggest that improving the norms that women face may not be enough to achieve household equity.

JEL Codes: J16, B55

*We thank Julian Martinez Correa and Mélyne Nzabonimpa for excellent research assistance.

1 Introduction

Gender equality is a stated goal of many governments and agencies.[1] A large literature uses large-scale surveys, such as the World Values Survey, to document a correlation between attitudes toward women and gender equality (Bertrand, 2011, 2020; Fortin, 2005). However, the focus of surveys meant to capture gender attitudes, as well as the work seeking to bring about gender equality, has largely been on women. There has been much less discussion or measurement of the norms and constraints that men face.

Why measure the norms that men face when women have historically been constrained or discriminated against? Along with growing concerns about men's education and labor market outcomes (Reeves, 2022), a significant portion of the remaining gender gap, particularly in Western countries, is attributed to the impact of children on women's wages and labor market participation (Kleven, 2022; Goldin, 2021). When couples make decisions regarding who should stay home, the norms that both men and women face matter. For example, even if people are supportive of women working, they may find it strange for men to be stay-at-home fathers. The relative strength of the norms that men and women face are therefore important, and only measuring views toward women might miss part of the picture.

This paper attempts to bridge this gap. We use a large-scale survey with a representative sample of U.S. Americans to document how participants respond to common survey questions designed to elicit gender attitudes. Many of the standard survey questions surrounding gender ask about perceptions of women in historically non-traditional roles, such as being a full-time worker. By contrast, none of the standard questions ask about perceptions of men in non-traditional roles, such as providing childcare or being a stay-at-home father. To understand how respondents think about both women's and men's role in society, we take 19 of the most commonly asked questions regarding gender from the World Values Survey and the General Social Survey and ask each question in two ways. We both ask questions using the "standard wording" and then we replace "women" with "men" (and vice versa)

[1]For example, the Bill and Melinda Gates Foundation as an entire division devoted to achieving gender equality. The U.S. Agency for International Development has a "Gender Equality and Women's Empowerment" division and fund.

everywhere. For example, we ask participants how strongly they agree or disagree with the statement "Children suffer when their mother works", as well as "Children suffer when their father works". This provides us with a measure of relative beliefs about men and women in the same situation.

To build our survey, we collected all questions about women and men that were commonly asked in the World Values Survey (WVS) and the General Social Survey (GSS).[2] Focusing on these standard questions allows us to demonstrate the additional value of measuring attitudes towards men with minimal other changes, and respondents' answers provide us with an individual measure of agreement with traditional female and male norms. Using the responses, we measure norms in two ways. We first conduct a data-driven principal component analysis to summarize the data in two indices, one for the standard questions and one for the flipped questions. We then code each response as being in line with what the traditional belief about men's or women's roles would be. For example, if someone agrees with the statement, "It is much better for everyone involved if the man is the achiever outside the home", they are *agreeing* with the traditional male norm that men should be breadwinners. If someone agrees with the statement "It is much better for everyone involved if the woman is the achiever outside the home", they are *disagreeing* with the traditional female norm of women taking care of the home. We average over all responses to create this second measure of male and female norms.

Our survey yields three facts. First, respondents are relatively progressive when it comes to female norms, with a minority of individuals agreeing with traditional female norms. For example, only 30% of individuals agree that preschool children suffer when women work, and roughly 20% believe that women need to have children to be fulfilled.[3] However, the majority of respondents agree with traditional male norms. Nearly 60% of respondents agree with the traditional view that it is a man's job to earn money for the home and family. Only

[2]See Section 2.2.1 for details.

[3]These responses are in line with what has been found in the actual World Values and General Social surveys, with many states becoming more progressive over time. For example, in 1985, 53% of respondents agreed or strongly agreed with the statement "A preschool child is likely to suffer if his or her mother works." In 2022, 26% agreed or strongly agreed. Changes in responses to these prompts over time can be explored at https://gssdataexplorer.norc.org/.

37% of respondents agree or strongly agree with the non-traditional view that a being a stay-at-home husband is just as fulfilling as working for pay. In addition, there is less variation in attitudes toward men than there are in attitudes toward women.

Second, responses to the traditional and flipped questions are correlated, but are far from being perfectly correlated with large differences by question. For many questions, there are dispersed views in response to one version of the question but the response to the flipped version is the same. This typically corresponds to individuals having more dispersed views about women, but having similar views about men. For example, the majority of individuals believe that a man's job is to earn money for the home and family. Among these respondents, some disagree that a woman's job is to earn money for the home and family. Yet a significant fraction agree that a woman's job is also to earn money for the home and family. Asking only one version of this statement misses this nuance.

Our third fact establishes why this nuance matters. We show that attitudes toward men and women are both important in explaining respondents' own labor market and marriage decisions, as well as state and county-level differences in labor force participation and voting patterns. We asked participants of the representative survey a host of questions regarding both their and their spouse's employment, earnings, marriage and family choices, and dating history. Using these responses, we construct measures of within-couple inequality, such as differences in earnings or take-up of parental leave. Across a wide variety of these outcomes we find that both measures of norms independently predict behavior, and that the male norms are particularly helpful in predicting men's behavior. We also find that our norms indices, particularly those pertaining to men, correlate with female labor force participation and Trump voting.

This paper relates to the large literature on the role that gender norms play in determining women's labor market outcomes. A smaller literature has focussed on the role that the gender norms that *men* face affect economic outcomes. Grosjean and Khattar (2019) and Baranov et al. (2023) are among the seminal papers in economics that show that show the importance of measuring masculinity norms. These papers show that men in areas with historically

male-biased convict populations exhibit behavior that is strongly associated with masculinity norms today and that this has consequences for men's and women's outcomes. Boys in these areas are more likely to be bullied (Baranov et al., 2023) and women work less and, conditional on working, work in less prestigious jobs (Grosjean and Khattar, 2019). As per usual, a larger literature outside of economics identified the importance of measuring the norms that both men and women face, but typically using qualitative methods and smaller samples.[4] Our paper expands on the above work, showing that the norms that women face are not a sufficient statistic to understand gender norms writ large, and documenting the correlation between norms and the implications they have for within-household decision-making.

The paper proceeds as follows. Section 2 outlines our survey design, including the sample used, choice of questions, and construction of our gender norms measures. We present descriptive results that discuss the relationship between the different question versions in section 3 and correlate our measures of norms with measures of gender inequality in section 4. Section 5 concludes.

2 Methodology

2.1 Samples

Our primary dataset comes from a survey we ran with a representative sample of 5000 U.S. Americans. Participants were drawn from NORC at the University of Chicago's AmeriSpeak panel, a probability-based, representative panel of the United States.[5] Our survey targeted Americans 18 years and older currently living in the United States. It was fielded between March 19, 2024 and April 24, 2024.

For robustness, we also ran our survey with a sample of 5300 individuals on Prolific in spring 2024. Using a Prolific sample allowed us to incentivize responses: We could ask re-

[4]See for example, Kimmel (2014); **?**); **?**. More recently, Reeves (2022) has argued that greater attention should be paid to the pressures and norms that men face.

[5]For more information on NORC, a nonpartisan research organization based at the University of Chicago, and the AmeriSpeak panel, see AmeriSpeak.norc.org.

spondents on Prolific to guess what they thought respondents in the representative panel would say. We can then both match this to the representative sample to see how accurate beliefs are, and see how different responses are when they are not incentivized in this way. Specifically, we randomly asked half of the respondents to give their best guess as to how a random sample of Americans would respond and paid them based on how close they were to the responses in the representative sample. More details on this procedure are in the next section.

Appendix Table BI shows the demographic characteristics for the representative and Prolific samples, as well as the average difference between the two. The representative sample is older, less white, higher income, and more Republican than the Prolific sample.

2.2 Survey Design

2.2.1 Gender Norm Questions

Our questions to measure gender norms are adapted from the World Values Survey (WVS), an internationally administered survey designed to measure norms, values, and beliefs across countries; and the General Social Survey (GSS), a similar survey administered within the United States. These two surveys are widely used in the social sciences to measure societal views on a variety of topics, and are especially common for measuring gender norms.[6] To narrow the number of questions for our survey, we began with the full set of WVS/GSS questions that are tagged as being about "gender". GSS questions are tagged based on their topic, and so this amounted to including any questions with a tag of either "men", "women", or "gender". For the WVS, we took all questions from the gender module, as well as any other questions containing at least one of the following terms: men, women, man, woman, husband, wife, boy, girl.

From this set of questions, we made the following five restrictions. First, we restricted to questions that were asked at least three times for the GSS or at least twice for the WVS so that we are including questions that have been used to measure norms over time. Second,

[6] See, for example, Bertrand (2011, 2020); Fortin (2005).

we restricted to questions that were asked at least once in the past twenty years so that we are not including questions that are outdated. Third, we excluded any questions for which it did not make sense if the sexes were reversed.[7] Fourth, we excluded questions that are country-dependent and, because our survey is run in the United States, excluded questions that would not be relevant for the U.S.[8] Fifth, because we wanted to ask all questions in the format of providing respondents with a statement and then asking them to agree or disagree with it, we included only WVS/GSS questions that were also measured in this format.[9] Finally, we merged the remaining GSS and WVS questions and dropped one instance of any questions that were common across the two surveys. In one case, we additionally altered wording to make it more natural. Specifically, we changed the statement "Being a stay-at-home housewife is just as fulfilling as working for pay" to "Being a stay-at-home wife is just as fulfilling as working for pay" so that we could replace the word "wife" with "husband" rather than "househusband".

Our final list of questions is shown in Appendix Table DI. For two of the questions, question 1 ("It is much better for everyone involved if the main is the achiever outside the home and the woman takes care of the home and family") and question 3 ("A man's job is to earn money; a woman's job is to look after the home and family"), we took the first part of the statement and included it as a separate question (questions 18 and 19 in the same table). We include these additional statements as it is easier to see whether and how respondents change their answer depending on the subject of the statement (male or female) if the statement only references one sex. It was not possible to split the other questions that reference both sexes (questions 2, 9, 13, 14, 16, and 17) as they involve direct comparisons of men and women).

We chose to ask the questions in a way that elicited what the respondent thought *others* believe about the statement. This a departure from the standard phrasing in the WVS/GSS. For example, in the WVS, respondents are prompted with: "Do you agree, disagree or neither

[7]For example, the statement "Both the husband and the wife should contribute to household income" cannot be flipped. It would be asking the same thing.

[8]This excludes questions such as "Is it justifiable for a man to beat his wife" where there is near universal agreement in the U.S.

[9]This mainly excluded questions that asked the respondent about traits that are important for a successful marriage.

agree nor disagree with the following statements". We instead asked "Please indicate the extent to which you think society agrees, in general, with the following statement". We opted to ask about society as we believe that one's beliefs about others' beliefs are what is relevant when it comes to social norm constraints.[10]

For each of these questions, we created a mirror question in which we flip the sex referenced in the statement. For example, "Family life suffers when the woman has a full-time job" becomes "Family life suffers when the man has a full-time job". All respondents see every statement, meaning that they respond to a total of 38 statements.

To ensure that participants gave responses that accurately reflect their beliefs about how society thinks, we also randomized participants from the Prolific sample into an incentive treatment.[11] Half of the participants were told that we would randomly select one of their answers to a question and that they would receive a bonus payment of $1 if their answer is the most commonly chosen answer by a group of other survey-takers from the nationally representative sample. The other half of participants did not receive such a participant. Responses were similar regardless of whether respondents received the incentive.

We chose to build on common WVS and GSS questions asking about women in order to demonstrate the additional value of measuring norms facing men with minimal other alterations. We could have alternatively used measures of masculinity developed by psychologists and other social scientists and compared the predictive power of these scales to the WVS/GSS questions about women's norms. However, this would have introduced additional variation such as introducing the topics of violence and public displays of emotion that are not part of the standard WVS/GSS scales. These additional topics are of interest for future research on the role of gender norms, but for this exercise we chose to minimize the number of alterations to demonstrate that asking about men's roles, even in the context of the traditional questions, is important to understanding the full picture of gender norms.

[10]We do not specify what one should think about when they are considering "society". As such, individuals could be considering different parts of society.

[11]We did not run the incentive treatment with the representative sample as it was administratively more difficult to contact them again with the incentive payments.

2.2.2 Demographic and Labor Market Questions

In addition to the questions on gender norms, we asked a set of demographic and labor market questions. Some demographic questions were provided by NORC. Others were added so that we could see how answers vary by respondents' characteristics.

To test whether individuals themselves respond to these norms, we asked a series of labor market questions to test whether the norms that men and woman face are reflective of respondents' own labor market and relationship choices. These questions were only asked of the representative sample because of the selection into Prolific, specifically of men. Many male respondents on Prolific are unemployed or earn less than their spouse, which is highly unrepresentative of the United States.

We asked questions about whether the individuals' partner or spouse works, allowing us to look at differences in men's and women's labor supply; whether respondents and their partners took parental leave; whether the respondent had ever dated someone who earns more or less than they do; and their partner/spouse's income, allowing us to measure income differences within the household. Questions on prior dating experiences were asked of every everyone. Questions regarding parental leave and partner labor supply and income were only relevant to respondents currently in a long term partnership.

Average responses to these questions are reported in Table 1. There is an income gap between men and women for those who have partners, with men earning more than women.[12] 80% of men have ever dated someone who earns less than they do, while 60% state they have ever dated someone who earns more than they do. This mirrors women's responses, 56% of whom have dated someone who earns less while 80% have dated someone who earns more. Note that we ask whether the respondent has *ever* dated someone who earns more or less, meaning that some of these could be cases where one person was in school. 3% of men have a partner who stays at home full-time while 1.5% of women have a partner who stays home. Of respondents who have children, 85% of women took the full amount of parental leave

[12] Income is measured in bins so that a one unit difference between men's and women's incomes represents a $5000 gap.

they were offered while only 57% took full leave. Similarly, 66% of male respondents report having a partner who took the full amount of leave offered while only 24% of women report having such a partner, lower than the amount that men state for themselves. These gaps are similar when looking at whether a respondent or their partner took any amount of parental leave.

2.3 Measuring Norms

2.3.1 PCA Indices

We would like to reduce our questions into indices we can use to summarize and analyze responses. To do so in a "hands off' way, we first conduct a principal component analysis and estimate the first component scores for each respondent. This provides us with two indices: one corresponding to the standard wording and one corresponding to the flipped wording.

Appendix Table BII shows which questions each of the indices loads onto. For the standard questions, the questions are fairly uniformly positively correlated with the first component, except for questions 10 and 11 ("It is acceptable if a woman wants to have a child as a single parent but she doesn't want to have a stable relationship with a man" and "A working mother can establish just as warm and secure a relationship with her children as a mother who does not work"), which are negatively correlated, and questions 6 and 12 ("Family life often suffers because men concentrate too much on their work" and "Being a stay-at-home wife is just as fulfilling as working for pay"), which are relatively unimportant. A higher value of the standard index therefore corresponds to more traditional gender views in response to the standard questions. For the flipped questions, questions 6, 10, and 12 are relatively unimportant while question 9 ("When jobs are scarce, women should have more right to a job than men") is both negatively correlated and relatively unimportant. A a higher value of the flipped index can therefore be interpreted as a respondent having *less* traditional gender views in response to the flipped questions.[13] To make the two indices consistent in

[13]If someone states that they agree with the statements "A woman's job is to earn money; a man's job is to

their interpretation, we take the negative of the flipped index so that a higher value also corresponds to more traditional gender norms in response to the flipped questions.

Because many of the traditional questions are about women and many of the flipped are about men, we can also interpret the standard index as representing respondents' views about gender norms that women face and the flipped index as representing gender norms that men face. However, we create a second measure of gender norms that removes the questions referring to both men and women, discussed next.

2.3.2 Average Agreement

The benefit of the PCA indices is that they do not involve any human judgment and simply summarize the data. However, it is difficult to know whether some of the questions most heavily pertain to norms that women face, norms that men face, or both. Our second measure of gender norms therefore attempts to measure traditional gender norms pertaining to men and traditional gender norms pertaining to women.

To create this measure, we first redefine all variables to indicate agreement with the "traditional" norm. We define the traditional norm as how someone with traditional gender views would respond to the question. As an example, consider the statement "Family life suffers when the woman has a full-time job". Here, the traditional view is that family life would indeed suffer if women work full-time. Therefore, someone who strongly agrees with this statement is coded as strongly agreeing with the gender-congruent norm for women. When the subject is instead male ("Family life suffers when the man has a full-time job"), the traditional view is that family life would *not* suffer if the man works full-time. Someone who strongly agrees with the statement when the subject is male is therefore strongly *disagreeing* with the gender-congruent norm.[14]

For each respondent, we then average across their answers to the questions about women

look after the home and family" or "Family life suffers when the man has a full-time job", two of the more important questions, that respondent is expressing views that are opposite of traditional gender norms.

[14]For this exercise, we exclude all statements that contain references to both men and women. These are the non-bolded questions in Appendix Table DI. These statements are excluded as agreement or disagreement is jointly determined by views of both male and female norms making them unsuitable for a summary index.

and men respectively. This provides us with our two gender norms measure, each of which takes a value from 1 to 5, with 1 indicating that a respondent strongly disagrees with the traditional, gender-congruent norm for men/women, and 5 indicating that the respondent strongly agrees with the traditional, gender-congruent norm for men/women.

3 Descriptive Results

We begin by analyzing the raw responses before they have been recoded as agreement with gender norms. We then document three facts. First, people have more traditional attitudes about men than about women. This is true regardless of a respondents' sex, political affiliation, and age. There is also less variation in attitudes toward men than in attitudes toward women both overall and across space. Second, attitudes toward men and women are correlated, but imperfectly. Finally, attitudes toward both men and women matter as they predict individuals' own choices. Therefore, by collecting data only on women, we miss the full set of norms that may affect couples' decisions.

3.1 Gender Norms

We begin by analyzing the raw responses to each statement, which are summarized in Figure 1. This figure shows how participants respond when the statement is presented in the "standard" format and when the sexes are flipped. The statements are ordered by how different the responses are when the statement was flipped. For example, the statement "A [man/woman]'s job is to earn money for the home and family", shown in the first panel in the top left corner, had the most disagreement in responses when the statement was phrased as being about men versus women. Roughly 55% of respondents agree that a man's job is to earn money whereas only 10% agree that it is a woman's job to earn money. At the other extreme, respondents are roughly in agreement about the statement "A university education is more important for a [boy/girl] than for a [girl/boy]" regardless of how the question is asked: 13% agree that it is more important for boys and 8.6% agree that it is more important

for girls.

From Figure 1, it is unclear whether those who agree with a statement posed with the standard statement are also those who disagree with the statement posed with the flipped wording. For example, we do not know whether the respondents who agree that a man's job is to earn money for the home and family disagree that it is a woman's job to earn money for the home and family, or whether some also agree that it is a woman's job to earn money for the home and family. Those who disagree that a preschool child will suffer if his or her father works may agree that the child will suffer if the mother works, indicating that they believe in traditional gender norms, or they may also disagree that the child will suffer when the mother works, indicating that they think parents can both work and take care of kids.

As a first step to understanding the relationship between the statements, figure 2 shows the correlation between the standard and "flipped" wording for each statement. Agreement with the standard wording is shown on the x-axis while agreement with the flipped wording is shown in the y-axis. "1" indicates strong disagreement with the statement while "5" indicates strong agreement.

We see that individuals' beliefs are not fully captured by their responses to the standard wording. Going back to the statement ,"A [man/woman]'s job is to earn money for the home and family", we again saw that just over 50% of respondents agree with the statement "A man's job is to earn money for the home and family", and that just under 50% disagree with the statement "A woman's job is to earn money for the home and family". These answers may seem consistent, such that asking only one version is enough to capture individuals' beliefs. However, looking at the correlation between the standard and flipped versions in the top left figure of Figure 2 shows that this is not the case. In fact, there is a weak positive correlation between the standard and flipped versions of the statement. Some of individuals who agree that a man's job is to earn money for the home and family are neutral or agree that it is a woman's job to earn money for the home and family, while others disagree with the latter. Those who state that it is both a man and a woman's job to earn money for the home are very different than those who believe it is only the man's job. Asking only one version of

this question misses this nuance.

Other statements exhibit a stronger relationship between the standard and flipped questions, but the slope is never one. For example, individuals who strongly agree that it is acceptable for a woman to have a child as a single parent on average agree that it is also acceptable for men. There is still a difference, though, in the average response. Overall, the patterns in Figure 2 suggests that the individuals who gave the same response to the standard version of a question may give very different responses to the flipped version of the question. For example, take individuals who strongly disagree with the statement "On the whole, men make better business executives than women do". Their average response to the flipped version (women make better business executives than men) is neutral. This could mean that some believe that men in fact do make better business executives, whereas others disagree with both versions of the question, believing that neither sex is better. Figure 3 shows the correlation between the standard and the flipped versions of the questions, averaging across all questions but including individual fixed effects. The correlation between the two is 0.32.

We also see that respondents are relatively progressive about women in a way that is not always matched in their views about men. Figure 1 shows that the minority of respondents agree with statements that women should take on a traditional role. Only 30% of respondents agree that it is better that the man be the achiever while the woman takes care of the home and family. Less that 25% agree that children suffer when women work for pay. Nearly 70% believe that a working mother can establish just as warm and secure a relationship with their children as a mother who does not work. These responses can be compared to results from early waves of the General Social Survey to see how society has progressed. For example, in 1990, over 40% of respondents agree that the man should be the achiever outside of the home while women take care of the home and family. Just over 50% of respondents agreed that preschool children suffer if their mother works.[15]

If we again only look at responses to the standard way of asking the questions, we might conclude that society is becoming more progressive in its view toward women. However, a

[15]See https://gssdataexplorer.norc.org/variables/706/vshow and https://gssdataexplorer.norc.org/variables/705/vsho

more nuanced story emerges when the statements are flipped. Even though a minority agree that men should be the achiever and women should take care of the home, virtually no one believes that women should be the achievers and that men should take care of the home. Similarly, although less that 25% of respondents think that children will suffer if women work, less than 5% believe children will suffer if men work. Looking at Figure 2, we see that for many questions, respondents provide similar answers about men but different answers about women. This is especially true of statements regarding the home and family.

Documenting and understanding these relative views are important as a couple does not make decisions based on the social norms facing only one person in the couple. Even if the majority of people think that children will be fine if women work, nearly everyone thinks that children will be fine if men don't work. If social norms affect decisions that couples make, which research suggests is true (Bertrand et al., 2021, 2015; Folke and Rickne, 2020; Jayachandran et al., 2023), it may be the case that the woman spends more time with kids given the relative strength of the two norms.

3.2 Differences in Views Toward Men and Women

We now use our PCA indices and measures of gender norms described in Section 2.3 to further understand differences in how respondents view men's and women's roles. The measures of agreement with traditional gender norms are measured on a scale from 1 to 5, with 1 indicating strong disagreement, 2 indicating disagreement, 3 indicating neither agreement nor disagreement, 4 indicating agreement, and 5 indicating strong agreement with the "traditional norm" for each sex. For example, if someone disagrees with the statement "Family life suffers when the man has a full-time job", they are agreeing with the traditional norm for men and our measure would take the value 4 (agree) for that question.

Panel (a) of Figure 4 shows the distribution of the standard and flipped indices. The flipped index is more tightly distributed around zero, suggesting that there is less variation in views toward men's role in society compared to women's. Indeed, the variance of the standard index is 8.08 while the flipped index is 5.23. Panel (b) shows the distributions of the

traditional male and female norms. Individuals are more likely to agree with the traditional norm for men than for women: The mean of the male norm is 3.53 (between neutral and agree) and the mean of the female norm is 2.88 (between neutral and disagree). The means are statistically significantly different from one another. There is again less variation in attitudes toward men than about women: The variance of the male norms measure is 0.169, while the variance of the female norms measures is 0.33. This suggests that there is more disagreement about what is acceptable for women to do while individuals generally agree that men should act like "traditional men".

Figure 5 shows the level of agreement with traditional male and female norms aggregated to the state level. At first glance, there does not appear to be a strong relationship between male and female norms. States with strong traditional male norms, like Montana and North Dakota, have relatively weak norms about women.

3.3 Respondent Types

Another way of thinking about the data is to put respondents into bins based on whether they think that men and women should maintain traditional gender roles, men and women should flip gender roles, both men and women should act like traditional men, or both men and women should act like traditional women. This helps understand whether capturing only one norm (the norm that women face) is enough to understand household dynamics. For example, if someone disagrees with the statement "Family life suffers when the woman has a full-time job", it is not clear whether they would also disagree with the same statement about a man, in which case they believe family life is fine regardless of which parent works, or whether they would agree with the same statement about a man, in which case they think that women should work more and men should work less. These norms would have very different implications for the decisions couples make.

Table 2 shows the fraction of respondents that fall into each of these four categories by question and overall. The first column shows the percent of respondents who believe that

men and women should act according to traditional norms.[16]. The second column shows the percent of respondents who disagree with traditional norms for both men and women, thinking that men and women should take on each other's traditional roles. The third column shows the fraction of respondents who agree (or strongly agree) with traditional female norms but disagree (or disagree) with traditional male norms, and the fourth column shows the percent of respondents who disagree with traditional female norms but agree with traditional male norms. In this sense, the respondents in column 3 believe that women should continue to act traditionally (for example, they believe that children will suffer if their mother works), but that men should act in the same way (that children will also suffer if their father works). The respondents in column 4 believe the opposite: children will not suffer if either parent works.[17]

Although responses vary by question, the majority of respondents fall into the fourth category, believing that both men and women can act like traditional men. However, respondents are split across categories with 13.5% believe that men and women should act traditionally. These results suggest that measuring only female gender norms provides an incomplete picture. For example, one might think that respondents who agree with traditional female gender norms also agree with traditional male gender norms. Table 2 suggests that this is not the case. Although 13.5% of respondents agree that men should also abide by traditional norms, 15% believe that men should act more like traditional women. For example, 20% of respondents do not think that it is better for everyone if women are the achievers outside of the home, but they also do not think that it is better if men are the achievers outside of the home. Of the 40% of respondents who believe that it is not good if women are the achievers outside of the home, only 20% believe that men should be the achiever outside of the home. This again has implications for household decision-making, which we explore in Section 4.

[16]These are individuals who agree/strongly agree with the traditional norm for both men and women

[17]The rows do not add up to 100% because respondents who answer that they "neither agree nor disagree" with a statement do not fall into one of the given categories.

3.4 Robustness: Prolific Sample

We also ran our survey with a sample of Prolific respondents. As mentioned, this allowed us to incentivize responses in the following way. Half of the respondents were randomized into a treatment in which they were asked to give the response that they think would be the most common response to the prompt in the representative survey. Participants were received a bonus payment of $1if their answer to a randomly selected question matched the most commonly chosen answer by the representative sample.

Appendix Figure AII shows the results for the Prolific sample when we split by whether the respondent was in the incentive treatment. The results are similar across treatment. The mean of the male and female norms in the incentivized treatment are 3.48 and 3.26 respectively. In the unincentivized treatment, they are 3.50 and 3.21.

4 Do Gender Attitudes Matter?

We have argued that measuring the social norms that both men and women face is important for intra-household decision-making. If society believes that children do not suffer when women work, we do not know whether they believe that it is equally fine when men work, or whether children suffer more (or less) when women work relative to when men work. We now test whether respondents' beliefs about societal views correspond to their marriage and labor market decisions.

4.1 Self-Reported Outcomes

In our survey, we asked respondents questions about their own labor market participation, income, whether they have dated someone who earns more or less than they have, and whether they took parental leave. To test whether individuals' beliefs are related to their own relationship and work dynamics, we correlate these self-reported outcomes with our standard and flipped indices. We focus on the PCA indices as they include all questions and have less researcher input.

We estimate

$$Y_i = \beta_1 Standard_i + \beta_2 Flipped_i + \gamma' X_i + \epsilon_i \quad (1)$$

separately for male and female respondents. $Standard_i$ and $Flipped_i$ are the standard and flipped indices from the PCA analysis, measured at the individual level, and X_i is a vector of individual controls.[18] To account for differences in adherence to norms by demographic characteristics, political views, and geography, we include education, race, and state fixed effects; control linearly for age, and include indicators for identifying as Democrat, Republican, or adhering to any religion.

Most of our outcome variables are only measured for couples or individuals with children. For our main analysis, we focus on married individuals as they have likely made a choice of whom to marry based on their views. However, for the questions concerning dating history (whether an individual has ever dated someone whose occupations typically earns more or less than theirs), we include all respondents in the sample. To maximize power, we also create a summary index of respondents' answers. To do so, we standardize each outcome such that a higher value corresponds to a *more traditional* outcome.[19] We then create a summary index, which is the average over all of the standardized outcome variable. A higher value of this index therefore indicates that an individual abides by more traditional gender norms.

The results are presented in Figure 6. From the summary index reported at the bottom of each figure, we immediately see that both the standard and flipped indices are correlated with a respondent abiding by more traditional norms. This is true for both female (panel A) and male (panel B) respondents. Importantly, it is not just the standard index, which largely represents views about women, that matters. Women who believe that society has more traditional views about men, corresponding to a higher value of the flipped index, are less likely to be employed full-time, have a larger income gap with their spouse, are more

[18] Appendix Figure AI presents the results excluding controls.

[19] For example, one outcome is an indicator for whether an individual is employed. We standardized this variable so that a higher value corresponds to a *women* being *less* likely to be full-time employed and a *man* being *more* like to be full-time employed. Similarly, a higher value of the standard outcome "Stay-at-home spouse/parent" corresponds to a woman being more likely to stay at home and a man being less likely to stay at home.

likely to be a stay-at-home spouse, and are more likely to have dated someone who earns more than they do. The standard index is also correlated with many of these outcomes.

For male respondents, we again see that both the standard and flipped indices matter. For some outcomes, though the flipped index matters significant more. Men who believe society has more traditional views about men are less likely to have taken full or any amount of parental leave, and are more likely to have dated someone who earns less than they do.

4.2 Labor Force Participation

In addition to self-reported outcomes, we collect data on two other outcomes that have shown to correlate with female gender norms: labor force participation and political views.

Male gender norms may affect female labor force participation if couples are constrained by the norms that men face (in addition to the norms that women face). For example, if people believe that children will suffer less when men work compared to women, female labor force participation may be lower in counties with more traditional male gender norms regardless of the female norm. We collect male and female labor force participation from the 2024 March Current Population Survey. Because we only know survey respondents' state of residence, we have limited geographic variation. We therefore collect labor force participation at the state $\times$ race $\times$ age group level. These data are merged to the survey data, which is collapsed to the same state $\times$ race $\times$ age category level.

We estimate

$$LFP_{g,r,a,s} = \beta_1 Standard_{r,a,s} + \beta_2 Flipped_{r,a,s} + \epsilon_s$$

where the outcome variable is labor force participation for gender g in state s is measured at the age-race level, and norms are aggregated to the state-race-age group level. We weight by the total population in each race-age-state cell.

The results are shown in Table 3. Panel A uses the responses of all individuals to construct the PCA indices while Panels B and C use women's and men's responses respectively. In Panel A, we see that the standard index is uncorrelated with overall, female, and male labor

force participation. However, the flipped index is significantly correlated with female labor force participation: A one standard deviation increase in the index is associated with a 2.2 percentage point decline in FLFP relative to a mean of 63%. The flipped index is also weakly negatively correlated with MLFP, but the correlation is not statistically significantly different from zero.

When the indices are created based on men's and women's responses, we see that the results in Panel A are largely driven by men's beliefs about society. When men believe that society has more traditional views about men, women are less likely to work.

4.3 Trump Voting

Research suggests that extreme forms of masculinity are associated with Trump voting (Vescio and Schermerhorn, 2021). To test whether there is a relationship between male gender norms and voting behavior, we collect voting data for the 2020 election from the MIT Election Lab. The vote share going to Trump is constructed at the county level. We then estimate

$$\text{TrumpVoteShare}_{c(s)} = \beta_1 FemNorm_s + \beta_2 MaleNorm_s + X_p^{'}\gamma + \epsilon_{c(s)}$$

where Trump vote share is measured at the county (c) level and the male and female norms are aggregated to the state level (s). We include county-level demographic controls (X_c) including the share of the county that has a college degree, median age, and median income, and weight by the county population. Standard errors are clustered at the state level.

The results, shown in Table 4, show that overall, the standard index is more strongly associated with Trump vote share. A one standard deviation increase in the standard index is associated with a 6.6 percentage point increase in Trump's vote share (relative to a mean of 65%). In contrast, a one standard deviation increase in the flipped index is associated with a 3.5 percentage point increase percentage point increase in Trump's vote share. However, when we again create the indices based on women's responses (column 2) and men's responses (column 3), we see that both the standard and flipped indices predict Trump's vote

share for men.

5 Conclusion

Gender norms have been cited as a major factor contributing to gender inequality (Fortin, 2005; Bertrand, 2020). Norms bound behaviour of both men and women. Indeed, when it comes to household decisions, the norms that both men and women face matter. Even if individuals think that it is ok for a woman to pursue a high-paying, demanding career, it is unlikely that a woman who wants a family could pursue such a career if it is not ok for men to step back from their careers. Yet surveys largely ask questions regarding what is appropriate for women. This paper takes a first step toward understanding what is missed when we do not ask similar questions of men. We show that individuals who have relatively progressive views about women do not necessarily have progressive views about men. Further, we find that the norms that both women and men face explain within-couple inequality as well as labor force participation and voting patterns across geographies. Although we used standard survey questions to measure norms, future work might tailor questions to more precisely identify and measure the gender norms that men face to further assess their importance.

References

Baranov, V., R. De Haas, and P. Grosjean (2023). Men. male-biased sex ratios and masculinity norms: evidence from australia's colonial past. Journal of Economic Growth 28(3), 339–396.

Bertrand, M. (2011). New perspectives on gender. Handbook of Labor Economics 4, 1543–1590.

Bertrand, M. (2020). Gender in the twenty-first century. AEA Papers and Proceedings 110, 1–24.

Bertrand, M., P. Cortes, C. Olivetti, and J. Pan (2021). Social norms, labor market opportunities, and the marriage market penalty for skilled women. Review of Economic Studies 88(4), 1936–1978.

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

Folke, O. and J. Rickne (2020). All the single ladies: Job promotions and the durability of marriage. American Economic Journal: Applied Economics 12(1), 260–287.

Fortin, N. (2005). Gender role attitudes and women's labour market outcomes across oecd countries. Oxford Review of Economic Policy 21, 416–438.

Goldin, C. (2021). Career and family: Women's century-long journey toward equity. Princeton University Press.

Grosjean, P. and R. Khattar (2019). It's raining men! hallelujah? the long-run consequences of male-biased sex ratios. The Review of Economic Studies 86(2), 723–754.

Jayachandran, S., L. Nassal, M. Notowidigdo, M. Paul, H. Sarsons, and E. Sundberg (2023). Moving to opportunity, together. Technical report, Working paper.

Kimmel, M. (2014). After fifteen years: the impact of the sociology of masculinity on the masculinity of sociology 1. In Men, Masculinities and Social Theory (RLE Social Theory), pp. 93–109. Routledge.

Kleven, H. (2022). The geography of child penalties and gender norms: Evidence from the united states. Technical report, National Bureau of Economic Research.

Reeves, R. V. (2022). Of boys and men: Why the modern male is struggling, why it matters, and what to do about it. Brookings Institution Press.

Vescio, T. K. and N. E. Schermerhorn (2021). Hegemonic masculinity predicts 2016 and 2020 voting and candidate evaluations. Proceedings of the National Academy of Sciences 118(2), e2020589118.

Tables

Table 1: Self-Reported Labor Market and Family Outcomes

	Male Respondent	Female Respondent	Difference
Diff. income (M-F)	2.17	1.69	0.48**
			(0.20)
Dated earns less	0.80	0.56	0.23***
			(0.014)
Dated earns more	0.58	0.83	-0.24***
			(0.013)
Partner stays home	0.073	0.016	0.057***
			(0.0057)
Full parental leave	0.57	0.85	-0.28***
			(0.024)
Partner, full parental leave	0.66	0.24	0.42***
			(0.017)
Any parental leave	0.80	0.96	-0.16***
			(0.016)
Partner, any parental leave	0.74	0.32	0.42***
			(0.017)
Observations	5070		

Note: This table reports the means for the listed variables for male and female respondents, as well as the difference between the two. The difference between male and female income (*Diff. income (M-W)*) is the difference between the income bucket of the man and the woman, where one unit corresponds to a $5000 difference. For women, this variable is measured as the difference between her husband's minus her income. For men, it is measured as the difference between his income and his wife's. It is only defined for heterosexual couples. All other variables are indicator variables that take the value one of the statement is true.

Table 2: Responses Categorized by Belief

	Be Traditional	Men Women Flip	Be like Traditional Woman	Be like Traditional Man
Panel A: Representative				
Question 4	17.70	0.89	4.06	40.56
Question 5	10.63	2.31	10.66	18.32
Question 6	6.62	10.61	17.88	17.36
Question 7	17.21	0.65	3.20	45.36
Question 8	6.67	1.30	5.87	50.71
Question 11	3.17	8.16	55.00	5.12
Question 12	13.44	2.29	25.18	7.98
Question 15	13.11	0.86	2.30	47.85
Question 18	19.75	1.14	20.99	2.51
Question 19	26.72	1.07	10.62	7.43
Mean	13.50	2.93	15.58	24.32
Panel B: Prolific				
Question 4	34.39	0.83	4.01	38.90
Question 5	21.74	2.74	17.76	19.10
Question 6	14.39	14.02	24.47	19.07
Question 7	33.42	0.73	3.51	41.50
Question 8	16.59	2.00	11.92	50.75
Question 11	3.37	12.89	59.83	6.74
Question 12	27.15	2.47	33.92	9.22
Question 15	27.85	1.04	2.90	47.95
Question 18	40.97	1.34	18.16	4.04
Question 19	46.04	1.00	8.48	9.28
Mean	26.59	3.91	18.50	24.65

Note: This table groups respondents based on how they respond to each question, excluding questions that are asked about both men and women. "Be Traditional" means that the respondent thinks that men should act according to traditional male gender norms while women should act according to traditional female gender norms. "Men Women Flip" means that the respondent thinks that men should act according to traditional female gender roles and women should act according to traditional male gender roles. "Be like Traditional Woman (Man)" means that the respondent thinks that both women and men should act according to traditional female (male) gender roles.

Table 3: Gender Norms and Labor Force Participation

	(1) LFP	(2) Female LFP	(3) Male LFP
	Panel A: Index based on All Responses		
Standard Index	-0.002	-0.001	-0.001
	(0.005)	(0.006)	(0.006)
Flipped Index	-0.015	-0.022	-0.009
	(0.006)	(0.006)	(0.007)
	Panel B: Index based on Women's Responses		
Standard Index	0.009	0.006	0.013
	(0.006)	(0.006)	(0.006)
Flipped Index	-0.001	-0.003	0.001
	(0.007)	(0.007)	(0.007)
	Panel C: Index based on Men's Responses		
Standard Index	-0.012	-0.008	-0.012
	(0.006)	(0.007)	(0.007)
Flipped Index	-0.017	-0.023	-0.013
	(0.007)	(0.007)	(0.007)
Mean Dep. Variable	0.676	0.632	0.722
Obs.	327	326	327

Note: This table shows the correlation between the standard and flipped PCA indices and labor force participation.

Table 4: Gender Norms and Trump Voting

	Trump Vote Share		
	(1)	(2)	(3)
Index Based On:	Full Sample	Women	Men
Standard Index	0.066	0.016	0.059
	(0.022)	(0.016)	(0.023)
Flipped Index	0.035	-0.017	0.048
	(0.026)	(0.026)	(0.019)
Mean Dep. Variable	0.654	0.654	0.654
Obs.	3,005	3,005	3,005

Note: This table shows the correlation between the standard and flipped PCA indices and Trump's vote share, measured at the county level, in 2020.

Figures

Figure 1: Full Raw Data Ordered by Average Disagreement

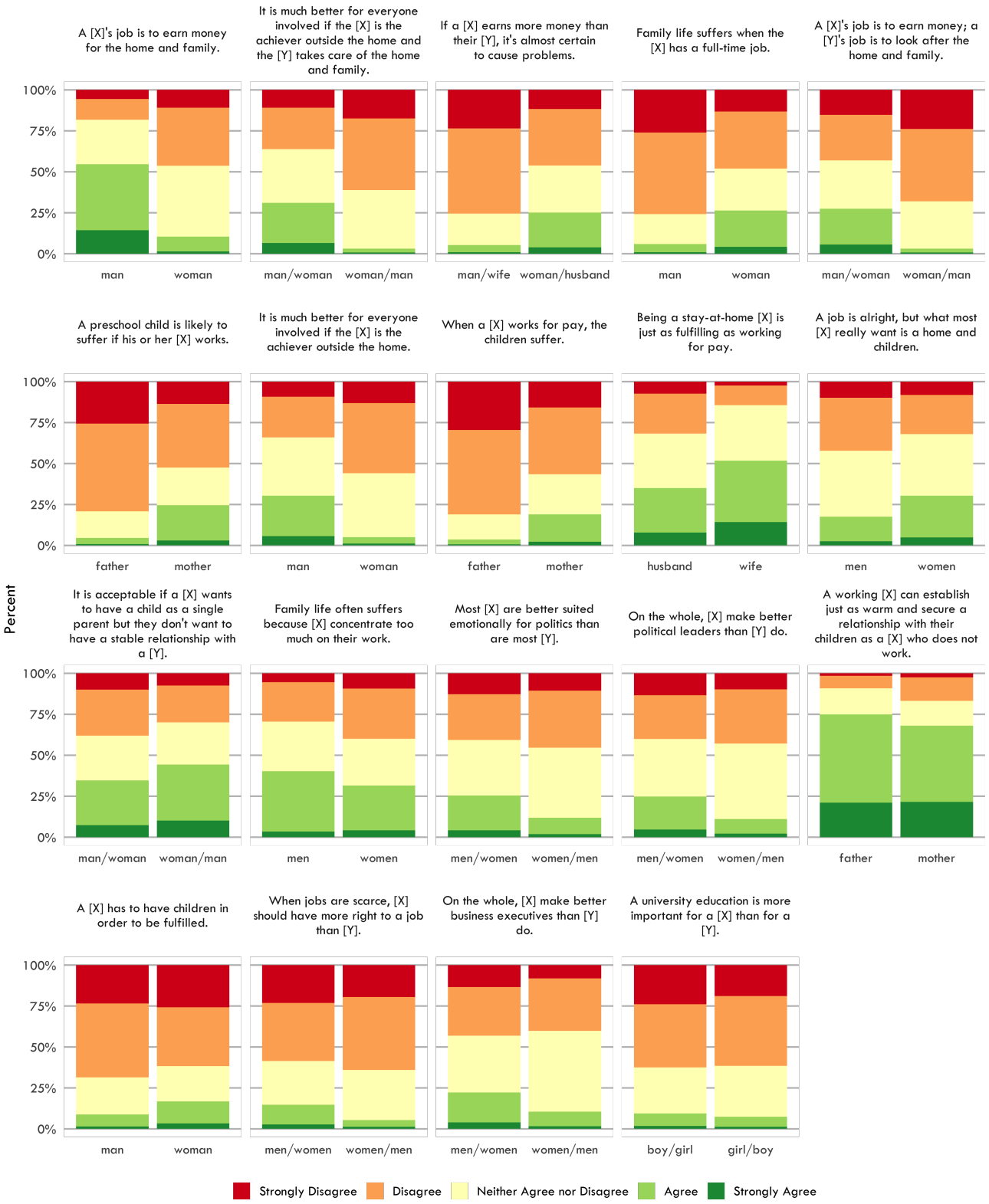
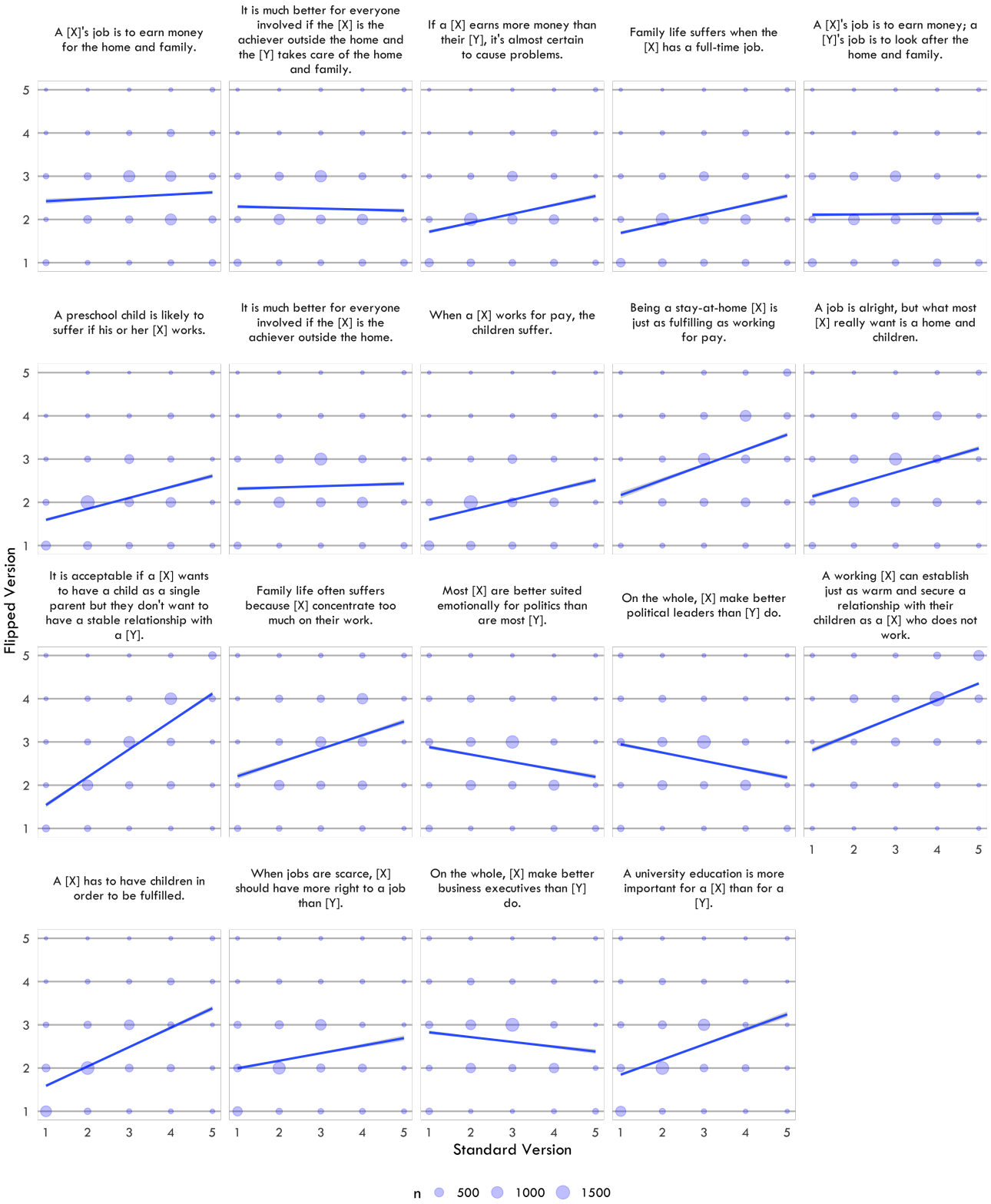
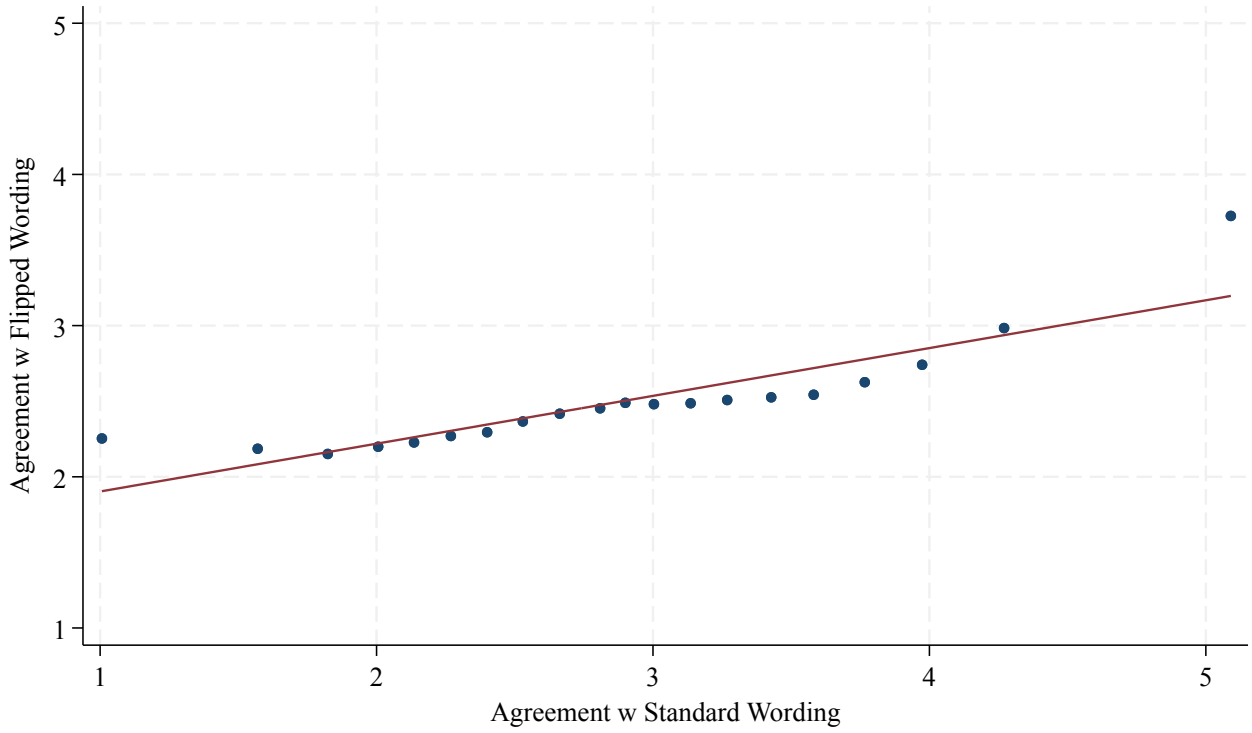


Figure 2: Correlation between Original and Flipped Questions



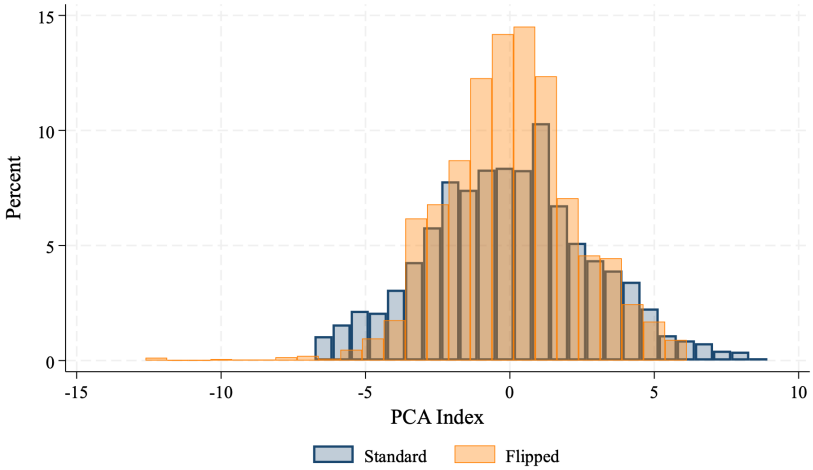
Note: 1 corresponds to strongly disagree and 5 corresponds to strongly agree.

Figure 3: Average Correlation between Statements with Individual FEs

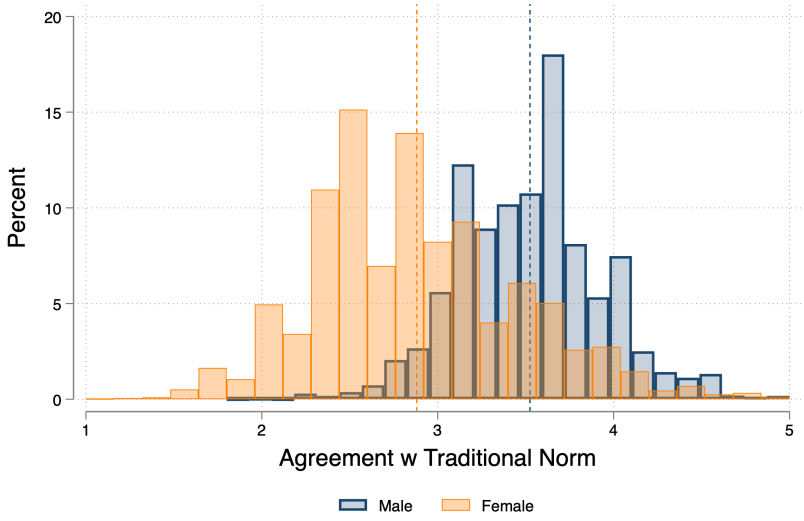


Note: This binned scatter plot shows the correlation between the standard and flipped questions when including individual fixed effects. The line of best fit is plotted and has a slope of 0.32 (s.e. = 0.003).

Figure 4: Agreement with Traditional Gender Norms



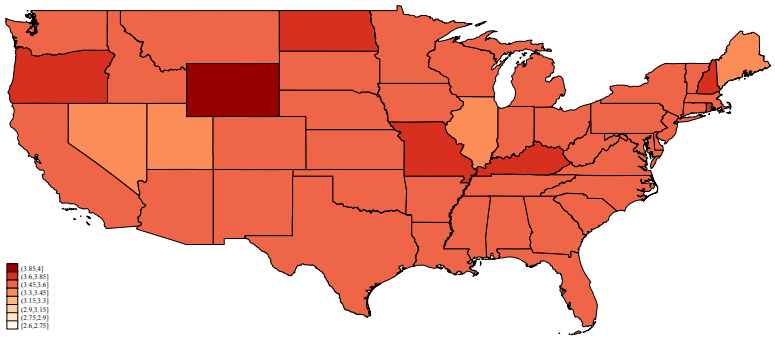
(a) PCA Index



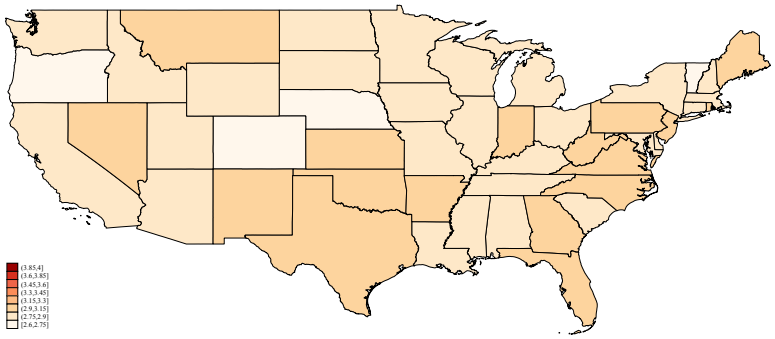
(b) Agreement with Traditional Norms

Note: Panel (a) shows the distribution of the PCA index for the standard and flipped questions. We take the negative of the flipped index so that a higher value indicates more agreement with a more traditional norm. Panel (b) show the distribution of agreement with traditional male norms (blue bars) and traditional female norms (orange bars). The mean for the male norm is 3.35 and for the female norm is 3.02, indicated by the dashed lines. Agreement measures are constructed as described in section 2.3.

Figure 5: Maps of Male and Female Norms



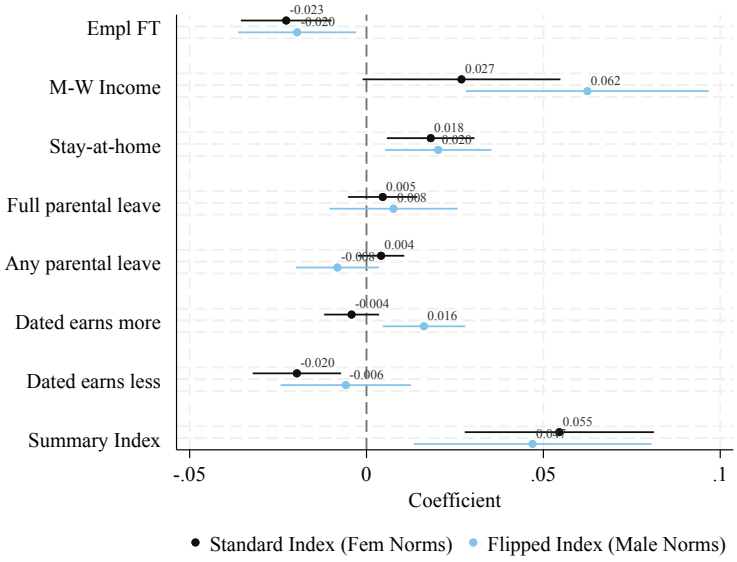
(a) Agreement with Traditional Male Gender Norms



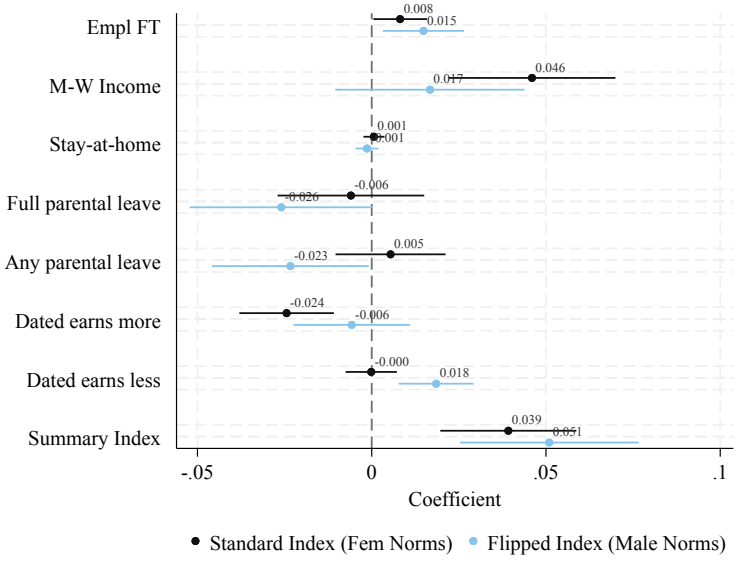
(b) Agreement with Traditional Female Gender Norms

Note: Maps show the geographic distribution of agreement with male (Panel A) and female (Panel B) norms. For comparison, the scale is chosen such that the cut points are the same for the two norms. A lighter color indicates less agreement with the traditional norm and a darker color indicates more agreement.

Figure 6: Correlation between Norms and Self-Reported Outcomes



(a) Female Respondents

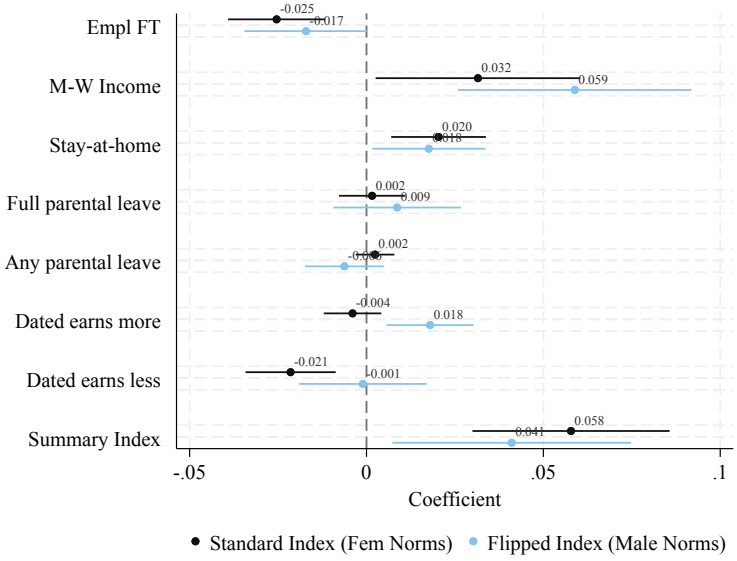


(b) Male Respondents

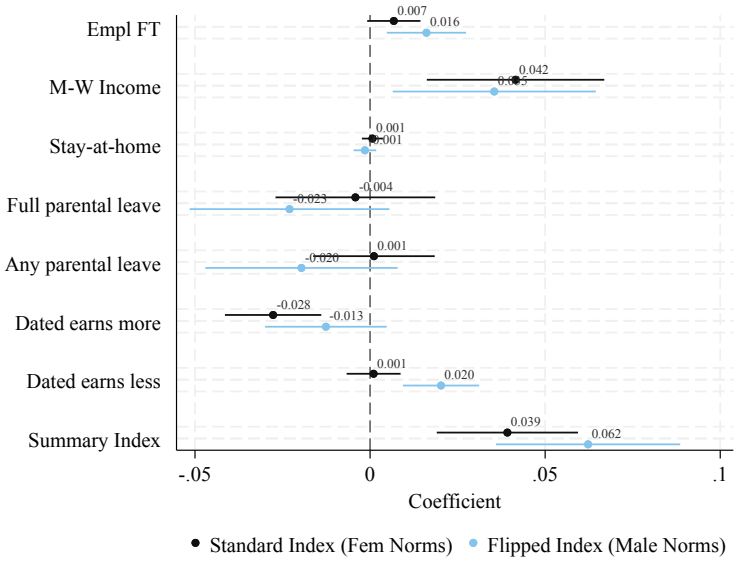
Note: These figures plot the coefficients from regressions relating the standard and flipped indices with various self-reported outcomes from respondents. The summary index is an average of all variables after they have been transformed such that a larger value indicates a more traditional action (such as being less likely to be employed full-time for women). A larger value of the summary index therefore indicates more actions consistent with a traditional gender norm. We take the z-score of the difference between male and female income and of the summary index. We include controls for age, political views, religiosity, state of residence, and education.

Appendix A Appendix Figures

Figure AI: Correlation between Norms and Self-Reported Outcomes, No Controls



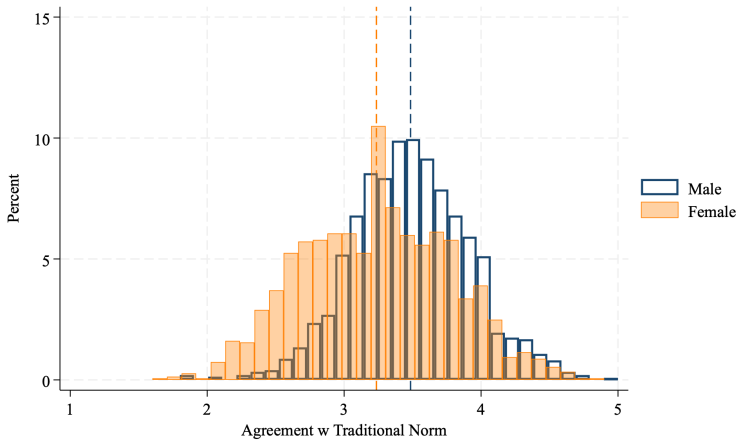
(a) Female Respondents



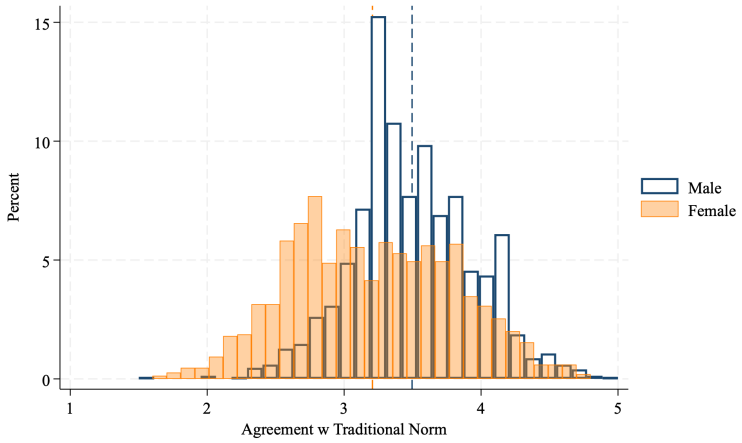
(b) Male Respondents

Note: These figures plot the coefficients from regressions relating the standard and flipped indices with various self-reported outcomes from respondents. The summary index is an average of all variables after they have been transformed such that a larger value indicates a more traditional action (such as being less likely to be employed full-time for women). A larger value of the summary index therefore indicates more actions consistent with a traditional gender norm. We take the z-score of the difference between male and female income and of the summary index. No controls are included.

Figure AII: Agreement with Traditional Gender Norms by Incentivization



(a) Incentive



(b) No Incentive

Note: The above figures show male norms (blue bars) and traditional female norms (orange bars) for the Prolific sample, split based on whether the respondents were incentivized. The mean for the male norm in panel (a) is 3.48 and for the female norm is 3.26. The mean for the male norm in panel (b) is 3.50 and for the female norm is 3.21. The agreement measures are constructed as described in section 2.3.

Appendix B Appendix Tables

Table BI: Sample Demographics

	(1) Prolific	(2) Representative	(3) Difference
Female	0.50	0.51	-0.015
			(0.012)
Non-white	0.26	0.40	-0.14***
			(0.011)
Age	41.4	48.1	-6.68***
			(0.37)
Income Bucket	7.89	10.5	-2.60***
			(0.10)
Num. Children	1.06	2.50	-1.44***
			(0.035)
High School of Less	0.12	0.24	-0.12***
			(0.0089)
College or More	0.036	0.34	-0.31***
			(0.0090)
Employed FT	0.61	0.47	0.14***
			(0.011)
Employed PT	0.14	0.10	0.040***
			(0.0073)
Democrat	0.49	0.34	0.14***
			(0.011)
Republican	0.21	0.26	-0.048***
			(0.0098)
Observations	2995	5294	

Note: This table shows the characteristics of the Prolific and Representative samples. Column 3 shows the difference between the two samples.

Table BII: Principal Component Analysis

Question Number	Standard Questions	Flipped Questions
1	0.2776	0.2980
2	0.2530	0.2410
3	0.2804	0.3001
4	0.2604	0.2856
5	0.2171	0.1703
6	0.0669	0.0446
7	0.2521	0.2759
8	0.2421	0.1452
9	0.2457	0.2730
10	-0.1625	0.0699
11	-0.2019	-0.0996
12	0.0257	0.1020
13	0.2614	0.2419
14	0.2167	0.2156
15	0.2541	0.2877
16	0.2016	0.2546
17	0.2674	0.2514
18	0.2637	0.2912
19	0.2214	0.2302

Note: This table shows the eigenvectors for the first component, for the standard questions (column 2) and flipped questions (column 3)

Appendix C Data Appendix

Table DI: Question List

1	It is much better for everyone involved if the man is the achiever outside the home and the woman takes care of the home and family.
2	Most men are better suited emotionally for politics than are most women.
3	A man's job is to earn money; a woman's job is to look after the home and family.
4	Family life suffers when the woman has a full-time job.
5	A job is alright, but what most women really want is a home and children.
6	Family life often suffers because men concentrate too much on their work.
7	A preschool child is likely to suffer if his or her mother works.
8	A woman has to have children in order to be fulfilled.
9	When jobs are scarce, men should have more right to a job than women.
10	It is acceptable if a woman wants to have a child as a single parent but she doesn't want to have a stable relationship with a man.
11	A working mother can establish just as warm and secure a relationship with her children as a mother who does not work.
12	Being a stay-at-home wife is just as fulfilling as working for pay.
13	On the whole, men make better political leaders than women do.
14	A university education is more important for a boy than for a girl.
15	When a mother works for pay, the children suffer.
16	If a woman earns more money than her husband, it's almost certain to cause problems.
17	On the whole, men make better business executives than women do.
18	It is much better for everyone involved if the man is the achiever outside the home.
19	A man's job is to earn money for the home and family.

Notes: For each question, agreement with the statement indicates adherence to gender-congruent norms, except for questions 6 and 11, where disagreement indicates adherence to gender-congruent norms. Questions with a bolded number are those used in the primary analysis as they reference only men or only women.