

Measuring Social Norms About Female Labor Force Participation in Jordan

Varun Gauri
Tasmia Rahman
Iman Sen



WORLD BANK GROUP

Development Economics
Development Research Group
June 2019

Abstract

This study conducted a large-scale, representative survey of social norms for female labor force participation in three governorates of Jordan. The social norms measures are disaggregated into thematic clusters, empirical and normative expectations, and interpersonal expectations within the household. The measurements satisfy reasonable tests for internal consistency, external validity, and test-retest reliability. The survey shows that the great majority of men and women favor women's labor force participation, although support falls under specific scenarios. Most non-working

women would like a job. Among married women, the strongest correlates of working are the woman's expectations of her husband's views and the husband's personal beliefs. Among unmarried women, empirical expectations of the number of women working correlate strongest with labor force participation. The study findings indicate that information campaigns highlighting hidden support for women working could be effective, although distinct messages for men, married women, and unmarried women may be useful.

This paper is a product of the Development Research Group, Development Economics. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at <http://www.worldbank.org/prwp>. The authors may be contacted at vgauri@worldbank.org.

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

Measuring Social Norms About Female Labor Force Participation in Jordan

Varun Gauri, Tasmia Rahman, and Iman Sen

Keywords: Survey methods, Social norms, Female labor force participation

JEL Classification Codes: C83, D91, J29

Mind, Behavior, and Development (eMBeD) team, World Bank. E-mails: vgauri@worldbank.org, trahman4@worldbank.org, isen@worldbank.org. We thank Mariana Felicio, Samantha Constant, and Anoud Allouzi for inviting us to collaborate on the project. We are grateful to D3 Systems (<https://www.d3systems.com/>) and their local partner in Jordan for advice on and implementation of the survey that underlies this work. Caren Grown, Elena Ianchovichina, Julian Jamison, Nina Mazar, Ryan Muldoon, Maria Beatriz Orlando, Betsy Paluck, and Tara Vishwanath provided valuable advice. We appreciate the feedback received from seminar participants at the University of Toronto, Oxford University, the Woodrow Wilson Center for Scholars, and the World Bank. All errors remain our own. The views and findings expressed in this paper are those of the authors alone and do not necessarily represent those of the World Bank or its Executive Directors.

Measuring Social Norms About Female Labor Force Participation in Jordan

Varun Gauri, Tasmia Rahman, and Iman Sen

It is widely understood that social norms constrain women's labor force participation worldwide (Alesina, Giuliano, & Nunn, 2013; Bertrand, Kamenica, & Pan., 2015; Chamlou, Muzi, & Ahmed, 2016; Fernandez, 2013; World Bank, 2012). The role of women in childcare and domestic tasks, their visibility in public and mixed gender environments, segregation into particular jobs and industries, and women's decision making power relative to men are all thought to bear on women's labor force participation. What is not as well understood are the salience and importance of these various behaviors and beliefs, the relative influence of importance of behavior versus normative views, and the influence of husbands and fathers in comparison to the views and behavior of the society at large.

This study develops methods for measuring and validating social norms and beliefs regarding women's access to and participation in labor markets. Using a large-scale quantitative survey conducted across three governorates of Jordan, the study measures the components of social norms in a sample of 1,000 households; conducts tests of the consistency, validity, stability, and reliability of the measures; and examines the influence of inter-personal expectations within the household on women's beliefs and choices.

Jordan's female labor force participation is the lowest in the world of a country not at war. Though Jordanian women are among the most educated in the MENA region (53% of university graduates are women), female labor force participation (FLFP) was 14 percent in 2008, compared to 64 percent among men (The World Bank, 2018). Our study finds that social norms play a key role in these labor market outcomes for women. But social norms are not so deeply internalized that there is little scope for interventions – most non-working women would like to work. More than 95% of Jordanians, men and women, support women's participation in the labor force, though they underestimate the extent to which others are supportive. At the same time, support for FLFP declines when people are asked about specific scenarios. The most important influences on a married woman's orientation toward work appear to arise from her beliefs about what her husband expects and believes, rather than from concerns regarding generalized social norms. Our study shows that a careful examination of social norms can orient the targets and messages of information campaigns to increase FLFP.

What Are Social Norms and How Do We Measure Them?

Social norms, which refer to widely shared beliefs about how others in our social group behave and how they ought to behave, are a product of human sociality. They arise from social interdependence, and are the product of a rule for behavior and common knowledge of that rule. Social norms are informal governance mechanisms and exert a powerful influence of our decision-making and behavior. In consequence, norms have been called the “glue” or “cement” of society (Elster, 1989). This tendency to associate and behave as members of groups- what we call human sociality- can cause groups or societies to get stuck in and perpetuate negative or harmful collective patterns of behavior (World Bank, 2015).

Social norms related to gender are pervasive in every human society. They prescribe a wide range of practices, behaviors, and even thoughts and feelings that are considered appropriate for men and women to have and exhibit. These include inter-personal behaviors (e.g., when and how many children to have), social practices (e.g., dress, speech, dominance, child rearing), political actions (e.g., holding and exercising public office, voting), and economic decisions (e.g., participating in the labor force, opening a bank account, starting a business, employing others). Although these various behaviors usually draw on a common set of mental models of the world and human relationships (including the very idea that society is

composed of people of either one or another gender), they are somewhat separable. There can be changes in social norms in one area but not in others. For example, in many societies it has become more acceptable for women to work and hold office, but there has been less change in the division of household labor and responsibility for child-rearing.

Social norms consist of empirical and normative beliefs. These have been characterized as “social empirical” and “social normative” expectations (Bicchieri 2006, 2016) or as “descriptive” and “injunctive” norms (Cialdini, Kallgren, & Reno, 1991). Another way of stating this dual aspect is to say that social norms consist of a rule for behavior and common knowledge of that rule (Brennan, Eriksson, Goodin, & Southwood, 2013). While Cialdini & Trost (1998) and Fishbein & Aizen (2010) classify observed behavior of others as “descriptive norms”, others classify this as “collective behavioral norms” (Lipiniski & Rimal, 2005; Storey & Schoemaker, 2006) or “empirical expectations” (Bicchieri, 2006). Beliefs about what others approve of, on the other hand, are called “injunctive norms”, “collective attitudinal norms”, or “normative expectations” respectively.

Similarly, the force of social norms can arise from social proof or conformity. While Cialdini defines descriptive norms in terms of social proof (or one-way dependence where we follow something because others are doing it by assuming that they know what they are doing), Bicchieri also takes into account the role of social convention, which is more interdependent (i.e. each individual does something with the expectation that others will also do the same- also known as many-way interdependence). Similarly, Bicchieri’s definition of normative expectations goes beyond simple expectations of acceptable behavior within one’s reference group to also incorporate the role of conformity- i.e. beliefs about others’ expectation of conformity from us, and our obligation to conform- and the subsequent expectations that others would sanction conforming and non-conforming behavior (Mackie, Moneti, Shakya, & Denny, 2015).

Social norms can be internalized. Women’s economic participation may, for instance, be influenced by mental models about how the world works, or by internalized cultural schemata that operate below the level of explicit articulation and recall. These cultural schema shape perceptions and filter the “facts” that people believe and are able to understand (DiMaggio, 1997). For example, women may believe (correctly or incorrectly) that some industries and not others are female friendly. They may draw on beliefs about safety, transportation, and the level of skills required. They may also have visceral, habitus-like reactions to purity and order that affect their reactions to mixed gender environments or to supplanting their husbands as breadwinners (Dildar, 2015; Evans, 2016).

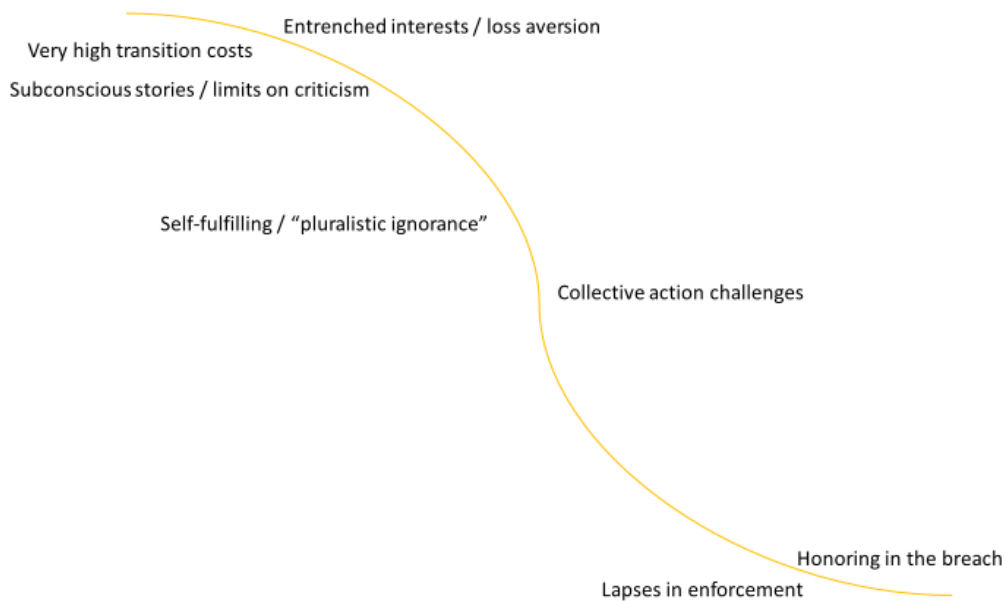
Most theorists agree that social norms are rules for behavior within a particular group. In other words, they hold for a relevant reference group. For example, women’s labor force participation in Europe probably does not matter much for women’s LFP decisions in Jordan, and even the LFP in Amman may not matter for women in, say, Zarqa. One of the major challenges in measuring social norms is to identify and ask about the relevant reference group for the behavior and social norm in question, a problem made more complex by the fact that the reference group for a social norm may be implicit or variable.

Norm Change

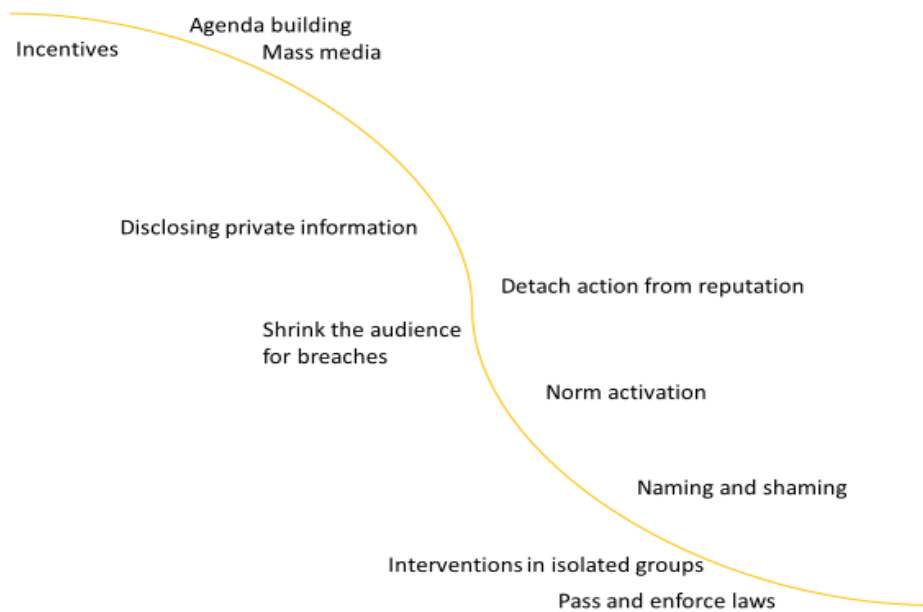
Sometimes social norms unravel as a result of historical, economic, or social factors. New production technologies, migration, conquest, and cultural and religious revolutions can destabilize social norms and create new ones (Bicchieri & Mercier, 2014; Connell, 2016; Ocheje, 2018; Paluck, 2009). The success of explicit efforts to change social norms depends on the extent of norm internalization and the mechanisms that sustain norms. One can graph social norms on an kind of inverted S-curve, with a tipping point in

which the cost of explicit norm change efforts is high when behaviors are widely shared and sustained by deeply internalized beliefs, but falls dramatically during the vertical segment of the S, where steep declines in norm following can occur at low cost. At the tail of the S, norm following is very low and remains low even when interventions formerly used to sustain social norms are withdrawn. Figure 1 presents a stylized version of an S curve for social norms change with respect to FLFP, with adherence to the social norm of women not working on the y-axis, the cost of interventions in the x-axis, and candidate mechanisms that sustain the social norms at different points of the curve labeled. Figure 2 presents the same curve with potential interventions targeting the norm maintenance mechanisms, along different points of the curve. The figures illustrate that the diagnostics of social norms can help identify: a) the prevalence of social norms and norm-following behavior; b) the likely mechanisms maintaining the social norms at different levels of prevalence; and therefore c) the most promising and cost-effective potential interventions to weaken social norms, given the social norms maintenance mechanisms at work.

Figures 1. S-curve (mechanisms sustaining social norms)



Figures 2. S-Curve (interventions to address barriers)



When norms are deeply entrenched, either because they are highly internalized or because of mobilized interests, expensive norm change mechanisms, such as incentives and agenda building through political mobilization, may be necessary. In the vertical portion of the S-curve, pluralistic ignorance (a discrepancy between widespread private support for norm breaching and what gets expressed publicly) may be a key norm maintenance mechanism. In that case, simply disclosing privately held but unshared views could unravel the norm. For example, a recent study in Saudi Arabia showed that when married men were provided information about preferences of other married men within their social group with regards to FLFP, it significantly improved the former’s willingness to let their wives work (Bursztyn, Gonzalez, & Yanagizawa-Drott, 2018). The experiment revealed that around 72% of participants underestimated other men’s support for FLFP (a case of pluralistic ignorance). By disclosing private information about others’ actual support for FLFP, the experiment corrected the misperceived social norm that was preventing men from supporting their wives to find jobs.

Similarly, activating a countervailing social norm in a new context can be another means of unraveling a social norm. A study in Kenya addressed collective action challenges when it comes to reducing traffic accidents by encouraging passengers on buses to speak up against reckless driving (Habyarimana & Jack, 2009, 2011). This norm activation- communicated through stickers posted inside buses- was remarkably successful in getting passengers to play an active role in improving safety, and reduced insurance claims involving injury or death by half.

At the bottom of the S-curve, most individuals are opposed to the social norm and publicly express that opposition. But there can be lapses in that public expression, leading to occasional retrogression. Here, authoritative actions, such as laws, can make the social norm salient to transgressors, and naming and

shaming techniques can use widespread common knowledge of opposition to the social norm to inflict reputational costs on transgressors.

Measuring Social Norms

For Bicchieri, Lindemans, & Jiang (2014), the identification of a social norm requires measuring five components: a) individual behavior; b) empirical expectations about the behavior of others (e.g., how many other women work, and what they sacrifice when they work); c) personal beliefs; d) normative expectations concerning the moral beliefs of others and their sanctioning behavior (e.g., how many other people disapprove of women working, and whether they speak badly of women who do); and e) cost-benefit calculations on the part of individuals (in order to distinguish rational from social norms driven behavior).

In addition to these, it is also important, particularly in the context gender norms, to consider the role of power relations in the perpetuation and enforcement of social norms. Norms influence behavior through expectations of what others in the reference group do or approve. However, not all groups of individuals within the reference group may exert the same amount of influence in maintaining a norm. Coleman (1994) classifies the relevant agents into two groups- targets and beneficiaries. Targets refer to those to whom the norm applies, while beneficiaries are those who benefit from and are usually the ones to enforce or sanction the norm. In the case of female labor force participation, this distinction may be relevant, given the power dynamics between spouses within a given household. While men may or may not be explicit “beneficiaries,” their role in enforcing the norm of women not working may be more potent than the role of others in the larger reference group. Another way to think of this is that men might be the “targets” and “beneficiaries” of the male breadwinner norm, and they might enforce that norm by sanctioning women who join the labor force. In any case, it may be useful in the case of FLFP not only to assess the generalized social norm but the beliefs and behaviors of the husbands, fathers, and other influential males for a given woman.

If normative and empirical expectations are consistently reported in a social group, there is strong prima facie evidence that a social norm exists. Similarly, if a sufficient number of respondents express mutually consistent views about what others should do, normative expectations are strong. Normative expectations are necessary, but not sufficient, for the existence of a strong social norm. In order for a social norm to be strong, the normative expectations need to cause conformity, or change behavior, in sufficiently large numbers. To demonstrate the causal efficacy of the norm and see if individuals, in fact, do conform to the social norm, it is thus necessary to measure individual behavior as well, either directly or through counterfactual vignettes.

Methodology and Instrument Design

Fourteen focus group discussions (FGDs), with working and non-working men and women, across two age groups (21-30 and 31-59), in urban areas in Amman and Mafraq, informed the instrument design. The FGDs revealed that although women are becoming more publicly visible in Jordanian society, women, whether working or not, are still expected to handle childcare responsibilities and household chores while men continue to be seen as household financial providers. There were concerns in the FGDs that working women may undermine men’s role as the source of the family’s financial status. Despite wanting educated and even working wives, men expressed concerns about women becoming less tolerant and less patient as a result of working. Men also worried about women’s safety due to increased risk of sexual harassment.

FGD participants named several obstacles to women working, including the influence of male family members (especially husbands), limited childcare options and the stigma of leaving young children to work,

long work hours, low pay, and the risk of harassment, particularly when jobs require regular interaction with men. Interestingly, generalized social norms (understood as empirical expectations of whether other women work or as expectations that people frown upon women working) appeared to be less of a barrier, in the FGDs, than women’s interpersonal expectations of specific male family members. Both male and female participants consistently estimated FLFP in Jordan to be above 50% (in reality, it is around 14%).

The theoretical account of social norms, described above, and the FGDs formed the basis of the quantitative survey instrument. The survey instrument examined normative influences across four related themes identified earlier: i) general views on whether women should work (which we called “working women”); ii) beliefs about gender roles as they relate to women’s ability and availability to work (“gender roles”); iii) beliefs about publicness and mixing of genders in the work environment (“public-ness and mixing”); and iv) beliefs about the link between women working and family status (“status”). All four themes were explored using a sub-group of questions, which were adapted as needed for male and female respondents (see Table 1).

Table 1. Themes

Working Women (WW)	Gender Roles (GR)
<i>WW1. Women working from home</i>	<i>GR1. Married women working</i>
<i>WW2. Women working outside the home</i>	<i>GR2. Married women returning home after 5 PM</i>
<i>WW3. Women working in Jordan (only SE)</i>	<i>GR3. Leaving child <5 with relative during work</i>
<i>WW4. Women working if husband not comfortable (only PB/SN/IB)</i>	<i>GR4. Appropriate age to leave child and go to work</i>
<i>WW5. Necessary for both husband and wife to work to live comfortably (only PB)</i>	
<i>WW6. Women can work if other women are too (only IB)</i>	
Publicness and Mixing (PM)	Status (ST)
<i>PM1. Working in mixed-gender environment</i>	<i>ST1. Family does not abide by traditions</i>
<i>PM2. Exposure to harassment</i>	<i>ST2. Husband cannot provide for her</i>
<i>PM3. Reputation of working women</i>	<i>ST3. Husband not in charge of family</i>
	<i>ST4. Family suffering from financial need</i>

For each question under a theme, the instrument first elicited the respondent’s own behavior and direct beliefs, followed by their corresponding social empirical and normative expectations. Table 2 provides an illustrative example this. This structure allowed an examination of how closely associated the main outcome of interest- women working outside their homes- was to expectations of whether others in one’s social group also work outside their homes, and to expectations of the acceptability of working outside the home.

Table 2. Structure

Personal behavior	Personal normative beliefs
Do you/your spouse work?	Is it okay for women to work outside of their homes?
Social empirical expectations	Social normative expectations

Take a moment to think about the adult women where you live. These could include your family members, friends, neighbors, and others. Out of 10 such women, how many work outside their home?

Take a moment to think about all the people where you live. These could include your family members, friends, neighbors, and others. Out of ten such people how many would think or speak badly about married women who, because of work, return home after 5pm in the evening?

For women only:

Think now for a moment about your husband/ father/ brother, and his views Does he think or speak badly about women who work outside their homes?

To examine the cultural mental models and “schemata,” respondents were asked to record open-ended questions inspired by classical accounts of how narratives shape behavior (Bruner, 1990; Crossley, 2002; Sarbin, 1986). Each respondent was asked a random set of narrative questions that explored their views of how typical or atypical they perceived themselves or their relationship with a specified male or female relative to be, how they explained their or their spouse’s motivation to work to others, how they talked to others about work related decisions, and why they thought education is important for women’s future.

The draft quantitative survey was field tested to determine how effectively the questions elicited social norms and beliefs, improve the language and instructions provided to enumerators on questions, and refine guidance to enumerators on how to elicit informative but concise narratives from respondents. The pre-test was conducted with 50 households in Amman and included households where both the husband and wife were interviewed. The purpose of interviewing both spouses within a given household was to test the feasibility of interviewing couples in a household separately.

Some variables explored during the pre-test were excluded in the final questionnaire. An implicit attitude test (IAT) was designed but excluded once it became clear that FLFP in Jordan is likely driven by explicit rather than implicit views and because the IAT proved costly to administer. A series of vignettes to examine the conditionality of FLFP preferences were designed but excluded once the number of circumstances, contingencies, and conditions to the social norms around FLFP became evident (e.g., visibility of women’s work, family circumstances, age of children, gender mixing at work, work hours).

In addition, the pre-test helped refine the reference group, which is theoretically crucial in the link between social norms and behavior. Respondents did not appear to know about FLFP in Jordan as a whole, and the questions about social norms in the neighborhood or community appeared to be variously and inconsistently interpreted in the peri-urban settings. The final questionnaire instead asked about social norms “where you live,” in order to cue the respondents’ own sense of place. This flexibility was preferred, we believe, because the actual reference group varied, and it allowed individuals to imagine the group that mattered to them. We avoided further clarification questions regarding reference groups so that the reference group “priming” remained consistent across respondents.

In measuring sensitive beliefs and social norms, it was important to be wary of social desirability bias- the tendency of respondents to answer questions in a way that is “appropriate” or “correct”, rather than a true reflection of their beliefs. In our case, respondents might guess that the enumerators would prefer to hear answers that support women’s economic empowerment. Or perhaps respondents have an internalized

judgment about women who do not work. Our enumerators were all Jordanians, which may have gone some way toward mitigating social desirability bias. In addition, we incentivized responses to some questions by offering rewards if they were able to estimate correctly the number of respondents who expressed certain beliefs or behaved in certain ways. Incentivizing accuracy should have mitigated social desirability bias. The pre-test also made clear the importance of examining the expectations couples had of each other. The final questionnaire focused on the mutual expectations of male-female couples in order to explore whether a) social norms are enforced primarily in households, rather than through generalized disapproval, gossip, or harassment; and b) intra-household power dynamics, or interpersonal expectations within the household, could constrain women's economic participation even independently of social norms.

Two additional findings from the pre-test were of interest. First, it became clear that social norms regarding women working in the private sector, though important in the minds of policy makers, did not appear to be prominent to respondents, who did not know enough women working in the private sector to be able to adequately distinguish between jobs in the public and private sectors. Accordingly, the final questionnaire dropped a set of questions regarding social norms of women working in the public vs. private sectors. Second, there appeared to be some continuity between normative and empirical expectations questions. For instance, the harassment of women working is both a circumstance that affects the empirical likelihood a Jordanian woman works (like having young children) and an expression of normative disapproval in Jordan society. But rather than deciding in advance whether questions like this were to be understood as social empirical or social normative expectations, we decided to examine this in the exploratory factor analysis after the data were collected.

Approximately two weeks after the completion of the main survey, a follow-up survey was conducted with 1,000 respondents. The follow-up survey included selected questions from each of the themes from each social norms module. It was conducted over the phone with a random subset of respondents who agreed to share their phone numbers for the short follow-up survey. The purpose of this survey was to assess if responses to social norms and beliefs related questions stayed consistent over time.

Sample and Data Collection

The main survey was conducted with a sample of 2,007 respondents from low to middle-income families, in urban and peri-urban areas of Amman, Zarqa, and Mafraq, between March and April 2018. Given the low rate of female labor force participation in Jordan, working women were oversampled at each location to obtain a sufficiently large sample size for comparison and analysis. Targeting at least 30% working women, the sampling methodology was a stratified cluster sample based off an enumeration exercise. With Amman comprising of 70% of the sampling frame, the governorate was under-sampled by 15 clusters, which were transferred to the Mafraq stratum. Localities within each governorate served as the primary sampling unit; grids within each locality as the secondary sampling unit; and within each randomly selected grid, census blocks acted as the tertiary sampling unit. Census blocks were enumerated to locate working women, and priority of selection was given to working women within each household. Lastly, a stratified (by work status and gender) simple random sample was drawn from eligible households (fourth sampling unit).

Quantitative data from the pre-test, quantitative survey, and follow-up survey were collected using tablets. Enumerators were trained extensively by research team members and the local data firm, and the survey was pre-tested to check for comprehension, proper administration, and survey fatigue. Enumeration data were used to pre-assign respondents to enumerators, and replacement households were also assigned to minimize enumerator discretion in the field. Data were analyzed using Stata to generate descriptive tables and figures and to conduct validity, reliability and regression analysis, and Python for narrative analysis.

The final sample was split evenly between male and female respondents. Across all three governorates, 30% of female respondents in the sample were working (by design), and 70% of male respondents were working. Of male respondents, 29% had working female relatives (either wife, sister, or daughter), and 71% had non-working female relatives. The sample included 828 households where both a male and a female respondent was interviewed. Nearly 83% of the respondents were interviewed in pairs- father and daughter (7%), brother and sister (17%), or spouses (76%) living in the same household- resulting in 1,179 unique households altogether. Table 3 presents a detailed breakdown of the key demographic information of the survey sample, separated by gender and work status of respondents (for women) or respondent's female counterpart (for men).

Table 3. Demographic Information

Variables	All	Women		Men	
N	2007	1061		946	
Gender		53%		47%	
Working		30%		70%	
	All	Working Women	Non-working Women	Men with working counterpart	Men with non-working counterpart
N	2007	315	746	275	671
Location					
<i>Amman</i>	57%	67%	55%	65%	52%
<i>Mafraq</i>	13%	13%	13%	13%	14%
<i>Zarqa</i>	29%	20%	32%	21%	34%
Married (%)	81%	63%	84%	73%	89%
% w/ child's age<6	42%	37%	47%	35%	42%
Age (years)	39.4 (12.34)	34.7 (8.95)	36.0 (9.92)	42.0 (14.48)	44.3 (13.17)
Years of education	11.90 (3.52)	13.72 (3.25)	11.32 (3.10)	12.87 (3.66)	11.29 (3.63)
Avg. no. of children	1.94 (1.77)	1.57 (1.67)	2.17 (1.79)	1.51 (1.67)	2.02 (1.79)
Avg. HH income (JD/month)	540 (499)	717 (731)	448 (316)	769 (733)	463 (329)

**Standard deviation in brackets*

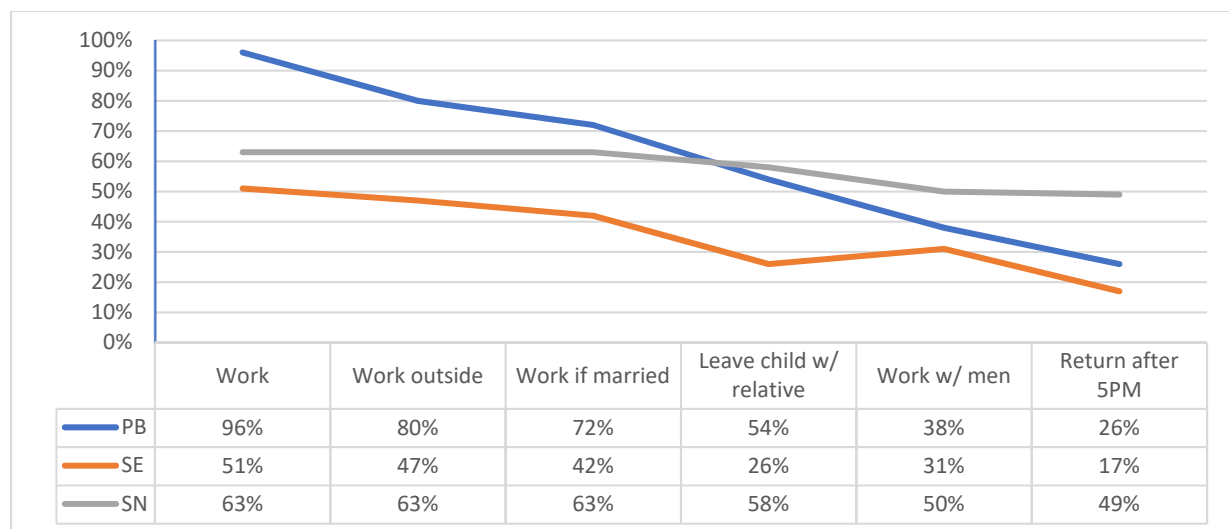
There are notable demographic differences between working and non-working women, especially when it comes to marital status and children. Across all groups, working women have the highest level of education on average (13.7 years), while men with non-working counterparts have the lowest (11.3 years). While the difference is small, it is worth noting that the educational attainment of working women surpasses that of working men. Non-working women have levels of education similar to their male counterparts. Sixty-three percent of working women in the sample are married compared to 84% of non-working women. Working women also tend to have lower number of children on average (1.57 compared to 2.17 for non-working women), with fewer having children below the age of six (37% compared to 47%). Similar trends are also observed among male respondents with working and non-working female counterparts. Male respondents in the sample are older on average (42-44 years) compared to women (around 35 years) because our sample includes father-daughter.

Descriptive Findings

Despite the prevailing low labor force participation rate among women in Jordan, there was widespread support for FLFP among survey respondents. Ninety-six percent of respondents (both men and women) were okay with women working (see Figure 3 below, personal beliefs). However, this general support for FLFP came with important caveats, and declined steadily as additional information about the nature of the job or working women are specified. Eighty percent were okay with women working outside the home, while slightly less, 72%, were okay with married women working. Working in mixed gender-environments, which is a common feature of many jobs, brought down support for (i.e. acceptability of) working women to 38%. Working women returning home from work after 5 PM further lowered the acceptability of women working to 26%. Given that many jobs, particularly jobs in the private sector, in which 72% of working women in our sample were employed, frequently involve working in mixed-gender environments and working till 5 PM, these views could be severely constraining the job options women can realistically consider. More generally, these findings make clear that prevailing social norms may be nuanced, and that, as a result, responses to simple yes/no questions regarding support for FLFP can be misleading.

Figure 3. Is it okay for a woman to work?

Personal Beliefs, Social Empirical Expectations, and Social Normative Expectations



* PB= Personal Beliefs; SE= Social Empirical Expectations; SN= Social Normative Expectations

If accurate, social expectations are simply the aggregate of personal beliefs of all respondents in the sample. In our sample, although 96% of respondents believed it was fine for women to work, they estimated, on average, that 63% of the people where they lived believed the same. This divergence opens the possibility that public knowledge regarding the personal beliefs of others diverges from the views that others privately hold (“pluralistic ignorance”). In this situation, it is possible, that disclosing the strong but hidden support for women working could by itself change social norms and FLFP (Bursztyn et al., 2018). Personal beliefs were also more “liberal” than social normative expectations regarding women working outside the home and women working if married. On the other hand, personal beliefs were more “conservative” than social normative expectations regarding women working with men and women returning after 5 pm. Overall, a picture emerges in which Jordanian respondents saw themselves as more “liberal” than others when it came to supporting women’s right to work, but more “conservative” than others when it came to protecting women from physical and reputational risks. This suggests that while there may be scope for dispelling

pluralistic ignorance about the general attitude towards women working, the practical/structural dimensions of work appear to be less influenced by what others think.

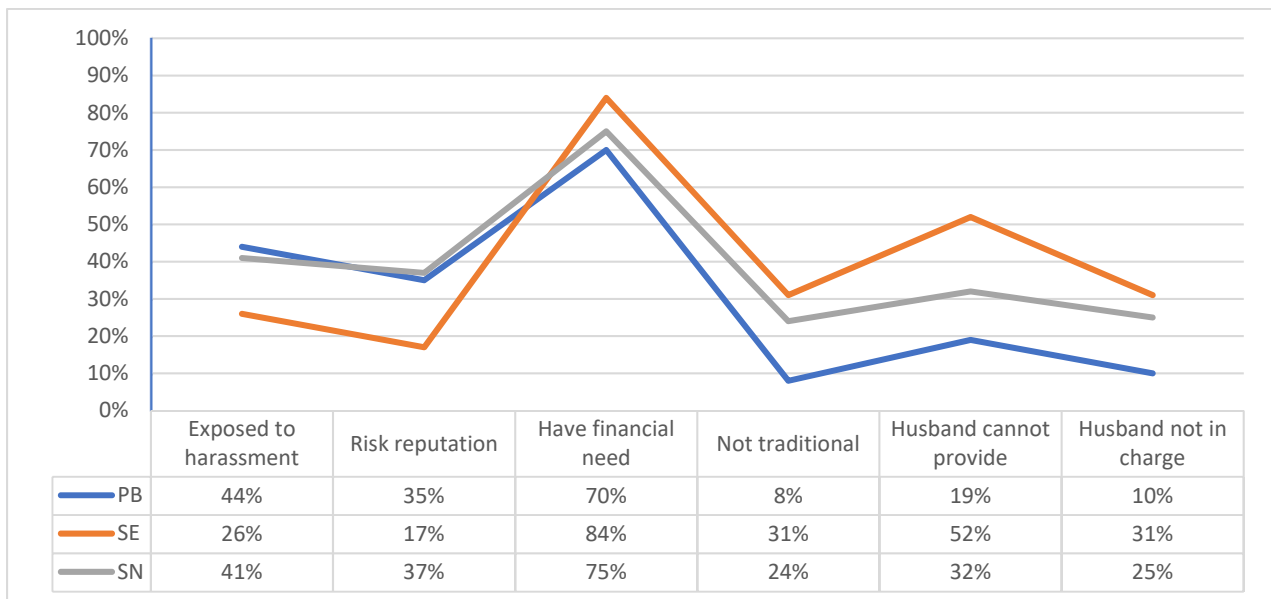
Social empirical expectations (e.g., how many women where you live work, work if married, work and return after 5 pm, etc.) were lower than social normative expectations across the board. Respondents also tended to believe that women worked, in the various scenarios, at lower rates than the respondents themselves endorsed (in their own personal beliefs). Interestingly, respondents significantly overestimated the actual level of FLFP around them (51% compared to 14% in reality).

A second set of questions explored prejudices and concerns regarding working women and their families. For the most part, personal beliefs and social expectations closely tracked (Figure 4). But when it came to the effect of a woman working on the family’s status and the husband’s reputation (e.g., when a woman works, it shows that her husband cannot provide, is not in charge, shows the family is not traditional), social empirical expectations were far more negative than personal beliefs. This suggests that the pluralistic ignorance regarding the “breadwinner norm” may inhibit FLFP in Jordan- women and men worry that a woman working will make the husband appear to be less “masculine,” even if they themselves do not believe it should.

Putting this together with generalized pluralistic ignorance and the “protection norm” above, the three domains ripe for social norms messaging regarding FLFP in Jordan may be: a) most people support FLFP, contrary to their expectations regarding the conservatism of others; b) women can work safely and avoid putting their reputations at risk, because this is a widely held concern; and c) most people overestimate the reputational cost to the family and the husband when a woman works. At the same time, people seem to have widely shared personal beliefs regarding the circumstances and conditions under which it is okay for a woman to work, which suggests that job design is an important part of the FLFP agenda, as well.

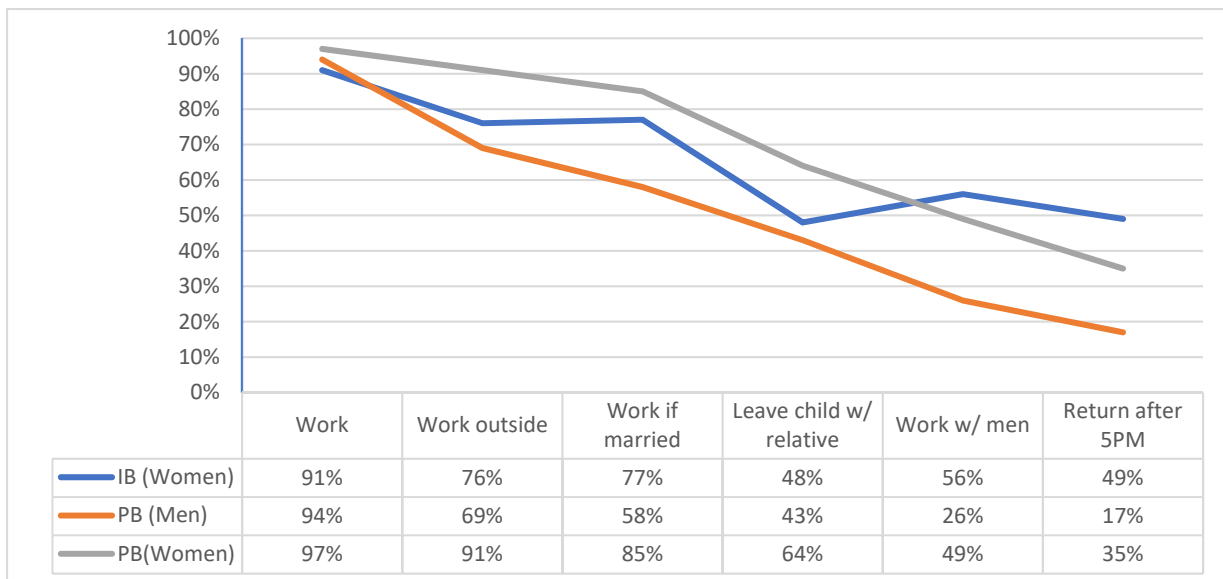
Figure 4. Do working women risk safety and reputation?

Personal Beliefs, Social Empirical Expectations, and Social Normative Expectations



As discussed above, the crucial social norms enforcers for FLFP may be intimate men (husbands, fathers, brothers), rather than friends, acquaintances, strangers, and society in general.¹ To assess this, we compared personal beliefs and interpersonal expectations at the household level. For the most part, the interpersonal expectations of women were not far from the actual personal beliefs of the men they were asked about, but there were notable exceptions (Figure 5). While 56% of women expected their counterparts to not disapprove of women working in mixed-gender environments, only 26% said they did not disapprove. Similarly, 49% of women expected their intimate counterparts to be okay with a married woman returning from work after 5PM, but in reality 17% of those men were okay with this. Women in these cases appeared to overestimate how liberal their counterparts were. These overestimations seemed to match more closely with their own beliefs, rather than with the actual beliefs of their household male partners. This suggests that household discussions regarding FLFP may be limited, that men hide their views from their partners, or some women are “in denial” about the views of their male counterparts.

**Figure 5. Is it okay for a woman to work?
Interpersonal Expectations and Personal Beliefs**



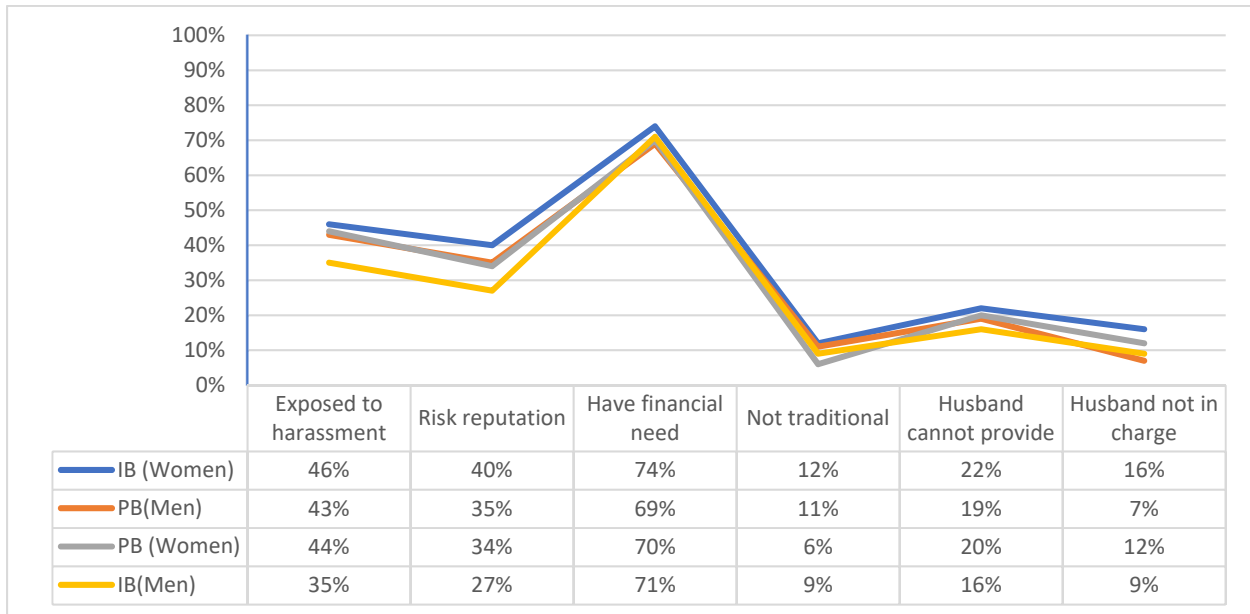
Note: Graph shows proportion okay with women doing this. For these questions, only women were asked about interpersonal expectations of household intimates.

In the domain of safety and reputation, interpersonal expectations, on the part of men and women, were largely aligned with the personal beliefs of their household intimates (see Figure 6). Notably, personal beliefs and interpersonal expectations all suggested that when a woman works, it indicates that a household is in “financial need.” This suggests that finding ways to counteract the perception of financial neediness on the part of households in which a woman works may be another fruitful avenue for social norms messaging.

¹ Of these couples in our sample, 623 were husbands and wives (75%), 145 were brothers and sisters (17%), and 60 were fathers and daughters (7%).

Figure 6. Do working women risk safety and reputation?

Interpersonal Expectations and Personal Beliefs



Note: Graph shows proportion holding these beliefs about working women. For these questions, both men and women were asked about interpersonal expectations.

All things considered, how accurate are interpersonal expectations in the households? We measured accuracy by examining whether a respondent’s expectation of his or her counterpart’s belief matched exactly with the counterpart’s reported personal beliefs. Because the majority of survey respondents were paired from the same household, our data allow us to make direct comparisons of expectations and reality. On average, interpersonal expectations were accurate for around 60% of the questions, for both male and female respondents (see Table 1.1 in Annex 1). In other words, men and women in our sample were more or less aware of how liberal or conservative their counterparts were about 60% of the time.

We developed a standardized aggregate social norms index for each respondent (see Annex 2). After dividing survey respondents into four groups- working women, men with working female counterparts, non-working women, and men with non-working female counterparts – we were able to examine the “conservatism” or “liberality” of each group. The distribution of the standardized aggregated social norms index across these four categories (see Figure 7a. below) shows that working women are generally the most liberal in their beliefs and expectations. They are followed by men with working counterparts, non-working women, and, finally, men with non-working counterparts. Figure 7b shows the difference in the standardized aggregated social norm index distribution between the most liberal and the most conservative demographic groups.

Figure 7a. Standardized Social Norm Index across all four categories

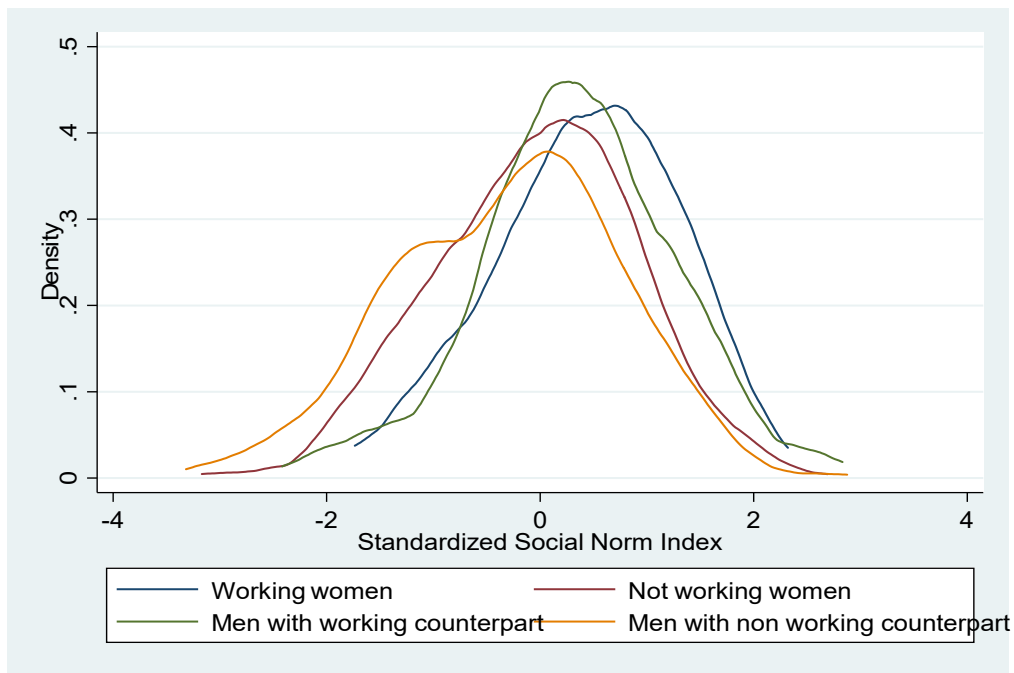
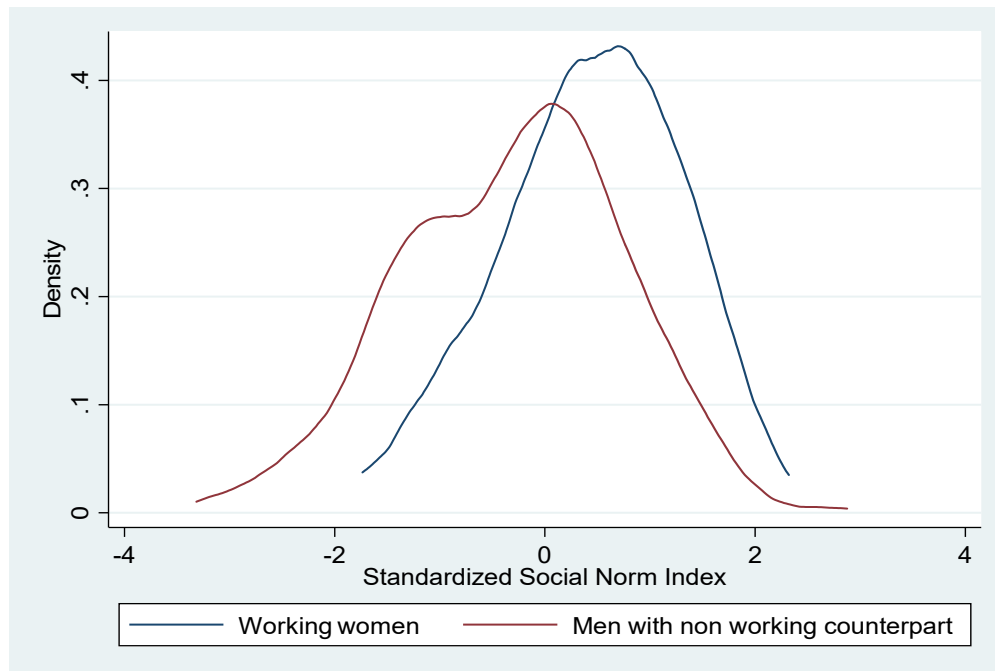


Figure 7b. Standardized Social Norm Index for conservative and most liberal



Men with non-working female counterparts have the most conservative and patriarchal personal beliefs regarding FLFP, while working women are the most liberal across almost all dimensions of normative beliefs asked about in the survey. This is also true for social empirical expectations, with the exception of

expectations around publicness and mixing, where non-working women have the most conservative estimates. There was less variation in the social normative expectations index, across the four groups, than in social empirical expectations and personal beliefs. This is consistent with an account in which social empirical expectations and personal beliefs drive FLFP, and FLFP in turn provides knowledge about how working women behave. Social normative expectations may be harder to discern, and slower to change, absent regular conversations with others regarding FLFP.

We also developed social norms indices with different components for personal beliefs, social normative expectations, social empirical expectations, and interpersonal expectations (for women only). The details of the method are in Annex 3. We use these indices to examine FLFP in Jordan.

Reliability and Validity of Norm Indices

Ex-ante, we classified all items into four distinct components- social empirical expectations, social normative expectations, personal beliefs, and interpersonal expectations, each covering a similar set of thematic questions outlined in Table 1. However, instead of relying on these pre-determined structures, we conducted a series of reliability and validity checks, including calculations of Cronbach's Alpha, inter-item correlation, Kaiser-Meyer-Olkin (KMO) statistic, exploratory factor analysis, and construct and narrative validity, to inform the formation of indices that group similar items together (see Annex 2 for details).

Internal consistency of the ex-ante components was lower than optimal, with Cronbach's Alpha scores ranging from 0.37 to 0.87 for each of the four components (only one above 0.7). However, inter-item correlation within each component shows that most items have significant, non-zero inter-item correlations, especially within each of the four thematic areas outlined in Table 1 (WW, PM, GR, and ST). Separate exploratory factor analysis (EFA) was then conducted with each of the components to understand which items can be grouped together into indices (both component indices and indices of the four thematic areas within each component).² The rotated factor loadings,³ though not very dissimilar, exhibit differences with the items and thematic areas or constructs within each component, as compared to our ex ante constructs.

Based on these findings, each of the four components were revised slightly. For example, the items like financial status of household (ST4), childcare and appropriate age of child (GR3 and GR4) and women's right (WW4) to work were dropped from the interpersonal expectations component, while two status (ST1 and ST4) items as well as appropriateness of women working outside (WW2) were dropped from the personal beliefs component. WW4, GR4, working women in Jordan (WW3) and WW3, ST4 were excluded from the revised social empirical expectations and social normative components respectively. Internal consistency improves significantly with the revised components, with all having Cronbach's alpha values greater than 0.8.

Two sets of indices were created based on the original and the revised components, and a third set of indices were created with thematic constructs within each component. Indices reflected the standardized average of standardized responses across all items within each component and thematic construct. Some responses required reverse-coding. In the final versions, higher scores for each item and index indicated more liberal views.

² The KMO measures for sample adequacy for each of the four components ranged from 0.72 to 0.88 (i.e. Middling or Meritorious), which indicates suitability of data for factor analysis.

³ Two criteria- Kaiser's eigenvalue rule and Cattell's scree plots- were used to determine the number of latent factors to extract from each component. Factor loadings were rotated using oblique rotation for ease of interpretation.

Validity of these indices was further explored through construct and narrative validity. The former examines how closely related indices were to relevant real-life outcomes, such as women's labor force participation, willingness to work, and whether they have looked for work. Results show statistically significant positive relationships between liberal beliefs and expectations and all three outcomes using the revised indices, confirming that working women and their male counterparts were more likely to have more liberal perceptions about women working. Narrative analysis uses the open-ended narrative questions to perform checks using basic text analysis (e.g. if working women and men with working female counterparts were more likely to define their relationships as "not traditional") and calculate aggregated "sentiment" scores for each response,⁴ to understand how they correlated with the indices and a full index combining all four components. The mean aggregated sentiment scores across all the narrative responses are significantly correlated with both the full and the individual component indices, and also mean scores for a subset of narratives are correlated with relevant variables.

Test-Retest Reliability

A short follow-up survey was conducted with a random subset of 1,000 of the original respondents, between 2-3 weeks after the quantitative survey to test the reliability of the social norms measures. Respondents were asked a subset of the social norms and personal belief questions across the different themes as shown in Table 4 below.

The comparison of responses across the two periods assesses the stability of social norms and personal beliefs over time. Table 4 shows the percentage of responses that were the same, the mean of the absolute difference, and the standard deviation of the difference. Items measuring personal beliefs were measured on a shorter Likert scale, and had the highest percentage matches, with an average match rate of 62%. Social empirical and normative questions were typically measured on a scale of 0-10, and had higher discrepancies between the two rounds. For example, the average absolute difference for questions on the longer scale for social empirical questions is 2.46 and for social normative questions is 2.73.

We also report the Pearson's correlation statistic, and Cohen's Kappa score in Table 4. The average Pearson's correlation is 0.25, which indicates weak positive linear relationship between the variables from the two rounds (scores can vary from -1 to 1). The average for Cohen's Kappa, an interrater agreement score, was 0.19. These additional tests confirm that the test retest reliability for these measures is weaker than expected.

We do not believe that measures of test-retest reliability for the stability of social norms measures exists in the literature, and so our study contributes the first benchmark. Given that social norms often influence choice without the need for explicit articulation, and that inquiries into normative and empirical expectations are novel for many respondents, it is not surprising that test-retest reliability for explicit social norms measures are not perfect. Although other benchmarks do not exist, we believe that these social norms measures performed reasonably well in the test-retest exercise. The average difference of 2-3 points on a 10-point scale, in measuring social norms, suggests that a person's social expectations can move from, say, a view in which only 20% of people in Jordan agree with statement X, to one in which about half agree with X, but not to a view in which nearly all Jordanians agree with statement X.

⁴ The Vader lexicon is utilized to calculate these scores, which range from -1 to 1, and indicate the following: >0 is positive sentiment, 0 is neutral and <0 is a negative sentiment.

Table 4: Test-Retest Reliability

Question Type	Variable	Scale	% Same	Mean Diff	SD Diff	Pearsons	Kappa
<i>Social Empirical</i>	<i>work outside home</i>	0-10	17.64	2.54335	3.36	0.21	0.20
<i>Social Empirical</i>	<i>harassment</i>	0-10	27.95	2.38129	3.33	0.30	0.15
<i>Social Empirical</i>	<i>married women work</i>	0-10	14.96	2.44085	3.09	0.21	0.18
<i>Social Empirical</i>	<i>financial need</i>	1-4	48.21	0.75141	1.15	0.37	0.16
<i>Personal Belief</i>	<i>work outside home</i>	1-3	80.80	0.32400	0.76	0.35	0.35
<i>Personal Belief</i>	<i>married women work</i>	1-3	72.30	0.44900	0.88	0.21	0.32
<i>Personal Belief</i>	<i>harassment</i>	1-3	41.83	0.78435	1.09	0.16	0.17
<i>Personal Belief</i>	<i>financial need</i>	1-3	52.40	0.62900	0.97	0.26	0.13
<i>Social Normative</i>	<i>women work</i>	0-10	21.34	2.72445	3.62	0.29	0.18
<i>Social Normative</i>	<i>married women work</i>	0-10	16.58	2.69548	3.52	0.29	0.19
<i>Social Normative</i>	<i>harassment</i>	0-10	16.11	2.77341	3.59	0.07	0.19
<i>Social Normative</i>	<i>financial need</i>	1-3	54.55	0.59259	0.93	0.21	0.06

The correlation between social norms and FLFP

Decisions about women’s labor force participation or willingness to work may be influenced by a combination of social norms in the community, personal beliefs, and expectations of relevant family members within the household. Working women, in turn, may update such beliefs through their experiences in the labor market. We explored how these beliefs vary between working women and non-working women (as well as their male counterparts) across the four themes. For non-working women we further explored the relationship between beliefs and their willingness to work. We wanted to determine which of the components, aggregated across all themes (after factor analysis) mattered most for female labor force participation and willingness to work. We present results for specific sets of beliefs (or themes), which may help identify the most fruitful messaging interventions.

Equation (i) regresses women’s work status (which includes both work status of female respondents and of female counterparts of male respondents), willingness to work (among women not currently working), and whether women have looked for work on indices of social empirical expectations, social normative expectations, personal beliefs, and normative expectations of counterpart (i.e. interpersonal expectations). Equation (ii) separates out the female and male perceptions from the same household.

$$(i) \quad Y_i = \beta_0 + \beta_1 SE_index_i + \beta_2 SN_index_i + \beta_3 PB_index_i + \beta_4 IH_index_i + \beta_i X_i + \epsilon_i$$

$$(ii) \quad Y_{ip} = \beta_0 + \beta_1 SE_index_f_{ip} + \beta_2 SN_index_f_{ip} + \beta_3 PB_index_f_{ip} + \beta_4 IH_index_f_{ip} + \beta_1 SE_index_m_{ip} + \beta_2 SN_index_m_{ip} + \beta_3 PB_index_m_{ip} + \beta_4 IH_index_m_{ip} + \beta_{ip} X_{ip} + \epsilon_{ip}$$

Y_i is the work status for individual i (i.e. whether the female respondent or female counterpart of male respondent is working or wants to work). In equation (ii), the outcome variable is indexed to household p . SE_index_i and $SE_index_f_{ip}$ are indices of relevant social empirical questions for both female and male in equation (i) and the female or male i in household p in equation (ii). Similarly, $SN_indices$ includes social normative variables, $PB_indices$ includes personal beliefs, and $IH_indices$ includes perceptions of male or female counterparts. X_i is a set of control variables which include age, years of education, marital status, number of children, dummy variables indicating whether the respondent is from Amman or Zarqa, and if the male or male counterpart works.

Table 5 below shows results with revised indices after exploratory factor analysis.⁵ We first show the relationship between the revised indices and three outcome variables- women’s labor force participation women’s willingness to work, and whether women have looked for work. The results predict that a one standard deviation increase in the index including social empirical, normative, and interpersonal expectations⁶ is linked to a 1.4 to 1.7 times increase in the odds of women working (Columns 1 and 2). Given that female labor force participation in our sample is about 30%, this represents a fairly large increase. For non-working women, the index does not increase the odds of wanting to work, but does predict a 1.3 times increase in looking for work (Columns 3 and 4). Columns 5 and 6 show how the four components are linked to female work status for both women and men. For the full sample, working status is correlated with social empirical expectations, as well as interpersonal and personal beliefs. Columns 7 and 8 show that for non-working women and the partners of non-working women, personal beliefs appear to be the main driver of wanting to work and having looked for work, rather than social or interpersonal expectations.

Table 5: Individual level regression of revised indices after factor analysis

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample	Full	Full	Non-working	Non-working	Full	Full	Non-working	Non-working
	Sample	Sample	women	women	Sample	Sample	women	women
Dependent Variable	Women’s work status	Women’s work status	Want to work	Looked for work	Women’s work status	Women’s work status	Want to work	Looked for work
Norms index	1.679*** (0.0912)	1.444*** (0.0936)	1.137 (0.0968)	1.276* (0.153)				
SE index					1.315*** (0.0696)	1.267*** (0.0777)	1.005 (0.0988)	0.988 (0.130)
SN index					1.031 (0.0606)	0.972 (0.0705)	0.854 (0.0774)	0.945 (0.111)
PB index					1.459*** (0.0914)	1.373*** (0.0965)	1.533*** (0.159)	1.450* (0.219)
IH index					1.242*** (0.0804)	1.146* (0.0767)	1.094 (0.0988)	1.230 (0.160)
Controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Observations	2,007	2,004	745	746	2,005	2,002	745	746

Note: *** denotes significance at 1%, ** at 5%, and * at 10% level. Standard errors are in parentheses. All columns report results from Logistic regressions. Control variables include age, years of education, marital status, number of children, dummy variables for Amman and Zarqa and if male counterpart works. The standardized indices includes items after factor analysis, as specified in Annex 2, Table 2.8.

⁵ All items as shown in Annex 2 Table 2.8. A standardized index is created with items across all the themes that load onto the first 4 or 5 factors for each of social empirical, social normative, personal beliefs, and interpersonal expectations, but ensuring that the item is represented in each of these components to be consistent with theory.

⁶ Excluding personal belief items. See Annex 2 Tale 2.10 for regressions of an index including all 4 components.

In Table 6, we use equation (ii) to explore if these relationships hold when using the household specification. The patterns are largely similar to individual regressions in Table 5, in spite of the smaller samples.⁷ For the full sample of couples sampled from the same household, social empirical expectations and the male counterpart's personal beliefs are relevant (Column 1). We show results next separating out by married and unmarried women to show that the effects are driven by these sub-samples. For married women, interpersonal expectations of women and men's personal beliefs are relevant, suggesting the important role played by men after marriage (Column 3). However, we do not observe much influence of male counterparts (father or brother) towards unmarried women's work decisions. Instead, social empirical expectations are relevant here (Columns 5). For non-working women, willingness to work is determined by her personal beliefs, as well the male counterpart's beliefs and these effects are driven by married women (Columns 2 and 4). Annex 3 Table 3.1 shows the results separated out by for the four themes (i.e. women working, publicness and mixing, gender roles, and status), but including all items asked for each of these themes. Columns 1-4 show results for married women and Columns 5-8 for unmarried women.⁸

Table 6: Household level regression of revised indices after factor analysis

	(1)	(2)	(3)	(4)	(5)	(6)
Sample	Full Sample	Full Sample	Married Women	Married Women	Unmarried Women	Unmarried Women
Dependent Variable	Women's work status	Want to work	Women's work status	Want to work	Women's work status	Want to work
SE index (female)	1.297** (0.128)	1.044 (0.109)	1.167 (0.140)	1.087 (0.124)	1.536* (0.316)	0.700 (0.223)
SN index (female)	0.817 (0.0906)	0.859 (0.0987)	0.778 (0.102)	0.788 (0.0980)	1.017 (0.229)	3.059* (1.460)
PB index (female)	1.023 (0.116)	1.380** (0.154)	0.948 (0.130)	1.452** (0.176)	1.361 (0.308)	1.460 (0.638)
IH index (female)	1.188 (0.145)	0.909 (0.103)	1.384* (0.206)	0.894 (0.109)	0.772 (0.193)	0.588 (0.304)
SE index (male)	1.151 (0.110)	0.813 (0.0866)	1.213 (0.142)	0.809 (0.0938)	1.044 (0.194)	0.840 (0.278)
SN index (male)	1.076 (0.126)	0.925 (0.111)	1.072 (0.152)	0.865 (0.114)	1.094 (0.257)	1.416 (0.571)

⁷ Because of reducing to the household level, and looking at further sub-samples.

⁸ The results for married women highlight the importance of women's perception of spousal beliefs, and the men's own beliefs, for the different themes as seen before. In particular, the coefficients on interpersonal beliefs for women on acceptability of women's work and gender role items are highest. It is interesting to note that social normative beliefs are also relevant (at the 10% level) when it comes to gender roles – showing that the confluence of societal, personal, and interpersonal normative beliefs may play a role here. For publicness and mixing, men's personal beliefs appear to be relevant towards women's work. Overall the items under the status theme do not appear to be different for working and non-working women and the male counterpart.

Results for unmarried women are different as before. These show that before marriage, societal expectations of women's work and gender roles are different for unmarried woman who are working. For unmarried women, in the context of gender roles, it appears that both social empirical and personal beliefs are significant (at the 10% level). The expectations by a daughter or sister, of their father or brother respectively, do not appear to be relevant and neither are the personal beliefs of the father or brother.

PB index (male)	1.617*** (0.192)	1.535*** (0.193)	1.729*** (0.257)	1.576*** (0.215)	1.465 (0.324)	0.777 (0.345)
IH index (male)	0.893 (0.104)	1.079 (0.122)	0.965 (0.145)	1.115 (0.137)	0.692 (0.150)	1.906 (0.937)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
# of households	825	573	621	468	199	103

Note: *** denotes significance at 1%, ** at 5%, and * at 10% level. Standard errors are in parentheses. All columns report results from Logistic regressions. Control variables include age, years of education, marital status, number of children, dummy variables for Amman and Zarqa and if male counterpart works. The standardized indices include all items asked for each component after factor analysis, as specified in Annex 2, Table 2.8, and separated for men and women from the same household.

The overall finding for married women is that the expectations of their husband’s views, and the husbands’ actual views, seem to be correlated with decision making about women’s work, and especially by perceptions of acceptability of woman’s work and traditional gender roles. Husbands are also uncomfortable with women working with other men (the publicness and mixing index). Interventions targeting married women need to include their husbands as well. It may be the “breadwinner and protector norm,” rather than social empirical expectations of women working per se, that drive the choices of married women.

For unmarried women, these results suggest that social empirical expectations, i.e. expectations of what other women do, are likely to be important perceptions for women’s labor force participation. The influence of male counterparts (fathers and brothers) may not be as binding. This prescribes a different set of policy tools messages when targeting unmarried women.

What explains liberal social norms?

Table 7 presents OLS results on the revised social norms index consisting of social empirical and social normative items after factor analysis.⁹ We regress the index on a number of observables in the survey data as shown in Columns 1-3. Results are presented for all respondents before we show separate results for women and men.

Overall, women have more liberal social norms towards women’s work, although this is driven largely by working women. Being married leads to increasingly negative perceptions towards women working, compared to men. Women in Amman appear to also have more negative perceptions of women working and consequences. Age squared is positively correlated with liberal social norms, and somewhat counter intuitively, having a younger child is also positively correlated for women. As expected, education and household income are positively and significantly correlated for both men and women. These variables, together, explain between 9% to 14% of the variation in social norms.

⁹ The results are robust to using an index of all of the original social empirical and social normative items asked to respondents.

Table 7: Explaining social norms

	(1)	(2)	(3)
	SN index after factor analysis		
	Standardized SN index	Standardized SN index	Standardized SN index
	All	Female	Male
female	0.113** [0.0486]		
age	-0.000618 [0.0105]	-0.0326 [0.0251]	-0.00810 [0.0162]
age_squared	0.000144 [0.000117]	0.000659* [0.000340]	0.000176 [0.000166]
years of education	0.0612*** [0.00690]	0.0766*** [0.00985]	0.0500*** [0.00997]
whether married	-0.171** [0.0783]	-0.178* [0.0940]	-0.129 [0.138]
Has young children	0.0524 [0.0557]	0.167** [0.0762]	-0.00546 [0.0857]
Amman	-0.117*** [0.0437]	-0.231*** [0.0563]	0.00499 [0.0674]
mother worked	0.00757 [0.0672]	0.0631 [0.0789]	-0.111 [0.119]
ever worked	0.0380 [0.0490]	0.0779 [0.0609]	-0.0436 [0.0868]
log household income	0.275*** [0.0393]	0.237*** [0.0509]	0.310*** [0.0610]
Observations	1,919	1,010	909
Adjusted R-squared	0.108	0.143	0.087

Note: *** denotes significance at 1%, ** at 5%, and * at 10% level. Standard errors are in parentheses. All columns report results from OLS regressions. The dependent variable is the standardized index of items asked for the social empirical and normative components after factor analysis, as specified in Annex 2, Table 2.8.

A rational model of female labor force participation?

We present OLS regressions for a framework that incorporates “rational” incentives for women to join the labor market. In Table 8, the left-hand side variable is whether women are working, or want to work, and the explanatory variables are women’s education, household income, the reservation wage, and the revised social empirical, social normative, and personal belief indices after factor analysis.¹⁰ The specification attempts to ascertain whether social norms and personal beliefs contribute to a woman’s decision to work, controlling for “rational” determinants such as earning power (proxied by years of education), the labor-leisure trade-off (proxied by responses to the reservation wage question in the survey), and substitution effects (proxied by household income without women’s earnings). The main finding is that social norms (in particular, social empirical expectations) add explanatory power to a simple “rational” model of FLFP for women currently working. As seen before, the variable “want to work” is better predicted by personal

¹⁰ Results are similar when using the social empirical and normative index with all original items.

beliefs. The results further indicate that, as expected, higher education and lower household income drive more women into the labor force. The positive relation between reservation wages for women and labor force participation is interesting, indicating possibly that women only enter the work force once their expectations of higher wages are met, or that working women have higher reservation wages. Column 2 shows that the reservation wage calculation is different and lower for non-working women who want to work compared to those who do not.¹¹

Table 8: Determinants of whether women work

	(1)	(2)
	Working female	Want to work
Education	0.0404*** [0.00455]	0.0225*** [0.00637]
Log HH Income (minus female income)	-0.144*** [0.0222]	-0.0713** [0.0350]
Reservation wage for female	0.000601*** [0.000123]	-0.000995*** [0.000240]
PB index	0.0317** [0.0148]	0.115*** [0.0193]
SE index	0.0398*** [0.0140]	0.00720 [0.0192]
SN index	-0.00501 [0.0144]	-0.0476** [0.0188]
Observations	949	685
Adjusted R-squared	0.167	0.085

Note: *** denotes significance at 1%, ** at 5%, and * at 10% level. Standard errors are in parentheses. All columns report results from OLS regressions. The PB, SE, and SN indices are the standardized index of items for the personal belief, social empirical, and social normative components after factor analysis, as specified in Annex 2, Table 2.8.

Discussion

Social norms clearly influence FLFP all over the world, but they have not been systematically studied nor measured. We show that social norms can be measured in ways that exhibit reasonable features with respect to internal consistency, external validity, test-retest reliability, and predictive usefulness.

There are multiple social norms influencing FLFP. Social norms are best measured not through yes/no binary questions, but as expectations regarding a diverse set of roles and expectations, including women's role in public life, occupational segregation, women's role in the household, and the male role in the household. Although respondents express support for FLFP in general, support declines when respondents are asked about specific scenarios, such as when a working woman can return from work and the age at which it is appropriate for a mother to leave a child in order to work. For married women, the strongest correlate of FLFP is not the generalized social norm but rather the interpersonal expectations within the household and the husband's personal beliefs. For unmarried women, social empirical expectations appear important. Personal beliefs, and not only social norms strictly understood, affect the extent to which women

¹¹ This is consistent with the results in general, for example as seen in Table 5, personal belief items, but not social expectations, appear to be correlated with willingness to work.

are interested in working and have looked for a job. Overall, the findings of this study suggest that social norms “campaigns” require detailed contextual knowledge and diagnostic surveys to help tailor messages to their appropriate targets.

In Jordan, it would be useful to update social norms on general views towards FLFP. Although 96% of people believe that it is okay for women to work, they expect that only 63% of people will approve. Any campaign to promote or encourage FLFP could use this as a starting point. The views of husbands, and wives’ perceptions of those views, are critical. Given the critical role men play in women’s labor force participation decisions, interventions could usefully focus on and target men as well. Campaigns might be developed to engage men as active partners in women’s empowerment. Finally, aspects of job design such as working hours and childcare, remain important barriers. Promoting flexible work hours or remote-work options might enable more women to work.

REFERENCES

- Alesina, A., Giuliano, P., & Nunn, N. (2013). On the origins of gender roles: Women and the plough. *The Quarterly Journal of Economics*, 128(2), 469-530.
- Bertrand, M., Kamenica, E., & Pan, J. (2015). Gender identity and relative income within households. *The Quarterly Journal of Economics*, 130 (2), 571-614.
- Bicchieri, C. (2006). *The Grammar of Society: The Nature and Dynamics of Social Norms*. New York: Cambridge University Press.
- Bicchieri, C., & Mercier, H. (2014). Norms and Beliefs: How Change Occurs. In *The Complexity of Social Norms* (pp. 37-54). Springer, Cham.
- Bicchieri, C., Lindemans, J. W., & Jiang, T. (2014). A structured approach to a diagnostic of collective practices. *Frontiers in psychology*, 5, 1418.
- Bicchieri, C. (2016). *Norms in the wild: How to diagnose, measure, and change social norms*. Oxford University Press.
- Brennan, G., Eriksson, L., Goodin, R.E., & Southwood, N. (2013). *Explaining Norms*. Oxford, UK: Oxford University Press.
- Bruner, J.S. (1990). *Acts of Meaning*. Cambridge, MA: Harvard University Press.
- Bursztyjn, L., González, A. L., & Yanagizawa-Drott, D. (2018). *Misperceived social norms: Female labor force participation in Saudi Arabia* (No. w24736). National Bureau of Economic Research.
- Chamlou, N., Muzi, S., & Ahmed, H. (2016). The Determinants of Female Labor Force Participation in the Middle East and North Africa Region: The Role of Education and Social Norms in Amman, Cairo, and Sana'a. In *Women, Work and Welfare in the Middle East and North Africa: The Role of Socio-demographics, Entrepreneurship and Public Policies* (pp. 323-350).
- Cialdini, R.B., Kallgren, C.A., & Reno, R.R. (1991). A Focus Theory of Normative Conduct: A Theoretical Refinement and Reevaluation of the Role Of Norms in Human Behavior. *Advances in Experimental Social Psychology*, 24(20), 201-234.
- Cialdini, R.B. & Trost, M.R. (1998). Social Influence: Social Norms, Conformity and Compliance. In Daniel Gilbert, Susan T. Fiske, Gardner Lindzey, eds., *The Handbook of Social Psychology, fourth edition*, pp. 151-192.
- Coleman, J.S. (1994). *Foundations of Social Theory*. Cambridge, MA: Belknap Press.
- Connell, R. (2016). Masculinities in global perspective: hegemony , contestation , and changing structures of power. *Theory and Society*, 303–318
- Crossley, M. L. (2002). *Introducing Narrative Psychology*. University of Huddersfield.
- Dildar, Y. (2015). Patriarchal Norms, Religion, and Female Labor Supply: Evidence from Turkey. *World Development*, 76(2012), 40–61.
- DiMaggio, P. (1997). Culture and Cognition. *Annual Review of Sociology*, 23(1), 263–87.
- Elster, J. (1989). *The Cement of Society: A Study of Social Order*. New York: Cambridge University Press.

- Evans, A. (2016). The Decline of the Male Breadwinner and Persistence of the Female Carer: Exposure , Interests , and Micro–Macro Interactions. *Annals of the American Association of Geographers*, 106(5), 1135-1151.
- Fernandez, R. (2013). Cultural change as learning: The evolution of female labor force participation over a century. *American Economic Review*, 103(1), 472-500.
- Fishbein, M. & Ajzen, I. (2010). *Predicting and Changing Behavior: The Reasoned Action Approach*. New York: Psychology Press.
- Habyarimana, J., & Jack, W. (2009). *Heckle and Chide: Results of a Randomized Road Safety Intervention in Kenya*. Working Paper 169, Center for Global Development, Washington, DC.
- Habyarimana, J., & Jack, W. (2011). Heckle and Chide: Results of a Randomized Road Safety Intervention in Kenya. *Journal of Public Economics*, 95(11), 1438–46.
- Hutto, C.J. & Gilbert, E.E. (2014, June). VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. In *Eighth International Conference on Weblogs and Social Media*.
- Jordan Department of Statistics. (2016) Employment and Unemployment. Retrieved from <https://data.worldbank.org/indicator/SL.TLF.CACT.FE.ZS?locations=ZQ>
- Lapinski, M.K. & Rimal, R.N. (2005). An Explication of Social Norms. *Communication Theory*, 15(2),127-147.
- Mackie, G., Moneti, F., Shakya, H., & Denny, E. (2015). What are social norms? How are they measured. *University of California at San Diego-UNICEF Working Paper, San Diego*.
- Ocheje, P. D. (2018). Norms, law and social change: Nigeria’s anti-corruption struggle, 1999–2017. *Crime, Law and Social Change*, 70(3), 363-381.
- Paluck, E. L. (2009). Reducing Intergroup Prejudice and Conflict Using the Media: A Field Experiment in Rwanda. *Journal of Personality and Social Psychology*, 96(3), 574–587.
- Sarbin, T. R. (1986). The narrative as root metaphor for psychology. In T. R. Sarbin (Ed.), *Narrative psychology: The storied nature of human conduct* (pp. 3-21). New York: Praeger.
- Storey, D. & Schoemaker, J. (2006, June). Communication, Normative Influence and The Sustainability Of Health Behavior Over Time: A Multilevel Analysis Of Contraceptive Use In Indonesia, 1997-2003. In *Annual Meeting of the International Communication Association*.
- World Bank. (2012). *World Development Report 2012: Gender Equality and Development*. World Bank Publications.
- World Bank. (2014). *World development report 2015: Mind, Society, and Behavior*. World Bank Publications.
- The World Bank. (2018) Data Bank. Retrieved from <https://data.worldbank.org/indicator/SL.TLF.CACT.FE.ZS?locations=ZQ>.

Annex 1: Interpersonal Expectations

Table 1.1: Expectation vs. Reality of Counterpart's Beliefs Within the Household

Variables	Women's Beliefs	Men's Expectation	Men's Accuracy	Men's Beliefs	Women's Expectations	Women's Accuracy
Working Women						
Women Work *						
Yes	2.05		N/A	4.71	8.34	85.49
Sometimes Y/N, No	97.95			95.29	91.66	
Work Outside Home*						
Yes	6.52		N/A	23.67	18.72	69.32
Sometimes Y/N, No	93.48			76.33	81.28	
Women Right to Work						
Yes	29.95	21.77	63.12	16.57	19.37	67.03
Sometimes Y/N, No	70.05	78.23		83.43	80.63	
Publicness & Mixing						
Work with Men*						
Yes	42.03		N/A	59.78	34.95	45.34
Sometimes Y/N, No	57.97			40.22	65.05	
Harassment						
Yes	39.47	32.97	57.35	41.50	42.44	54.56
Sometimes Y/N, No	60.53	67.03		58.50	57.56	
Reputations at Risk						
Yes	31.64	26.51	56.17	34.50	37.73	51.39
Sometimes Y/N, No	68.36	73.49		65.50	62.27	
Gender Roles						
Married Women Work*						
Yes	10.02		N/A	30.19	17.87	62.32
Sometimes Y/N, No	89.98			69.81	82.13	
Return After 5pm*						
Yes	59.78		N/A	70.41	42.15	45.65
Sometimes Y/N, No	40.22			29.59	57.85	
Leave Child w/ Relatives						
Yes	62.80		N/A	41.55	48.43	47.09
Sometimes Y/N, No	37.20			58.45	51.57	
Status						
Do Not Follow Traditions						
Yes	5.92	8.83	81.74	11.11	11.76	77.46
Sometimes Y/N, No	94.08	91.17		88.89	88.24	
Husband Cannot Provide						
Yes	19.81	16.06	60.87	18.26	19.83	62.1
Sometimes Y/N, No	80.19	83.94		81.74	80.17	
Husband Not in Charge						
Yes	12.32	8.82	72.95	7.49	14.77	73.73
Sometimes Y/N, No	87.68	91.18		92.51	85.23	
Financial Need						

Yes	69.81	71.01	60.26	68.36	73.52	59.73
Sometimes Y/N, No	30.19	28.99		31.64	26.48	

*These questions were asked in the reverse order, i.e. they asked if counterpart would think or speak badly of women who do work/work outside/return after 5, etc. The answers to the personal beliefs subsequently had to be reversed to match the responses to the interpersonal expectation questions.

Annex 2: Reliability and Validity of Norm Indices

Social Norm Indices

i. Ex-ante Classification of Constructs

The following factor structure was conceived ex-ante, based on findings from the qualitative data, to classify all items under each component (social empirical expectations, social normative expectations, personal beliefs, and intra-household normative expectations) into four themes: women working; publicness and mixing; gender roles; and status.

Table 2.1. Ex-ante Thematic Areas

Working Women (WW)	Gender Roles (GR)
<i>WW1. Women working from home</i>	<i>GR1. Married women working</i>
<i>WW2. Women working outside the home</i>	<i>GR2. Married women returning home after 5 PM</i>
<i>WW3. Women working in Jordan (only SE)</i>	<i>GR3. Leaving child <5 with relative during work</i>
<i>WW4. Women working if husband not comfortable (only PB/SN/IB)</i>	<i>GR4. Appropriate age to leave child and go to work</i>
<i>WW5. Necessary for both husband and wife to work to live comfortably (only PB)</i>	
<i>WW6. Women can work if other women are too (only IB)</i>	
Publicness and Mixing (PM)	Status (ST)
<i>PM1. Working in mixed-gender environment</i>	<i>ST1. Family does not abide by traditions</i>
<i>PM2. Exposure to harassment</i>	<i>ST2. Husband cannot provide for her</i>
<i>PM3. Reputation of working women</i>	<i>ST3. Husband not in charge of family</i>
	<i>ST4. Family suffering from financial need</i>

ii. Cronbach's Alpha

Cronbach's alpha measures internal consistency or reliability, and indicates how closely a set of items within a group or scale are related. Internal consistency appears to be relatively low for the components. Average alphas for each of the four ex-ante components specified above range from 0.37 to 0.87, with social normative expectations having the highest level of internal consistency, and social empirical expectations having the lowest (see Table 2.2 below).

Table 2.2. Cronbach's Alpha

Component	Score
<i>Social empirical expectations</i>	0.3761
<i>Personal beliefs</i>	0.6562
<i>Social normative expectations</i>	0.8709
<i>Interpersonal expectations</i>	0.6200

iii. Inter-item Correlation and Kaiser-Meyer-Olkin (KMO) Statistic

Exploratory factor analysis is valuable when items (i.e. variables) in question are sufficiently correlated. Inter-item correlation within each component shows that most items have significant, non-zero inter-item correlations, especially within each of the four thematic areas.

Sufficiency of item correlation can also be tested using a measure of sampling adequacy, called the KMO statistic. The KMO statistic provides an indication of how well other items in the module explain the correlation between items, and is a reliable indication of whether factor analysis is appropriate given the data structure.

The overall KMO measures for sample adequacy of the four components range from 0.72 to 0.88 (i.e. Middling or Meritorious), indicating that the data are suitable for factor analysis. All individual items, except one,¹² within each module have KMO values above 0.50 (which is the threshold for acceptable). However, since this item has relatively high KMO value (between 0.73 to 0.89) in the remaining three components, and was also deemed as relevant based on findings from the focus group discussions, the item was not dropped for the factor analysis.

Table 2.3. Kaiser-Meyer-Olkin Statistic- All

Social empirical		Personal beliefs		Social normative		Interpersonal	
Variable	KMO	Variable	KMO	Variable	KMO	Variable	KMO
WW1	0.719195	WW1	0.685505	WW1	0.808241	WW1	0.714581
WW2	0.729659	WW2	0.789447	WW2	0.814945	WW2	0.839129
WW3	0.763245	WW4	0.802254	WW4	0.750436	WW4	0.500814
PM1	0.909413	WW5	0.805853	PM1	0.952637	WW6	0.513222
PM2	0.81014	PM1	0.805087	PM2	0.880868	PM1	0.896412
PM3	0.839929	PM2	0.617262	PM3	0.87897	PM2	0.722185
GR1	0.878625	PM3	0.633532	GR1	0.94414	PM3	0.750731
GR2	0.87772	GR1	0.784099	GR2	0.927059	GR1	0.878984
GR3	0.892537	GR2	0.769167	GR3	0.93156	GR2	0.849761
GR4	0.638272	GR3	0.734793	GR4	0.794925	GR3	0.461736
ST1	0.812601	GR4	0.699507	ST1	0.90063	GR4	0.791671
ST2	0.785831	ST1	0.800644	ST2	0.790097	ST1	0.874209
ST3	0.788275	ST2	0.667085	ST3	0.796487	ST2	0.737587
ST4	0.810876	ST3	0.65239	ST4	0.850898	ST3	0.722219

¹²KMO value of 0.46 for interpersonal expectations: GR3.

		ST4	0.650007			ST4	0.737823
Overall	0.807913	Overall	0.720667	Overall	0.877947	Overall	0.783398

iv. Exploratory Factor Analysis

Exploratory factor analysis (EFA) can tell us how many factors, or latent constructs, can be extracted from a given set of data. This can, subsequently, inform the formation of indices that group similar items together instead of relying on pre-determined scales. In this case, separate factor analysis was conducted for each of the components (which included 14-15 items each) to determine the factor structure, and subsequently, the items to include in each component.

The first step of factor analysis is to determine how many factors to extract from a given set of items. Two criteria- Kaiser’s eigenvalue rule and Cattell’s scree plots- are used to determine the number of latent factors to extract. The results, which are identical across both criteria, suggest two factors each to be retained from the social empirical expectations and personal beliefs modules, and one factor each from the social normative expectations and interpersonal expectations components. For ease of interpretation, the factor loadings are rotated using oblique rotation. Factor loadings change substantially after rotation, with personal beliefs and social normative expectations loading on to five factors (with Eigenvalues greater than 1) and social empirical expectations and interpersonal expectations loading on to four factors.

The rotated factor loadings, though not completely dissimilar, do appear to have some key differences with the ex-ante thematic constructs. While certain items within ex-ante thematic questions are similarly grouped in the rotated factor loadings, some factors appear to have a mix of various items, making them difficult to interpret. Interpreting and identifying factors become particularly challenging when we take into account the variation across the four modules. Tables 2.4-2.7 show the factor loadings for each component.

Table 2.4. Rotated factor loadings for personal belief items

Variable	Factor1	Factor2	Factor3	Factor4	Factor5	Factor6
work outside home (WW2)	0.6417					
married women work (GR1)	0.6368					
appropriate age of child (GR4)		0.5982				
leave children with relative (GR3)		0.5899				
Harassment (PM2)			0.7376			
reputations at risk (PM3)			0.7244			
work with men (PM1)				0.5061		
return after 5pm (GR2)				0.4637		
women right to work (WW4)				0.4116		
husband cannot provide (ST2)					0.615	
husband not in charge (ST3)					0.5895	
do not follow traditions (ST1)						
women work (WW1)						
need work to live comfortably (WW5)						
financial need (ST4)						

(blanks represent abs(loading)<0.3)

Table 2.5. Rotated factor loadings for social empirical expectation items

Variable	Factor1	Factor2	Factor3	Factor4	Factor5	Factor6
reputation at risk (PM3)	0.8733					
Harassment (PM2)	0.8361					
married women work (GR1)	0.4614					
return after 5pm (GR2)	0.3595				0.3339	
work with men (PM1)						
women work (WW1)		0.9211				
work outside home (WW2)		0.8821				
leave children with relative (GR3)			0.4865			
appropriate age of child (G)R4)			0.4246			
husband cannot provide (ST2)				0.7271		
husband not in charge (ST3)				0.6892		
do not follow traditions (ST1)				0.4897		
financial need (ST4)					0.3009	
women's right to work (WW4)						

(blanks represent abs(loading)<0.3)

Table 2.6. Rotated factor loadings for social normative expectation items

Variable	Factor1	Factor2	Factor3	Factor4	Factor5	Factor6
reputation at risk (PM3)	0.8733					
harassment (PM2)	0.8361					
married women work (GR1)	0.4614					
return after 5pm (GR2)	0.3595				0.3339	
work with men (PM1)						
women work (WW1)		0.9211				
work outside home (WW2)		0.8821				
leave children with relative (GR3)			0.4865			
appropriate age of child (GR4)			0.4246			
husband cannot provide (ST2)				0.7271		
husband not in charge (ST3)				0.6892		
do not follow traditions (ST1)				0.4897		
financial need (ST4)					0.3009	
women's right to work (WW4)						

(blanks represent abs(loading)<0.3)

Table 2.7. Rotated factor loadings for interpersonal expectation items

Variable	Factor1	Factor2	Factor3	Factor4	Factor5	Factor6
reputation at risk (PM3)	0.8733					
harassment (PM2)	0.8361					
married women work (GR1)	0.4614					
return after 5pm (GR2)	0.3595				0.3339	
work with men (PM1)						
women work (WW1)		0.9211				

work outside home (WW2)	0.8821	
leave children with relative (GR3)		0.4865
appropriate age of child (GR4)	0.4246	
husband cannot provide (ST2)		0.7271
husband not in charge (ST3)		0.6892
do not follow traditions (ST1)		0.4897
financial need (ST4)		0.3009
women's right to work (WW4)		

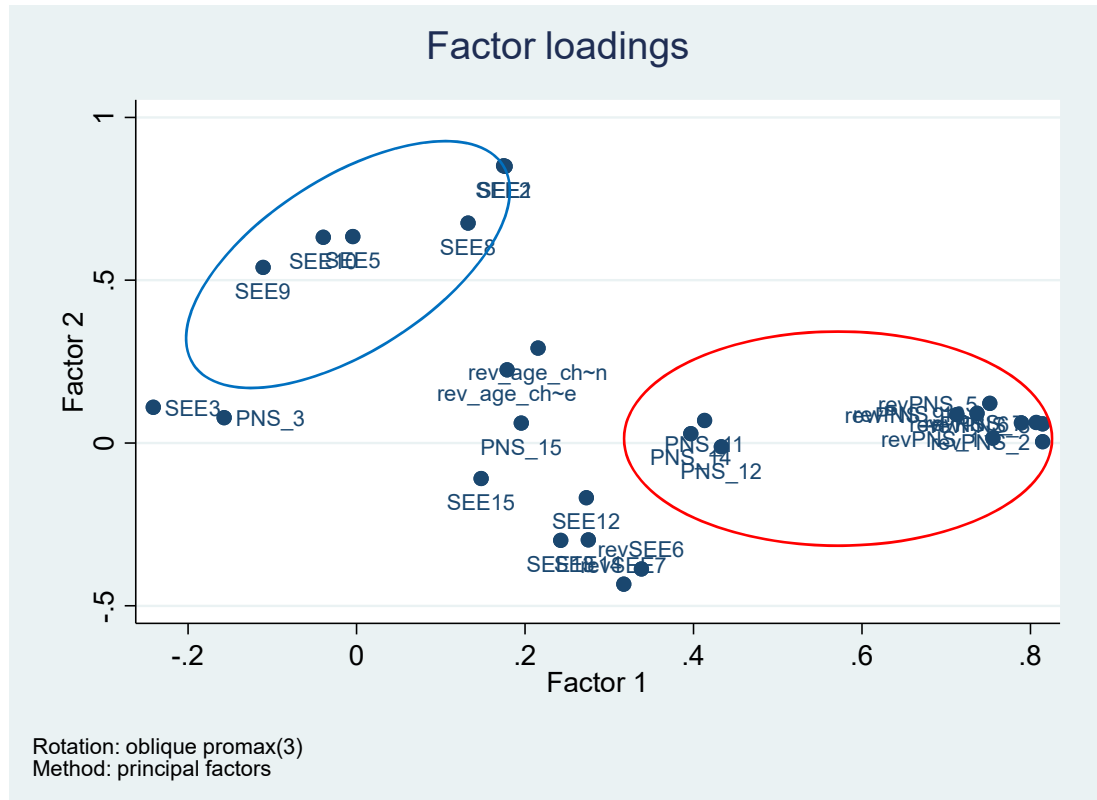
(blanks represent abs(loading)<0.3)

PM2 and PM3, generally uniquely, load on a single factor, confirming that grouping these items together is supported by the data. PM1, however, does not seem to have any singular pattern in its factor loadings across modules, and is therefore more difficult to group under any specific latent factor. Similarly, ST2 and ST3 also load on to the same factor, though ST1 and ST4 appear to be more scattered. WW1 and WW2, not surprisingly, tend to load on to the same factor, often with GR1, but the remaining items in the WW construct are either scattered or do not load on to any factor. GR items have the least consistency in their factor loadings. This does raise some concerns about whether certain groups of items belong together, as expected ex-ante, and suggests the need for rearranging the items across constructs.

Are Social Empirical Expectations Different from Social Normative Expectations?

Because we expect social empirical and social normative expectations to be driven by two distinct latent factors or constructs, we conducted separate exploratory factor analysis with all variables in the social empirical and social normative expectation components. Rotated factor loadings (forcing a 2 factor solution) confirm that the latent factors driving social empirical and social normative expectations are, in fact, distinct.

Figure 2.1. Graph of Factor Loadings for Social Empirical and Social Normative Expectations



All social empirical expectation items (SEE in Figure 2.1) have high factor loadings for Factor 2, while all social normative expectation items (PNS in Figure 2.1) have high factor loadings for to Factor 1. When restricted to factor loadings higher than 0.3, all items in Factor 1 belong to the Social Normative Expectations module. Similarly, all items with factor loadings higher than 0.3 in Factor 2 are Social Empirical Expectations variables. All “Status” items for social empirical expectations, however, have low factor loadings, i.e. less than 0.3.

v. **Revised Components**

Based on the exploratory factor analysis above, the following are revised components. Each revised component includes all items from the respective component that load on to the extracted factors for that components (i.e. factors with Eigenvalues>1). Items that did not load on to any of the extracted factors were dropped from the component. One exception to this was the social normative expectation component, in which Factor 5 (which has eigenvalue greater than 1) was dropped because both items in the factor had relatively low factor loadings (0.30 and 0.33).

Table 2.8. Revised Components

Personal Beliefs	Social Empirical Exp.	Social Normative Exp.	Interpersonal Exp.
<i>WW1. Working from home</i>	<i>WW1. Working from home</i>	<i>WW1. Working from home</i>	<i>WW1. Working from home</i>
<i>WW4. Right to work</i>	<i>WW2. Working outside</i>	<i>WW2. Working outside</i>	<i>WW2. Working outside</i>
<i>PM1. Mixed gender workplace</i>	<i>PM1. Mixed gender workplace</i>	<i>PM1. Mixed gender workplace</i>	<i>PM1. Mixed gender workplace</i>
<i>PM2. Exposure to harassment</i>	<i>PM2. Exposure to harassment</i>	<i>PM2. Exposure to harassment</i>	<i>PM2. Exposure to harassment</i>
<i>PM3. Risk reputation</i>	<i>PM3. Risk reputation</i>	<i>PM3. Risk reputation</i>	<i>PM3. Risk reputation</i>

<i>GR1. Married women working</i>	<i>GR1. Married women working</i>	<i>GR1. Married women working</i>	<i>GR1. Married women working</i>
<i>GR2. Home after 5 PM</i>	<i>GR2. Home after 5 PM</i>	<i>GR2. Home after 5 PM</i>	<i>GR2. Home after 5 PM</i>
<i>GR3. Leave child w/ relative</i>	<i>GR3. Leave child w/ relative</i>	<i>GR3. Leave child w/ relative</i>	<i>ST1. Family not traditional</i>
<i>GR4. Appropriate age of child</i>	<i>ST1. Family not traditional</i>	<i>GR4. Appropriate age of child</i>	<i>ST2. Husband cannot provide</i>
<i>ST2. Husband cannot provide</i>	<i>ST2. Husband cannot provide</i>	<i>ST1. Family not traditional</i>	<i>ST3. Husband not in charge</i>
<i>ST3. Husband not in charge</i>	<i>ST3. Husband not in charge</i>	<i>ST2. Husband cannot provide</i>	<i>ST4. Family has financial need</i>
	<i>ST4. Family has financial need</i>	<i>ST3. Husband not in charge</i>	

Internal consistency (as measured by Cronbach’s alpha value) improves significantly with the revised components, with three components having alpha values greater than 0.8 (see Column 1 in Table 2.9 below). The alpha value for the personal beliefs component, however, still remains below 0.7. Analysis of how individual items within the module affect the Cronbach’s alpha value reveals that the internal consistency of this module can be further improved if GR4 is excluded. To improve internal consistency (as seen in Column 2 below), GR4 is therefore dropped from the personal beliefs component.

Table 2.9. Cronbach’s Alpha Revised Constructs

Construct	Alpha (1)	Alpha (2)
<i>Social empirical expectations</i>	0.8323	0.8323
<i>Personal beliefs</i>	0.6359	0.7226
<i>Social normative expectations</i>	0.8678	0.8678
<i>Interpersonal expectations</i>	0.8169	0.8169

These revised components allow for items within each of the four components to vary substantially. Because the thematic constructs also vary as a result, analysis using the revised components does not include the thematic breakdown used in the original components.

vi. Creating Indices

Two sets of indices are created based on the original and the revised components, and a third set of indices are created with thematic constructs within each component. Indices reflect the standardized average of standardized responses across all items within each component and thematic construct. Some responses needed to be reverse-coded to make sure all items within each index are measuring the same thing, i.e. either conservative or liberal views.¹³

Construct Validity

Another measure of validity examines how well the indices (or construct indices) are related to relevant real-life outcomes. In this case, we analyze the relationship between the original and revised indices and three outcome variables- women’s labor force participation, women’s willingness to work, and whether women have looked for work. Results (see Table 2.10 below) of the logistic regression show a statistically significant positive relationship between women’s labor force decisions and liberal beliefs and expectations This validates that working women and their male counterparts are more likely to have more liberal perceptions about women working. The results predict that a one standard deviation increase in the index is linked to a 1.6 times increase in the odds of women working (Column1). Given that

¹³ Items were reverse coded prior to conducting the validity and reliability analysis.

female labor force participation in our sample is about 30%, this represents a fairly large increase. Similarly, for non-working women, a one standard deviation increase in the overall index predicts a 1.3 times increase in the odds of wanting to work, and 1.4 times increase in looking for work (Columns 4 and 5).

Table 2.10. Construct Validity

	All original items			Revised items		
	(1) All respondents	(2) Non-working Female	(3) Non-working Female	(4) All respondents	(5) Non-working Female	(6) Non-working Female
	Working female	Want to work	Look for work	Working female	Want to work	Look for work
Standardized Index of items and components (SE, SN, PB, IH)	1.590*** (0.107)	1.317** (0.118)	1.416** (0.181)	1.564*** (0.106)	1.288** (0.115)	1.444** (0.185)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2003	744	745	2,003	744	745

Note: *** denotes significance at 1%, ** at 5%, and * at 10% level. Standard errors are in parentheses. Standard errors are clustered at the household level for regressions that include 2 members from the same household (Column 1 and 4). All columns report results from Logistic regressions. Control variables include age, years of education, marital status, number of children, dummy variables for Amman and Zarqa and if male counterpart works. The standardized index includes all items asked for all themes and components as specified in Table 1 for Columns 1-3 and all items after factor analysis for each of the components as specified in Annex Table 2.8 in Columns 4-5.

Validity Using Narratives

Respondents were also asked a set of randomly chosen open-ended narrative questions, for example, in what ways was their relationship with their spouse typical or atypical, based on their work status and gender. We perform some basic text analysis including some simple checks, for example, that working women and men with working female counterparts are more likely to define their relationships as “not traditional” in response to the above question.¹⁴

Next, we calculate aggregated sentiment scores for each set of narrative responses provided by the respondent. The distribution of these scores is shown in Figure 2.2 below. We utilize the Vader lexicon, a part of the Natural Language Toolkit in Python to calculate these scores (Hutto & Gilbert, 2014). Most words that commonly occur in social media and other text have been tagged by several raters where scores range from -1 to 1, and indicate the following: >0 is positive sentiment, 0 is neutral and <0 is a negative sentiment. We appreciate that sentiments attached to different words may be different in the Jordanian context, and ideally a training set is developed by the researchers assigning relevant sentiment

¹⁴The question we asked specifically was: “Do you consider your relationship with your husband to be typical or atypical? Tell us about ways in which it is typical or atypical.”

scores for the context and questions. However, we only utilize the Vader lexicon in this case to efficiently develop another test of validity for the social norms measurement.

Figure 2.2. Distribution and Correlation of sentiment analysis scores

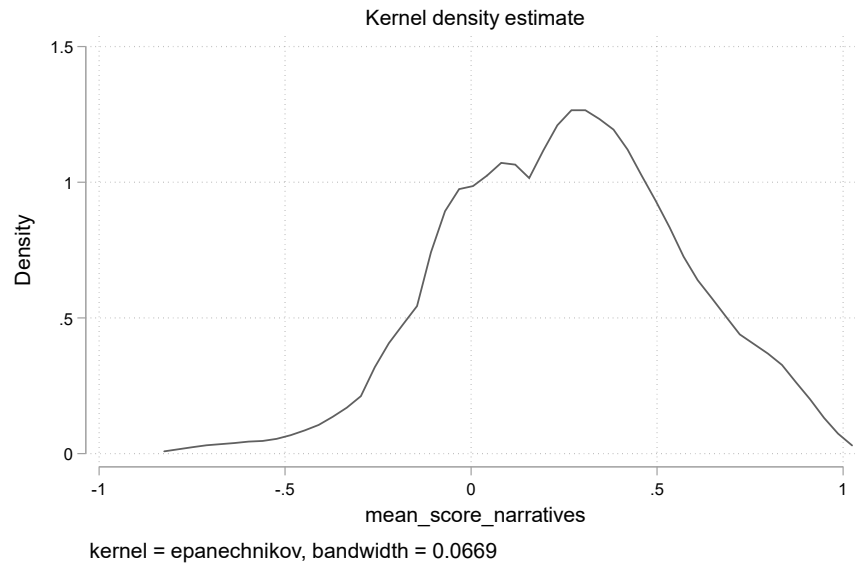


Table 2.11 shows the correlations of the sentiment scores with the full index, social norms index, and other outcomes. The mean aggregated sentiment scores across all the narrative responses are significantly correlated with both the full and the social norms indices (Columns 1 and 2). Next, we show that mean scores for a subset of narratives are correlated with relevant variables. Column 3 shows the full index is correlated with positive sentiments in response to the typical/atypical question above. Column 4 shows that positive sentiments in response to relevant questions about wanting to work (discussed work with spouse, the value of education for women in Jordan¹⁵) asked to non-working women and men with a non-working female counterpart, was correlated with whether the female member wanted to work.

Table 2.11. Sentiment Analysis

	(1)	(2)	(3)	(4)
Dependent Variable	All indices	SN indices	All indices	Want to work
Mean Sentiment Score	0.228** [0.0941]	0.164** [0.0740]		
Mean Sentiment Score (Typical/Atypical)			0.134* [0.0748]	
Mean Sentiment Score (Discussed work, value of education)				0.117*** [0.0393]
Observations	1,235	1,235	1,228	844
Adjusted R-squared	0.004	0.003	0.002	0.009

¹⁵ The specific questions were as follows: “If you ever had a discussion with your husband/father/brother about whether to work, what did you discuss about salary, household work, childcare, transport, etc.?” and “Women in Jordan are very well educated. Why do you think women decide to pursue further education? How do you think it helps them in the future?”

Note: *** denotes significance at 1%, ** at 5%, and * at 10% level. Standard errors are in parentheses. All columns report results from OLS regressions. The dependent variable is the standardized index for each component including all original items (Column 1 and 3), and Social Empirical and Normative components (Column 2).

Annex 3: Regression tables on female work status and willingness to work

Table 3.1: Household level regression of original themed indices

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Married Women				Unmarried women			
Theme	Women's work	Publicness/Mixing	Gender Roles	Status	Women's work	Publicness/Mixing	Gender Roles	Status
Dependent Variable	Working female	Working female	Working female	Working female	Working female	Working female	Working female	Working female
SE index (female)	1.139 (0.128)	1.016 (0.121)	1.258 (0.158)	0.905 (0.109)	1.495* (0.285)	0.977 (0.186)	1.638* (0.334)	0.994 (0.204)
SN index (female)	0.910 (0.108)	0.938 (0.122)	0.726* (0.0964)	0.814 (0.103)	0.835 (0.159)	0.830 (0.191)	1.532 (0.368)	0.898 (0.172)
PB index (female)	1.149 (0.139)	0.866 (0.108)	1.076 (0.152)	1.015 (0.126)	1.134 (0.197)	1.186 (0.237)	1.599* (0.355)	1.022 (0.205)
IH index (female)	1.341* (0.161)	1.236 (0.158)	1.721*** (0.261)		0.956 (0.180)	1.045 (0.241)	0.693 (0.159)	
SE index (male)	1.074 (0.130)	1.059 (0.125)	1.266 (0.159)	1.331* (0.188)	0.934 (0.167)	1.158 (0.233)	1.005 (0.201)	1.006 (0.226)
SN index (male)	1.065 (0.124)	0.827 (0.121)	1.345* (0.193)	1.203 (0.148)	0.865 (0.162)	1.309 (0.325)	0.975 (0.229)	1.198 (0.255)
PB index (male)	1.475** (0.192)	1.534** (0.211)	1.422* (0.208)	0.985 (0.121)	1.234 (0.228)	1.127 (0.231)	1.225 (0.280)	1.233 (0.270)
IH index (male)	0.787* (0.0940)	1.127 (0.151)	0.892 (0.117)	1.123 (0.154)	0.932 (0.147)	0.868 (0.172)	0.807 (0.186)	0.742 (0.163)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
# of households	618	618	554	622	198	198	188	200

Note: *** denotes significance at 1%, ** at 5%, and * at 10% level. Standard errors are in parentheses. All columns report results from Logistic regressions. Control variables include age, years of education, marital status, number of children, dummy variables for Amman and Zarqa and if male counterpart works. The standardized indices include all items asked as specified in Table 1, and separated for men and women from the same household.