## Article Title

Information Provision to Parents on Women Working Outside the Home in Rural Pakistan: Baseline and Endline Dataset


## Authors

Momoe Makino

## Affiliations

Institute of Developing Economies, JETRO, Nihon Boeki Shinko Kiko Asia Keizai Kenkyujo

## Corresponding author(s)

Momoe Makino (momoe_makino@ide.go.jp)



## Abstract

The dataset contains raw panel data of 400 households obtained by structured survey before and after the randomized controlled trial. The survey was conducted in rural Punjab in Pakistan in 2016–2017 to examine whether the provision of information about young women's income earning opportunities to their parents affects their attitude toward women working outside the home, in particular, women working in factories (Makino 2021). The questionnaire was uniquely designed to collect information requisite to capturing change in attitudes toward women working outside the home. To conduct the survey, the area within commuting distance of export-oriented garment factories that actively hire women was first identified in the districts of Faisalabad, Hafizabad, Nankana Sahib, and Chiniot. Forty villages were randomly selected in a commutable rural area. The village questionnaire consists of questions about basic village information, such as demography, land holdings, caste structure, income earning opportunities, infrastructure, access to export-oriented garment factories, and educational opportunities. Ten households were randomly selected from the list of all eligible households in each of the 40 villages. The household questionnaire consists of a household roster concerning age, enrollment status, work status, and educational attainment level of all household members; typical socioeconomic questions; and unique questions specific to this survey including those asking for personal and general opinions pertaining to women working outside the home. The questionnaire also contains questions about gender relations, such as who has the authority to make decisions, degree of mobility, and level of son preference. The same questionnaire was utilized for the baseline survey conducted before the intervention and the endline survey conducted one year after the intervention. The attrition rate is 0%, which is achieved by the careful survey design, a relatively small sample size, i.e., 400, and a relatively low household migration rate in the survey area.

## Keywords

Household survey, Women working outside the home, Parental decision, Social norm, Pakistan

## Specifications Table

| | |
|---|---|
| **Subject** | Economics and Econometrics |
| **Specific subject area** | Development Economics, Household Economics, Population Economics |
| **Type of data** | STATA data table |
| **How data were acquired** | Survey<br>See survey ReadMe |
| **Data format** | Raw panel data |
| **Parameters for data collection** | The interviewer started interview only after obtaining the interviewee's signed agreement to participate in the interview. Prior to the interview, the interviewees were notified of the purpose of this research and the interviewees' rights. Because the survey outcome was the interviewee's attitude, the interviewers did not reveal the exact objective of the intervention, i.e., whether the provision of information to young women's parents affects their attitude toward women working outside the home, to avoid experimenter-expectancy effect. However, the interviewers revealed the broader purpose of the research reflected in the questions included in baseline/endline survey questionnaires by reading aloud as "to academically understand people's perception about working opportunities and to improve the economic condition of households in rural Punjab, Pakistan." The interviewees' rights included confidentiality, and asking questions/withdrawing the interview at any time. The interviewers assured confidentiality by reading aloud "all information gathered from you will be treated as strictly confidential and will be used only in an aggregate form. No individual name will be used in any document prepared based on this survey." |
| **Description of data collection** | The data were collected by face-to-face interviews using the structured village and household questionnaires. The field survey was conducted by Faizan Data Collection and Research Centre (FDCRC) in Pakistan. Under the supervision of chief enumerator, Tariq Munir, the survey team was formed, which consisted of two female and two male interviewers who completed minimum higher secondary education (12 years) and can speak and read both Punjabi and Urdu. The interviewers were practically trained for 1 week in the field to be able to conduct interviews by themselves. The village survey was conducted in the 40 randomly chosen villages in the survey area, i.e., within commuting distance of export- |

| | |
|---|---|
| | oriented garment factories that actively hire women in the districts of Faisalabad, Hafizabad, Nankana Sahib, and Chiniot. The respondents of the village questionnaire were village chiefs or equivalent informants in the villages. In collaboration with the village survey respondents and other village informants, all households that were eligible for inclusion in the survey were listed in each village. The eligible households were defined as follows: either landless or own no more than 3 acres of land, presence of both father and mother, and family comprising at least one unmarried daughter aged between 15 and 30 years. Ten households were randomly selected from the list of all eligible households in each of the 40 villages. Among these 40 villages, 20 were randomly chosen with mothers as respondents (and intervention targets for the treatment group) and the other half had fathers as respondents (and intervention targets for the treatment group). Female interviewers conducted interviews in the villages where mothers were assigned as respondents. Male interviewers conducted interviews in the villages where fathers were assigned as respondents. |
| **Data source location** | Districts: Faisalabad, Hafizabad, Nankana Sahib, and Chiniot<br>Country: Pakistan<br>10km radius of 31.895782, 73.276062 |
| **Data accessibility** | Repository name: Economic Development and Cultural Change Dataverse<br>Direct URL to data: https://doi.org/10.7910/DVN/JNR8WM |
| **Related research article** | Makino, M. Labor market information and parental attitudes toward women working outside the home: Experimental evidence from rural Pakistan. *Economic Development and Cultural Change,* forthcoming. |

## Value of the Data

- The dataset is the panel data of 400 households in rural Punjab, Pakistan. The questions include not only standard socioeconomic characteristics of the households, but unique questions such as those about personal and general opinions pertaining to women working outside the home, gender relations, daughters' marriage, time inconsistency, risk aversion, and social network.
- The data can be of interest to the researchers who are interested in cultural and social topics in the context of South Asia, such as low women's paid-work participation rate, lack of women's agency, son preference, marriage practices, and social pressure, and how these variables are related to standard socioeconomic characteristics.
- The effects of the intervention, i.e., information provision to parents, on various outcomes such as gender relations and daughters' marriage can be examined. A future field experiment as an

extension can be developed to examine the effectiveness of providing information and intervening parents in a different or more general contexts.

- The dataset contributes to our extensive and intensive understanding about the importance of cultural and social contexts in achieving specific development goals such as enhancing women's paid work participation, and girls' education and empowerment.

## Data Description

- wwoh_pak.dta
  - The raw data are constructed as panel data. Individual information is coded so that no information that would identify an interviewee is included.
  - Data dictionary and pref_wwohPakcode.do explain how the variables are generated and constructed.
  - HHID: unique household identifier from 1–400
  - No: unique household-round identifier, = HHID for the baseline survey, and = HHID + 400 for the endline survey
  - Year2: an indicator variable taking the value 1 for data obtained in the endline survey
  - Treated: an indicator variable taking the value 1 for the treated household
  - Vpercent: village-level randomized % of treated households in the village, taking the value either 0, 0.4, 0.6 or 1
  - v_treatment_type: village-level randomized treatment arms taking the value either 1= group treated, 2= individual treated, or 3= controlled
  - sex: an village-level randomized indicator which was treated and/or interviewed, father =1, or mother =2

- roster_baseline.dta (long)
  - Household roster data at baseline
  - HHID: unique household identifier from 1–400
  - Treated: an indicator variable taking the value 1 for the treated household

## Experimental Design, Materials and Methods

***Survey area selection***

A randomized controlled trial (RCT) was conducted in rural Punjab in Pakistan in 2016–2017. The objective was to examine whether the provision of information on working opportunities in factories that are suitable for young women to their parents affects parental attitude toward women working outside the home (WWOH) (Makino 2021). Because working opportunities for women, especially in factories, are generally limited in Pakistan, it was necessary to implement the RCT in an area where it was at least

realistic for young women to work in factories. Such a rural area within commuting distance of export-oriented garment factories that actively hire women was first identified in the districts of Faisalabad, Hafizabad, Nankana Sahib, and Chiniot. The census data of these districts were used to randomly select 40 survey villages and 10 substitutes in a commutable rural area. Substitute villages were ordered by replacement priority number.

***Village and household surveys***

Both village-level and household-level data were collected by face-to-face interviews using the structured village and household questionnaires. The field survey was conducted by Faizan Data Collection and Research Centre (FDCRC) in Pakistan. Under the supervision of chief enumerator, Tariq Munir, the survey team was formed, which consisted of two female and two male interviewers who completed minimum higher secondary education (12 years) and can speak and read both Punjabi and Urdu. The interviewers were practically trained for one week in the field to be able to conduct interviews by themselves.

In June 2016, the village survey was conducted using a structured village questionnaire. The respondents of the village questionnaire were village chiefs or equivalent informants in the villages. The village questionnaire consists of questions about basic village information, such as demography, land holdings, caste structure, income earning opportunities, infrastructure, access to export-oriented garment factories, and educational opportunities. To elicit cooperation of the village chiefs and influential persons, information was also collected on the village's cooperative atmosphere. Two initially selected villages were replaced with substitute villages in a pre-determined order of priority due to concerns about cooperation. Note that the replacement took place before random assignment of households to treatment or control groups.

In collaboration with the village survey respondents and other village informants, all households in each village were profiled. Households that were eligible for inclusion in the survey were defined as follows: either landless or own no more than 3 acres of land, presence of both father and mother, and family comprising at least one unmarried daughter aged between 15 and 30 years.

Ten households were randomly picked from the list of all eligible households in each of the 40 villages. Among these 40 villages, 20 were randomly chosen with mothers as respondents and intervention targets and the other half had fathers as respondents and intervention targets. The questionnaire was uniquely designed to collect information requisite to capturing change in attitudes toward WWOH. It consists of a household roster concerning age, enrollment status, work status, and educational attainment level of all household members; typical socioeconomic questions; and unique questions specific to this survey including those asking for personal and general opinions pertaining to WWOH. The same questionnaire was utilized for the baseline survey conducted before the intervention and the endline survey conducted 1 year after the intervention. The attrition rate is 0%, which is achieved by the careful survey design, a relatively small sample size, i.e., 400, and a relatively low household migration rate in the survey area. The data were constructed as panel data and the variable *Year2* was generated as an indicator variable taking the value 1 for data obtained in the endline survey.

***Experiment***

The RCT involved providing parents with information on working conditions and environments in the export-oriented garment factories. The intervention was conducted soon after the baseline household

survey was completed in the village. In households randomly allocated into the treatment group, either the father or mother received a lecture on the working conditions and environments of these factories. Whether the father or the mother received the lecture depended on the gender of the respondent who was randomly selected at the village level. The same information was not provided in a control group.

There were three randomizations at the village level: the first was whether the father or mother was the survey respondent and intervention target, the second was the number of treated households in the village, and the third concerned whether the treatment was delivered in a group session or individually. The variable *sex* was generated to capture the first-level randomization, i.e., whether the father or mother was the survey respondent and intervention target. The variable *Vpercent* was generated to capture the second-level randomization, i.e., taking the value 0, 0.4, 0.6, or 1 in the villages with the number of treated households being zero, four, six, or ten, respectively. The variable v_treatment_type was generated to capture the third-level randomization, i.e., taking the value 1 for the group treated, 2 for the individual treated, and 3 for the controlled. There were 164 households in the control group and 236 in the treatment group.

Below is the list of code files, and data dictionary to analyze the raw data wwoh_pak.dta (roster_baseline.dta (long) can be made from wwoh_pak.dta). Village and household questionnaires are also provided.

- wwoh_pak_code.do
  - Codes to replicate Tables 1 to 4 in the manuscript and Tables A1 to A6 in appendix in Makino (forthcoming) on wwoh_pak.dta and roster_baseline.dta files.

- prep_wwohPakcode.do
  - Codes to clean and prepare the variables in the wwoh_pak.dta file.

- DataDictionary_PakWWOH.pdf
  - Data dictionary about how the variables were constructed and correspond to the questions in both household and village questionnaire

- Vquestionnaire.pdf
  - Baseline/endline village questionnaire to be answered by village chiefs or equivalent informants in the village

- Hquestionnaire.pdf
  - Baseline/endline household questionnaire to be answered either fathers or mothers, which were randomly determined at the village level

- ReadMe_PakWWOH.pdf
  - Summary of all files, i.e., raw data, code files, data dictionary, and questionnaires

**Ethics Statement**

Informed consent

Prior to the interview, the interviewees were notified of the broad purpose of this survey and the interviewees' rights, i.e., confidentiality, and asking questions/withdrawing the interview at any time. The interviewer started an interview only after obtaining the interviewee's signed agreement to participate in the interview.

Institutional Review Board (IRB)
IRB Name: The Research Ethics Committee, Institute of Developing Economies, Japan External Trade Organization (IDE-JETRO)
IRB Approval Number: RPA170704018

## Acknowledgments

I thank Kyoji Fukao, Miki Hamada, Lukas Hensel, Yuki Higuchi, Yuya Kudo, Keijiro Otsuka, Tomohiro Machikita, Andreas Menzel, Hector Niehues-Jeuffroy, Kazushi Takahashi, Dainn Wie, Junichi Yamasaki, and seminar/conference participants at the Sustainability and Development Conference, IZA/World Bank/NJD Conference, Conference on Applied Development Economics, JADE/GRIPS Development Workshop, and IDE for their valuable comments and suggestions. I extend special thanks to Tariq Munir and his assistants at the Faizan Data Collection and Research Centre for their sincere efforts in conducting the field survey in Punjab, Pakistan. Financial support from the Japan Society for the Promotion of Science (Grant-in-Aid for Scientific Research, KAKENHI, 15K17065) is gratefully acknowledged. Any errors, omissions, or misrepresentations are my own.

## Declaration of Competing Interest

The collection of data was funded by Japan Society for the Promotion of Science (Grant-in-Aid for Scientific Research, Kakenhi-15K17065).

The author declares that she has no known competing financial interests or personal relationships which have, or could be perceived to have, influenced the work reported in this article.

## References

Makino, Momoe. 2014. "Pakistan: Challenges for Women's Labor Force Participation." In *The Garment Industry in Low-Income Countries: An Entry Point of Industrialization*, edited by Takahiro Fukunishi and Tatsufumi Yamagata, 132–76. New York: Palgrave Macmillan. https://doi.org/10.1057/9781137383181.

———. 2021. "Labor Market Information and Parental Attitudes toward Women Working Outside the Home: Experimental Evidence from Rural Pakistan." 826. IDE Discussion Paper.