

Pre-analysis plan

Kailash Rajah and Ishaana Talesara

October 2023

1 Sample

Our sample consists of married women who recently graduated from our partner organization’s vocational training program. We present women with 30 hypothetical job offers and ask whether they are interested in each job one-by-one. We measure their job choices at three decision stages:

1. Initial choices: their choices after hearing each job ad for the first time
2. Shared choices: their choices after being told they will have a few days to reflect on their decisions
3. Final choices: their choices after having time to consider the jobs and potentially discuss them with their family

We will run the analysis below on the initial, shared, and final choices. The final choices will be the main sample in Sections 2-6. For our main analysis, we will drop participants who initially reject all jobs. We will also drop participants for whom half of husband’s wage is more than the highest real job wage, or four times the husband’s wage is less than the lowest real job wage (at a given vocational training center).

2 Graphical presentation of data

We will plot locally smoothed interest in jobs against the relative wage with and without job and individual fixed effects.

3 Regression Discontinuity

For our main specification, we will test for a discontinuity at the husband’s income in the style of the evidence found in Bertrand, Kamenica, and Pan (2015). We will estimate:

$$Interested_{ij} = \alpha \mathbb{1}[RelativeWage_{ij} > 1] + f(RelativeWage_{ij}) + \gamma_i + \delta_j + g(d_{ij}) + \varepsilon_{ij}$$

Where i indexes participants, j indexes jobs and d_{ij} is a measure of the distance between individual i 's home and job j . The function $f(RelativeWage_{ij})$ denotes a flexible local linear function of the running variable. In our main specification, we will estimate this using the optimal bandwidth. The coefficient of interest is α .

For robustness, we will present results for:

- Different bandwidths, including 0.1 and 0.2 times the husband's income
- Alternative definitions of the cutoff – $\mathbb{1}[RelativeWage_{ij} \geq 1]$ including a relative wage of 1 on the right, and $\mathbb{1}[RelativeWage_{ij} > 1]$ excluding a relative wage of 1 from analysis.
- Regressions with and without controls

3.1 Analysis Plan Amendment: December 2023

Upon further consideration of the econometrics, we have updated our main specification. When we offer women jobs, we randomly draw the wage to be between 0.5 and 1.5 times the husband's wage. We then round the job offer to the nearest 100 rupees to ensure the offers appear natural. This rounding leads to the bunching of observations from lower income husbands at the cutoff, since they are more likely to be given an offer that exactly equals their husband's income. This may result in the expectations of potential outcomes being discontinuous at the cutoff. Our new specification ensures that the expectations of potential outcomes are continuous at the cut-off, thus yielding an unbiased estimate of the discontinuity.

Instead of running our regression discontinuity on the relative wage of the job offer (i.e. as a multiple of the husband's income), we will instead run it on the absolute wage (i.e. as the level difference from the husband's income). We will then weight each data point using inverse probability weights to correct for different treatment probabilities at different husband incomes.

While data collection has commenced, our updated specification is not based on preliminary data analysis, given only 30% of the target sample has been collected. Instead, we are using this specification given it is the econometrically preferred one.

Our new specification is:

$$Interested_{ij} = \alpha \mathbb{1}[RelativeWageLevel_{ij} > 0] + f(RelativeWageLevel_{ij}) + \gamma_i + \delta_j + g(d_{ij}) + \varepsilon_{ij}$$

Where *RelativeWageLevel* is the wage of the job offer relative to the husband in levels. We will weight the regression using inverse probability weights.

For robustness, we will present results for:

- Different bandwidths of [-1000, 1000], [-2000, 2000], and [-3000, 3000] around the husband's income. We must drop participants whose randomized wage offers do not span the full support of the bandwidth. Given this,

we will not use the optimal bandwidth for our main specification. We will instead use $[-2000, 2000]$ around the husband’s income as the bandwidth in our main specification.

- Alternative definitions of the cutoff – $\mathbb{1}[RelativeWageLevel_{ij} \geq 0]$ including an equal husband and wife wage on the right, and $\mathbb{1}[RelativeWageLevel_{ij} > 0]$ excluding wage offers equal to the husband’s wage from analysis.
- Regressions with and without controls

4 Slope of the labor supply curve

Next, we will estimate the slope of the woman’s labor supply near the husband’s income. If there is any noise, for example, from participants’ beliefs about their current or future incomes or mismeasurement, one may observe a slope change around the husband’s income, rather than a discontinuity. As such, we will run the following specification:

$$Interested_{ij} = \alpha + \beta RelativeWage_{ij} + \gamma_i + \delta_j + g(d_{ij}) + \varepsilon_{ij}$$

where we restrict the window of interest to be a small band around the husband’s income. We will estimate this equation for relative wage in $[0.9, 1.1]$. We will also test:

- Alternative bands around the husband income including $[0.8, 1.2]$ and $[0.95, 1.05]$ given the bandwidth for which breadwinning norms are binding may be highly sample specific.
- Regressions with and without controls

4.1 Analysis Plan Amendment: December 2023

Previously, we did not specify how *RelativeWage* was defined. We clarify this in this amendment. To deal with the same econometric issue (described in section 3.1) arising from rounded incomes bunching at the ends of our pre-specified bounds $[0.9, 1.1]$, $[0.8, 1.2]$ and $[0.95, 1.05]$, we will run the analysis where *RelativeWage* is defined as the relative wage before any rounding has been performed (i.e. on the randomized relative wage). Using the randomized wage ensures that treatment probability does not vary with husband’s income.

5 Difference in means

We will estimate the difference in means on either side of the husband’s income.

$$Interested_{ij} = \alpha + \beta WageAboveHusband_{ij} + \gamma_i + \delta_j + g(d_{ij}) + \varepsilon_{ij}$$

We will estimate this equation for a relative wage in $[0.9, 1.1]$.

We will also test results for:

- Alternative bands around the husband income including [0.8,1.2] and [0.95, 1.05]
- Regressions with and without controls

5.1 Analysis Plan Amendment: December 2023

We will run this same specification in levels at bandwidths of [-500,500], [-1000, 1000], [-2000, 2000], and [-3000, 3000] around the husband's income. Previously, we did not specify how *WageAboveHusband* was defined. Here, we specify that *WageAboveHusband* = 1 when the offered (rounded) wage exceeds the husband's wage.

6 Change in slope of labor supply curve

As exploratory analysis, we will test for a change in slope around the husband's income.

First we will test for a change in the slope of labor supply curve on either side of husband's income by estimating the following equation. β_2 tests for a change in the slope.

$$\begin{aligned} Interested_{ij} = & \alpha + \beta_0 WageAboveHusband_{ij} + \beta_1 RelativeWage_{ij} \\ & + \beta_2 WageAboveHusband_{ij} \times RelativeWage_{ij} + \gamma_i + \delta_j + g(d_{ij}) + \varepsilon_{ij} \end{aligned}$$

In a second test, we will test whether the slope of the labor supply curve around the husband's income is less positive than other regions of the supply curve. We will run the specification below where β_2 tests for a change in the slope:

$$\begin{aligned} Interested_{ij} = & \alpha + \beta_0 WageNearHusband_{ij} + \beta_1 RelativeWage_{ij} \\ & + \beta_2 WageNearHusband_{ij} \times RelativeWage_{ij} + \gamma_i + \delta_j + g(d_{ij}) + \varepsilon_{ij} \end{aligned}$$

These regression estimates will primarily be useful alongside the graphical analysis: a change in slope may simply reflect the curvature of labor supply and not a breadwinning norm.

6.1 Analysis Plan Amendment: December 2023

For the same reasons as specified in section 4.1, we will again run the analysis using the relative wage before any rounding has been performed (i.e. on the randomized relative wage). Previously, we did not specify how *RelativeWage* and *WageAboveHusband* were defined. We clarify this in this amendment. *RelativeWage* is defined as the relative wage before any rounding has been performed (i.e. on the randomized relative wage). *WageAboveHusband* = 1 if the unrounded, randomized relative wage is > 1 .

7 Second treatment: husband involvement

We also test how the husband's involvement in decision making affects the presence of the breadwinning norm. We do this by pooling initial and final decisions and running the same tests as listed above, fully interacted with with $Post_{ij}$, an indicator equal to 1 for participants' final choices. We will do this on the full sample of participants, on the subsample randomly assigned to have the husband receive a list of job choices, and the control subsample. We will use a question about who was involved in the decision to understand whether the second treatment affected husband or family involvement. Women who share a phone number with their husband or whose husbands decline to provide their numbers are not included in this experiment.

As exploratory analysis, we will also compare shared choices with initial and final choices.

We will also explore the effect of husband involvement on labor supply levels by comparing initial to final labor supply levels (for all jobs and certain types of jobs) across all relative incomes.

8 Heterogeneity analysis

We will analyze heterogeneity by conservative gender attitudes and behavior. In particular:

- Beliefs about whether it would cause problems if a woman earns more than her husband (standardized)
- Beliefs about whether it would cause problems if they earned more than their husband (standardized)
- Second-order beliefs about whether it would cause problems if a woman earns more than her husband (standardized)
- Practice of purdah (standardized)
- An index of the above – taking the average and standardizing
- Whether the husband was in the room or conference called in

We will also analyze heterogeneity by whether or not the husband's income varies month to month.

9 Standard Errors

We will estimate heteroskedasticity robust standard errors.