Gender-Specific Informal Sector Heterogeneity

Hyung Joon Chung, Jin Ho Kim, Benjamin Williams

July 10, 2022

1. Introduction

In developing countries, the majority of workers work in the informal sector, where economic agents are not registered and do not pay taxes. Despite the well-documented negative effect of informal sector economic activities on long-term economic growth, the informal sector is a significant source of income and employment to majority population in developing countries. As critical as informal sector economic activities are to the majority of the population in developing countries, a substantial amount of literature has been devoted to determining the nature of informal sector employment and what motivates informal sector economic activity. The literature has established that informal sector employment can be roughly classified into two tiers: those who are rationed out of formal sector jobs and those who voluntarily work in the informal sector (Radchenko 2014, Ulyssea 2018). Depending on the development stage of the nation, it is found that one class of informality has greater proportion than the other: upper-middle-income economy like Indonesia tend to have larger proportion of voluntary informal sector workers, whereas lower-middle-income economy like Egypt tend to have larger proportion of rationed-out informal sector workers. Though the research on informality have matured to identify these different layers of informal sector economy and underlying driving mechanism, there is still a dearth of research on the nature of informality experienced by different gender.

It is critical to investigate the nature of informality as experienced by different genders. Depending on the gender-specific characteristics of informality, labour regulations such as the minimum wage may exacerbate gender inequality and womens welfare (Kim and Williams 2021). A thorough understanding of labour market policies affecting household welfare, then, requires an understanding of the gender-specific features of informal sector labour market. According to previous research, the informal sector is particularly well-suited to female workers' needs because it offers greater flexibility in terms of working hours, work tempo, and job proximity to one's home (Chen, 2001; Manning, 1998; World Bank, 2009, p.xiv). Indeed, studies of developing countries highlight the overwhelming presence of women in the informal sector, particularly as self-employed workers in household enterprises (for an overview of the literature, see Chant & Pedwell, 2008). However, these research has rarely conducted a quantitative study that compares the welfare status of female workers in the formal and informal sectors whose individual characteristics are similar.

We attempt to close this research gap by analysing the Indonesian labour market from 2000 to 2014. Indonesia is a natural case study for this research question. As is the case in other developing countries, the majority of Indonesian workers are employed in informal sector economic activities, with women accounting for a disproportionate share of informal sector workers. Like other developing nations, Indonesia's formal sector labour market is characterised by employment without contracts or non-compliance on minimum wage payment, an excessive reliance on short-term contracts, and monopsonistic behaviour (Allen, 2016; Kim and Samaniego, 2022). Due to widespread non-compliance with existing labour market regulations and exploitative behaviour on the part of employers, it is debatable whether working in the formal sector necessarily improves welfare across all workers. Regarding female workers' plight, it is indicated that self-employed female workers often face financial constraints and difficulty in loaning compare to male self-employed (Babbitt et al. (2015)). When it comes to the formal sector, female workers earn significantly less than male workers. All of these characteristics are typical of the labour market in developing countries, which makes the Indonesian case study a significant case study.

This paper analyzes the nature and driving mechanism of informality faced by different gender, especially focuses on the remification of informality on women's welfare. We use the 3 rounds of Indonesian Family Life Survey (IFLS) data and assess treatment effect heterogeneity through marginal treatment effects (MTE) framework (Björklund and Moffitt (1987); Heckman, Urzua, and Vytlacil (1999, 2005, 2007)). There are several advantages to use MTE framework to infer the nature of informal sector employment. To begin, MTE framework estimates the heterogeneity in the treatment effect across population with individual-specific observed and unobserved characteristics. Second, the method allows to relate the distribution of outcome gap (between treatment and control status) across individual's innate preference over treatment status with allocation over the treatment status. In other words, the framework offers an inference on behavioral response of individuals' choice on the treatment status based on their expected outcome, and thus offers an useful framework to make an inference on individuals' choice on treatment status. Especially, typical way to present MTE curve across individuals' unobserved preferences allows us to infer the driving mechanism for in/formal sector employment.

We consider employment status—formal, informal, or non-employed worker—as the treatment. Dependent variables are income and women's involvement in household decision-making. The latter is a dummy variable that takes one if married women are involved with household decision-making; if not, the variable takes zero. Since individuals endogenously choose employment status, we construct instrument variables related to employment status but do not directly affect dependent variables. One of the instruments is a dummy variable that takes the value one if there is a household member working in the formal sector and zero otherwise. Also, we use district-level average ratio formal sector employment for the other instrument. These variables are strong deter-

minants of employment status. However, one could be concerned that household makes decisions together on who work formal sector and who makes how much remunerations. Likewise, district-level formality may directly affect dependent variables, not through employment status. To reduce potential endogeneity, we control for different mechanisms through which instruments may affect dependent variables: the number of household members, respondent's working hours per week and working weeks per year, and income brought by other household members. Likewise, we control province-level GDP and unemployment rate to control the spillover effect of formal sector employment.

With a female and male sample, we find substantial heterogeneity in returns to formal sector employment. However, the characteristic of heterogeneity varies by gender. Whereas we observe clear positive sorting on gain in the male sample, implying rational choice of in/formal sector employment, we do not see this pattern in the female sample; at best, our analysis suggests that regardless of individuals' innate preference for in/formal sector employment, females earn more in the formal sector. These findings indicate that employment in the informal sector is not always the last resort for male workers; rather, our findings indicate that, on average, male workers rationally choose between the formal and informal sectors. By contrast, our results with female sample do not suggest positive sorting on the gain into formal sector employment; rather, our findings suggest a segmented labour market for women, and that female workers' lives improve when they work in the formal sector. These intuitions are further tested with household bargaining indicators with married female sample. Our dependent variable on household decision-making reflects the deprivation of women's decision-making authority in the home. Our findings from the married female sample corroborate our findings: married women working in the formal sector have greater decision-making authority, regardless of their innate preference for in/formal sector work. These findings show the critical role of formal sector employment in enhancing female workers' welfare.

The contribution of this paper to the existing literature is two-fold. Firstly, our paper contributes to a body of knowledge about labour market informality by providing empirical evidence about the gender-specific nature of informal sector employment. Previous literature on the topic focused on the overall view of informal sector economy and the mechanism for informal sector employment. Traditional view of informal employment argues that informal sector is the last resort for workers who are rationed out in the formal sector labor market (Harris & Todaro, 1970; Fields, 1990,2005a). The alternative view is that informal employment is a result of workers' voluntary selection and thus conforms to the competitive market framework (Maloney, 1999, 2004). These works are more concerned with the population as a whole than with the informal sector's gender-specific nature. Recent policy reports address the informality's gender dimension (OECD 2019, ILO 2018). According to reports, women are more frequently found working in the most vulnerable segments of the informal economy, such as unpaid family workers, while a disproportionately

higher proportion of male workers are self-employed, which is considered the competitive segment of the informal economy. These and other studies compare the gender dimension of informality using income comparisons. While these remuneration comparisons provide important information about women's welfare status, these analysis cannot fully reflect female workers' welfare status. Female employees, for example, may opt for unpaid family work while remaining involved in major household decisions. Our findings address this knowledge gap by examining the informality experienced by different genders and also conducting a welfare comparison across different groups of people (paid wage work vs self-employment vs unpaid family work) in terms of income and household decision-making index.

Second, our paper contributes to the growing literature that estimates marginal treatment effects in different contexts. The research on benefits to formal sector employment typically focuses on individuals' observable characteristics that enables workers to sort into formal sector employment and gains from it. Though recent literatures on the topic emphasized the heterogeneity of formal and informal sector in developing countries, majority literature focuses on the role of policy, education, and other observed characteristics rather than studying the marginal treatment effects of formal sector working, as we do. The only two recent recent studies we know of that use an MTE framework to estimate heterogeneity in returns to working in the formal sector with respect to unobserved characteristics are Radchenko (2014,2017). These paper evaluate an overall population in several developing countries in the region of Africa and their focus is on overall populations. However, these studies investigate for the behavioral mechanism of informal sector employment for overall populations, not specifically focusing on the gender-specific informal sector employment. Also, the outcome is limited to earned income, so it cannot infer further on the implication of employment status on women's welfare. Our research not only investigates the mechanism of informality by different gender, but further study the welfare of female workers through investigating on household decision making. To the best of our knowledge, this is the first attempt to study heterogeneous treatment effect with the framework of MTE on household decision making. Previous literature on women's intrahousehold bargaining power using decision-making indicators only identified the important observables that affect household decision making (Friedberg and Webb(2006), Antman (2014), Majlesi (2016) in the average treatment effect.

The paper is organized as follows. Section 2 outlines the empirical framework and method to infer the mechanism to work in the specific sector. Section 3 introduces data and the main features of the formal sector and informal sector employment by gender. Section 4 presents our empirical results. Section 5 concludes.

2. Estimating Marginal Returns to Treatment

A. Model

The framework to describe the MTE method in the literature is a generalized Roy model based on the potential outcomes model and latent index model, as in Heckman and Vytlacil (2005,2007). We regard the formal sector workers as the treated population and workers in the informal sector as the untreated population. Earnings from formal and informal sector jobs represent the outcomes of being treated and untreated, respectively. Let W_1 be the potential earning of an individual in the formal sector (treated group; $D = 1$) and W_0 denote the potential earning for the individual in the informal sector (untreated group; $D = 0$). The observed income W can be linked to the potential incomes through the switching regression model

$$W = (1 - D)W_0 + DW_1 \tag{1}$$

We model the potential incomes W_j as a function of observed individual characteristics, X, and unobserved individual characteristics, U_j.

$$lnW_j = X\beta_j + U_j \tag{2}$$

where U_j is normalized to $E(U_1|X = x) = E(U_0|X = x) = 0$ for $j = 0, 1$, and for all x in the support of X. By applying the switching regression model, we can express the above equation as

$$lnW = X\beta_0 + \triangle * D + U_0 \tag{3}$$

where $\triangle = (\beta_1 - \beta_0)X + (U_1 - U_0)$, an individual treatment effect. We can use the following latent index model to describe selection into treatment group:

$$I_D = Z\gamma - V_D \tag{4}$$

That is, individual's net benefit of working in the formal sector, I_D, depends on observed variables Z and an unobserved component V_D, where $Z = (X, \widetilde{Z})$ implies that Z includes all the same covariates X in the outcome equation, and also contains instruments $\widetilde{Z}$, excluded from the outcome equation but enters the selection equation. Note that finding instruments is a necessary condition to identify MTE as the treatment heterogeneity comes from the correlation between anticipated gains from treatment choice and workers' prferences to formal sector employment. More formally, the necessary conditions on instruments for the identification of MTE are the following: (i) the instruments, $\widetilde{Z}$, should be independent on U_0, U_1, V_D conditional on $X\left(\widetilde{Z} \perp (U_0, U_1, V)|X\right)$. (ii) instruments should affect to the decision on formal sector employment$\left(cov(\widetilde{Z}, D) \neq 0\right)$. In other words, the instrument variables should affect to the income only through its effect on treatment choice, and the instruments should be as good as randomly assigned given other control variables

X. Let us interpret V_D as "unobserved resistance" or "distaste" to the formal sector work as V_D enters into the equation with a negative sign.[1] Individuals select the formal sector work if the benefit, which is explained by the observables Z, is greater than the unobserved resistance. We can re-write this selection equation as follows:

$$Z\gamma - V_D \geq 0 \Longleftrightarrow Z\gamma \geq V_D \Longleftrightarrow \Phi(Z\gamma) \geq \Phi(V_D) \tag{5}$$

where Φ denotes the cumulative distribution function of V_D. The term $\Phi(Z\gamma)$, also denoted as $\Phi(Z\gamma) = P(Z)$, is the propensity score, and $\Phi(V_D)$ is normalized as a uniform distribution on the unit interval. Let us define the quantiles of the distribution of unobserved resistance to formal sector employment as $U_D(\equiv \Phi(V_D))$. Thus, individuals whose propensity to work in the formal sector based on observables are higher than the unobserved distaste for formal sector job sort into the formal sector ($D = 1$ if $P(Z) > \Phi(V_D)$). MTE is then defined as the expected treatment effect among individuals whose observable and unobservable controls are $X = x$ and $U_D = u_D$, respectively.

$$MTE(X = x, U_D = u_D) = E(lnW_1 - lnW_0|X = x, U_D = u_D) \tag{6}$$

As U_D is proxied by the propensity score when it is estimated, we can express MTE as $MTE(X = x, U_D = p) = E(lnW_1 - lnW_0|X = x, U_D = p)$. Then the definition of MTE implies the average treatment effect of individuals whose probability of sorting into the formal sector based on observables is p and who are indifferent between participation and non-participation into the formal sector.

Following, Heckman, Urzua, and Vytlacil (2006), we estimate MTE with a semi-parametric approach. The approach assumes (i) additive separability between an observed and an unobserved component in the expected potential outcomes conditional on U_D, and (ii) conditional independence of instruments $Z \perp (U_0, U_1, V)|X$. Under these two assumptions, the marginal treatment effect can be decomposed into an observed and unobserved component in the additively separable way.

$MTE(X = x, U_D = u_D) = E(lnW_1 - lnW_0|X = x, U_D = u_D)$
$\qquad\qquad = x(\beta_1 - \beta_0) + E(U_1 - U_0|U_D = u_D)$

One can describe this equation from the switching regression model, controlling for the unobserved gains from formal sector employment non-parametrically. From $lnW = X\beta_0 + (X(\beta_1 - \beta_0) + (U_1 - U_0))*D + U_0$, unobserved component in the earning equation and its relation with treatment choice equation can be non-parametrically modelled with polynomial of propensity score so that the equation becomes

$$E(lnW|X = x, P(z) = p) = x\beta_0 + x(\beta_1 - \beta_0)p + K(p) \tag{7}$$

where $K(p)$ is a polynomial of the propensity score. We take a derivative with respect to p, which

---

[1]The distaste on formal sector employment, or preference for informal sector work can be anything that are not captured in our controlled observables. For instance, this can be having a boss, or losing independence, or preferences for entrepreneurial activity.

then generates $MTE(X = x, U_D = p)$ (Carneiro et al. (2011))

$$\frac{\partial E(lnW|X = x, P(Z) = p)}{\partial p} \big|_{p=u_D} = x(\beta_1 - \beta_0) + \frac{\partial K(p)}{\partial p} = MTE(X = x, U_D = p) \qquad (8)$$

From this equation, we note that the observed controls, $x(\beta_1 - \beta_0)$, determines the intercept of MTE, and the slope of MTE trajectory are driven by the polynomial terms of the propensity score. As the trajectory of the MTE curve is driven by the relationship between U_D and $U_1 - U_0$, the shape of the MTE curve illustrates the sorting behaviors of individuals, which then become an inference for the characteristics of the informal sector labor market.

B. Model Implication on the Characteristics of the Informal Sector

It is a standard practice to put the trajectory of the MTE along with the resistance to the treatment (U_D). In this section, we discuss how the shape of the MTE curve relates to the inference on the characteristics of informal sector economiy activities. This section follows the discussion of Radchenko (2014, 2017), and we focus on the three different types of MTE curve.

(1) MTE decreasing with U_D.
A negative sloping MTE curve along U_D means that a group of workers with the least resistance toward formal sector work earn the most by moving into the formal sector. We can observe this pattern of the graph when the unobserved resistance that discourages individuals from sorting into formal sector work is negatively related to the person's relative efficiency in the formal sector. That is, a person who has high prowess in the wage-earning environment and thus has a low resistance to the formal sector wage-earning job tends to get a formal sector job and gain a higher-earning compared to the counter-factual earning in the informal sector. Though the negative sloping MTE curve indicates the positive gain on the sorting, that does not necessarily mean that the informal sector labor market is mostly integrated. For instance, if the negative sloping MTE curve is greater than 0 across all U_D, this may indicate that even individuals whose distaste for formal sector jobs is highest due to her/his ineptitude in the wage-earning environment still get better off by taking a formal sector job. In other words, if the downward sloping MTE curve shows only positive values, the whole sample whose propensity score is estimated may get better off by taking formal sector work if we assume the cost of working in the formal sector is negligible. Thus this MTE curve may indicate a segmented labor market. However, if the downward sloping MTE curve shows negative values for some population whose resistance to formal sector work is high, it indicates that there is a portion of workers who are better off by working in the informal sector. Then MTE function decreasing with U_D and where the MTE function becomes negative at considerable resistance indicates the existence of a competitive informal sector for some segment

of the labor market, though, at another segment of the informal sector, informal sector workers are rationed out from the formal sector. Overall, this pattern could indicate the interposing of a competitive informal sector and marginalized informal sector.

(2) MTE increasing with U_D

A positive sloping MTE curve along U_D implies that a group of workers with the highest resistance to formal sector jobs earn the highest relative gain by working in the formal sector. Upward sloping MTE curve can occur if non-pecuniary benefits dominates monetary gain for individuals' occupational choice. To be more concrete, let us suppose that workers in the informal sector may have greater independence, flexibility, and other fringe benefit. Relative attractiveness of these non-monetary benefits may attract workers in the informal sector, and these workers may be willingly to sacrifice the potential financial gains from formal sector for their non-monetary gain. Other possibility is when the unobserved distaste on the formal sector work represents entry barrier. In this scenario, the upward sloping MTE curve suggests that workers who face higher entry barrier into the formal sector employment can get better off, had they are employed in the formal sector. If the upward sloping MTE curve is greater than 0 across all U_D, then it indicates all workers in the market are better off by working in the formal sector, which refers to the segmented labor market.

(3) MTE orthogonal to U_D

This shape of the MTE curve indicates either (i) no relationship between expected gains from the formal sector and worker's preference or (ii) the gap between formal and informal sector workers across the resistance is monotone. Both interpretations infer labor market segmentation: despite heterogeneous resistance, the average formal-informal earning gap due to the resistance is the same across the whole population of workers. In short, among the possible scenarios discussed, the clear indicator for the existence of labor segments for integrated formal-informal labor market can be found in a negative (positive) sloping MTE curve where the value of MTE goes negative with a large (small) resistance.

As shown by Heckman (2007), we can also integrate MTE to recover the the average treatment effect ATE, treatment effect on the treated TT, and the treatment effect on the untreated $ATUT$. In our analysis, TT estimates the average gains of working in the formal sector compared to the informal sector among formal sector workers. ATE calculates the effect of working in the formal sector relative to informal one among the overall population, and TUT measures the counter-factual earnings of working in the formal sector relative to those in the informal sector among informal sector workers. As discussed, the downward sloping MTE curve with a negative value of MTE at some resistance level gives a clear inference on the existence of a competitive informal sector labor market. We focus our discussion on the relationship among aggregated estimators when the MTE curve is downward sloping. The negative sloping MTE curve is related to $TT > ATE > TUT$. This means that the effect of working in the formal sector among

formally employed workers is greater than the earning effect among the overall population. Also, the earning effect of being in the formal sector among informal sector workers are less than the earning effects among the overall population. Especially, $TUT < 0$ indicates that the average counter-factual earning in the formal sector is less than what the informal sector workers would earn in their current job. If this is the case, on average, informal employment is voluntary and chosen based on comparative advantage considerations. When $TUT > 0$, the opposite is true. The average counter-factual earning in the formal sector among informal sector workers is greater than their average actual earning. As such, TUT in the informal population depends on the relative importance of the upper versus lower tier in the informal labor market and the average gains and losses in the two sub-populations.

3. Data

We use the Indonesian Family Life Survey (IFLS) in 2000,2007 and 2014. The IFLS covers 83 percent of the total population living in 13 out of the 27 provinces, primarily on the west side of the country. We follow previous literature to define formal sector workers as those employed in the private or public sectors and self-employed and casual workers as informal sector workers (ADB 2010). [2]

The sample for the analysis is drawn from the working population, ranging in age from 15 to 64 years. Additionally, we limit our sample to working individuals whose annual earnings and household assets fall between the 3st and 97th percentiles of real income and the real value of household assets. This leaves us with 50,759 valid observations (19,262 female workers and 31,497 male workers). Table 1 shows that the majority of male workers (60 percent) are involved in informal sector employment, and informal sector workers have less education and earned income. We also observe the majority of female workers in the informal sector (52 percent), which increase more if we account for non-paid family worker in the sample (XX). We also observe significant overlaps between the informal and formal sectors, even within narrowly defined industries, suggesting the possibility of transitioning between formal and informal sectors.

---

[2] We do not include non-paid family work in the category of informal sector as their recorded income is often zero. We cannot conduct a meaningful income comparison across formal and informal sector workers had we include non-paid family workers in the category of informal employment. Note that majority of non-paid family workers are found among female sample. When we conduct an analysis on household decision making with female sample, we do not exclude non-paid family workers from the sample.

	Female Worker		Male Worker	
	Formal Sector	Informal Sector	Formal Sector	Informal Sector
Employment				
Working Hours per Week	41.619	41.85	53.807	41.964
	[18.512]	[28.465]	[14.714]	[22.335]
Working Weeks per Year	40.834	40.207	42.421	40.234
	[15.596]	16.716]	[15.190]	[16.076]
Log Real Earning	14.858	14.164	15.531	13.211
	[3.221]	[3.180]	[2.911]	[5.458]
Job Size Category[1]	2.49	1.167	2.589	1.236
	[1.336]	[0.549]	[1.340]	[0.617]
Composition across Industries				
Agriculture, Forestry, Fishing	0.154	0.205	0.122	0.428
	[0.361]	[0.403]	[0.328]	[0.495]
Mining and Quarrying	0.009	0.007	0.029	0.025
	[0.096]	[0.085]	[0.167]	[0.156]
Manufacturing	0.244	0.171	0.242	0.12
	[0.429]	[0.376]	[0.428]	[0.325]
Electricity, Gas, Water	0.011	0.006	0.014	0.005
	[0.106]	[0.074]	[0.117]	[0.070]
Construction	0.008	0.003	0.072	0.036
	[0.092]	[0.050]	[0.258]	[0.186]
Wholesale, retail, restaurants	0.149	0.441	0.132	0.177
	[0.356]	[0.497]	[0.339]	[0.381]
Transportation, storage, communication	0.026	0.015	0.053	0.053
	[0.160]	[0.120]	[0.225]	[0.224]
Finance, Insurance, real estate, and business services	0.112	0.072	0.127	0.076
	[0.315]	[0.259]	[0.333]	[0.266]
Social services	0.286	0.082	0.210	0.081
	[0.452]	[0.274]	[0.407]	[0.272]
Individual Characteristics				
Another member of the household	0.441	0.413	0.445	0.402
working in the formal sector	[0.096]	[0.089]	[0.101]	[0.092]
Age	32.525	40.866	33.413	37.945
	[10.871]	[11.508]	[10.240]	[12.778]
Education Level[2]	2.517	1.644	2.453	1.828
	[1.238]	[1.083]	[1.042]	[1.028]
Share of Urban Pop	0.697	0.515	0.706	0.41
	[0.459]	[0.500]	[0.456]	[0.492]
Log Household Asset	21.331	21.588	21.361	21.352
	[2.060]	[1.759]	[1.970]	[1.781]
Number of Household Members	2.336	2.262	2.079	2.288
	[1.131]	[1.038]	[1.195]	[1.140]
Share of Sample	0.474	0.526	0.395	0.605
Sample Number	9,116	10,132	11,445	17,498

Source: Indonesian Family Life Survey (2000,2007,2014)

Notes: In each survey year, top and bottom 1 percentile of earnings and household assets are winsorized.

[1] Job size data is divided into 5 different categories: 1.Between 1 and 4; 2.Between 5 and 19; 3.Between 20 and 99; 4.Between 100 and 199; 5.Mover than 200.

[2] Education data is divided into four categories: 0.No education; 1.Elementary 2.Middle School 3.High Scholl 4.University or Above.

A standard requirement of any treatment evaluation estimator is to ensure the comparability of the treated and untreated. In our analysis, we need to have a reasonably large sample of people in both the formal and informal sectors whose observable characteristics are similar. Formally, the requirement is written as a non-zero probability of being in the formal or informal sector with the same observable characteristics (Heckman et al., 2006). As such, we have to choose the right observable characteristics to ensure that there is a large portion of both formal and informal sector workers with similar observable characteristics. Taking these factors into account, we choose to control for the number of household members, the income of other family members, the number of working hours per week, the number of working weeks per year, the log of household assets, the log of minimum wage, dummies for living in an urban area and being a female, education

dummies, log of age[3], provincial macro variables, and dummy variables for a year, district, and occupation.

We instrument formal sector employment with (i) a dummy variable that indicates whether the respondent has a household member working in the formal sector[4], and (ii) average formal sector employment within the province. For these to be valid instruments, they should satisfy both relevance and conditional exogeneity conditions. We can show statistically for a relevance condition, but it needs to be explained how these variables meet exogeneity conditions. The rationale for using the first instrument is two-fold. First, respondents are directly influenced by other household's employment choices in that it affects respondents' taste for wage-earning occupations. Second, we control any potential route that a household member's formality status may affect the respondent's income. Households likely decide which household member makes how much income and who takes care of children and household works. In that sense, having a household member working in the formal sector may affect the respondent's income other than through the respondent's decision on formal sector employment. For instance, having a spouse working in the formal sector affects the respondents' time and effort put in the workplace. Thus, we control the number of household members, respondent's working hours per week and working weeks per year, and income brought by other household members. Likewise, the second instrument, average provincial formal sector employment, has a strong rationale for relevance: the number of formal sector workers relative to informal sector workers in one's surrounding affect her/his employment status choice. We control provincial macro variables such as GDP and unemployment to absorb the spillover effect of overall formality within the province. Also, after controlling for other individual characteristics and dummy variables for industry, year, and district, it is plausible that both instrument variables satisfy exogeneity condition.

For the second set of results on household decision making with married female sample, we use the same instruments. We use the same logic for the validity of our instruments. Having a household member in the formal sector can affect household decision making other than through the mechanism of respondents' choice on the employment type. For instance, wives whose husbands working in the formal sector can have relatively less decision-making power as the relative income brought by the husbands may affect the decision making. By controlling income brought by other household members and also the number of household members, we can reduce the potential violation of exogeneity condition.

---

[3]The empirical work employs the logarithm form of the age level. With data from Indonesia, this specification fits the nonlinear effect of age on wages better than the traditional quadratic form of Mincer's equation.

[4]This instrument has been used by Radchenko (2014).

5. Results

5.1. Income Equation

We first estimate the propensity score with the probit model to investigate the density function for male and female workers, respectively. The dependent variable is an indicator taking value one if an individual work in the formal sector. Figure 1 shows the distributions of the predicted propensity score and confirms that the two distributions' supports overlap almost everywhere for both genders. We trim observations for which the estimated P(Z) is below 0.02 or above 0.98.
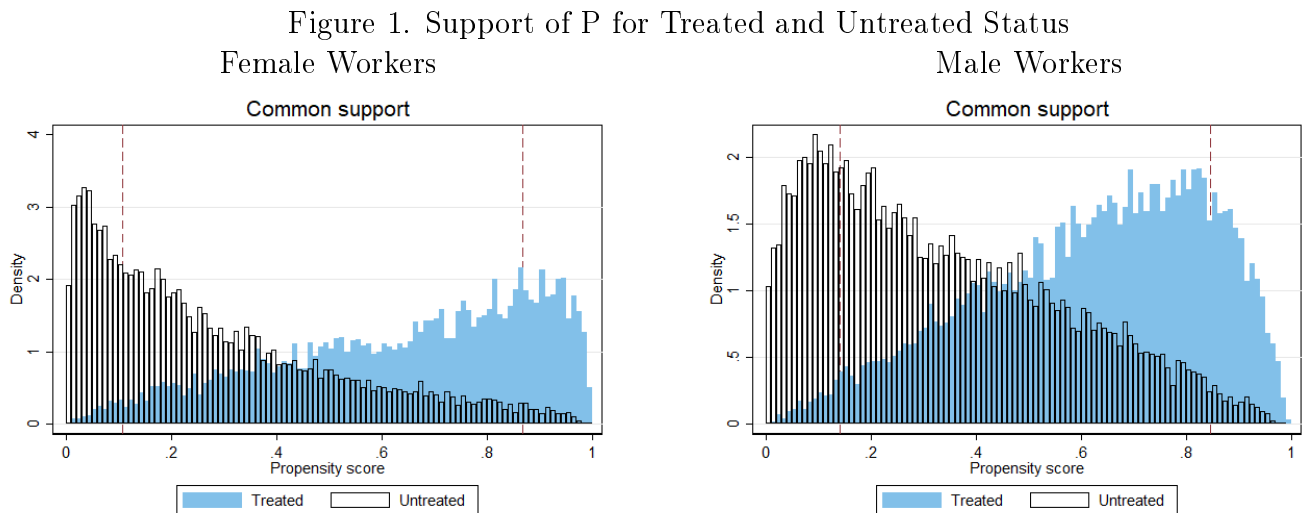
Figure 1. Support of P for Treated and Untreated Status

Female Workers	Male Workers

Table 1: Selection Equation: Propensity Score Analysis

	Formal Work	
	Female Sample	**Male Sample**
Formal Sector Working Among HH Members	0.494***	0.433***
	(0.033)	(0.024)
Provincial Average Formal Sector employment	6.657***	2.893***
	(0.893)	(0.500)
Number of HH Members	-0.335***	-0.435***
	(0.039)	(0.020)
Other Family Income	0.020***	0.025***
	(0.003)	(0.002)
Working Hours (Per Week)	0.004***	0.004***
	(0.001)	(0.000)
Working Weeks(Per Year)	0.011***	0.006***
	(0.001)	(0.000)
Log Household Asset	-0.135***	-0.035**
	(0.008)	(0.005)
Log Real Minimum Wage	0.123	0.490***
	(0.151)	(0.094)
Education	-0.690***	0.013***
	(0.054)	(0.040)
Education^2	0.268***	0.061**
	(0.012)	(0.009)
Urban/Rural	0.229***	0.175***
	(0.037)	(0.026)
Age	-0.107***	-0.022***
	(0.009)	(0.006)
Age^2	0.001***	0.000***
	(0.000)	(0.000)
Provincial Macro Variables	Yes	Yes
Year FE	Yes	Yes
District FE	Yes	Yes
Occupation FE	Yes	Yes
Observations	15,188	26,791

Sources **:** Indonesian Family Life Survey (2000, 2007, 2014)

Notes **:** The table reports average marginal effects from a probit selection model in which the dependent variable is equal to one for the respondents working in the formal sector. We use a dummy variable for the existence of formal sector workers in the household and province average formality as instruments. For female sample, we trimmed at 5 percent; with male sample, we trimmed at 1 percent. Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1.

Table 2 reports the results of the selection equation (5). It shows that both of our instruments are strong predictors of formal sector employment. Consistent with the literature, all our controlled variables, except minimum wage, show a statistically significant effect at 1 percent for both genders. It is interesting to note that minimum wage is positively related to male workers' formal sector employment, whereas we do not find a statistically significant relationship with the female sample. This is consistent with Kim and Williams (2021).
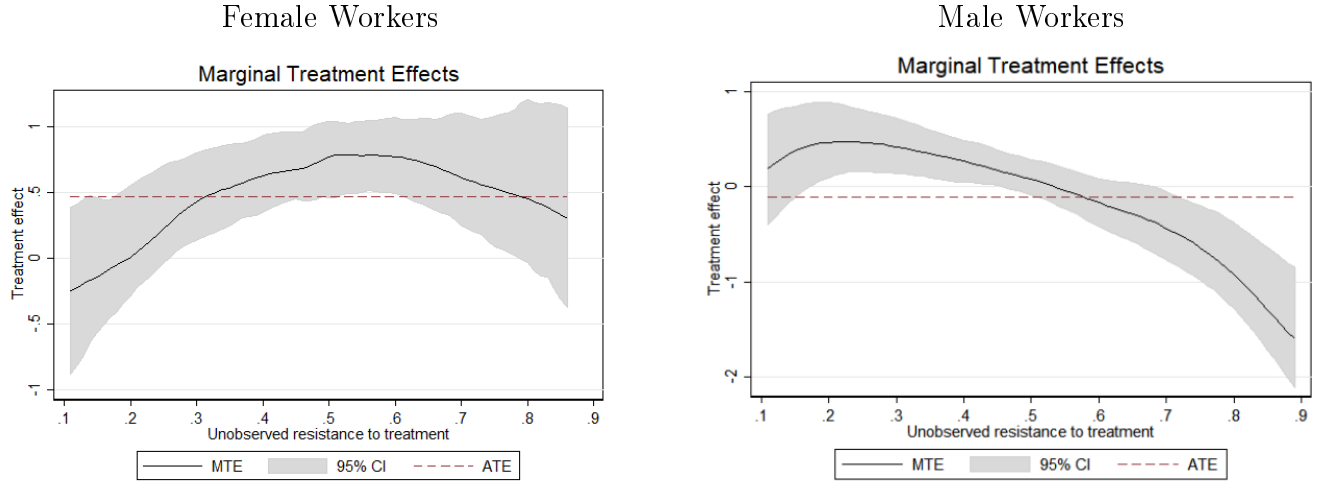
Figure 2. MTE curve

Female Workers   Male Workers

Marginal Treatment Effects (Female Workers)

Treatment effect vs Unobserved resistance to treatment
— MTE   95% CI   - - - - ATE

Marginal Treatment Effects (Male Workers)

Treatment effect vs Unobserved resistance to treatment
— MTE   95% CI   - - - - ATE

Table 2: Income Equations: Semi-Parametric Regression Estimates

Treatment	Full-time Formal			
	Female Sample		Male Sample	
Dependent Variable	Return from informal employment β_0 (SE)	Gap in formal-informal return $(\beta_1 - \beta_0)$ (SE)	Return from informal employment β_0 (SE)	Gap in formal-informal return $(\beta_1 - \beta_0)$ (SE)
	(1)	(2)	(3)	(4)
Number of HH Members	0.086***	-0.223***	0.089***	-0.308***
	(0.030)	(0.070)	(0.021)	(0.057)
Other Family Income	-0.006***	0.014***	-0.007***	0.019***
	(0.002)	(0.005)	(0.001)	(0.004)
Working Hours (Per Week)	0.005***	0.010***	0.005***	0.000
	(0.001)	(0.001)	(0.001)	(0.001)
Working Weeks(Per Year)	0.035***	0.004*	0.018***	0.023***
	(0.001)	(0.002)	(0.001)	(0.002)
Log Household Aseet	0.132***	-0.037**	0.217***	-0.155***
	(0.011)	(0.016)	(0.008)	(0.015)
Log Real Minimum Wage	0.114	0.492*	-0.158	0.978***
	(0.156)	(0.262)	(0.125)	(0.208)
Education	-0.056	0.515***	0.125***	-0.185*
	(0.061)	(0.107)	(0.047)	(0.096)
Education^2	0.051***	-0.062**	0.000	0.089***
	(0.017)	(0.026)	(0.013)	(0.022)
Urban/Rural	0.155***	-0.145*	0.109***	-0.021
	(0.050)	(0.078)	(0.041)	(0.073)
Age	0.063***	-0.049***	0.061***	0.019
	(0.010)	(0.017)	(0.008)	(0.013)
Age^2	-0.001***	0.001***	-0.001***	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Provincial MV	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Occupation FE	Yes	Yes	Yes	Yes
Observations	15,188	15,188	26,791	26,791
P1	1.713		2.608*	
	(1.883)		(1.522)	
P2	-0.589		-4.877***	
	(1.796)		(1.407)	
ATE	0.615***		-0.110	
	(0.091)		(0.102)	
ATT	0.674***		0.344***	
	(0.123)		(0.133)	
ATUT	0.568***		-0.507***	
	(0.114)		(0.152)	

Sources: Indonesian Family Life Survey (2000, 2007, 2014)

Notes: This table reports estimated coefficients of the income equations separtely for female and male samples. The coefficients are obtained from the residual semi-parametric regression detailed in the appendix. The table also provides the estimates of various returns to working in a formal sector for the semiparametric model estimated on several samples: average treatment effect (ATE), treatment on the treated (TT), treatment on the untreated (ATUT). The ATE, ATT, ATUT estimates are computed such that the weights integrate to one in the intervarl [0.01; 0.99]. Standard errors are bootstrapped (200 replications) and reported in parentheses *** p<0.01, ** p<0.05, * p<0.1.

The difference in the marginal treatment curves between male and female samples is illustrated in Figure 2. Whereas Panel B of Figure 2 depicts a downward sloping MTE curve for male workers, the MTE curve for female workers does not exhibit a clear sloping pattern. We can infer from the male worker's MTE curve that there is a positive sorting on the gain, and a sizable portion of male workers would earn less in the formal sector than in the informal sector. Numerous aggregate measures of treatment effect corroborate this conclusion. In Table 2, the treatment effect is greater among formal sector workers, TT, than it is among the general population, ATE. The treatment effect among informal sector male workers, TNT, is the smallest of the three measures and has a negative value, implying that informal sector workers earn a higher average wage than they would if they worked in the formal sector. Note that the average TNT in the informal population depends on the relative importance between upper versus lower-tier informal sectors and the average gain or loss in these two sub-populations (Radchenko (2017)). Thus our analysis suggests that the driving mechanism for the male laborer to work in the informal sector is primarily cost-benefit analysis.

On the other hand, we are not able to infer the positive sorting with the female sample. Even though aggregate statistics show that, on average, there is a positive sorting on the gain (TT>ATE>TNT), the MTE curve rather shows a flat or upward sloping trajectory. However, there is one thing that our estimation infers. Along with the positive values of ATE and TT for female workers, the results on TNT for informal sector female workers imply that female workers would earn higher wages in a formal sector, regardless of whether they are currently working in the formal or informal sector.

5.2. Household Decision Making Equation

d

6. Conclusion

The purpose of this article is to examine the distribution of the individual treatment effect (the effect of taking a formal sector job on earnings) and its relationship to job allocation for male and female workers, separately. A positive correlation between unobservables and formal sector employment indicates that individuals understand the potential benefits of formal sector employment and take action. A zero correlation between individual characteristics and treatment effect indicates no cost-benefit analysis has been conducted to determine worker assignment. The shape

of the MTE curves for male and female workers suggests that the labor market in Indonesia is relatively integrated for male workers, whereas female workers face a more segmented labor market. Coupled with aggregate treatment measures, the negative slope of the MTE curve for male workers indicates that individuals' cost-benefit analysis has primarily driven informal sector employment in Indonesia. On the other hand, there is no discernible pattern in the MTE curve for female workers. Although it is somewhat flat and even slopes upward at times, the MTE curve exhibits positive values across nearly all common support. This implies that female workers would earn more in the formal sector, regardless of their preference for formal sector employment. These findings lead to two policy implications: (1) While informal sector work (mostly self-employment) can be appealing to male workers, it is not so for female workers. (2) policies that encourage formal sector employment can narrow the income gap across gender.

References

[1] OECD 2019, Tackling Vulnerability in the Informal Economy

[2] ILO 2018, Women and men in the informal economy: a statistical picture (third edition)).