



PUEY UNGPHAKORN INSTITUTE
FOR ECONOMIC RESEARCH

Determinants, wage inequality, and occupational risk exposure of informal workers: A comprehensive analysis with the case study of Thailand

by

Upalat Korwatanasakul

September 2021

Discussion Paper

No. 160

The opinions expressed in this discussion paper are those of the author(s) and should not be attributed to the Puey Ungphakorn Institute for Economic Research.

Determinants, wage inequality, and occupational risk exposure of informal workers: A comprehensive analysis with the case study of Thailand

Upalat Korwatanasakul (PhD)

Research Fellow, United Nations University Institute for the Advanced Study of Sustainability

Abstract:

This study provides a comprehensive analysis of informal workers in Thailand by utilising the 2006-2019 Thai Informal Employment Survey data. The estimated results reveal the adverse effects of informal employment on workers' economic and social conditions as follows: 1) the wages gap working against informal employment, confirming that informal employment is not a choice but rather an unavoidable constraint (Oaxaca-Blinder decomposition), 2) a negative relationship between informal employment and wages, particularly among workers in the lowest tail of the wage distribution (quantile regression), and 3) a positive association between informal employment and occupational risks, particularly injury with high severity (logit and probit models). Therefore, policies to smooth informal workers' mobility to the formal sector is crucial. Furthermore, the analyses manifest the importance of schooling in reducing the tendency to work in the informal sector, narrowing the wages gap, and lowering occupational risks and injury severity. However, the estimated results from the pseudo-panel fixed effects regression show no relationship between schooling and informal workers' wages but a positive relationship between their wages and working experience. Thus, policymakers may adopt schooling-related policies to improve informal workers' welfares and mobility to the formal sector. On the other hand, to help workers who inevitably remain in the informal sector, the government may resort to policies regarding working experience, e.g. on-the-job training programmes, to help informal workers earn more wages and, in turn, become less vulnerable.

Keywords: Informal worker, Occupational risk, Sectoral transition, Thailand, Wage inequality

JEL: J16, J21, J31, J71, O17

Acknowledgement:

The author would like to express his sincere appreciation to the National Statistical Office of Thailand for the data used in this paper.

Disclaimer:

The views expressed in this publication are those of the author and do not necessarily reflect the views of the United Nations University.

1. Introduction

The issue of informal employment is common among developing economies, and Thailand is not an exception. In Thailand, large informal employment has persisted over the years. Recent estimates (the National Statistical Office of Thailand (NSO), 2020) reveal that the informality rate, measured as the percentage of the labour force not registered in the social security system, was as high as 53.8% or 20.4 million workers. As informal employment accounts for a large part of the labour market but may have adverse effects on economic growth and social well-being, the issue is worth closer examination.

The study of the informal economy has gained momentum in the labour economics literature since the 1970s. Thenceforth, there have been theoretical debates over the nature and composition of the informal economy. The empirical research in this lacuna has revolved around two dominant schools of thought, namely 'dualist' and 'legalist' (Chen, 2012; Maloney, 1999, 2004; Perry et al., 2007).¹ Based on the exclusion hypothesis, the dualist school argues that formal and informal sectors are fundamentally different due to productivity heterogeneity between the two sectors (Harris and Todaro, 1970; Hart, 1973; International Labour Office, 1972; Lewis, 1954; Sethuraman, 1976; Tokman, 1978). With insufficient human capital or limited productivity, economic agents (e.g. labour and micro-firms) find it difficult to survive in the formal sector and, therefore, inevitably enter the informal sector to make a living. Hence, the informal sector exists to absorb these economic agents and provide a safety net to them in times of crisis. The informal sector will become unnecessary and cease to exist when the formal sector can provide enough jobs to all economic agents.

In contrast, the legalist school is based its thought on the exit hypothesis and therefore sees the informal sector as the collection of firms that choose to operate informally to avoid taxation and other regulations or to enjoy the benefits of property rights (de Soto, 1989, 2000; Levy, 2008). Thus, the legalist school argues that informality is a choice, while the dualist approach views informality as an unavoidable constraint.

Recent studies empirically test the dualist and legalist views through three strands of the literature, namely wage inequality between formal and informal workers, the determinants of informal workers, and occupational risk exposure of informal workers. It is possible to presume that workers voluntarily work in the informal sector if one of the following conditions holds. First, there are no differences (e.g. in terms of wages, determinants, and working conditions) between the formal and informal sectors. Second, the differences between the two sectors are in favour of informal workers.

Against these theoretical and empirical backdrops, the objectives of this study are as follows. First, this study aims to identify the determinants of informal employment and characterise the livelihoods of informal workers. The second objective is to observe the earning and gender disparities among and between formal and informal workers. Lastly, this study aims to provide policy recommendations that enhance the welfare and livelihoods of informal workers and promote the inclusion of informal workers into the formal sector.

This study provides a comprehensive analysis of informal workers in Thailand. First, the study utilises cross-sectional probit regression analysis to examine the determinants of informal employment. Second, this study applies the Oaxaca-Blinder (OB) decomposition technique to investigate the earnings gap between formal and informal workers and to see how much of the disparity can be explained by differences in observable and non-observable characteristics over time. This study then estimates Mincer wage equations at the mean (ordinary least squares (OLS) regression) and at

¹ For more details on theoretical discussion, see Chen (2012).

different quartiles of the wage distribution (quantile regression). Moreover, the analysis employs a pseudo-panel approach and a fixed-effect model to take into account unobservable characteristics and identify the factors that explain differences in earnings. Lastly, a cross-sectional probit model assesses whether informal workers face higher occupational risks than formal workers. All the analyses utilise the 2006-2019 Thai Informal Employment Survey (IES) data.

Based on cross-sectional probit regression analysis, the study finds that personal and job-related characteristics, i.e. sex, age, marital status, area of residence, years of schooling, working hours per week, and wage type, are the determinants of informal employment. Furthermore, the estimated results from the OB decomposition confirm the legalist approach's argument that informal employment is not a choice but an unavoidable constraint since the wage gap favours formal workers. The estimated results from the quantile regression model reveal that the wage gaps between formal and informal employment exist in general and in all quartiles of the wage distribution, implying the absence of internal duality within the informal sector. The relative wages penalty of working in the informal sector is the largest for the lowest tail of the wage distribution (20th quartile). Regarding occupational risks, the results show that informal employment, the working environment (non-behavioural risk factor), and safety issues at work (behavioural risk factor) are the main factors that significantly increase the probability of occupational injury and illness. In general, the analyses point out that informal employment adversely affects workers' economic and social conditions. Therefore, policies to smooth informal workers' mobility to the formal sector is crucial. The analyses also manifest the importance of schooling in reducing the tendency to work in the informal sector, narrowing the wages gap, and lowering occupational risks and injury severity. Hence, policymakers can use education to improve the mobility and workers' welfares. However, the estimated results from a pseudo-panel approach with a fixed-effect model find that education does not impact informal workers' wages regardless of gender but show a positive relationship between their wages and working experience. In other words, working experience is more relevant than formal education for workers who inevitably remain in the informal sector. Thus, policies regarding the training of necessary skills can help these informal workers earn more wages and, in turn, become less vulnerable to economic shocks.

This study significantly contributes to the long-standing policy debates on determinants and wage inequality of informal workers with the unexplored dataset and the comprehensive estimation methods and research design. This study also introduces a new research area and policy debate concerning informal workers' livelihoods and occupational risks. In addition, it provides policy recommendations that enhance informal workers' welfare and livelihoods, a smooth transition of workers from the informal sector into the formal sector, and gender equality and equal income distribution within and between the formal and informal sectors.

The paper is structured as follows. Section 2 provides a brief overview of different strands of literature regarding informal employment. Section 3 describes the data, provides descriptive statistics of the variables, and presents various estimation models used for the analyses. Section 4 offers empirical results, while Section 5 provides discussion and policy implications. Section 6 concludes.

2. Literature Review

This section presents three strands of studies empirically testing the dualist and legalist views, including the determinants of informal workers, wage inequality between formal and informal workers, and occupational risk exposure of informal workers.

2.1 Determinants of Informal Workers

The determinants of informal workers involve personal characteristics, family background, and job-related characteristics. In general, the literature is centre on developing countries such as Latin American and African countries, and its findings are consistent among different studies.

In terms of personal characteristics and family background, the determinants of informal workers include age, education and skills, gender, household and marital status, migrant status, religion, and ethnicity. The studies observed a U-shaped pattern between age and informality (e.g. De Vreyer and Roubaud, 2013; García, 2017; Lehmann and Zaiceva, 2013; Marcouiller, Ruiz de Castilla, and Woodruff, 1997). In other words, as opposed to prime-aged workers, young and old workers tend to work in the informal sector. Education also has a significant impact on a sector choice. Less-educated workers are concentrated in the informal sector (De Vreyer and Roubaud, 2013; García, 2017; Gimpelson and Zudina, 2011; Marcouiller, Ruiz de Castilla, and Woodruff, 1997; Wahba, 2009). Wahba (2009) found that highly educated male workers use informal employment as a steppingstone to acquire a better job in the formal sector, while this argument is not valid in the case of less-educated workers. Skills also determine the sector to which a worker belongs. Similar to education, less-skilled or unskilled workers are likely to be in the informal sector (Lehmann and Zaiceva, 2013). Non-household-head workers (García, 2017), unmarried workers (García, 2017), and immigrant workers (De Vreyer and Roubaud, 2013) have higher probabilities of working in the informal sector. Regarding gender, there is disagreement within the literature whether women or men are more likely to be employed as informal workers. García (2017) observed that there are higher possibilities for women to work in the informal sector in Colombia, while Gimpelson and Zudina (2011) and Lehmann and Zaiceva (2013) found contradictory results in Russia. Moreover, the probability of being an informal worker of women also depends on other personal characteristics such as the number of children in a household, ethnicity, religion, among others (De Vreyer and Roubaud, 2013). Concerning family background, De Vreyer and Roubaud (2013) observed that the father's occupation and the father's education significantly impact sector choice.

Regarding job-related characteristics, retail trade and construction are mainly informal-sector activities (Gimpelson and Zudina, 2011; Lehmann and Zaiceva, 2013; Marcouiller, Ruiz de Castilla, and Woodruff, 1997). Businesses such as hotels and restaurants (Gimpelson and Zudina, 2011), transportation, and personal services (Marcouiller, Ruiz de Castilla, and Woodruff, 1997) are also significantly related to the high probability of being an informal worker. In addition, Marcouiller, Ruiz de Castilla, and Woodruff (1997) showed that jobs with an irregular length of working hours (e.g. very short or very long) are likely to be in the informal sector.

2.2 Wage Inequality between Formal and Informal Workers²

A large body of research investigates the wage differentials between the formal and informal sectors in various developing countries, especially those in Latin America. However, the results are highly country-specific and inconclusive. The first wave of the literature was primarily based on comparing the mean wages between the two sectors. The studies found that formal workers are better rewarded than informal workers (e.g. Blunch, 2015; Daza and Gamboa, 2013; Heckman and Hotz, 1986; Marcouiller, Ruiz de Castilla, and Woodruff, 1997; Mazumdar, 1981; Pradhan and van Soest, 1995; Roberts, 1989). The endogeneity problem owing to unobserved variable and self-selection biases was among the most prevalent issues of the first wave literature.

² For more comprehensive survey of the literature, see Leontaridi (1998), Perry et al. (2007), and Tansel, Keshin, and Ozdemir (2015).

To deal with the problem, relatively more recent studies utilised either 1) a Heckman two-stage procedure to explicitly correct for the possibility of self-selection of workers into a particular sector and earnings levels or 2) a fixed effects model and a quantile regression (QR) technique. The QR technique helps reveal more detailed patterns of wage differentials along the earnings distribution. Some studies found that the informal sector wage penalty persists even after taking into account the endogeneity problem (e.g. Arias and Khamis (2008), Bargain and Kwenda (2014), Botelho and Ponczek (2011), Carneiro and Henley (2001), Gong and van Soest (2002), Magnac (1991), Tansel and Kan (2016), and Tansel, Keskin, and Ozdemir (2015, 2020)). In contrast, Badaoui et al. (2008), Falco et al. (2010), Nordman, Rakotomanana, and Roubaud (2016), and Pratap and Quintin (2006) found that the informal wage penalty disappeared after fixing the endogeneity problems, while the studies such as Braithwaite (1995), Marcouiller, Ruiz de Castilla, and Woodruff (1997), and Staneva and Arabsheibani (2014) even found the wage premium in the informal sector.

The literature further extended its analyses to test whether the informal sector has its internal duality with the help of the QR technique and to decompose the wage differentials between the two sectors by the OB decomposition technique. Tannuri-Pianto and Pianto (2002) applied the QR technique to examine the wage differentials between the formal and informal sectors in Brazil. Their results showed that even though the wage penalty exists at low quantiles of the wage distribution, the penalty disappears at high quantiles. Similarly, Bargain and Kwenda (2014), Lehmann and Zaiceva (2013), and Nguyen, Nordman, and Roubaud (2013) found that the wage penalty tends to disappear at the high quantiles after controlling for time-invariant unobservable variables. A recent study in Egypt (Tansel, Keskin, and Ozdamir, 2020) showed a contrasting result that the wage penalties did not significantly vary across different wage quantiles. In other words, the internal duality within the informal sector was not observed in Egypt. Staneva and Arabsheibani (2014) and Nordman, Rakotomanana, and Roubaud (2016) also did not find evidence to support the duality within the informal sector. However, they found an informal sector wage premium across the income distribution.

In the case of Thailand, Dasgupta, Bhula-or, and Fakthong (2015) employed a cross-sectional quantile regression method with the 2011 Household Socio-Economic Survey data to estimate the earnings gap between the formal and informal sectors. They found that informal workers experience lower earnings across the earnings distribution and the earnings differentials get wider at higher quantiles. Pooittiwong Ramirez (2017) utilised Heckman maximum likelihood estimation and the OB decomposition with the pooled cross-sectional household survey data from 2005-2012. Consistent with Dasgupta, Bhula-or, and Fakthong (2015), Pooittiwong Ramirez (2017) found that informal workers earn less than formal workers, whereas the gender gap is relatively wider in the informal sector.

2.3 Occupational Risk Exposure of Informal Workers

In general, the literature regarding occupational risk is still underdeveloped and unmethodical. The literature covers two main areas, including the determinants of injury (e.g. Adane et al., 2013; Duric et al., 2018; Williams, Adul Hamid, and Misnan, 2018) and economic cost of injury (e.g. Corso et al., 2006; de la Fuente et al., 2014; Leigh, 2011). This section only focuses on the emerging literature in the realm of the determinants of injury, which help us understand the factors behind work-related injuries and, in turn, specify the estimation model to examine the likelihood of the occupational risk exposure of formal and informal workers.

According to Van den Broucke and Colémont (2011), work-related injuries among farmers in Belgium occur mainly due to behavioural factors such as machinery usage, chemical handling, and nonbehavioral risk factors, e.g. demographic characteristics and working environment. A common

finding is that behavioural factors are the major causes of work-related injuries, especially for informal workers. Santana and Loomis (2004) examined the relationship between demographic characteristics in urban Brazil. They found that work-related injuries are a common health problem regardless of the type of employment (formal or informal) and gender. Furthermore, their results showed that job training helps to reduce the likelihood of non-fatal work injuries. Williams, Adul Hamid, and Misnan (2018) carried out surveys in Uganda and showed that behavioural risk factors such as inadequate supervision, use of incompetent personnel, and use of inappropriate construction techniques are the major causes of accidents in construction projects.

On the other hand, a different set of studies identified nonbehavioral risk factors as the factors for work-related injuries. Arphorn et al. (2016) conducted a descriptive cross-sectional study on the working conditions of informal workers in Thailand. They found that working conditions were the main factors contributing to occupational accidents and health problems. Van den Broucke and Colémont (2011) argued that, apart from behavioural factors, the incidence of injuries varies accordingly to education level and the number of employees. In addition, the study of informal work in Serbia (Duric et al., 2018) revealed that age was the leading cause of work-related injuries regardless of gender.

To summarise, what we know so far is the following. First, even though a large body of research has comprehensively examined the wage differentials between the formal and informal sectors in various countries, the results are highly country-specific and mixed. Second, the literature on the determinants and the occupational risk exposure is scarce, and the findings are inconclusive. Third, in the context of Thailand, there are a minimal number of studies in all areas mentioned above. Lastly, the existing studies are likely to have estimation problems, e.g. endogeneity issues, resulting in inconsistent estimates and noncausal interpretation.

3. Data and Methodology

3.1 Data

The micro-level data set used in this study combines IES data from 2006 to 2019. The IES conducted by NSO is the only national dataset that comprehensively includes information both on demographic and labour-related characteristics of Thai informal workers and, to the best of the author’s knowledge, has not been used in any studies. NSO carries out the IES annually, collecting data on approximately 80,000 random households for a total of around 200,000 individual observations, representing 0.1–0.5 per cent of the total Thai population. Apart from typical variables collected by a labour force survey, the IES provides unique data concerning informal workers’ occupational risks, i.e. safety issues at work (behavioural risk factors) and working environmental problems (non-behavioural risk factors). Table 1 provides summary statistics.

Table 1. Summary statistics

Variable	Observation	Mean	SD	Min	Max
Dependent variable					
Model 1: Informal employment 1 = informal worker; 0 = otherwise	1,688,444	0.62	0.49	0	1
Model 2 - 4: Natural log of wages	711,545	9.00	0.84	2.08	13.82
Model 5: Injury 1 = sustained injury; 0 = otherwise	1,684,746	0.14	0.34	0	1
Model 5: Injury severity 1 = minor injury without medical treatment, 2 = minor injury requiring medicine intakes, 3 = medical treatment less than three days, 4 = medical	234,994	1.49	1.15	1	5

treatment for three days or more, and 5 = medical treatment due to loss of an organ					
Independent variable					
Individual characteristic					
Female: 1 = female; 0 = male	1,688,444	0.48	0.50	0	1
Age	1,688,444	43.33	13.32	15	98
Single: 1 = single; 0 = married	1,688,444	0.27	0.45	0	1
Household head: 1 = household head; 0 = otherwise	1,688,444	0.43	0.49	0	1
Rural area: 1 = rural area; 0 = urban area	1,688,444	0.44	0.50	0	1
Years of schooling	1,679,396	8.26	4.64	0	21
Work-related variable					
Total working hour	1,688,444	44.93	14.19	0	98
Wage type: Monthly wage 1 = monthly wage; 0 = otherwise	711,545	0.59	0.49	0	1
Primary sector: 1 = primary sector; 0 = otherwise	1,688,444	0.41	0.49	0	1
Manufacturing sector 1 = manufacturing; 0 = otherwise	1,688,444	0.12	0.33	0	1
Services sector: 1 = services sector; 0 = otherwise	1,688,444	0.47	0.50	0	1
Working environmental problems 1 = experienced the problems; 0 = otherwise	1,683,266	0.11	0.31	0	1
Safety issues at work 1 = experienced the issues; 0 = otherwise	1,682,969	0.08	0.28	0	1
Fixed effect					
Industry 1 = primary sector; 2 = manufacturing; 3 = services	1,688,444	2.06	0.94	1	3
Cohort	1,688,444	1969.23	13.57	1908	2004
Year	1,688,444	2012.56	4.00	2006	2019

Source: Author.

3.2 Methodology

This research conducts five different analyses with an emphasis on a gender dimension. First, this research utilises a linear probability model (LPM), a logistic model, and a probit regression with pooled cross-sectional IES data to identify the determinants of informal employment. The estimation model is as follows:

$$informal_i = \delta_0 + \delta_1 X_i + \varepsilon_i \quad (1)$$

where $informal_i$ is a dummy variable showing an employment type of individual i (informal worker = 1, otherwise = 0), while the vector X_i is a set of workers' characteristics, e.g. age, levels of education, sex, among others. ε_i represents the disturbance term. All models control for industry, region and year fixed effects.

Second, this study applies the standard OB decomposition technique to see how much of the earning differentials between formal and informal workers can be explained by differences in observable and non-observable characteristics over time. Utilising socio-economic indicators observed between two groups of workers (formal and informal workers), the OB decomposition method (Blinder 1973; Oaxaca 1973) decomposes the wage inequality into three components, including i) endowment accumulation (explained component) accounting for individual socio-economic characteristics; ii) return to endowments (unexplained component) referring to informal workers' ability to convert endowments into labour market outcome as measured by wages, and iii) the interaction between endowment and its return. Moreover, the unexplained component also involves factors such as the

nature of the work contract, labour market discrimination, the quality of the labour market and its legislation, and unobservable characteristics (Dasgupta, Bhula-or, and Fakhong, 2015). The estimation model is as follows:

$$wage_i = \begin{cases} \beta^{formal}x_i + \varepsilon_i^{formal} & \text{if formal} \\ \beta^{informal}x_i + \varepsilon_i^{informal} & \text{if informal} \end{cases} \quad (2)$$

$$wage^{formal} - wage^{informal} = \Delta x \beta^{informal} + \Delta \beta x^{informal} + \Delta x \Delta \beta \quad (3)$$

$$wage^{formal} - wage^{informal} = E + C + I \quad (4)$$

where

$$\Delta x = x^{formal} - x^{informal} \text{ and } \Delta \beta = \beta^{formal} - \beta^{informal} \quad (5)$$

$wage^{formal}$ and $wage^{informal}$ represent the mean wages of formal and informal workers, respectively, while x is the vector of determinants, e.g. age, marital status, sex, years of schooling, among others. The gap of mean wages between formal and informal workers come from a gap in endowments (E), a gap in coefficients (C), and a gap from the interaction of endowments and coefficients (I).

Third, this study then estimates Mincer wage equations (Mincer, 1974) both at the mean and at different quantiles of the wage distribution by utilising an OLS regression model and a quantile regression model (QRM), respectively. The latter model captures significant heterogeneity of the wages distribution among informal workers.

$$\ln wage_i = \beta_0 + \beta_1 \text{years of schooling}_i + \beta_2 \text{age}_i + \beta_3 \text{age}_i^2 + X_i + \varepsilon_i \quad (6)$$

where $\ln wage_i$ is the natural log of the hourly wage of an individual i , $\text{years of schooling}_i$ is the number of years of education, and age_i is a proxy of the number of years of experience. X_i represents the vector of control variables, including an area of residence and sex. Moreover, all models control for industry, region and year fixed effects.

However, both models are possibly subject to endogeneity problems such as omitted variables and may not establish a causal relationship. A pseudo-panel approach and a fixed-effect model take into account unobservable characteristics and identify the factors that explain differences in earnings. According to Deaton (1985) and Warunsiri and McNown (2010), tracking birth-year cohorts and averaging the Mincer equation over the cohort members eliminates the individual heterogeneity such as the different abilities or motivations across individuals:

$$\ln \overline{wage}_{ct} = \beta_0 + \beta_1 \overline{\text{years of schooling}}_{ct} + \beta_2 \overline{\text{age}}_{ct} + \beta_3 \overline{\text{age}}_{ct}^2 + \bar{a}_{ct} + \bar{\varepsilon}_{ct} \quad (7)$$

where α_{ct} captures unobserved individual heterogeneity, c indexes cohorts, and t indexes time periods. $\bar{\alpha}_{ct}$ is treated as the true cohort effect or the unobserved cohort fixed effect when the sample size in each cohort is sufficiently large or greater than 100 observations per cell (Verbeek and Nijman, 1992, 1993). Thus, employing a pseudo-panel approach with cohort means eliminates components within the disturbance term correlated with explanatory variables, resulting in a consistent fixed effect estimation. Moreover, cohort means can also mitigate individual measurement errors (Antman & McKenzie, 2007).

Lastly, this study uses an LPM, a logistic model, and a probit regression to assess whether informal workers face higher occupational risks than formal workers. The study also adopts ordered logistic and probit models for estimation with ordinal dependent variables. The estimation model is as follows:

$$injury_i = \gamma_0 + \gamma_1 informal_i + \gamma_2 X_i + \varepsilon_i \quad (8)$$

where $injury_i$ represents two different variables, namely injury and injury severity. Injury is a dummy variable (sustained injury = 1, otherwise = 0), while injury severity refers to different levels of injury seriousness, ranging from one to five (Table 1). The vector X_i contains behavioural risk factors and non-behavioural risk factors. Non-behavioural risk factors involve working environment characteristics and workers' characteristics, such as age, sex, residence area, and years of schooling. In contrast, behavioural risk factors proxied by safety issues at work contain components such as machinery usage, chemical handling, among others.

4. Estimated Results

4.1 Determinants of informal employment

The estimation results of Table 2 indicate the set of determinants of informal employment. Consistent with the previous studies, all personal and job-related characteristics specified in the estimation model serve as the determinants of informal employment. All variables are statistically significant at one per cent and express the expected sign and direction.

First, the results reveal that being a woman reduces the probability of working as an informal worker, supporting the previous findings in the case of Russia (Gimpelson and Zudina, 2011; Lehmann and Zaiceva, 2013). Russian and Thai women possibly have higher opportunities to receive better education and secure formal employment. In contrast, women in Cambodia (García, 2017) may have less access to formal jobs, given its economic and social structure.

Second, the age variables show a statistically significant negative relationship with the probability of being in the informal sector. The estimated results confirm the U-shaped pattern between age and informality (De Vreyer and Roubaud, 2013; García, 2017; Lehmann and Zaiceva, 2013; Marcouiller, Ruiz de Castilla, and Woodruff, 1997). In addition, being a household head, living in rural areas, and having more years of schooling are associated with less probability of being an informal worker, whereas being single increases the likelihood, consistent with the findings of the previous studies presented in Section 2.1.

Finally, the job-related variables, namely working hours per week and receiving a monthly wage, are negatively related to the tendency of being an informal worker. The results show that long working hours are associated with less likelihood of working as an informal worker. Hence, the results

somewhat contradict those of Marcouiller, Ruiz de Castilla, and Woodruff (1997), which argue that jobs with very short or very long working hours are likely to belong to the informal sector.

Table 2. Determinants of informal employment

Dependent Variable	Informal employment		
	LPM	logit	probit
Independent Variables	(1)	(2)	(3)
Female	-0.0162***	-0.147***	-0.0700***
	-0.00108	-0.0137	-0.00751
Age	-0.00240***	-0.0301***	-0.0202***
	-0.000321	-0.00294	-0.00162
Age squared	4.13e-05***	0.000430***	0.000274***
	-3.87E-06	-3.39E-05	-1.88E-05
Single	0.0206***	0.157***	0.0991***
	-0.00116	-0.0147	-0.00802
Household head	-0.0119***	-0.146***	-0.0825***
	-0.00111	-0.0143	-0.00787
Rural	0.0294***	0.356***	0.190***
	-0.000972	-0.012	-0.00667
Years of schooling	-0.00590***	-0.0790***	-0.0457***
	-0.000127	-0.00197	-0.00102
Working hours per week	-0.000940***	-0.0141***	-0.00717***
	-5.84E-05	-0.000577	-0.000316
Wage type: Monthly	-0.164***	-2.097***	-1.046***
	-0.00145	-0.0218	-0.0106
Constant	0.550***	0.265***	0.307***
	-0.00762	-0.0846	-0.0445
Observations	705,065	705,065	705,065
R-squared	0.277	0.359	0.352

LPM = linear probability model

Notes: Robust standard errors are below coefficients. All coefficients are statistically significant at one per cent. All models control for industry, region, and year fixed effects. The logit and probit models report pseudo R-squared.

Source: Author.

4.2 Oaxaca-Blinder decomposition of wages between formal and informal employment

Table 3 shows the mean predictions of the natural log of wages (hereafter wages) of formal and informal workers, the difference in wages between them, estimates for endowment, estimates for returns from endowments (coefficients), and estimates of the interaction between endowments and its returns. The results manifest that, on average, informal workers earn less than formal workers. The mean wage of informal workers is 8.26, whereas that of formal workers is 9.12. The difference between their mean wages yields a wage gap of 0.86.

Table 3. Oaxaca-Blinder decomposition of wages between formal and informal employment

Dependent variable: Wage gap						
	Without fixed effects			With fixed effects		
	Coefficient	Robust standard error	Contribution to total difference (%)	Coefficient	Robust standard error	Contribution to total difference (%)
Mean prediction						
Group 1: Formal employment	9.119	0.00155		9.119	0.00159	
Group 2: Informal employment	8.262	0.00414		8.262	0.00418	
Difference of wages	0.857	0.00442		0.857	0.00448	
Endowments (Explained)	0.320	0.00633	37.3	0.351	0.00670	41.0
Coefficients (Unexplained)	0.380	0.00444	44.3	0.375	0.00403	43.7
Interaction	0.158	0.00642	18.4	0.132	0.00652	15.3
Observations	705,065			705,065		

Notes: All coefficients are statistically significant at one per cent. The fixed-effect model controls regions and years.

Source: Author.

The results from the OB decomposition method further reveal that the effects of endowment accumulation and endowment return comparably contribute to the wage gap between the formal and informal workers. The results are in sharp contrast to Dasgupta, Bhula-or, and Fakhong (2015) as they found the considerably larger effects of explained factors, accounting for 67.9 per cent of the wage disparity. The difference between the results possibly comes from their data limitations, including a shorter analysis period (one year) and a smaller number of observations (73,382 observations).

Without controlling regions and years (without fixed effects), the estimates for coefficients are slightly greater than those of endowments. The coefficients (unexplained factors) explain 44.3 per cent of the wage gap, whereas 37.3 per cent come from the difference in endowments. The discrepancy between the explained and unexplained factors become narrower after controlling region and year variables. Nevertheless, the estimates for coefficients are still more prominent than those of endowments, accounting for 43.7 per cent and 41 per cent of the wage gap, respectively.

The decomposition results indicate that the mean wages of informal workers would increase at least 0.38 or approximately 4.5 per cent when applying the formal workers' coefficients or environment to the informal workers' characteristics. Furthermore, if informal workers had the same characteristics as formal workers, their mean wages would rise 4.2 per cent. The result also manifests a relatively small contribution of the interaction term (15.3 per cent), measuring the simultaneous effect of differences in endowments and coefficients to the wage differential. All of the components are statistically significant at one per cent.

The detailed OB decomposition offers more insights from analysing the gaps in socio-economic variables (Table 4). The results show that differences in endowment returns from worker characteristics, such as schooling, age, and economic sector, drive the wage gap. Schooling is the largest contributor in terms of endowment and total effects. To provide informal workers with similar education and opportunities as formal workers would raise their mean wages by 8.4 per cent. When considering only the estimates for returns to endowments, the results highlight the importance of age as one of the main factors widening the wage disparity. If informal and formal workers were of the same age and environment, informal workers' mean wages would increase by 5.7 per cent. In addition, working in the primary sector also contributes to the disparity with relatively milder effects compared to schooling and age.

Table 4. Detailed Oaxaca-Blinder decomposition of wages between formal and informal employment

Dependent variable: Wage gap										
Independent variables	Endowments (Explained)			Coefficients (Unexplained)			Interaction			Total Effect
	Coefficient		Robust standard error	Coefficient		Robust standard error	Coefficient		Robust standard error	
Years of schooling	0.134	***	0.00536	0.349	***	0.00888	0.214	***	0.00552	0.697
Age	-0.055	***	0.00368	0.474	***	0.06196	-0.027	***	0.00363	0.391
Age squared	0.068	***	0.00414	0.064	**	0.03185	-0.008	**	0.00421	0.123
Female	-0.014	***	0.00077	0.013	***	0.00292	0.003	***	0.00067	0.002
Single	-0.001	***	0.00021	0.004		0.00281	0.0002		0.00016	0.003
Rural area	0.031	***	0.00177	-0.029	***	0.00482	0.011	***	0.00187	0.013
Primary sector	0.103	***	0.00268	0.085	***	0.00395	-0.060	***	0.00282	0.128
Manufacturing sector	-0.030	***	0.00308	0.006	***	0.00097	0.020	***	0.00313	-0.004
Regions	0.050	***	0.00201	-0.003		0.01142	0.001		0.00202	0.048
Years	0.065	***	0.00191	-58.680	***	1.68512	-0.022	***	0.00087	-58.637
Constant				58.093	***	1.68580				58.093
Aggregate effect	0.351	***	0.00670	0.375	***	0.00403	0.132	***	0.00652	0.857

Notes: ** and *** represent statistically significant at 0.05 and 0.01, respectively.

Source: Author.

4.3 Wages and informal employment

The estimated results (Table 5) indicate that informal employment, on average, negatively affects wages (Column 1), but its coefficients vary disproportionately across different wage quartiles (Columns 2-4). The results show that the wage gaps between formal and informal employment exist in general and in all quartiles of the wage distribution, implying the absence of internal duality within the informal sector. The wages penalty of working in the informal sector is the largest for the lowest tail of the wage distribution (20th quartile). Due to several constraints, informal workers in the lowest quartile may have no other choice but to face exploitation and accept unfair wages. In contrast, employers must guarantee formal workers the minimum wages. Therefore, the gap is possibly most extensive in the lowest quartile. The results are in sharp contrast with those of Dasgupta, Bhula-or, and Fakthong (2015), which did not find a statistically significant difference in monthly wages between formal and informal employment in the 20th quartile.

Years of schooling and age, a proxy of working experience, on average, have a positive relationship with wages, consistent with previous studies, e.g. Korwatanasakul (2017), Korwatanasakul (2019), Warunsiri and McNown (2010). The estimated results manifest that workers in higher quartiles enjoy greater benefits from education. The jobs assigned to workers who earn more wages possibly require knowledge and skill set acquired from a higher level of formal education. In contrast, the same knowledge and skill set may not be relevant to the tasks occupied by workers from the lower quartiles since the tasks are usually labour-intensive and use skills gained outside formal education. Therefore, workers in the lower quartiles enjoy a lesser education premium. On the other hand, the premium of more working experience is the largest for the lowest tail of the wage distribution, supporting the previous reasoning that tasks occupied by workers from the lowest tail require skills gained from on-the-job training rather than formal education. García (2017) and Dasgupta, Bhula-or, and Fakthong (2015) find consistent results of the different working experience premiums among workers from different wage groups.

Table 5. Wages and informal employment

Dependent variable: Natural log of wages				
	OLS	QRM	QRM	QRM
		Q20th regression	Median regression	Q80th regression
	(1)	(2)	(3)	(4)
Informal employment	-0.305	-0.361	-0.309	-0.326
	0.004)	0.004	0.003	0.002
Years of schooling	0.083	0.077	0.088	0.092
	0.000	0.000	0.000	0.000
Age	0.040	0.045	0.040	0.025
	0.001	0.001	0.000	0.001
Age squared	-0.0003	-0.0004	-0.0003	-0.00003
	0.000	0.000	0.000	0.000
Female	-0.141	-0.124	-0.137	-0.141
	0.002	0.001	0.001	0.001
Rural	-0.094	-0.071	-0.068	-0.080
	0.002	0.002	0.001	0.001
Constant	6.833	6.838	7.095	7.683
	0.012	0.010	0.009	0.010
Observations	705,06 5	705,065	705,065	705,065
R-squared	0.560	0.334	0.349	0.389

OLS = ordinary least squares; Q = quantile; QRM = quantile regression model

Notes: Standard errors are below coefficients. OLS reports robust standard errors, while QRM utilises standard errors using a bootstrap method. All variables are statistically significant at one per cent. Pseudo R-squared is reported for QRM. All models control for industry, region and year fixed effects. Source: Author.

The estimated results support that gender inequality exists in the Thai labour market. In general, female workers earn less than male workers in all wage quartiles. Women in the upper tail of the distribution face a higher penalty of gender discrimination, possibly due to the glass ceiling effect. Similarly, other studies find that the inequality is more pronounced at the higher wage quartiles (e.g. Christofides Polycarpou, Vrachimis, 2013; Kaya, 2017; Piazzalunga and Di Tommaso, 2019). The penalty of discrimination posed to women is substantial even in the lower tail of the distribution. Although the glass ceiling effect may be less relevant at the lower tail, women may face other forms of discrimination and receive lower wages than men.

Workers who live in rural areas, on average, earn less than those living in urban areas, approximately seven to eight per cent. The penalty of living in rural areas is the highest among workers from the highest quartile as workers who live in rural areas may have limited access to higher-paid jobs.

Among all variables specified in the estimation models, the informal employment variable manifests the most considerable effect on wages. Informal workers, on average, earn less than formal workers, around 30-36 per cent. Moreover, gender inequality in the labour market plays a significant role in determining wages. Female workers earn less than male workers, approximately 12-14 per cent. The

adverse effects of the residential area variable (rural areas) are relatively moderate (seven to eight per cent), while years of schooling show a comparable but positive impact on wages (eight to nine per cent). In addition, age or working experience shows the smallest effects on wages (3-5 per cent).

In addition, this study employs a pseudo-panel approach with a fixed-effect model to take into account unobservable characteristics and identify the factors that explain differences in wages. However, the fixed-effect model poses analytical restrictions as the model eliminates time-invariant variables such as type of employment, sex, and area of residence. As the previous analysis (Table 5) points out the importance of informal employment and gender in determining wages, the fixed-effect model also includes disaggregated analysis by type of employment and gender to enrich the analysis. Table 6 shows the estimates from the regressions with individual data (Columns 1-3) and pseudo-panel with one-year cohort means (Columns 4-10). Columns 5-6 provides the estimates disaggregated by type of employment, while Columns 7-10 further break down the estimations by gender.

The estimated results of the OLS model are possibly subject to the endogeneity problem, resulting in a downward bias of the estimated returns to education and working experience. The estimated coefficients of years of education and age from the OLS regressions (Column 1-3) are generally lower than those of weighted least squares (WLS) regressions, consistent with the optimal choice of schooling model (Grilliches, 1977) and the recent empirical findings (e.g. Warunsiri and McNown, 2010).

The first finding from WLS regressions (Column 5, 7, and 8) is that schooling does not affect informal employees' wages as the coefficients are not statistically significant across the estimation models. The estimated results show no relationship between schooling and informal workers' wages since the jobs assigned to informal workers are usually labour-intensive and may not require formal education's knowledge and skill set. Instead, skills gained outside schools, such as on-the-job training, is more relevant to tasks performed by informal workers.

The estimated effects of age or working experience support the above argument. The coefficients of working experience for informal workers are statistically significant at one per cent. They indicate that an additional one year of working experience induces a nine-to-ten per cent increase in informal workers' wages. The results are robust across the estimation models (Columns 5, 7, and 8).

The results of the WLS regressions with the formal worker sample (Columns 6, 9, and 10) follow the typical pattern that schooling and working experience have positive impacts on wages. Moreover, the estimated results from WLS regressions are considerably greater than those of OLS regression (Column 3), approximately two to six per cent, emphasising the downward bias in estimates from the individual data regression. For example, the coefficient of schooling obtained from OLS regression is 0.101, whereas those of WLS regressions are 0.165, 0.117, and 0.151 (Column 6, 9, and 10, respectively).

In addition to the estimated results by type of employment, disaggregation by sex also provides important policy implications. Columns 7-10 show the regression results stratified by employment type and gender. The coefficients on years of schooling for female and male formal workers are 0.117 and 0.151, respectively (Columns 9-10), while the coefficients on age for female and male workers, regardless of the type of employment, are about 0.09 and 0.10, respectively. This disaggregation manifests the rates of return to education and working experience for women are lower than those of men, showing evidence of gender inequality in the labour market. The estimated results are consistent with previous studies from Thailand and other developing countries (e.g. Chiswick 1976, Psacharopoulos and Patrinos 2004).

Table 6. Returns to education and age (working experience) by type of employment and gender

	Dependent variable: Natural log of wages									
	Individual data (cross-sectional regression)			Pseudo-panel (one-year cohort means)						
	All workers	Informal workers	Formal workers	All workers	Informal workers	Formal workers	Informal female workers	Informal male workers	Formal female workers	Formal male workers
	OLS	OLS	OLS	WLS	WLS	WLS	WLS	WLS	WLS	WLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Constant	6.258	6.937	6.409	3.445	4.206	4.133	3.991	4.315	4.463	4.253
	0.010	0.029	0.010	0.117	0.198	0.083	0.236	0.185	0.086	0.079
Years of schooling	0.107	0.060	0.101	0.131	-0.015	0.165	0.014	-0.008	0.117	0.151
	0.000	0.001	0.000	0.014	0.021	0.007	0.024	0.022	0.009	0.009
Age	0.065	0.042	0.061	0.141	0.103	0.093	0.094	0.102	0.095	0.100
	0.0005	0.001	0.001	0.005	0.005	0.005	0.007	0.006	0.005	0.005
Age squared	-0.001	-0.0004	-0.0005	-0.001	-0.0002	-0.0003	-0.0001	-0.0002	-0.0002	-0.0004
	0.00001	0.00001	0.00001	0.00006	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
Cohort dummies	-	-	-	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual observations	656,145	85,475	570,670	656,145	85,475	570,670	34,928	50,547	272,186	298,484
Cohort-year observations	-	-	-	714	714	608	703	709	547	574
Individual observations per cohort										
Max	-	-	-	4,665	2,898	2,128	1,469	1,441	1,055	1,073
Min	-	-	-	122	118	100	100	100	100	100
R squared	0.419	0.069	0.427	0.881	0.862	0.943	0.826	0.838	0.948	0.946

OLS = ordinary least squares; WLS = weighted least squares

Notes: Robust standard errors are below coefficients. All variables are statistically significant at one per cent except those in highlighted cells. Adjusted R-squared is reported for WLS.

Source: Author.

4.4 The occupational risk of informal workers

The estimated results reveal that behavioural and non-behavioural risk factors are related to the likelihood of sustaining an injury and injury severity (Table 7). The coefficients of all specified variables (except age in Column 5) are statistically significant at one per cent and robust across different model specifications (Columns 1-5). The results show that informal workers have a higher probability of sustaining an injury (Columns 1-3) and have a higher chance of facing greater injury severity (Columns 4-5), possibly due to the nature of their tasks. In general, informal workers' task nature includes labour intensiveness, the usage of chemicals, machines, and equipment that may cause an injury, an unfavourable working environment with low safety and high risk (Guadalupe, 2003; Levine, Toffel, and Johnson, 2012; Rongo et al., 2004), among others. Furthermore, living in rural areas is positively associated with higher occupational risk. This finding could signal that workers in rural areas have relatively limited occupational choices and inevitably end up with jobs that pose a higher risk.

Table 7. Determinants of accident exposure and accident severity

Dependent Variable	Injury			Injury severity	
	LPM	Logit	Probit	Ordered logit	Ordered probit
Independent Variables	(1)	(2)	(3)	(4)	(5)
Non-behavioural risk factor					
Worker's characteristic					
Informal worker	0.0401	0.471	0.248	0.486	0.216
	0.0009	0.00985	0.00504	0.00914	0.00503
Rural	0.00964	0.0973	0.0487	0.119	0.0533
	0.000682	0.00697	0.00374	0.00661	0.00369
Female	-0.0208	-0.229	-0.122	-0.241	-0.135
	0.000711	0.00706	0.00377	0.00697	0.00368
Years of schooling	-0.00476	-0.0591	-0.032	-0.06	-0.0306
	8.70E-05	0.000989	0.000516	0.000962	0.000511
Age	0.00009	-0.00109	-0.000532	-0.000928	-0.00029
	0.00003	0.000293	0.000157	0.000288	0.000153
Working environmental problem	0.131	0.888	0.504	0.885	0.473
	0.00174	0.0101	0.00577	0.00997	0.00545
Behavioural risk factor					
Safety issue at work	0.181	1.032	0.599	1.031	0.557
	0.00207	0.0108	0.00626	0.0106	0.00589
Constant	0.103	-2.208	-1.271		
	0.00263	0.0291	0.0151		
Cut 1				2.279	1.238
				0.0275	0.0149
Cut 2				3.635	1.919
				0.028	0.0151
Cut 3				5.045	2.517
				0.0302	0.0158
Cut 4				6.665	3.083
				0.0379	0.0173
Cut 5				9.976	4.025
				0.134	0.0381
Observations	1,673,446	1,673,446	1,673,446	1,673,437	1,673,437
R-squared	0.0854	0.094	0.0952	0.074	0.0716

LPM = linear probability model

Notes: Robust standard errors are below coefficients. All coefficients are statistically significant at one per cent except those in highlighted cells which are statistically significant at ten per cent. All models control for industry, region, and year fixed effects. The logit and probit models report pseudo R-squared.

Source: Author.

On the contrary, gender and years of schooling negatively correlate with the probability of sustaining an injury. Job segregation by education and gender among workers is common as firms generally assign less risky tasks to female and better-educated workers. Moreover, workers with a higher level of education can cope with risky circumstances better (Oh and Shin, 2003; Piha et al., 2012). Even though the age variable shows a negative relationship with the occupational risk, their coefficients are not robust across different model specifications. In addition, the coefficients are very small and, therefore, negligible. This finding sharply contrasts with Duric et al. (2018), who found that age was the leading cause of work-related injuries.

Apart from demographic characteristics, the working environment (non-behavioural risk factor) and safety issues at work (behavioural risk factor) are the two main factors that significantly increase the probability of occupational injury and illness. The estimated results are consistent with Van den Broucke and Colémont (2011). However, the results are somewhat in contrast to other previous studies as they usually found the significant effect of either non-behavioural risk factors (e.g. Arphorn et al., 2016) or behavioural risk factors (e.g. Williams, Adul Hamid, and Misnan, 2018), but not both.

5. Discussion and policy implications

Based on the estimated results in Section 4, Table 8 comprehensively provides the entangled relationship among worker's characteristics, informal employment, wages, and occupational risks. The estimated results point out 1) the wages gap working against informal employment, 2) a negative relationship between informal employment and wages, especially among workers in the lowest tail of the wage distribution, and 3) a positive association between informal employment and occupational risks, particularly injury with high severity. Thus, informal employment possibly has adverse effects on workers' economic and social conditions.

Overall, the results support the dualist approach as the wages gap between formal and informal workers persists across different analyses. Therefore, informal employment is an unavoidable constraint imposed on a group of vulnerable economic agents in Thailand. To alleviate this issue, the government may reduce informal workers' hurdles to move to the formal sector, such as decreasing contribution to the social security system for workers with low and unstable income and reforming laws and regulations regarding formality.

Furthermore, recognising informal workers' characteristics helps policymakers pinpoint the areas that require improvements. The analysis of determinants of informal employment identifies a set of characteristics commonly found among informal workers. First, informal workers tend to work relatively short hours and receive wages on an hourly or daily basis without a formal contract, implying relatively low and unstable income streams and, in turn, inadequate savings. These characteristics make informal workers vulnerable, particularly during the time of crisis such as the COVID-19 pandemic. As it is typically the case that there is no formal agreement between employers and informal workers, informal workers are the first group to be laid off when the crisis hits an economy. Therefore, it is imperative to grant informal workers access to the social security system by facilitating informal workers' mobility to the formal sector. The second characteristic of informal workers is that they tend to live in rural areas where individuals have less access to quality education and formal jobs. Thus, policies to create more formal job opportunities in rural areas, e.g. incentive programmes for firms to establish branches and plants outside urban areas, may help improve the situation.

The analyses also manifest the importance of schooling in reducing the tendency to work in the informal sector, narrowing the wages gap between formal and informal workers, and lowering occupational risks and injury severity. Policymakers can use measures related to formal education to

smooth the transition of workers from the informal sector to the formal sector. Workers with higher levels of education have better job opportunities (less likely to engage in risky tasks) and more chances to enter the formal sector. As discussed above, the promotion of education policies is particularly relevant to people in rural areas who tend to work as informal workers and have limited access to the same quantity and quality of education compared with urban areas. Hence, an education policy can be one of the policy options.

However, the estimated results from the fixed effect model reveal that schooling does not induce higher wages for informal workers but show a positive relationship between their wages and working experience. Hence, training programmes of necessary skills possibly help workers who inevitably remain in the informal sector secure higher income and become less vulnerable to economic shocks. As the discrimination against informal workers is present in the labour market, the government may also help these workers by, for example, implementing a social welfare programme specific to informal workers, enforcing laws and regulations against exploitation and unfair wages, among others.

Regarding the occupational risks, the results manifest that the working environment and safety issues at work are the main factors that significantly increase the probability of occupational injury and injury severity. Thus, with regular inspections, stricter measures on the working environment and safety at work may help reduce the occupational risks and injury severity among workers.

Finally, the analyses point out the problem of gender inequality. Regardless of income quartile and employment type, gender discrimination persists in the Thai labour market. The estimated results manifest that female workers earn less than male workers in all wage quartiles, while women in the upper tail of the distribution face a higher penalty of gender discrimination. A positive action in public and private sectors to reduce the glass ceiling effect helps support female workers in the higher tail of income distribution, while women-specific support programmes, such as subsidies for expenses incurred from women-related necessary goods, women empowerment programmes, among others, possibly help female workers in the lower tail.

Table 8. Summary of selected estimated results

Independent Variables	Dependent Variables					
	Informal employment (1)	Wage gap (2)	Wages (OLS) (3)	Wages (QRM) (4)	Wages (WLS) (5)	Occupational risks (6)
Informal employment	na	+	-	-	na	+
Years of schooling	-	+	+	+	Informal: 0 Formal: +	-
Age (working experience)	-	+	+	+	+	0
Female	-	+	-	-	-	-
Rural	+	+	-	-	na	+

OLS = ordinary least squares, QRM = quantile regression model, WLS = weighted least squares

Notes: - = negative relationship, 0 = no relationship, + = positive relationship, na = not applicable

Source: Author.

6. Conclusion

This study provides a comprehensive analysis of informal workers in Thailand by utilising the 2006-2019 Thai IES data. The study covers multiple perspectives of the informal sector, including the determinants of informal employment, the wage gap between formal and informal workers, and the relationships between informal employment and wages and informal employment and occupational risks.

First, the study utilises cross-sectional logit and probit regression analyses to examine the determinants of informal employment. The study finds that personal and job-related characteristics, i.e. sex, age, marital status, area of residence, years of schooling, working hours per week, and wage type, are the determinants of informal employment.

Second, the study applies the OB decomposition technique to examine the earnings gap between formal and informal workers. As the wage gap favours formal workers, the results confirm that informal employment is not a choice but rather an unavoidable constraint. Therefore, the finding supports the legalist approach's argument.

Third, this study then estimates Mincer wage equations at the mean (OLS regression) and at different quantiles of the wage distribution (quantile regression) to identify the factors that explain differences in earnings. The estimated results reveal that the wage gaps between formal and informal employment persist in all quartiles of the wage distribution, implying the absence of internal duality within the informal sector. The relative wages penalty of working in the informal sector is the largest for the lowest tail of the wage distribution (20th quartile).

In addition, this study employs a pseudo-panel approach with a fixed-effect model to take into account unobservable characteristics and identify the factors that explain differences in wages. The estimation finds that education does not have any effect on informal workers' wages regardless of gender. In contrast, working experience is more relevant.

Lastly, this study examines occupational risks and injury severity between formal and informal workers. The results show that informal employment, the working environment (non-behavioural risk factor), and safety issues at work (behavioural risk factor) are the main factors that significantly increase the probability of occupational injury and injury severity.

So far, the analyses point out that informal employment has adverse effects on workers' economic and social conditions, making them vulnerable to economic shocks. Therefore, policies to smooth informal workers' mobility to the formal sector is crucial. According to the analyses, policymakers can use education as a tool to improve the mobility and also to reduce occupational risks and injury severity among workers. Nevertheless, within the informal sector, working experience is more relevant than formal education. Thus, policies regarding the training of necessary skills help workers who inevitably remain in the informal sector earn more wages and, in turn, become less vulnerable to shocks.

The limitations of this paper can pave the way for future research. For one, the analysis of determinants of informal employment does not reveal workers' motivations to remain in the informal sector and their difficulties in the transition to the formal sector. Combining quantitative and qualitative analyses possibly helps us further understand the informal sector. Future research may conduct follow-up focus group interviews based on the regression results to identify drivers to remain in the informal sector and difficulties to move to the formal sector. Moreover, the qualitative framework may holistically look at the issues from different perspectives, including workers,

employers, and the government. Second, the estimation models, except the fixed-effect model with a pseudo-panel approach, may suffer from endogeneity problems. Richer data of informal workers with more rigorous estimation models will provide better estimates. However, the primary motivation of this study is to understand the common characteristics of informal workers rather than aiming for establishing causal relationships. Therefore, the study sheds light on the current situation of the informal workers and draws comprehensive policy implications based on the directions of the estimated results rather than the magnitudes.

References

- Adane, M.M., Gelaye, K.A., Beyera G.K., Sharma HR, & Yalew, W.W. (2013). Occupational injuries among building construction workers in Gondar City, Ethiopia. *Occupational Medicine & Health Affairs*, 1(5). doi:<https://doi.org/10.4172/2329-6879.1000125>
- Antman, F., & Mckenzie, D. J. (2007). Poverty traps and nonlinear income dynamics with measurement error and individual heterogeneity. *Journal of Development Studies*, 56, 125–161.
- Arias, O., & Khamis, M. (2008). *Comparative advantage, segmentation and informal earnings: A marginal treatment effects approach*. (IZA Discussion Paper 3916). Retrieved from <http://ftp.iza.org/dp3916.pdf>
- Arphorn, S., Ishimaru, T., Noochana, K., Buachum, S., & Yoshikawa, T. (2014). Working conditions and occupational accidents of informal workers in Bangkok, Thailand: A case study of taxi drivers, motorbike taxi drivers, hairdressers and tailors, *The Journal of Science of Labour*, 90(5), 183-189.
- Badaoui, E., Strobl, E., & Walsh, F. (2008). Is there an informal employment wage penalty? Evidence from South Africa, *Economic Development and Cultural Change*, 56(3), 683–710.
- Bargain, O., & Kwenda, P. (2014). The informal sector wage gap: New evidence using quantile estimations on panel data. *Economic Development and Cultural Change*, 63(1), 117–153.
- Blinder, A.S. (1973). Wage discrimination: Reduced form and structural estimates. *The Journal of Human Resources*, 8(4), 436–55. <https://doi.org/10.2307/144855>.
- Blunch, N.H. (2015). Bound to lose, bound to win? The financial crisis and the informal-formal sector earnings gap in Serbia. *IZA Journal of Labor and Development*, 4(1), 13–34.
- Botelho, F., & Ponczek, V. (2011), Segmentation in the Brazilian labor market. *Economic Development and Cultural Change*, 59(2), 437–463.
- Braithwaite, J. (1995). *From second economy to informal sector: The Russian labor market in transition* (World Bank Working Paper 21228). Washington, DC: World Bank. Retrieved from <http://documents.worldbank.org/curated/en/503731468759888818/pdf/multi-page.pdf>
- Carneiro, F., & Henley, A. (2001). *Modelling formal vs. informal employment and earnings: Microeconomic evidence for Brazil* (Management and Business Working Paper 2001–2015). Aberystwyth: University of Wales. Retrieved from https://papers.ssrn.com/sol3/Delivery.cfm/SSRN_ID289826_code011112140.pdf?abstractid=289826&mirid=1
- Chen, M.A. (2012). *The informal economy: Definitions, theories and policies* (WIEGO Working Paper 1(26)). Manchester: Women in Informal Employment: Globalizing & Organizing. Retrieved from https://www.wiego.org/sites/default/files/publications/files/Chen_WIEGO_WP1.pdf
- Chiswick, C. (1976). On estimating earning functions for LDCs. *Journal of Development Economics*, 3, 67–78.
- Christofides, L.N., Polycarpou, A., & Vrachimis, K. (2013). The gender wage gaps, ‘sticky floors’ and ‘glass ceilings’ of the European Union, *Labour Economics*, 21: 86-102.
- Corso, P., Finkelstein, E., Miller, T., Fiebelkorn, I., & Zaloshnja, E. (2006). Incidence and lifetime costs of injuries in the United States. *Injury Prevention: Journal of The International Society for Child and Adolescent Injury Prevention*, 12(4), 212–218.

- Dasgupta, S., & Bhula-or, R., & Fakthong, T. (2015). *Earnings differentials between formal and informal employment in Thailand* (ILO Working Papers 994896403402676), Bangkok: International Labour Organization. Retrieved from https://www.ilo.org/wcmsp5/groups/public/---asia/---ro-bangkok/documents/publication/wcms_436903.pdf
- Daza, N., & Gamboa, L. (2013). *An approximation to the informal–formal wage gap in Colombia 2008–2012* (Serie Documentos de Trabajo 138), Bogota: Universidad del Rosario. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.864.7644&rep=rep1&type=pdf>
- De la Fuente, V.S., López, M.A., González, I.F., Alcántara, O.J., & Ritzel, D.O. (2014). The impact of the economic crisis on occupational injuries, *Journal of Safety Research*, 48, 77-85.
- De Soto, H. 1989. *The other path: The invisible revolution in the third world*. New York: Harper and Row.
- De Soto, H. 2000. *The mystery of capital: Why capitalism triumphs in the west and fails everywhere else*. New York: Basic Books.
- De Vreyer, P., & Roubaud, F. (Eds.). (2013). *Urban labor markets in sub-Saharan Africa*. Washington, DC: World Bank.
- Deaton, A. (1985). Panel data from a time series of cross-sections. *Journal of Econometrics*, 30, 109–126.
- Duric, P., Diaconu, K., O'May, F., Rybovic, J., & Stevenson, M. (2018). Informal work-related injuries: a one year prospective study in a rural community in Serbia. *International Journal of Injury Control and Safety Promotion*. 25(4), 378-386.
- Falco, P., Kerr, A., Rankin, N., Sandefur, J., & Teal, F. (2010). *The returns to formality and informality in Urban Africa* (CSAE Working Paper 2010-03). Oxford: Centre for the Study of African Economies. Retrieved from <https://www.csae.ox.ac.uk/materials/papers/2010-03text.pdf>
- García, G.A. (2017). Labor informality: Choice or sign of segmentation? A quantile regression approach at the regional level for Colombia. *Review of Development Economics*, 21(4), 985-1017.
- Gimpelson, V.E., & Zudina, A.A. (2011). Informal Sector in Russia: Dynamics, Structure, Determinants. In: E.G. Jasin (Ed.) *XII International Conference on Economic and Social Development, Proceedings, Book 1* (pp. 444-458). Moscow: Higher School of Economics.
- Gong, X., & van Soest, A. (2002). Wage differentials and mobility in the urban labor market: A panel data analysis for Mexico. *Labour Economics*, 9(4), 513–529.
- Grilliches, Z. (1977). Estimating the returns to schooling: Some econometric problems. *Econometrica*, 45(1), 1–22.
- Guadalupe, M. (2003). The hidden costs of fixed term contracts: The impact on work accidents. *Labour Economics*, 10(3), 339–357.
- Harris, J.R., & Todaro, M.P. (1970). Migration, unemployment, and development: A two-sector analysis. *American Economic Review*, 60(1), 126–142.
- Hart, K. (1973). Informal income opportunities and urban employment in Ghana. In: R. Jolly et al (Eds.) *Third world employment: problems and strategy* (pp. 113-125). Harmondsworth: Penguin.

- Heckman, J., & Hotz, V. (1986). An investigation of the labor market earnings of Panamanian males. *The Journal of Human Resources*, 21(4), 507–542.
- International Labour Office (ILO). (1972). *Employment, incomes and equality: A strategy for increasing productive employment in Kenya*. Geneva: ILO.
- Kaya, E. (2017). Quantile regression and the gender wage gap: Is there a glass ceiling in the Turkish labor market?, *Cardiff Economics Working Papers* No. E2017/5.
- Korwatanasakul, U. (2017). Analysis of returns to schooling: Empirical evidence from Thailand. Doctoral Dissertation, Tokyo: Waseda University.
- Korwatanasakul, U. (2019). Revisiting the returns to education during the rapid structural and rural transformation in Thailand: A regression discontinuity approach. WIDER Working Paper 2019/105. Helsinki: UNU-WIDER.
- Lehmann, H., & Zaiceva, A. (2013). *Informal employment in Russia: Incidence, determinants and labor market segmentation* (Working Paper DSE N930). Bologna: University of Bologna. Retrieved from <http://amsacta.unibo.it/3795/1/WP903.pdf>
- Leigh, J.P. (2011). Economic burden of occupational injury and illness in the United States. *The Milbank Quarterly*, 89(4), 728-772.
- Leontaridi, M. (1998). Segmented labour markets: theory and evidence. *Journal of Economic Surveys*, 12(1), 103-9.
- Levine, D.I., Toffel, M.W., & Johnson, M.S. (2012). Randomised government safety inspections reduce worker injuries with no detectable job loss. *Science*, 336(6083), 907–911.
- Levy, S. (2008). *Good intentions, bad outcomes: Social policy, informality, and economic growth in Mexico*. Washington, DC: Brookings Institution Press.
- Lewis, W.A. (1954). Economic development with unlimited supplies of labor. *Manchester School of Economic and Social Studies*, 22(2), 139–191.
- Magnac, T. (1991). Segmented or competitive labor markets. *Econometrica*, 59(1), 165–187.
- Maloney, W.F. (1999). Does informality imply segmentation in urban labor markets? Evidence from sectoral transitions in Mexico. *World Bank Economic Review*, 13(3), 275–302.
- Maloney, W.F. (2004). Informality revisited. *World Development*, 32(7), 1159-1178.
- Marcouiller, D., Ruiz de Castilla, V., & Woodruff, C. (1997). Formal measures of the informal sector wage gap in Mexico, El Salvador, and Peru. *Economic Development and Cultural Change*, 45 (2), 367–392.
- Mazumdar, D. (1981). *The urban labor market income distribution: A study of Malaysia*. Oxford: Oxford University Press.
- Mincer, J. (1974). *Schooling, experience, and earnings*. New York: National Bureau of Economic Research.
- National Statistical Office of Thailand (NSO). (2020). *The Informal Employment Survey 2020*. Bangkok: NSO. http://www.nso.go.th/sites/2014en/Survey/social/labour/informalEmployment/2020/Full_Report_2020.pdf

- Nordman, C.J., Rakotomanana, F., & Roubaud, F. (2016). Informal versus formal: A panel data analysis of earnings gaps in Madagascar. *World Development*, 86, 1-17.
- Nguyen, H.C., Nordman, C.J., & Roubaud, F. (2013). Who suffers the penalty? A panel data analysis of earnings gaps in Vietnam. *Journal of Development Studies*, 49(12), 1694–1710.
- Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review*, 14(3), 693–709. <https://doi.org/10.2307/2525981>.
- Oh, J.H. & Shin, E.H. (2003). Inequalities in non-fatal work injury: The significance of race, human capital and occupations. *Social Science & Medicine*, 57, 2173–2182.
- Perry, G.E., Maloney, W.F., Arias, O.S., Fajnzylber, P., Mason, A.D., & Saavedra-Chanduvi, J. (2007). *Informality: Exit and exclusion*. Washington DC: The World Bank.
- Piazzalunga, D. & Di Tommaso, L.M. (2019). The increase of the gender wage gap in Italy during the 2008-2012 economic crisis. *Journal of Economic Inequality*, 17(2), 171-193.
- Piha, K., Laaksonen, M., Martikainen, P., Rahkonen, O., & Lahelma, E. (2012). Socio-economic and occupational determinants of work injury absence. *The European Journal of Public Health*, 23(4), 693–698.
- Pooittiwong Ramirez, A. (2017). Informal workers' gender wage gap across Thailand's regions and economic sectors (Social Science Research Network). Retrieved from https://papers.ssrn.com/sol3/Delivery.cfm/SSRN_ID2972925_code2601080.pdf?abstractid=2972925&mirid=1
- Pradhan, M., & van Soest, A. (1997). Household labor supply in urban areas of Bolivia. *The Review of Economics and Statistics*, 79(2), 300–310.
- Pratap, S., & Quintin, E. (2006). Are labor markets segmented in developing countries? A semiparametric approach. *European Economic Review*, 50(7), 1817–1841.
- Psacharopoulos, G. & Patrinos, H.A. (2004). Returns to investment in education: A further update. *Education Economics*, 12(2), 111–134.
- Roberts, B.R. (1989). Employment structure life cycle and life chances: Formal and informal sectors in Guadalajara. In A. Portes, M. Castells, & L.A. Benton (Eds.) *The informal economy: Studies in advanced and less developed countries* (pp. 41-49), Baltimore, MD: Johns Hopkins University Press.
- Rongo, L., Barten, F., Msamanga, G., Heederik, D., & Dolmans, W. (2004). Occupational exposure and health problems in small-scale industry workers in Dar es Salaam, Tanzania: A situation analysis. *Occupational Medicine*, 54(1), 42–46.
- Santana, V.S., & Loomis, D. (2004). Informal jobs and non-fatal occupational injuries. *Annals of Occupational Hygiene*, 48(2), 147-157.
- Sethuraman, SV (1976). The urban informal sector: concept, measurement and policy. *International Labour Review*, 114(1), 69-81.
- Staneva, A., & Arabsheibani, G. (2014). Is there an informal employment wage premium? Evidence from Tajikistan. *IZA Journal of Labor and Development*, 3(1), 1–24.
- Tannuri-Pianto, M., & Pianto, D. (2002). *Informal employment in Brazil—A choice at the top and segmentation at the bottom: A quantile regression approach* (Serie de Textos para Discussao 236).

Brasilia: Universidade de Brasilia. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.454.4705&rep=rep1&type=pdf>

Tansel, A., & Kan, E.O. (2016). The formal/informal employment earnings gap: Evidence from Turkey. In: J.A. Bishop & J.G. Rodríguez (Eds.) *Inequality after the 20th century: Papers from the sixth ECINEQ meeting (research on economic inequality), vol 24* (pp. 121-154). Bingley: Emerald Group Publishing Limited.

Tansel, A., Keskin, H.I., & Ozdemir, Z.A. (2015). *Is there an informal employment wage penalty in Egypt?* (ERC Working Papers in Economics 15/08). Ankara: Economic Research Center. Retrieved from <https://erc.metu.edu.tr/en/system/files/menu/series15/1508.pdf>

Tansel, A., Keskin, H.I., & Ozdemir, Z.A. (2020). Is there an informal employment wage penalty in Egypt? Evidence from quantile regression on panel data. *Empirical Economics*, 58, 2949–2979.

Tokman, V.E. (1982). Unequal development and the absorption of labour: Latin America 1950–1980. *CEPAL Review*, 17, 121–133.

Van den Broucke, S., & Colémont, A. (2011). Behavioral and nonbehavioral risk factors for occupational injuries and health problems among Belgian farmers. *Journal of Agromedicine*. 16(4), 299-310.

Verbeek, M., & Nijman, T. (1992). Can cohort data be treated as genuine panel data?. *Empirical Economics*, 17, 9–23.

Verbeek, M., & Nijman, T. (1993). Minimum MSE estimation of a regression model with fixed effects from a series of cross-sections. *Journal of Econometrics*, 59, 125–136.

Wahba, J. (2009). *Informality in Egypt: A stepping stone or a dead end?* (ERF Working Paper 456) Cairo: Economic Research Forum. Retrieved from <https://erf.org.eg/wp-content/uploads/2014/08/456.pdf>

Warunsiri, S. & McNown, R. (2010). The returns to education in Thailand: A pseudo-panel approach. *World Development*, 38(11), 1616–25.

Williams, O., Adul Hamid, R., & Misnan, M. (2018). Accident Causal Factors on the Building Construction Sites: A Review. *International Journal of Built Environment and Sustainability*, 5(1). doi:<https://doi.org/10.11113/ijbes.v5.n1.248>