ARC Centre of Excellence in Population Ageing Research

Working Paper 2024/06

Sectoral choices and household welfare in emerging economies: Evidence from Vietnam

Huyen Hoang and George Kudrna

Sectoral choices and household welfare in emerging economies: Evidence from Vietnam[1]

Huyen Hoang[2] and George Kudrna[3]

May 2024

[1] Previous drafts of this project were presented at the Australian Conference of Economists (ACE) 2022 and the 36th PhD Conference in Economics and Business 2022 at the University of Western Australia and as an UNSW economic PhD seminar. We would like to thank participants at these events for their feedback. Special thanks to John Piggott and Isabella Dobrescu for their general comments on the paper, as well as to Giang Thanh Long, Peter Robertson, Erik Hernaes, Dandan Yu, Michele De Nadai, Claudio Labanca for their comments on the econometric methodology. This project is supported under the Australian Research Council (ARC) Linkage Grant LP190100681. We would also like to acknowledge support from the UNSW School of Economics and the ARC Centre of Excellence in Population Ageing Research (CEPAR).

[2] CEPAR and UNSW School of Economics, UNSW Business School, UNSW Sydney; Email: huyen.hoang@unsw.edu.au.

[3] CEPAR, UNSW Business School, UNSW Sydney; Email: g.kudrna@unsw.edu.au.

Abstract

This study examines the effects of sectoral choices between formal and informal labour on household consumption and welfare in emerging economies. Analysing data from the Vietnam Household Living Standards Survey (2014-2018), we investigate two main questions: (1) What factors influence sectoral labour choices? and (2) How do these choices impact household consumption and welfare? We use a multinomial logit model to show that sectoral choices are primarily influenced by education level, gender, and marital status. The analysis extends to propensity score matching, supplemented by instrumental variable and multinomial endogenous switching regression models. Our results indicate that entering informal employment, particularly by low-skill workers, significantly reduces spending on food, while high-skill employment induces higher consumption of non-durable goods. Interestingly, informal employment increases housing wealth compared to low-skill formal employment, suggesting that informal workers invest in safe assets to mitigate high employment risks, while formal workers diversify their assets portfolio. The findings highlight the need for improved professional training and social security measures to facilitate transitions from informal to formal employment, enhancing household welfare.

Keywords: Informality, Sectoral choice, Structural change, Welfare, Propensity score matching, Multinomial endogenous switching regression

JEL Classification: E26, J24, J26, O17

1. Introduction

In 1971, British anthropologist Keith Hart coined the term 'informal sector' to illustrate the part of the urban labour force that worked outside the formal labour sector. It is associated with many concepts, such as 'underground economy', 'grey economy', 'black economy', and 'non-observed economy' (Hart, 1973). The informal sector has gained great significance in providing a source of labour and and substantially contributing to the GDP of countries worldwide. The informal labour force represents a large element of many countries' economies. Specifically, it is a defining feature of emerging countries, which attracts 93% of the world's informal employment (International Labour Organization (ILO), 2018; Maloney, 1999). The literature shows that the informal sector is considered residual and coexists in parallel with the formal economy. Casual employment is usually associated with employment, income, and health risks, as its jobs are typically seasonal, unstable, precarious, and lack insurance against shocks; therefore, informal labourers are highly-vulnerable groups (Alcaraz et al., 2015; World Bank, 2016). Thus, informal labourers typically seek to secure better jobs in the formal sector. While the informal sector is often thought of as having low earnings and poor employment benefits, this assessment is questionable due to the highly-heterogeneous characteristics of informal employment (Adoho & Doumbia, 2018). Therefore, an analysis of the informal sector could provide an important foundation for the earnings structure and the impact of government policies on welfare maximization.

The focus of this study is on Vietnam – an interesting case study for an emerging aging economy with a large informal labour sector which currently comprises approximately 80% of the labour force.[4] The structural changes in Vietnam's labour market can be divided into two periods. From 1986 to 2000, Vietnam transformed from a command to a market economy. This transition from closed to open international trade promoted huge resource allocation from agricultural and other low-productivity sectors to higher productivity sectors and services. Alongside this process, the reform of state-owned enterprises rendered a significant number of jobs unnecessary. Furthermore, the role of private enterprises was enhanced. The agricultural sector could no longer provide

[4] These characteristics also apply to other Southeast Asian countries, see Kudrna et al. (2022).

jobs to all workers, leading many to join the informal sector (Phan & Coxhead, 2010).

However, structural changes in the later period (2000–present) can characterized by a working sector transition based on human capital (education and workers' skills) growth. Accordingly, it is believed that the accumulation of human capital explains the structural changes in this later development phase. There are two main approaches to explaining the informality in the labour market. The first relates to the market friction, suggesting that labour participation in the informal sector is explained by a surplus of labour supply in the formal sector (Joubert, 2015). The second explains this transition by the increase in human capital (Keane & Wolpin, 1997). Interestingly, both views hold some truth and are related to the two development periods of Vietnam. In the latter phase (2000-present), amid rapidly-increasing human capital, the transition has mainly been due to the self-selection of labour. While other impacts of labour market friction also occurred, these were not the main factors leading to the labour transition in this phase. Our views, with the main assumption of self-selection in the labour market, suitably match the later development phase of labour market in Vietnam.

There is much uncertainty about the impacts of sectoral choice on household welfare in Vietnam's later development phase (from 2000 to present). On the one hand, the informal working sector might be the worst option, due to the small size of firms operating within it. As such, they offer less competitive wages compared to formal sector firms. Formal firms face the risk of enormous penalties if they are caught defaulting; hence, the majority tend to be large (Bargain & Kwenda, 2014). On the other hand, many workers are self-employed or choose to work in the informal sector because they want to avoid registration and taxation. Moreover, a sustained number of immigrants with formal jobs in the rural areas are attracted by income earning opportunities in the informal sector in urban areas (Bhattacharya, 1996, 1998). The dynamics of immigration promote the self-selection into the informal sector from rural-formal people. Hence, on the whole, informal employment could lead to higher incomes or improved welfare, thus compensating for the lost value of benefits, such as medical insurance or pensions.

This background prompts the following questions: What are the determinants of sectoral labour choices in these economies? What are the impacts of sectoral choices on household consumption and welfare? Are there any potential governmental policy implications? To answer these questions, this study explores how sectoral choice improves household consumption and welfare through four indicators: (1) household non-durable food consumption, (2) household non-durable non-food consumption, (3) household durable consumption, and (4) household housing wealth. It also highlights how these impacts differ between counterfactuals. Analysing these issues can contribute to both the existing literature and government policy. Understanding the features of informality in such emerging countries as Vietnam could help governments better manage the informal sector and to design a social security scheme covering informal workers.

The remainder of this paper proceeds as follows. Section 2 reviews the literature. Section 3 briefly describes the dataset used. Section 4 outlines the conceptual framework. Section 5 presents the analytical frameworks for our models and methodology. The results and main findings are discussed in Section 6, and the concluding remarks are presented in Section 7.

2. Literature Review

This study connects to two main strands of empirical literature:

The first relates to studies on the determinants and impact factors of employment transitions. Several studies have investigated panel surveys on the determinants of sectoral transitions between formal and informal sectors in households. Adair & Bellache (2018) examined the determinants of mobility across formal and informal sectors in Bejaia, a Mediterranean port city in Algeria, and found that human capital, age, gender, and marital status significantly affect mobility patterns. On the one hand, informal employment is considered the only choice for people who cannot obtain positions in the formal sector. Informal workers are usually unskilled, poorly educated, and categorized as unproductive (Chandra & Khan, 1993; Harris & Todaro, 1970). However, one view states that informal jobs are completely voluntary, and that workers can and do transition into the formal sector. Of course, those with diverse skillsets tend to have more opportunities, but mobility is induced by differences between the two sectors in terms of wages and associated benefits. For example, informal employment can provide high flexibility and autonomy, particularly for the self-employed. Fajnzylber et al. (2011) stated that people who are comfortable with risks may prefer to run their own businesses. Additionally, individuals who choose to work in the informal sector may expect higher salaries to compensate for the lack of benefits, such as social insurance or pension (Maloney, 1999).

The second strand relates to studies on wage differentials between formal and informal groups. There is mixed evidence of wage gaps between both sectors and there are two competing stylized views on this aspect. In the traditional view, informal workers have lower wages than their formal counterparts. Fields (1975) introduced a traditional staging hypothesis wherein formal sector employment is rationed, and the motivation to transition is determined only by wages. This model is based on wage dualism, which is in equilibrium. In this model, an individual's wages in the informal sector are lower than their potential wages in the formal sector. This could be due to the effects of minimum wage and higher unionization, which would enable the wages of informal workers to increase above market levels.

Some, however, consider both sectors to be competitive and symmetrical. Indeed, while some works may be more productive in the formal sector, others are more so in the informal sector. Many studies have indicated that informal workers may receive higher remuneration, thus decreasing the salience of the wage differentials between the two sectors (Chong & Saavedra, 1999; Maloney, 1999, 2004). This fact may hold true for developed countries, where informal (or shadow) economies are small, or the emerging economies of Latin America or Mexico, but it is less likely to be true in Southeast Asia. Pradhan & van Soest (1997) asserted that wage differentials between the two sectors are more likely to be negative than positive, and such non-monetary job characteristics as stability, social security, and healthcare access are the main factors that explain why workers prefer formal over informal employment. Additionally, researchers have found significant differences in the wage gaps between different countries in the same region or of the same type. Marcouiller et al. (1997) ran wage regressions to assess unexplained wage gaps between both sectors. Their results indicated significant wage premiums associated with formal employment in El Salvador and Peru. However, in Mexico, wage premiums were found to be associated with informal work. Pratap & Quitin (2006) found no wage premium for Argentina. Furthermore, some scholars have investigated wage differences between the sectors regarding workers' backgrounds. Gong & van Soet (2002) showed that the wage gap is not substantial for workers with less education, although it increases in line with education level.

The mixed and inconclusive findings of these theoretical and empirical studies could be ascribed to the differences in the incorporation of unobserved heterogeneity and selection bias or to unobserved characteristics being insufficiently accounted for in the models (Hamilton, 2000; Heckman et al., 1998). As unobserved skills may be associated with both the choice of working sector and workers' income, studies have applied two-stage models in which selection is jointly determined by wage regressions. To deal with unobserved characteristics that affect both labour sectoral choice and the impacts on earnings levels, Carneiro & Henley (2001) introduced a conventional approach to the simultaneous modelling of a participation decision (in formal and informal working sectors) and earnings. In the first stage, a reduced-form probit model of the formal versus informal decision is estimated, thus aiming to construct a sample selection correction term. In stage two, the results from the first stage are incorporated into conventional Mincerian semi-log earnings functions for both the formal and informal sectors. This stage enables us to control for any comparative earnings that two groups would have, as well as to account for bias, from which the sample selection effect can influence the determinants of workers' earnings. In the third stage, the results of the earnings function obtained from the second are used to predict the earning differentials between both sectors. Some studies have attempted to investigate earning distributions using quantile regressions (Bargain & Kwenda, 2014), while others have sought to address the essential heterogeneity and selectivity issues using marginal treatment effects (Arias & Khamis, 2008).

Our paper contributes to these research strands in three main ways:

First, to the best of our knowledge, this study is the first to apply both the propensity score matching (PSM) and multinominal endogenous switching regression (MESR) models to investigate the impact of labour sector choice on household consumption and welfare. Fundamentally, the MESR approach follows the same direction as Carneiro & Henley's (2001) model or the Heckman (1976) 's selection model when correcting for selection bias through the construction of simultaneous modelling of a sectoral participation decision equation and outcome. A key feature of the MESR model is that it can correct the problems of endogeneity, especially selection bias in labour sectoral transition, by including the instrumental variable (IV) in the simultaneous modelling (Lokshin & Sajaia, 2004). The method also provides us with counterfactual analysis, incoherent differences, and heterogeneity effects for two working sectors that previous related studies on the impacts of sectoral choice and welfare may not have captured sufficiently. Di Falco et al. (2011) employed the ESR model to investigate the driving forces behind farmers' decisions to adapt to climate change, and the impact of adaptation on farmers' food production through two main equations (selection and outcome). We applied our model in the same spirit. In short, employing the PSM and MESR models enabled us to effectively account for statistical problems when dealing with the issues of sectoral choices and household consumption and welfare.

Second, our study contributes to the literature by examining more facets of household consumption and welfare accounting for the consumption of food, non-food, and durable goods, and the housing wealth of households. Empirical studies of household surveys in developing countries have showed that measuring and investigating income alone could result in measurement errors due to the respondents' (especially informal workers) reluctance of hesitancy in reporting their true incomes. Hence, the welfare approach covering consumption expenditure and housing wealth may bring us closer to the true welfare of a household, rather than just counting household members' income.

Third, we employed the most updated panel dataset of the Vietnam's Household Living Standards Survey (VHLSS), which enabled us to understand the most recent structural changes in the Vietnamese economy. Lastly, our study contributes to a more insightful understanding of the impacts of informality by dividing it into different categories. Accounting for these different layers of analysis on informality allowed for a more comprehensive understanding of the impact of sectoral choice on household welfare in developing economies.

3. Data

3.1. Data set

The study utilised data on household socioeconomic status from the VHLSS, which is an ongoing longitudinal survey managed by the General Statistics Office (GSO) in Vietnam. It is conducted through a randomly-stratified sampling method that guarantees a sample that is nationally and regionally representative of the whole population. Specifically, it includes detailed questionnaires on household expenditures, education, and labour force participation, as well as other subject-specific modules with a random sub-sample of over

30,000 individuals and approximately 9,000 households in each wave. We used three waves of VHLSS data from 2014, 2016, and 2018. The 2014 sample included 9,399 households and 36,094 individuals, the 2016 sample included 9,399 households and 35,798 individuals, and the 2018 sample included 9,396 households and 35,076 individuals. Since half of the surveyed households were continuously interviewed in each round, by extracting the data of those that had employment information, we were able to construct a 3-wave balanced panel data of 1,911 households and 6,418 individuals. After narrowing the sample of working-age population[5] for the 2014-2018 period, we arrived at a final panel data of 1,592 households and 3,457 individuals. Note that all the variables extracted from our dataset are listed and defined in Table A.1 of Appendix A.

3.2. Extracting the data on the informal sector

Since the VHLSS does not survey the informal sector, we extracted the microdata relating to informal workers following the informality definition from the International Labour Organization (ILO, 2013). The ILO defines the informal sector as a sector that comprises: (1) business and production activities where do not send notices in order to evade taxes, (2) illegal economic business prohibited by law, (3) unregistered businesses or enterprises (e.g. small businesses with no employee contracts or business licenses), (4) self-consumed (or non-marketed) activities by households (e.g. household self-accumulation and self-production), and (5) any economic activities without data or information. In the three waves of the VHLSS, two questions could be used to identify and categorize the informal sector. The first related to whether waged workers had (1) a signed labour contract, (2) social insurance, or (3) paid leave and holidays. The second related to whether the workers were self-employed or not. Therefore, according to the ILO's (2013) definition of informality, if a worker had either not received any of these benefits or was self-employed, they were identified as belonging to the informal sector.

3.3. Descriptive statistics

In this subsection, we provide descriptive statistics for the employment structure and all variables by employment sector in each of the three VHLSS waves and combined

[5] Working-age population is here defined as those aged between (16-58). We considered age 58 as the retirement age, which is the average of the male (60 years) and female retirement ages (55 years).

sample.[6] Table 1 shows that informal workers on average account for roughly 80% of the total working-age population, of whom the number of low-skill workers are more than double those categorized as high-skill in the informal sector. In contrast, although formal employment only represents a small proportion of the total labour force, there is a noticeable difference between high- and low-skill formal workers. In fact, formal workers belonging to the low-skill group account for half of their high-skill counterparts.

As seen from Table 2, household consumption and welfare of low-skill informal workers are much lower than the other groups. Conversely, high-skill formal workers spend more on non-durable and durable goods than others.[7] Regarding individual characteristics, men have a greater participation in low-skill sectors than women. High-skill informal workers are, on average, older than workers in other sectors. In terms of human capital, workers in the formal labour sectors have a much higher education among all workers, whereas those in the low-skill informal sector have far lower education levels, representing nearly half the number of years spent on education compared to high-skill formal workers. This reflects the huge gap of human capital between high-skill formal and low-skill informal sectors. The informal group mostly lives in rural areas, while formal workers usually live in urban areas.

A comparison of density estimates of household food consumption among different employment types is illustrated in Figure 1. In general, formal employment have higher food consumption than their informal counterpart. Unemployed appears to have lowest household food consumption.

---

[6] In Table A.2. of Appendix A, we also provide general descriptive statistics for the variables (across all sectors) for each wave and the combined sample.

[7] Note that a limitation of the datasets used in this study is the lack of information on the non-durable and durable consumption expenditure of each individual. Instead, we used the data of the entire household for that information.

Table 1: Employment structure in Vietnam (2014-2018)

	2014		2016		2018	
	Number of workers	**Percentage**	**Number of workers**	**Percentage**	**Number of workers**	**Percentage**
Formal	363	10.5%	381	12.24%	444	14.07%
High-skill	*255*	*7.38%*	*259*	*8.32%*	*291*	*9.22%*
Low-skill	*108*	*3.12%*	*122*	*3.92%*	*153*	*4.85%*
Informal	2633	76.16%	2651	85.16%	2619	83.01%
High-skill	*734*	*21.23%*	*734*	*23.58%*	*688*	*21.81%*
Low-skill	*1899*	*54.93%*	*1917*	*61.58%*	*1931*	*61.2%*
Unemployment	461	13.34%	81	2.6%	92	2.92%
Total	**3457**	**100%**	**3113**	**100%**	**3155**	**100%**

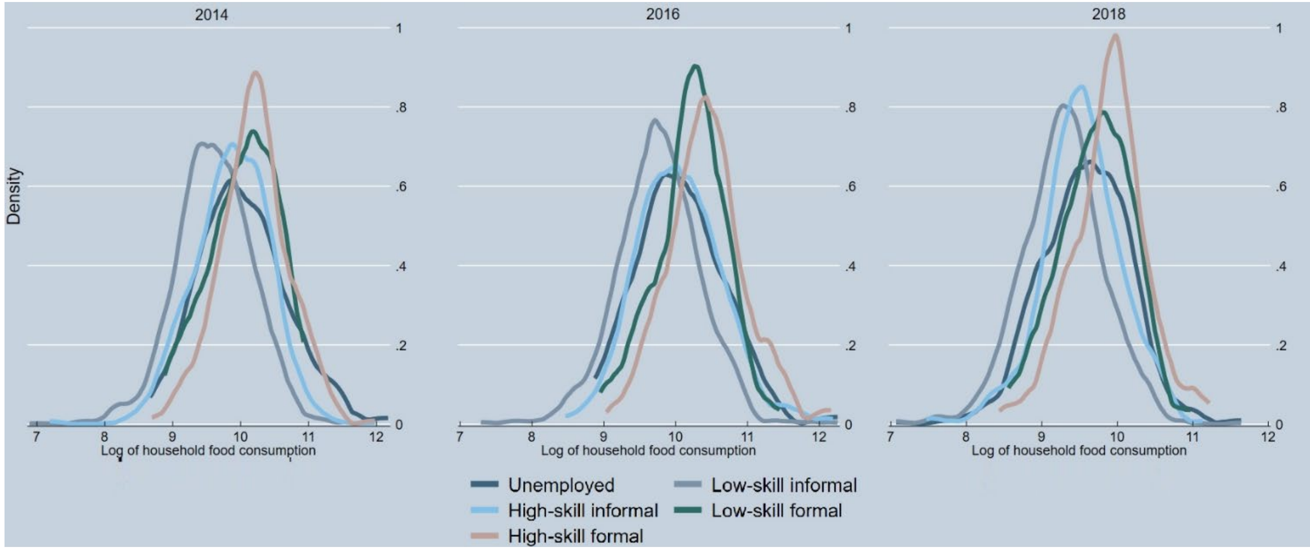
Source: Author's calculation from VHLSS (2014-2018).

Table 2: Descriptive statistics for each sector (2014-2018)

Employment sectors	Year	Age	Male	Married	Number of children	Years of Education	Urban	Health expenditure	HH non-durable consumption - food	HH non-durable consumption – non food	HH durable consumption	HH housing wealth	Informal networks
Unemployment	2014	25.11	0.39	0.28	1.11	9.22	0.41	219.54	17,139.59	5,342.35	11,266.86	210,635.50	0.41
	2016	38.58	0.11	0.75	1.10	6.98	0.47	364.57	28,983.09	10,054.27	15,245.39	361,427.80	0.41
	2018	42.21	0.08	0.80	1.38	7.71	0.48	448.43	28,555.21	9,291.42	23,877.88	391,752.60	0.39
	All	29.31	0.31	0.41	1.15	8.72	0.43	271.28	20,314.25	6,517.40	13,605.15	257,136.20	0.41
Low-skill informal	2014	37.57	0.55	0.82	1.34	6.83	0.14	167.48	12,447.17	4,037.40	7,015.00	97,163.02	0.58
	2016	39.01	0.54	0.82	1.29	6.82	0.15	302.61	20,647.16	6,435.72	12,456.09	166,884.30	0.59
	2018	40.90	0.54	0.83	1.20	7.02	0.15	348.68	17,326.27	5,718.61	13,475.41	168,213.90	0.58
	All	39.17	0.54	0.82	1.28	6.89	0.15	273.44	16,821.79	5,402.28	11,010.95	144,470.00	0.58
High-skill informal	2014	40.19	0.44	0.86	1.16	8.51	0.31	201.49	15,746.37	5,333.56	14,077.51	189,875.80	0.54
	2016	40.81	0.42	0.87	1.19	8.60	0.32	356.37	28,278.67	9,638.49	26,463.41	320,077.90	0.54
	2018	42.90	0.41	0.86	1.11	8.88	0.32	469.02	22,371.88	8,087.42	23,873.93	306,282.40	0.54
	All	41.27	0.42	0.86	1.15	8.66	0.31	339.59	22,127.19	7,677.93	21,413.33	271,624.10	0.54
Low-skill formal	2014	32.34	0.52	0.76	1.21	10.18	0.44	228.57	18,592.56	6,822.44	11,385.75	286,714.70	0.35
	2016	32.04	0.50	0.72	1.20	10.11	0.40	391.72	30,211.18	10,021.61	15,711.01	388,705.90	0.35
	2018	32.93	0.50	0.65	1.12	10.34	0.36	567.78	25,529.54	8,772.65	15,189.23	338,608.70	0.36
	All	32.48	0.50	0.70	1.17	10.23	0.39	416.05	25,064.70	8,620.56	14,282.91	339,879.00	0.35
High-skill formal	2014	35.25	0.46	0.79	1.10	14.75	0.64	290.83	22,223.68	8,295.02	19,814.50	332,152.80	0.31
	2016	36.02	0.51	0.81	1.12	14.83	0.64	510.69	39,786.83	14,535.65	37,155.78	478,527.80	0.30
	2018	35.99	0.49	0.80	1.20	15.05	0.63	626.62	30,542.78	12,452.83	32,907.39	527,397.30	0.30
	All	35.77	0.49	0.80	1.14	14.88	0.63	482.95	30,881.71	11,805.88	30,126.83	449,749.70	0.30

Source: Authors' calculation from VHLSS (2014-2018).

Figure 1: The Kernel density estimates of household food consumption for different sectors



Source: Authors' calculation from VHLSS (2014-2018).

4. Conceptual Framework: Informality and Sectoral Choice

4.1. Informality

There are two main approaches in the informal sector. One concerns the size of economic activities, while the other focuses on social security, where workers typically do have no social security protection.[8] We followed the second approach from the viewpoint of the ILO and used the VHLSS dataset to extract information on informality.

The majority of informal workers do not participate in social security systems and their daily life requires much exertion due to lower wages. Naturally, they are becoming a highly-vulnerable group in the face of idiosyncratic or covariate shocks. Therefore, this approach could promote a better understanding of of their vulnerability. Moreover, the social security view accounts for the proportion of formal labour income in income tax or social security participation, as well as the pension scheme, thereby offering an effective approach for examining the net income among formal and informal employees.[9]

Informal economic activities are popular, diverse, and easily observable in Vietnam, and range from the sale of street food to informal real estate activities. It is estimated that the share of the informal economy in Vietnam's total GDP has been between 15-27% since 2008 (Nguyen, 2019). According to the General Statistics Office of Vietnam (GSO, 2019), informal employment increased significantly (by 1.2 million) from 2014 to 2016. However, the full and systematic statistics on the contribution of informal activities to the entire economy remain unknown. Moreover, although informal activities have played an increasingly important role in the entire economy, most informal workers (97.9%) lack social insurance, with the remainder joining a social insurance scheme either on their own or with their employer's support.

As in other developing countries, the key feature of informal workers in Vietnam is their heterogeneity. Informal employment can be seen in a range of industries, such as

---

8 Informality can be defined in various ways, with numerous approaches available. For instance, when examining the extent or the margin of informality identification, Ulyssea (2020) highlighted situations where determining firm informality becomes less straightforward. In both developed and developing countries, numerous formally-registered businesses engage in tax evasion by underreporting their revenues, indicating partial adherence to tax regulations. Furthermore, when assessing informal employment, it is crucial to recognize that many formal companies hire a portion of their workforce informally to evade the expenses associated with labour regulations.

9 For a recent review of social security and pensions in emerging East and Southeast Asian countries, see Kudrna et al. (2023).

retail trading, construction industry sectors, and the real estate industry. The wages of informal workers vary depending on the region and industry. While a few informal workers receive substantial incomes, most rely on poor earnings (Cling et al., 2010). Informal work usually involves long hours under undesirable working conditions and a lack of protection from labour laws due to their unregistered employment. The majority possess possess low skills and have few opportunities to upgrade them. Moreover, regarding labour and employment, their voices are often ignored when designing policies.

4.2. Sectoral labour choice

Whether an individual decides to work formally or informally mostly depends on two factors: (i) 'push' factors, such as their skills, technology shocks, and work experience, and (ii) 'pull' factors, such as the expected value of the lifetime rewards, they can gain in the informal sector relative to the formal sector. Empirical data show that low-skill workers tend to jump into the informal sector, while high-skill workers gain employment in the formal sector. Those seeking to move from a formal to an informal job must consider the informality cost, such as the loss of a health insurance package or pension. In other aspects, skill accumulation encourages individuals to join the formal sector with the notion of getting a stable job, a chance of promotion, and benefits from social security policies. Moreover, the movement of workers from the informal to the formal sector could be driven by the spread of manufacturers in industrial zones, where labourers are attracted to work in local factories. This process can be promoted by governmental/provincial policies.

We divided formal and informal workers into high- and low-skill informal groups. The former included jobs requiring professional training. We investigated the determinants and the impacts of the working sectoral choices on household welfare.

5. Methodology

We developed an econometrics model to investigate (1) the determinants of sectoral labour choices and (2) the impacts of sectoral choices on household consumption and welfare. This section first discusses the main factors that determine the labour sectoral choice by employing the multinomial logit model (MNL). Then, we use the propensity score matching method to examine how these choices impact household consumption and welfare. Lastly, the multinominal endogenous switching regression model and the IV method are obtained to serve as a robustness check to confirm the former results.

5.1. Multinomial logit model (MNL)

The purpose of the first part of the study is to assess the determinants of sectoral labour choices. An advantage of our data is that it allowed us to categorize employment into high and low-skill groups – the former of which requires professional training, while the latter does not. We divided formal working employment into high- and low-skill, and similarly, informal employment into high- and low-skill. One of the most effective and widely used approaches to investigating the determinants of several categories is to employ the MNL, wherein the dependent variable can fall into these possible categories. In other words, it compares the probability of membership in other categories to the probability of membership in the reference (or baseline) category. The MNL has been widely applied to investigate educational or employment choices (Stratton et al., 2008; Cameron et al., 2023).

In this model, the dependent variables were the observed outcomes, which were the choices of working sector: having a high-skill formal job, having a low-skill formal job, having a high-skill informal job, having a low-skill informal job, and being unemployed. The explanatory variables were the individual's age, gender, marital status, number of children, years of education, urban area, health expenditure, household non-durable consumption (food and non-food consumption) and household durable consumption, and housing wealth.

To model working sectoral choices, we assumed that individuals could choose from a set of alternatives. Thus, the MNL model with both the conditional fixed-effects and random effects estimator can be specified in utility-maximization form as:

$$U_{jit} = X_{it}\beta_j + u_{ij} + \epsilon_{jit}, \tag{1}$$

where U_{ijt} denotes the utility of individual i in selecting the type of employment category j at time t, with $i = 1, \dots, N$, $j = 1, \dots, J$ and $t = 1, \dots, T_i$. $X_{it}\beta_j$ represents the observed component of utility in which X_{it} denotes a row vector of covariates, such as family background and individual characteristics; and β_j is the column vector of coefficients for outcome j. The unobserved components include u_{ij} and ϵ_{jit} which are the panel-level heterogeneity term and observation-level error term, respectively.

Although we could not observe U_{ij} (i.e., the latent variable), we could infer from the individual choices, which reflect how they rank these alternatives. We observed the polychotomous variable denoted as m, which has values ranging from 1 to 5 that correspond to the 5 types of employment categories that we extracted from the data, defined as follows:

- Unemployment ($m = 1$)
- Low-skill informal employment ($m = 2$)
- High-skill informal employment ($m = 3$)
- Low-skill formal employment ($m = 4$)
- High-skill formal employment ($m = 5$).

The probability that the i[th] individual would choose the outcome employment m at the time t is that:

$$\Pr(y_{it} = m | x_{it}, \beta_j, u_{ij}) = F(y_{it} = m, x_{it}\beta_j + u_{ij}) = \begin{cases} \frac{1}{1+\sum_{j=2}^{J=5} \exp(x_{it}\beta_j + u_{ij})} & if\ m = 1 \\ \frac{\exp(x_{it}\beta_m + u_{im})}{1+\sum_{j=2}^{J=5} \exp(x_{it}\beta_j + u_{ij})} & if\ m > 1 \end{cases} \quad (2)$$

where $F(.)$ presents the cumulative logistic distribution function. Following the formula in equation (2), the random- and fixed-effects estimators distinguish themselves in their assumption about the unobservables in u_{ij}.

5.2. Propensity score matching (PSM) method

The second part of our study focuses on evaluating the impacts of sectoral choice on the household's consumption and welfare. Accordingly, we employed the PSM as our main approach. The PSM method aims to reconstruct the counterfactuals based on a framework for estimating the probability of joining a programme that is conditional on the observed characteristics of different groups. It compares the expected outcomes between the samples of the comparison group and those with similar observable characteristics. The validity of this approach is based on two assumptions.

The first reflects conditional independence, thus indicating that given the observable characteristics in pre-treatment, the outcomes of both non-participants and participants are assumed to be independent of the treatment assignment (Lechner, 2002). The second is the overlap condition assumption, which assumes that a substantial overlap of the covariates between nonparticipants and participants guarantees that those with common covariate values obtain positive values of the probability of being participants or non-participants (Becker & Ichino, 2002). Fundamentally, PSM generates estimations by processing matching observations from the treated and control groups, based primarily on their estimated propensity scores. Based on these assumptions, our PSM method adhered to the following steps:

First, we categorised two main groups: the control group which included the unemployed and the treated group that had employments (low-skill, high-skill informal sector or low-skill, high-skill formal sector).

Second, we estimated the propensity score for each group of individuals. Commonly, a probit or logit model is employed for this approach. Based on that, we selected the observed covariate X that influenced the likelihood of being assigned to the treated group. We thus estimated the propensity score by a logit model.

Third, we obtained the propensity score as the conditional or predicted probability of having the treatment given pre-treatment characteristics X:

$$p(X) = prob(T = 1|X) = E(T|X), \tag{3}$$

where X denotes a vector of the individual's characteristics and $p(X)$ represents the propensity scores given X. Next, treatment T denotes a binary variable that is equal to 1 if individuals are employed (low- and high-skill formal and informal sectors) and 0 if they are unemployed.

Fourth, we conducted a matching process in which each participant belonging to the treated group was matched to one or more non-participants in the control group based on their propensity scores.

Fifth, we measured the treatment effects by comparing the outcomes y between the observations from the treated group and control group after matching:

$$Y = \begin{cases} Y_1 \; if \; T = 1 \\ Y_0 \; if \; T = 0 \end{cases} \tag{4}$$

After gaining the propensity scores, we estimated the ATT as the difference in mean between the outcome of the participants $(Y_1|T = 1)$ and non-participants $(Y_0|T = 0)$:

$$ATT = E(Y_1|p(X), T = 1) - E(Y_0|p(X), T = 0). \tag{5}$$

5.3. Multinominal endogenous switching regression (MESR) model

To check the robustness of the PSM method, we employed the MESR model to assess the impact of an individual's working sectoral choice on their household consumption and welfare. We advanced that working sectoral choices mainly depend on people's own decisions. Hence, the problem of selection bias arises from their selection. The MESR model is considered to be one of the most effective for accounting for the problem of selection bias (Malikov et al., 2018; Nahm et al., 2017; Di Falco et al., 2011; Lokshin & Sajaia, 2004). The model includes two main stages: the selection and outcome equation stages. The selection equation stage relates to the decision of working sectors that individuals choose to join. This stage is followed by the MNL (as in Section 5.1.) and applied to the cross-sectional data. The second stage aims to evaluate the impact of each sectoral choice on the outcome variables (household welfare). It should be noted that, in the MESR model, the IV must be included in the explanatory variables in the selection equation stage. We provide more details of how to we select the appropriate IV at the end of this section.

Stage 1: Selection equation of employment choice:

$$U_{ji}^* = X_{ji}\alpha + \epsilon_{ji}, \tag{6}$$

$$\text{with} \qquad \mathrm{U}_i = \begin{cases} 1 \quad if\ U_{ji}{}^* > \max\limits_{m \neq 1}(U_{mi}{}^*) \text{ or } \tau_{1i} < 0 \\ \quad \cdots \\ \quad \cdots \\ \quad \cdots \\ \mathrm{J} \quad if\ U_{ji}{}^* > \max\limits_{m \neq j}(U_{mi}{}^*) \text{ or } \tau_{ji} < 0 \end{cases} \tag{7}$$

Stage 2: Outcome equation:

Regime 1

(Individual who is unemployed): $Y_{1i} = Z_{1i}\beta_1 + \varepsilon_{1i} \quad if \quad \mathrm{U}_i = 1$ (8)

Regime 2

(Individual with sectoral labour choice) $Y_{ji} = Z_{ji}\beta_j + \varepsilon_{0i} \quad if \quad \mathrm{U}_i = \mathrm{J}$ (9)

j=2,3,4,5

The selection of the working sector in equation (6) reflects the difference in the expected value of taking the sectoral transition and the expected value the individuals would have if they remained in their previous working sector. $U_{ij}{}^*$ denotes the latent variable, which captures the expected benefits from the choice of working in labour sectors regarding the status of being employed. An individual will decide to work in the formal or informal sector if the expected value of working in this sector is higher than that of the other. U_i will equal 1 if an individual is unemployed and 0 if they are employed. τ_{1i} represents the difference $\tau_{ji} = \max\limits_{m \neq 1}\left(U_{mi}{}^* - U_{ji}{}^*\right)$. ϵ_{ij} is assumed to be independent and identically Gumbel distributed. Z_{ij} is a vector of observed exogenous covariates that demonstrate individual and household characteristics, including age, gender, marital status, number of children, years of education, urban area, health expenditure, and informal networks. Especially, the covariates in this case must also include the IV which is the variable of informal network.

Equations (8) and (9) demonstrate endogenous switching regressions that correct the selection biases and structural differences between working sectoral choices (Lokshin & Sajaia, 2004). The full-information ML estimation method (FIML) was used to estimate these equations. An individual encounters two regimes: transit to another working sector (regime 2) or remain employed (regime 1). Y_{ji} and Y_{1i} are the outcomes of household consumption and welfare. Z_{ji}, Z_{1i} represent the vectors of weakly exogenous variables. The sets of Z_{ji} from the outcome equation and variables X_{ji} from the selection equation could overlap, and one important condition is that the IV must be included in X_{ji}, but excluded from the Z_{ji}.

One of the main aims of the MESR model is to estimate ATT (Di Falco et al., 2011; Heckman et al., 1998). After the estimates from equations (8) and (9), its procedure of estimates is conducted by comparing the expected values of outcomes of those who participate in the labour markets and those who are unemployed in actual and counterfactual, which are respectively given in equations (10) and (11).

Actual case (employed workers with their chosen labour sectoral choices):

$$E(Y_{ji}|U = j, Z_{ji}, \hat{\lambda}_{ji}) = Z_{ji}\beta_j + \sigma_j\hat{\lambda}_{ji} \tag{10}$$

Counterfactual (employed workers had they decided not to work or if they are unemployed):

$$E(Y_{1i}|U = j, Z_{ji}, \hat{\lambda}_{ji}) = Z_{ji}\beta_1 + \sigma_1\hat{\lambda}_{ji} \tag{11}$$

From that, the ATT is calculated as the difference between equations (11) and (10):

$$ATT = E(Y_{ji}|U = j, Z_{ji}, \hat{\lambda}_{ji}) - E(Y_{1i}|U = j, Z_{ji}, \hat{\lambda}_{ji}) = Z_{ji}(\beta_j - \beta_1) + \hat{\lambda}_{ji}(\sigma_j - \sigma_1),$$

where $\hat{\lambda}_{ji}$ and $\hat{\lambda}_{1i}$ denote the inverse mill ratios derived from the selection equation to account for possible selection biases.

5.4. The selection of instrumental variables (IV)

To overcome the endogeneity problem, we employed the IV approach. A valid IV must satisfy two important conditions: relevance and exogeneity. In other words, a suitable IV must have an influence on the labour sector choices, but no direct impact on either household consumption or welfare. For the first condition, we tested the IV through the first-stage regression. However, the second required further explanation. In our study, we considered the network of informality as a potential IV for the model. This means the informality network was calculated as the proportion of workers in the informal sector to the population in the commune-level (district) units.

Concerning the condition of relevance, previous studies have shown that networks could significantly influence individuals' participation decisions in labour markets (Kajisa, 2007; Torezani et al., 2008). For example, concerning non-farm employment, Kajisa (2007) argued that non-farm networks reduce the search and transaction within job searches, meaning that they level-up the opportunities and individuals' accessibility to non-farm employment. Similarly, Brünjes & Revilla Diez (2016) show that family contacts play a crucial role in the process of discovering non-farm employment participation. Furthermore, Goncalves & Martins (2021) argued that the proportion of workers who are self-employed in a given district captures the structure of the labour market in the area or demographic group. For example, there might be a predominant industry in a district that relies on wage workers, or a new service could emerge which attracts young, self-employed workers. The higher or lower proportion of informal employment in the district would affect the opportunities to join the informal sectors based on the features and characteristics of the area's labour market. Therefore, for either informal or non-farm employment sectors, the case differs only slightly, with the fundamental argument being that the increase in individual networks in labour working sectors will affect the other individuals' decision to join them.

Regarding the condition of exogeneity, one might assume that the existence of an informal network reflects different levels of development and characteristics within the region or area to which this network belongs. Several studies have demonstrated the exogeneity of the informal networks as the IV in the case of developing countries (Goncalves & Martins, 2021; Noseleit, 2014; Torezani et al., 2008; Vu & Rammohan, 2022). These have indicated the salience of possible factors which relate to the area's characteristics and affect both a household's informal network and household consumption and welfare. To investigate this issue, we sought to find the correlation between informal networks with the other variables that related to the districts (or commune) characteristics and verify the strength of the relationship between this correlation and household consumption and welfare – with a weak relationship meaning that the IV satisfies the exogeneity condition. For example, to prove the exogeneity of the

IV, Goncalves & Martins (2021) controlled the district fixed effects. District fixed effects here should consider any district characteristics that correlate with both the instrument and the outcome, household income, and hospitalizations, as long as those characteristics are constant over time. To explore this issue further, the authors examined the evolution of certain district characteristics: general income index, general health index, and firm dimension index. They observed all of these indices to be fairly constant over time, meaning that such characteristics should be appropriately captured by the district fixed effects. Their results showed that, after controlling for the district characteristics, the informal networks had no direct impact on household income or welfare. Therefore, the variable was proven to satisfy the exclusion restriction conditions for the IV (Goncalves & Martins, 2021). Similarly, Vu & Rammohan (2022) used the variable "distance from commune centre to the nearest market" and "distance from commune centre to the nearest city" as the proxy for the commune characteristics in rural areas, and investigated whether the correlation between these variables and informal networks of non-farm employment significantly impacts the household food security. The authors found the correlation to be weak, indicating the exogeneity of informal networks of non-farm employment in rural areas. Following the approaches from Goncalves & Martins (2021) and Vu & Rammohan (2022), we controlled the district fixed effects and accounted for any district characteristics that correlated with both the instrument and the outcome – household consumption and welfare. We generated two variables as a proxy for the commune's characteristics, namely "distance from the commune to the market" and "distance from the commune to the city centre", and examined whether the correlation between the informal network and these variables affect household consumption and welfare. The results in Table B.1 (Appendix B) show a weak correlation, meaning that the IV satisfied the exclusion condition.

To further understand the IVs, we conducted several tests, including the Stock & Yogo (2005) test, to determine their strength. Table B.2 demonstrates the IV tests for the informality network (in the form of IV). The Wu-Hausman test indicates a small P-value, prompting us to reject hypothesis H_0 that all variables are exogenous, and the model has an endogeneity issue. The first-stage regression test targeted the strength of the IVs, aiming to check whether they correlated with the endogenous variable, with the result indicating no correlation with the disturbances. Stock & Yogo (2005) strongly emphasized the importance of this test because weak IVs might result in the distortion of hypothesis tests and the inconsistent estimation of the parameters. The results from the first-stage regression test (see Table B.2) show that our IV is not weak. This fact is confirmed by the Cragg & Donald (1993) minimum eigenvalue statistics, two-stage least squares (2SLS) size, and the LIML size of the nominal 5% Wald test. The eigenvalue statistics of the instrument are higher than the critical values of 16.38 at a 10% rejection rate and 8.96 at a 15% rejection rate, respectively. The LIML estimation also confirms this fact. Therefore, based on both empirical analysis and econometric theory, the informal networks variable proved to be appropriate for our model.

6. Estimation Results

This section presents the results. We discuss the key factors that influence the choice of working sector, before considering the impacts of sectoral labour choice on household consumption and welfare.

6.1. Determinants of labour sectoral choice

This section provides empirical evidence on the factors that affect an agent's working sectoral choice. It examines how the characteristics of the whole family and agent impact their sectoral choices using a MNL regression model for the panel dataset.

The MNL illustrates the estimated, conditional, and transition probabilities with the status of being unemployed as the baseline. Table 3 shows that gender has a significant impact on the choice of sectoral employment. Specifically, males are more likely to be willing to take on low-skill jobs, possibly due to their physical strength, which is often a requirement for such positions. Moreover, the family's financial pressure is often the responsibility of men, who are typically regarded as the breadwinners. This might force them into low-skill industries where employment can quickly be found. In contrast, married women who do not satisfy the requirements for a manual job could instead become homemakers. Moreover, marital status affects the choice of labour sector. Married people are those who most wish to obtain a high-skill, formal sector job, likely due to earning stability and a more secure work–life balance. Unmarried people, however, are more flexible and less restricted in seeking a job. It was also observed that education attainment has significant and positive impacts on the choice of employment sector. An additional year of education motivates people to seek a (high-skill) job in the formal sector. This seems reasonable as educated people have greater opportunities to secure high-paid positions and higher chances to be promoted. In rural areas, people are more likely to have low levels of education, which results in greater constraints on securing well-paid jobs. Consequently, these people are usually satisfied with a low-skill informal job. It should be noted that health expenditure, household consumption, and housing wealth seemed to have insignificant influences on sectoral choices.

Table 3: Determinants of sectoral choice

Multinominal logistic model								
Unemployed as the baseline								
	Low-skill informal		**High-skill informal**		**Low-skill formal**		**High-skill formal**	
Random vs. fixed effects	**R**	**F**	**R**	**F**	**R**	**F**	**R**	**F**
Age	**0.090*****	**0.182****	**0.111*****	**0.163****	-0.003	**0.180*****	**0.037*****	**0.176*****
	(0.008)	(0.039)	(0.009)	(0.042)	(0.014)	(0.049)	(0.013)	(0.0477)
Male	**2.058*****	2.636	**0.903*****	1.746	**1.355*****	**2.180*****	**1.151*****	**1.68****
	(0.162)	(0.593)	(0.168)	(0.593)	(0.255)	(0.671)	(0.240)	(0.676)
Married	**1.773*****	0.098	**1.820*****	-0.172	1.558***	-0.511	2.526***	0.365**
	(0.192)	(0.480)	(0.208)	(0.523)	(0.327)	(0.645)	(0.327)	(0.615)
Number of children	0.046	-0.745	0.043	-0.582	-0.125	**-0.971*****	-0.023	**-7.03****
	(0.077)	(0.222)	(0.082)	(0.229)	(0.137)	(0.32)	(0.132)	(0.318)
Years of education	**-0.093*****	**0.036***	**0.092*****	**0.147***	**0.182*****	**0.227****	**0.781*****	**0.368*****
	(0.020)	(0.063)	(0.021)	(0.065)	(0.035)	(0.089)	(0.047)	(0.0797)
Urban	**-2.435*****	-15.455	**-0.972*****	-0.909	**-0.876*****	-16.452	0.304	-77.813
	(0.193)	(657.432)	(0.185)	(1657.854)	(0.304)	(3055.943)	(0.274)	(2039.084)
Health expenditure	$-1.27.10^{-4}$***	$1.93.10^{-4}$***	$9.54.10^{-5}$***	$1.47.10^{-4}$***	$1.6.10^{-4}$***	$-1.2.10^{-5}$	$1.62.10^{-4}$	$11.73.10^{-5}$
	($1.08.10^{-4}$)	($1.72.10^{-4}$)	($1.09.10^{-4}$)	($1.71.10^{-4}$)	($1.29.10^{-4}$)	($1.00.10^{-5}$)	($1.2.10^{-4}$)	($18.56.10^{-5}$)
HH non-durable	$-1.09.10^{-5}$***	$1.09.10^{-5}$***	$-2.68.10^{-6}$***	$1.47.10^{-5}$	$9.04.10^{-6}$***	$3.22.10^{-5}$	$4.04.10^{-6}$	$-2.10.10^{-6}$
consumption -food	($5.48.10^{-6}$)	($9.88.10^{-6}$)	($5.26.10^{-6}$)	($9.88.10^{-6}$)	($7.84.10^{-6}$)	($1.78.10^{-5}$)	($6.81.10^{-6}$)	($11.9.10^{-6}$)
HH non-durable	$5.91.10^{-6}$***	$7.84.10^{-5}$***	$5.99.10^{-5}$***	$8.73.10^{-5}$**	$1.03.10^{-4}$***	$8.8.10^{-5}$***	$1.03.10^{-4}$***	$9.07.10^{-5}$**
consumption – non-food	($1.99.10^{-5}$)	($3.38.10^{-5}$)	($1.91.10^{-5}$)	($3.36.10^{-5}$)	($2.68.10^{-5}$)	($4.89.10^{-5}$)	($2.27.10^{-5}$)	($3.91.10^{-5}$)
HH durable consumption	$-5.97.10^{-6}$***	$3.58.10^{-6}$***	$2.95.10^{-6}$***	$1.30.10^{-5}$**	$-1.49.10^{-5}$***	$2.29.10^{-5}$	$-3.66.10^{-6}$***	$1.73.10^{-5}$*
	($2.94.10^{-6}$)	$4.89.10^{-6}$	($2.74.10^{-6}$)	($6.35.10^{-6}$)	($6.73.10^{-6}$)	($1.43.10^{-5}$)	($3.45.10^{-6}$)	($8.66.10^{-6}$)
Housing wealth	$-5.37.10^{-7}$***	$1.14.10^{-6}$***	$-2.58.10^{-7}$***	$8.16.10^{-7}$	$8.14.10^{-8}$***	$9.8.10^{-7}$***	$-2.84.10^{-7}$***	$1.14.10^{-6}$
	($2.19.10^{-7}$)	($5.58.10^{-7}$)	($2.04.10^{-7}$)	($5.31.10^{-7}$)	($3.01.10^{-7}$)	($7.96.10^{-7}$)	($2.59.10^{-7}$)	($6.92.10^{-7}$)
Observation	8975	3064	8975	3064	8975	3064	8975	3064
Wald chi2 (44)	1240	253	1240	253	1240	253	1240	253
Prob>chi2				0.0000				

Notes: *Mean statistically significant at 10%; **Mean statistically significant at 5%; ***Mean statistically significant at 1%; [a] Informality networks are calculated as the proportion of informal workers over the population at the commune-level.
Source: Authors' calculation from VHLSS (2014-2018).

6.2. Impacts of employment choice on household consumption and welfare

Tables 4 and 5 indicate the treatment effects of sectoral labour choice on household consumption and welfare by applying the PSM method. The results indicate that, if an individual moves to the formal sector, their household food consumption and welfare would increase significantly more than if they were to remain unemployed. Moreover, if the individual has low-skill informal employment, their household would spend less on food expenditure than if they stayed unemployed. However, moving from unemployed status to a low-skill informal sector would enable their household to accumulate more housing wealth. This indicates an interesting direction of consumption and welfare effects, reflecting the impact of low-skill employment on the unemployed group. If an unemployed individual moved to high-skill informal employment, their household would spend more on durable goods. However, if they moved to a low-skill formal sector, the durable goods consumption of their household would be less than if they were unemployed. Moreover, moving to either the informal or formal sector would help the household improve their non-durable (non-food) consumption. In particular, the highest impacts are found when high-skill formal employment is used as the treatment variable.

Table 4: Treatment effects of sectoral choice on household consumption and welfare by the PSM method (2014-2018)

Outcome: Household non-durable: food consumption			
	2014	**2016**	**2018**
Unemployed	-	-	-
Low-skill informal	423	2289	**-5868***
	(813)	(2106)	(2246)
High-skill informal	-891	4761	-354
	(1190)	(4203)	(2473)
Low-skill formal	-201	**5542**	-5672
	(1453)	(2241)	(5304)
High-skill formal	-1442	**13147***	-15585
	(4263)	(3827)	(14401)
Outcome: Household non-durable: non-food consumption			
	2014	**2016**	**2018**
Unemployment	-	-	-
Low-skill informal	**643**	-535	-1281
	(257)	(1528)	(2110)
High-skill informal	385	-26	1012
	(378)	(992)	(934)
Low-skill formal	**1522***	120	685
	(436)	(932)	(935)
High-skill formal	1543	**3494***	1233
	(1183)	(1055)	(1330)

Notes: *Mean statistically significant at 10%; **Mean statistically significant at 5%; ***Mean statistically significant at 1%; All specifications control for age, gender, marital status, number of children, years of education, urban area, health expenditure, and informal networks.
Source: Authors' calculation from VHLSS (2014-2018).

Table 5: Treatment effects of sectoral choice on household consumption and welfare by the PSM method (2014-2018)

Outcome: Household durable consumption			
	2014	**2016**	**2018**
Unemployment	-	-	-
Low-skill informal	301	-1368	-934
	(1996)	(3355)	(1494)
High-skill informal	-1061	**8708*****	5691
	(6078)	(2706)	(3665)
Low-skill formal	-2225	-328	**-15064****
	(3180)	(1939)	(7578)
High-skill formal	-30107	10602	-34105
	(20395)	(4627)	(23373)
Outcome: Housing wealth			
	2014	**2016**	**2018**
Unemployment	-	-	-
Low-skill informal	8150	-42947	**43421****
	(14498)	(47018)	(19629)
High-skill informal	-7115	36663	**94105*****
	(31102)	(65680)	(31587)
Low-skill formal	14157	69501	**13187*****
	(60957)	(65923)	(45070)
High-skill formal	-132312	**206874*****	**174275****
	(115248)	(56377)	(75348)

Notes: *Mean statistically significant at 10%; **Mean statistically significant at 5%; ***Mean statistically significant at 1%; All specifications control for age, gender, marital status, number of children, years of education, urban area, health expenditure, and informal networks.

Source: Authors' calculation from VHLSS (2014-2018).

6.3. Robustness checks

We verified the robustness of the results from the PSM model by employing the IV, ordinary least squares (OLS), and MESR methods. We begin by reviewing the results of the OLS and IV model before reviewing the results from the MESR model. In this study, IVs followed the 2SLS procedure. One of the main the advantages of MESR over IV is its ability to give ATT, while IV can only give LATE (local average treatment effect). It should also be noted that, in this case, the IVs must be included in the MESR model, specifically in the selection equation stage (as mentioned in Section 5.3.).

As shown in Table C.1 (Appendix C), the results from the OLS and IV models with the informality network as the IV indicate the positive impacts of the choice of working in the formal sector on household consumption and welfare. The impact of formal working sectoral choice is much higher when we apply the IV method, thus indicating this method's efficiency in dealing with the problem of endogeneity. Tables C.2, C.3, and C.4 illustrate the results of the ATT (in 2014, 2016 and 2018, respectively) from the MESR model, with the same implications as in the PSM results. However, despite these same implications, the magnitudes of the estimates differ. These differences could be explained by the fact that the PSM technique does not fully capture the unobserved heterogeneity between two comparison (treated and control) groups, while the MESR does take these unobserved heterogeneities into consideration. In other words, MESR and IV can be employed to address both selection on observables and unobservables. However, the PSM can only be used to tackle selection on observables. This is the most significant difference between the PSM and the other methods.

6.4. Summary and discussion

From the PSM, OLS, IV and MESR models, we arrive at the results in Table 6, which provides the whole picture of average treatment effects on the treated ATT of sectoral labour choices and the summary results with the signs and degree levels (magnitude orders) across the models used. The magnitude orders indicate the orders of comparative degree between the ATT of each sector for each category of household consumption and welfare, using the unemployed status as the baseline. They are real numbers and do not take the form of absolute values. The results indicate that, with the unemployed as the baseline, those with high-skill employment have higher non-durable consumption (including food and non-food consumption). This is understandable as high-skill workers typically receive high payments and tend to be associated with higher education attainment, which generally results in a greater focus on nutrition and general health. Hence, more of their income is invested in food consumption, as well as healthy and environmentally-friendly non-food products. In contrast, low-skill informal workers face greater disadvantages, meaning they are more likely to spend less money on consumption.

Table 6: Summary of the signs and degree levels (magnitude orders) of average treatment effect on the treated – ATT (2014-2018) from PSM, OLS, IV and MESR models

	Food consumption		**Non-food consumption**		**Durable consumption**		**Housing wealth**	
	Sign[a]	*Magnitude order[b]*	*Sign[a]*	*Magnitude order[b]*	*Sign[a]*	*Magnitude order[b]*	*Sign[a]*	*Magnitude order[b]*
Low-skill informal	-	1	+	1	+	3	+	2
High-skill informal	+	2	+	3	+	4	+	3
Low-skill formal	+	3	+	2	-	2	+	1
High-skill formal	+	4	+	4	-	1	+	4

Note: [a] It indicates the sign of ATT that we can get from the PSM and MESR; [b] The magnitude order indicates the orders of comparative degree between ATT of each sector for each category of household consumption and welfare, with unemployed status as the baseline; They are real numbers and do not take the absolute-value forms.
Source: Authors' calculation from VHLSS (2014-2018).

As indicated, with the unemployed as the baseline, workers who join the informal employment sector tend to spend more on durable consumption than those in the formal employment sector. This can be explained by the fact that formal workers are more likely to receive benefits from their employers. For example, they may commute to work by company bus, thus saving them travel expenses. Joining either formal or informal employment also helps in increasing housing wealth. Noticeably, those with high- and low-skill informal employment have higher housing wealth than those in low-skill formal groups. This finding is consistent with previous empirical results, as it shows that informal workers compensate for their high employment risk by investing in safe assets (e.g., a house), while those in the formal sector are more likely to diversify their asset portfolios. Granda & Hamann (2015) and Schclarek & Caggia (2015) supported the idea that, due to more uncertain and variable incomes, informal workers tend to have higher saving rates for precautionary reasons and invest in real estate as a safe method for preserving their finances for an unpredictable future.

7. Conclusion

This paper has studied the determinants and impacts of individuals' sectoral choices on household consumption and welfare. Using an MNL model, the results show that low education attainment pushes people into low-skill or informal jobs, which lead to low household welfare. The findings also reveal that the choice of employment largely depends on gender, marital status, and level of education. The average treatment effect results show that joining the low-skill informal sector leads to low household spending on food, while high-skill employment results in a high level of spending on non-durable goods. Interestingly, high-skill informal work leads to higher housing wealth compared to other groups, which shows that informal workers compensate their high employment risks by investing in safe assets (e.g., housing), while those in the formal sector tend to diversify their assets portfolios.

In terms of policy implications, low-skill informal workers are highly-vulnerable groups as they tend to be poor and lacking in social protection. This situation seems to create a vicious circle of low education, productivity, income, and investment in education. To break this cycle, the government has an important role in providing support in education and vocational training. Providing financial assistance to low-skill workers to improve their professional skills is key in enabling access to gainful jobs. This support should focus on rural areas and those with high levels of informal networks where the majority of low-skill and informal workers gather. Additionally, as informal workers are not covered by social insurance and are employed in high-risk jobs, there is tendency to invest in safe assets. We believe that extending social security coverage to informal workers would improve their access to essential services, such as health care and education, and also ensure access to necessary daily nutrition, which would contribute to both greater labour productivity and human welfare.

References

Adair, P., & Bellache, Y. (2018). Labor market segmentation and occupational mobility in Algeria: Repeated cross-sectional and longitudinal analyses (2007 to 2012). *Review of Development Economics* 22(4), 1765–1783.

Adoho, F.M., & Doumbia, D. (2018). Informal sector heterogeneity and income inequality: Evidence from the Democratic Republic of Congo. Policy Research Working Paper, No. WPS 8328. World Bank Group.

Alcaraz, C., Chiquiar, D., & Salcedo, A. (2015). Informality and segmentation in the Mexican labour market. Working Papers 2015-25, Banco de Mexico.

Arias, O., & Khamis, M. (2008). Comparative advantage, segmentation, and informal earnings: A marginal treatment effects approach. IZA Working Papers No. 3916.

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.

Becker, S.O., & Ichino, A. (2002). Estimation of average treatment effects based on propensity scores. *Stata Journal* 2(4), 358–77.

Bhattacharya, P.C., (1996). The role of the informal sector in structural transformation: Some Indian evidence. *Journal of International Development* 8(1), 83–94.

Bhattacharya, P.C., (1998). Migration, employment, and development: A three-sector analysis. *Journal of International Development* 10(7), 899–921.

Brünjes, J., & Revilla Diez, J. (2016). Obtaining non-farm wage employment in rural Vietnam. *Asia Pacific Viewpoint* 57(2), 263–279.

Cameron, L., Contreras Suarez, D., & Tseng, Y. (2023). Women's transitions in the labour market as a result of childbearing: The challenges of formal sector employment in Indonesia. IZA DP No. 16136.

Carneiro, F., & Henley, A. (2001). Modelling formal vs. informal employment and earnings: Micro-econometric evidence for Brazil. *Anais do XXIX Encontro Nacional de Economia*.

Chandra, V., & Khan, M.A. (1993). Foreign investment in the presence of an informal sector. *Economica* 60(237), 79–103.

Chong, A., & Saavedra, J. (1999). Structural reform, institutions and earnings: Evidence from the formal and informal sectors in urban Peru. *Journal of Development Studies* 35(4), 95–116.

Cling, J-P., Razafindrakoto, M., & Roubaud, F. (2010). The informal economy in Viet Nam. ILO Office in Vietnam. International Labour Organization.

Cragg, J.G., & Donald, S.G. (1993). Testing identifiability and specification in instrumental variable models. *Econometric Theory* 9(2), 222–240.

Di Falco, S., Veronesi, M., & Yesuf, M. (2011). Does adaptation to climate change provide food security? A micro-perspective from Ethiopia. *American Journal of Agricultural Economics* 93(3), 829–846.

Fajnzylber, P., Maloney, W.F. & Montes-Rojas, G.V. (2011). Does formality improve micro-firm performance? Evidence from the Brazilian SIMPLES program. *Journal of Development Economics* 94(2), 262–276.

Fields, G.S. (1975). Rural-urban migration, urban unemployment and underemployment, and job-search activity in LDCs. *Journal of Development Economics* 2(2), 165–187.

Goncalves, J., & Martins, P.S. (2021). Effects of self-employment on hospitalizations: Instrumental variables analysis of social security data. *Small Business Economics* 57, 1527–1543.

Gong, X. & Soest, A. (2002). Wage differentials and mobility in the urban labour market: a panel data analysis for Mexico, *Labour Economics* 9(4), 513–529.

Granda, C., & Hamann, F. (2015). Informality, saving and wealth inequality in Colombia. IDB Working Paper Series, No. IDB-WP-575, Inter-American Development Bank (IDB).

GSO (2019). Selected key indicators the Vietnam population and housing census 2019. General Statistics Office of Vietnam. Retrieved from https://www.gso.gov.vn/default en.aspx?tabid=515&idmid=5&ItemID=19449.

Hamilton, B.H. (2000). Does entrepreneurship pay? An empirical analysis of the returns to self-employment. *Journal of Political Economy* 108, 604–631.

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. *Journal of Modern African Studies* 11(1), 61–89.

Heckman, J. (1976). The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. *Annals of Economic and Social Measurement* 5, 475–492.

Heckman, J., Ichimura, H., Smith, J., & Todd, P. (1998). Characterizing selection bias using experimental data. *Econometrica* 66(5), 1017–1098.

ILO (2013). Measuring informality: A statistical manual on the informal sector and informal employment. International Labour Organization, Geneva.

ILO (2018). Informality and non-standard forms of employment. International Labour Organization.

Joubert, C. (2015). Pension design with a large informal labour market: Evidence from Chile. *International Economic Review* 56(2), 673–694.

Kajisa, K. (2007). Personal networks and non-agricultural employment: The case of a farming village in the Philippines. *Economic Development and Cultural Change* 55(4), 669–707.

Keane, M., & Wolpin, K. (1997). The career decisions of young men. *Journal of Political Economy* 105(3), 473–522.

Kudrna, G., Le, T., & Piggott, J. (2022). Macro-demographics and ageing in emerging Asia: The case of Indonesia. *Journal of Population Ageing* 15, 7–38.

Kudrna, G., O'Keefe, P., & Piggott, J. (2023). Pension policy in emerging Asian economies with population ageing: What do we know, where should we go? In Bloom, D., Sousa-Poza, A. & Sunde, U. (Eds.), Handbook on the economics of ageing, Routledge.

Lechner, M. (2002). Program heterogeneity and propensity score matching: An application to the evaluation of active labour market policies. *Review of Economics and Statistics* 84(2), 205–220.

Lokshin, M., & Sajaia, Z. (2004). Maximum likelihood estimation of endogenous switching regression models. *Stata Journal* 4(3), 282–289.

Malikov, E., Diego, A., Tobón, R., & Kumbhakar, S.C. (2018). Heterogeneous credit union production technologies with endogenous switching and correlated effects. *Econometric Reviews* 37(10), 1095–1119.

Maloney, W.F. (1999). Does informality imply segmentation in urban labour markets? Evidence from sectoral transitions in Mexico, *World Bank Economic Review* 13(2), 275–302.

Maloney, W.F. (2004). Informality revisited. *World Development* 32(7), 1159–1178.

Marcouiller, D., Castilla, R.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.

Nahm, D., Dobbie, M., & MacMillan, C. (2017). Union wage effects in Australia: An endogenous switching approach. *Applied Economics* 49(39), 3927–3942.

Nguyen, T.H. (2019). How large is Vietnam's informal economy? *Economic Affairs* 39(1), 81–100.

Noseleit, F. (2014). Female self-employment and children. *Small Business Economics* 43(3), 549–569.

Phan, D., & Coxhead, I. (2010). Inter-provincial migration and inequality during Vietnam's transition. *Journal of Development Economics* 91(1), 100–112.

Pradhan, M., & van Soest, A., (1997). Household labour supply in urban areas of Bolivia. *Review of Economics and Statistics* 79, 300–310.

Pratap, S., & Quintin, E. (2006). Are labor markets segmented in developing countries? A semiparametric approach. *European Economic Review* 50(7), 1817–1841.

Schclarek, A, & Caggia, M. (2015). Household saving and labor informality: The case of Chile. IDB Working Paper Series No. IDB-WP-581. Inter-American Development Bank (IDB).

Stock, J., & Yogo, M. (2005). Testing for weak instruments in linear IV regression. In D. W. Andrews (Ed.), Identification and inference for econometric models. Cambridge University Press.

Stratton, S.L., O'Toole, M.D., & Wetzel, N.J. (2008). A multinomial logit model of college stopout and dropout behaviour. *Economics of Education Review* 27(3), 319–331.

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 Econonomics* 58, 2949–2979.

Torezani, S., Colic-Peisker D. V., & Fozdar., F. (2008). Looking for a "missing link": Formal employment services and social networks in refugees' job search. *Journal of Intercultural Studies* 29(2), 135–152.

Ulyssea, G. (2020). Informality: Causes and consequences for development. *Annual Review of Economics* 12, 525–546.

Vu, L., & Rammohan, A. (2022). Is there an informal employment penalty in food security? Evidence from rural Vietnam. *European Journal of Development Research* 34(6), 2923–2947.

World Bank (2016). Live long and prosper: Aging in East Asia and Pacific. World Bank East Asia and Pacific Regional Report. Washington, DC: The World Bank.

Appendices

In the Appendices, we included several tables with details related to the data (Appendix A), IV selection and tests (Appendix B), and the results from our robustness checks (Appendix C). We have discussed (some of) these results in the main text.

Appendix A. Variable definitions and descriptive statistics

Table A.1 lists all the variables derived from the Vietnam Household Living Standards Survey (VHLSS) and provides their definitions. These variables are then used in our econometric models. Table A.2 then reports the corresponding general descriptive statistics in each of the three waves and the combined sample. As shown in each wave, the marriage rate is quite high and average years of education very low, compared to developed countries. However, similar observations apply to Indonesia and other Southeast Asian countries (as shown in Kudrna et al., 2022).

Table A.1: Variable definitions

Variable	Definition
Age	Age of individuals (year)
Gender	Male = 1; female = 0
Married	Married = 1; 0 = otherwise
Number of children	Household's total number of children aged below 16
Years of education[a]	Total number of years that an individual spent on education or vocational training
Urban	Urban = 1; rural = 0
Health expenditure	Total household's expenditure on medicines and medical facilities over the last 12 months
Informality networks	A proportion of workers working in informal sector over the population at the commune level
Household non-durable consumption – food[b]	The household's annual non-durable food consumption expenditure (in thousands Vietnam Dong (VND))
Household non-durable consumption – non-food[c]	The household's annual non-durable non-food consumption expenditure (in thousands Vietnam Dong (VND))
Household durable consumption[d]	The household's annual durable goods expenditure (in thousands Vietnam Dong (VND))
Household housing wealth[e]	The household's housing wealth which is measured as the expected value of the whole accommodation if it were for sale (in thousands Vietnam Dong (VND))
Low-skill informal employment	Low-skill informal employment includes those informal employments that do not require professional training.
High-skill informal employment	High-skill informal employment includes those informal employments that requires professional training.
Low-skill formal employment	Low-skill formal employment are those formal employments that do not require professional training.
High-skill formal employment	High-skill formal employment are those employments that require professional training

Notes: [a] We converted total years of education based on Vietnam's general education and vocational training systems; [b,c,d,e] We depreciated the value of household consumption and wealth with the CPI indexes (CPI in 2010 = 100) obtained from the World Bank (2022).
Source: Vietnam Household Living Standards Survey (VHLSS).

Table A.2: General descriptive statistics (2014-2018)

	2014		2016		2018		All sample	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
Age	43.02	16.66	44.31	16.61	46.25	16.67	44.53	16.70
Male	0.49	0.50	0.49	0.50	0.48	0.50	0.49	0.50
Married	0.74	0.44	0.74	0.44	0.75	0.43	0.74	0.44
Number of children	1.11	1.05	1.08	1.03	1.03	1.02	1.07	1.03
Years of Education	7.90	4.72	7.93	4.75	8.18	4.81	8.00	4.76
Urban	0.27	0.44	0.27	0.44	0.27	0.44	0.27	0.44
Health expenditure	207.26	377.78	403.21	918.15	467.05	1,358.19	359.17	977.52
Household non-durable consumption – food	14,404.97	9,502.60	24,386.75	19,596.45	20,000.00	13,853.33	19,597.97	15,451.62
Household non-durable consumption – non food	4,767.31	3,489.30	7,805.60	6,347.56	6,766.25	5,177.65	6,446.39	5,292.54
Household durable consumption	9,786.94	17,383.67	17,379.18	33,892.48	16,765.22	30,980.56	14,647.08	28,558.38
Household housing wealth	155,817.3	260,974.9	258,916.5	401,743.6	255,011.3	372,394.7	223,506.5	353,791.8
Informality networks	0.52	0.19	0.52	0.19	0.51	0.20	0.51	0.19

Source: Authors' calculation from VHLSS (2014-2018).

Appendix B. Selection of IVs and IV tests

This appendix relates to Subsection 5.4. Specifically, Table B.1 reports the correlation between the informal networks and two variables constructed as a proxy for the commune's characteristics, namely "distance from the commune to the market" and "distance from the commune to the city centre". As discussed, the results show a weak correlation, meaning that the IV satisfied the exclusion condition. In Table B.2, we then report on the findings of several tests, including the Stock & Yogo (2005) test, to determine the strength of our IV. As indicated, the informal networks variable is proved to be appropriate for our model.

Table B.1: The correlation between the informal networks and two variables of commune's characteristics

	2014	2016	2018
Commune's characteristic 1[a]	0.13*	0.143*	0.156*
	(0.000)	(0.000)	(0.000)
Commune's characteristic 2[b]	0.23*	0.196*	0.148*
	(0.0000)	(0.000)	(0.000)

Notes: [a] Commune's characteristic 1 is measured as the average distance from this commune to the market; [b] Commune's characteristic 2 is measured as the average distance from this commune to the city centre.
Source: Authors' calculation form VHLSS (2014-2018).

Table B.2: Tests of instrumental variable[a] for MESR model

	Informality networks as instrumental variable and food consumption[b] as outcome variable			
	Low-skill informal treatment	**High-skill informal treatment**	**Low-skill formal treatment**	**High-skill formal treatment**
Summary Statistics	R-Sq. = 0.2496 Adjusted. R-Sq. = 0.249 Partial R-Sq. = 0.0753 Prob > F = 0.0000	R-Sq. = 0.332 Adjusted. R-Sq. = 0.33 Partial R-Sq. = 0.0753 Prob > F = 0.0000	R-Sq. = 0.332 Adjusted. R-Sq. = 0.33 Partial R-Sq. = 0.0753 Prob > F = 0.0000	R-Sq. = 0.545 Adjusted. R-Sq. = 0.543 Partial R-Sq. = 0.01 Prob > F = 0.0001
Tests of endogeneity				
Durbin (score) chi2(1)	104.691(p = 0.0000)	49.1942 (p = 0.0000)	24.8991 (p = 0.0000)	34.7001 (p = 0.0000)
Wu-Hausman F(1,5887)	106.402(p = 0.0000)	49.9194 (p = 0.0000)	25.2855 (p = 0.0000)	35.3242 (p = 0.0000)
First Stage Regression Test IV				
Cragg and Donald Wald F Statistic Minimum eigenvalue statistic	479.173	216.552	22.576	14.502
Stock-Yogo weak ID test critical values	10% maximal IV size: 16.38 15% maximal IV size: 8.96 20% maximal IV size: 6.66 25% maximal IV size: 5.53			

Note: [a] Informality networks (calculated as the proportion of workers in the informal sector over the population at the commune-level); [b] Household food consumption is in natural log form.
Source: Authors' calculation.

Appendix C. Result tables for robustness checks

This appendix relates to Subsection 5.4, with the results in Table C.1 from OLS and IV models, and in Tables C.2, C.3 and C.4 from Multinominal endogenous switching regression model (MESR) with the ATT in different years.

Table C.1: OLS and IV models (2014-2018)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	**IV**	**OLS**	**IV**	**OLS**	**IV**	**OLS**	**IV**
Panel A: Household nondurable consumption: Food as outcome variable								
Low-skill informal	-0.038	**-0.919*****						
	(0.027)	(0.115)						
High-skill informal			**0.091*****	**-0.654*****				
			(0.032)	(0.132)				
Low-skill formal					**0.218*****	**1.519*****		
					(0.041)	(0.581)		
High-skill formal							**0.158*****	**2.385****
							(0.044)	(1.180)
Observations	5897	5897	2708	2708	985	985	1408	1408
Panel B: Household nondurable consumption: Non-food as outcome variable								
Low-skill informal	**0.014**	**0.793*****						
	(0.028)	(0.118)						
High-skill informal			**0.161*****	**0.500*****				
			(0.034)	(0.140)				
Low-skill formal					**0.336*****	**1.526****		
					(0.045)	(0.614)		
High-skill formal							**0.310*****	**2.269****
							(0.048)	(1.107)
Observations	5898	5898	2709	2709	986	986	1409	1409
Panel C: Household durable consumption as outcome variable								
Low-skill informal	0.012	**0.985*****						
	(0.047)	(0.191)						
High-skill informal			**0.287*****	0.010				
			(0.053)	(0.205)				
Low-skill formal					**-0.221*****	**-1.034***		
					(0.070)	(0.600)		
High-skill formal							**-0.308*****	1.073
							(0.071)	(0.742)
Observations	5883	5883	2707	2707	986	986	1409	1409
Panel D: Household housing wealth as outcome variable								
Low-skill informal	-0.134	0.227						
	(0.0487)	(0.388)						
High-skill informal			0.071	**0.649*****				
			(0.0557)	(0.372)				
Low-skill formal					**0.287*****	-0.094		
					(0.0798)	(2.007)		
High-skill formal							**0.159****	-7.953
							(0.0729)	(16.577)
Observations	5769	5769	2661	2661	957	957	1382	1382

Notes: *Mean statistically significant at 10%; **Mean statistically significant at 5%; ***Mean statistically significant at 1%; All specifications control for age, gender and marital status, number of children, years of education, urban area, health expenditure, and informal networks; Baseline level of employment: **unemployed**; Instrumental variable is measured as informality network which is calculated as the proportion of informal workers over the population at the commune-level.
Source: Authors' calculation from VHLSS (2014-2018).

Table C.2: Multinominal endogenous switching regression model (MESR) – ATT (2014)

Outcome variables	Sectoral choice		Employment status: Employment choice (j=1, 2,3,4)	Employment status: Unemployed (j=0)	ATT
			(1)	(2)	(3)=(1)-(2)
Household non-durable consumption - Food	Unemployment		-	-	-
	Low-skill informal	(j=1)	11261	11234	26
			(59)	(81)	(48)
	High-skill informal	(j=2)	14226	13353	**873*****
			(121)	(178)	(96)
	Low-skill formal	(j=3)	17708	19134	-1426
			(614)	(583)	(493)
	High-skill formal	(j=4)	20316	22726	-2410
			(315)	(397)	(274)
Household non-durable consumption – Non food	Unemployment		-	-	-
	Low-skill informal	(j=1)	3622	3104	**517*****
			(23)	(26)	(18)
	High-skill informal	(j=2)	4627	3785	**841*****
			(54)	(61)	(31)
	Low-skill formal	(j=3)	6178	5854	**323****
			(170)	(186)	(164)
	High-skill formal	(j=4)	7482	7425	57
			(100)	(130)	(91)
Household durable consumption	Unemployment		-	-	-
	Low-skill informal	(j=1)	4993	4361	**631*****
			(44)	(49)	(43)
	High-skill informal	(j=2)	8532	6060	**2471*****
			(160)	(150)	(105)
	Low-skill formal	(j=3)	8737	9223	**-485*****
			(395)	(426)	(302)
	High-skill formal	(j=4)	14395	16418	**-2023*****
			(248)	(419)	(320)
Household housing wealth	Unemployment		-	-	-
	Low-skill informal	(j=1)	66751	62960	**3791*****
			(1223)	(1451)	(709)
	High-skill informal	(j=2)	121357	110119	**11237*****
			(3077)	(3895)	(1482)
	Low-skill formal	(j=3)	185423	160629	24794
			(14167)	(11912)	(10172)
	High-skill formal	(j=4)	242392	251268	-8875
			(7773)	(7921)	(2506)

Notes: *Mean statistically significant at 10%; **Mean statistically significant at 5%; ***Mean statistically significant at 1%; All specifications control for age, gender, marital status, number of children, years of education, urban area, health expenditure, and informal networks; ATT: The effects of the treatment on the treated.
Source: Authors' calculation from VHLSS.

Table C.3: Multinominal endogenous switching regression model (MESR) – ATT (2016)

Outcome variables	Sectoral choice		Employment status		ATT
			Employment choice (j=1, 2,3,4)	Unemployed (j=0)	
			(1)	(2)	(3)=(1)-(2)
Household non-durable consumption - Food	Unemployment		-	-	-
	Low-skill informal	(j=1)	18123	20539	**-2415*****
			(99)	(164)	(125)
	High-skill informal	(j=2)	24375	24202	172
			(287)	(347)	(245)
	Low-skill formal	(j=3)	28126	25867	**2259*****
			(783)	(820)	(555)
	High-skill formal	(j=4)	35450	35311	336
			(654)	(830)	(377)
Household non-durable consumption – Non food	Unemployment				
	Low-skill informal	(j=1)	5767	8458	-2690
			(35)	(57)	(53)
	High-skill informal	(j=2)	8098	9005	-907
			(109)	(118)	(90)
	Low-skill formal	(j=3)	9139	11694	-2555
			(208)	(329)	(233)
	High-skill formal	(j=4)	12804	17609	-4808
			(222)	(412)	(288)
Household durable consumption	Unemployment				
	Low-skill informal	(j=1)	8477	12171	-3694
			(83)	(98)	(106)
	High-skill informal	(j=2)	14772	12840	**1931*****
			(246)	(193)	(203)
	Low-skill formal	(j=3)	12715	18514	**-5798*****
			(470)	(681)	(562)
	High-skill formal	(j=4)	23414	34367	**-10952*****
			(471)	(1010)	(800)
Household housing wealth	Unemployment				
	Low-skill informal	(j=1)	118831	153930	-35099
			(2125)	(2551)	(1747)
	High-skill informal	(j=2)	216763	193381	**23382*****
			(5823)	(5599)	(4791)
	Low-skill formal	(j=3)	303349	4226660	-119310
			(19020)	(30244)	(19052)
	High-skill formal	(j=4)	378691	708426	-329735
			(10859)	(27658)	(21731)

Notes: *Mean statistically significant at 10%; **Mean statistically significant at 5%; ***Mean statistically significant at 1%; All specifications control for age, gender, marital status, number of children, years of education, urban area, health expenditure, and informal networks; ATT: The effects of the treatment on the treated.
Source: Authors' calculation from VHLSS.

Table C.4: Multinominal endogenous switching regression model (MESR) – ATT (2018)

Outcome variables	Sectoral choice		Employment status		ATT
			Employment choice (j=1, 2,3,4)	Unemployed (j=0)	
			(1)	(2)	(3)=(1)-(2)
Household non-durable consumption - Food	Unemployment		-	-	-
	Low-skill informal	(j=1)	15655	19503	**-3848*****
			(89)	(207)	(175)
	High-skill informal	(j=2)	19951	23564	-3613
			(187)	(418)	(351)
	Low-skill formal	(j=3)	23619	28947	-5328
			(548)	(1489)	(1287)
	High-skill formal	(j=4)	27672	45030	-17357
			(364)	(1274)	(1211)
Household non-durable consumption – Non-food	Unemployment				
	Low-skill informal	(j=1)	5120	5463	523
			(28)	(216)	(209)
	High-skill informal	(j=2)	7090	6512	**577*****
			(84)	(113)	(77)
	Low-skill formal	(j=3)	8016	8862	-846
			(133)	(1286)	(1275)
	High-skill formal	(j=4)	11007	13965	-2958
			(163)	(1469)	(1459)
Household durable consumption	Unemployment				
	Low-skill informal	(j=1)	8707	13633	-4925
			(83)	(885)	(871)
	High-skill informal	(j=2)	14273	15233	-960
			(169)	(388)	(318)
	Low-skill formal	(j=3)	12008	23419	**-11410****
			(293)	(5432)	(5431)
	High-skill formal	(j=4)	21147	36522	**-15375*****
			(340)	(3692)	(3588)
Household housing wealth	Unemployment				
	Low-skill informal	(j=1)	115466	96783	**18683*****
			(2190)	(2062)	(1552)
	High-skill informal	(j=2)	208854	168970	**39884*****
			(4795)	(6973)	(5794)
	Low-skill formal	(j=3)	244363	245453	-1090
			(12074)	(20969)	(20279)
	High-skill formal	(j=4)	401412	671862	-270449
			(11101)	(26291)	(22526)

Notes: *Mean statistically significant at 10%; **Mean statistically significant at 5%; ***Mean statistically significant at 1%; All specifications control for age, gender, marital status, number of children, years of education, urban area, health expenditure, and informal networks; ATT: The effects of the treatment on the treated.
Source: Authors' calculation from VHLSS.