

Factors Associated with Participation of Vietnamese Informal Workers in Voluntary Social Insurance Scheme

Thi-Thu Do^a

Faculty of Economics, Banking Academy of Vietnam

Thanh-Long Giang^b

Faculty of Economics, National Economics University

Abstract: This study estimated impacts of demographic and employment factors on the voluntary social insurance participation of Vietnamese informal workers by using data from the Labour Force Survey and Heckman's two-step probit selection model for four groups of informal workers. The results showed that age, the highest education level, number of school-age children in household, number of household members participating in the social insurance system, job position and income level of informal workers had significant impacts on their decision to participate in the scheme. Moreover, these impacts were different for informal workers when considering their job positions. Based on the findings, this study recommended various policies considering the aforementioned characteristics of informal workers.

Keywords: Voluntary social insurance, informal workers, modelling, Viet Nam
JEL classification: C31, H55, I39, J48, J88

1. Introduction

Voluntary social insurance (VSI) is one of the most important pillars in a country's social protection system. VSI is based on voluntary contribution of employees and organised by government or private sector to expand both vertical and horizontal coverage of social security system of a nation (International Labour Organization [ILO], 2012). VSI usually covers one or more benefit schemes, such as sickness, maternity leave, occupational accident and occupational disease, retirement and survivorship.

In Viet Nam, the social insurance (SI) system has been operating since the early 1960s. Prior to 1995, the SI system in Viet Nam only covered public employees. However, dramatic changes in economic and social conditions along with fast develop-

^a Faculty of Economics, Banking Academy of Vietnam, 12 Chua Boc Street, Dong Da District, Hanoi, Vietnam. Email: thudt@hvn.edu.vn / orcid: <https://orcid.org/0000-0002-1027-4078> (First author)

^b Faculty of Economics, National Economics University, 207 Giai Phong Street, Hai Ba Trung District, Hanoi 11616, Viet Nam. Email: longgt@neu.edu.vn / orcid: <https://orcid.org/0000-0002-6296-0563> (Corresponding author)

* We would like to thank a number of colleagues at the National Economics University, and the Banking Academy of Vietnam for providing useful comments on the previous versions of this research. We are thankful to anonymous reviewers of the *Malaysian Journal of Economic Studies (MJES)* for insightful comments and suggestions so that we could improve our research.

ment of the private sector resulting from the *Doi moi* (economic renovation) forced the Government of Vietnam to reform the SI system to achieve the objectives of universal social protection. This reform was marked by the Law on Social Insurance in 2006, which took effect from 2007 with the compulsory social insurance, and then from 2008 with VSI. The Law on Social Insurance in 2006 was then amended in 2014. In the 2014 Law, the SI system has been offered to hitherto both formal and informal workers with three pillars: the compulsory social insurance (CSI), the VSI, and the supplementary retirement scheme (Vietnam's National Assembly, 2014). More particularly, CSI and the supplementary retirement scheme cover formal workers, while VSI covers informal workers. Although the VSI offers only benefits for retirement and survivorship, it can be seen as a solution to expand social protection coverage for informal workers, especially the "the missing middle" in Viet Nam (Do, 2023). As a result, the number of participants in the VSI increased sharply from 6,110 people in 2008 to 41,193 people in 2009 and then reached to approximately 1.01 million people in 2020, which accounted for 6% of the total number of CSI and VSI participants. However, there has been a great limitation in VSI's coverage. Among about 35 million informal workers who are legally entitled to the VSI scheme as estimated by the General Statistics Office (GSO) and ILO (2017), participation rate in VSI was only 3%. This situation has raised a critical question for policy makers about solutions to expand VSI coverage to the Vietnamese informal workers.

To provide empirical evidence for this topic, many studies have examined how individual, household and employment characteristics of informal workers affected their decision to participate in SI schemes. First, regarding individual characteristics, studies showed positive effects of age, education level, duration of current residency and urban living on probability to participate in VSI (Auerbach, et al., 2007; Bhan et al., 2014; Dang et al., 2003; Duong, et al., 2011; Giles et al., 2013; Heenkenda, 2016; Nivakoski, 2014; Roushdy & Selwaness, 2019; Yao & Kim, 2015). Studies also showed that in general, married workers are more actively engaged in SI than single workers (Collins-Sowah et al., 2013; Giles et al., 2013). More particularly, married female workers were less likely to participate in the SI than their male counterparts (Heenkenda, 2016; Nivakoski, 2014; Roushdy & Selwaness, 2019), and this could be explained by the household-related SI benefits in these countries. In contrast, for single female workers, Auerbach et al. (2007) showed that they tended to be more actively involved in SI than their male counterparts. Second, regarding household characteristics, Posel (2001) concluded that household head was the main decision-making person for SI participation. Third, considering employment characteristics, a number of studies found that public sector workers and white-collar were more likely to participate in SI system (Auerbach et al., 2007; Carpio, 2011; Nivakoski, 2014). Besides, Nivakoski (2014) and Roushdy and Selwaness (2019) concluded that the number of years of experience had a non-linear effect on SI participation. Income also had a positive impact on SI participation (Castel (2008) for Viet Nam; Heenkenda (2016) for Sri Lanka).

To the best of our knowledge, there have been no studies examining the impacts of three following characteristics which have been supported by the theoretical foundations. First, Pindyck and Rubinfeld's (2014) bandwagon effect was defined as "network externality in which a consumer wishes to possess a good in part because

others do" (ibid, p. 136), and accordingly a worker with higher number of household members participating in SI might be more likely to participate in VSI. Second, facing a budget constraint, workers having more school-age children might reduce or avoid contributions to SI, among others, for the current spending, and as such those with more school-age children might be less likely to participate in VSI. Third, different preferences toward risks affect a consumer's (worker's) decision to contribute to SI schemes as Pindyck and Rubinfeld (2014: p. 166) argued that "An individual who is risk averse not only buy life insurance, health insurance, and car insurance but also seek occupations with relatively stable wage". Therefore, job position, which reflects worker's preferences toward risk, might affect his/her decision to participate in VSI scheme.

This study aims to fill these research gaps by applying Heckman's two-step probit selection model to estimate impacts of these three factors along with other socio-demographic and economic factors on Vietnamese informal workers' decision to participate in VSI. The structure of this study is as follows. Section 2 describes data and methods. Section 3 presents the main findings. Policy implications and concluding remarks are presented in section 4.

2. Data and Research Methods

2.1 Data

This study used data from the Labour Force Survey (LFS) conducted by GSO in 2016. LFS has been conducted annually since 2007. The LFS is a sample survey, which is generated by a two-stage stratified sampling method by selecting areas (i.e., urban and rural places) in the first stage and then selecting households in the second stage. The sampling unit is household. All interviewees in the selected households are the survey's sample. LFS provided various pieces of information about socio-demographic, economic and employment characteristics of workers. Since 2014, LFS has provided information about workers' VSI participation. In LFS 2016, there were 225,138 households (of which 98,355 households were in urban areas, and 126,783 households were in rural areas) with a total of 814,611 respondents.

To select a sample of informal workers which were defined by GSO and ILO (2017), this study extracted the LFS 2016 data along with two main steps: i) classifying all respondents by economic sectors (including both formal and informal sectors), and ii) determining informal employment based on status of employment. Informal workers are those holding informal employment in both formal and informal economic sectors.

In the formal economic sector, informal workers include: i) employees in the formal sector without a labour contract or with a verbal agreement, lump-sum contract or short-term (less than three months) contract or a more than three-month contract without compulsory SI, ii) family business workers in the establishment belonging to the formal sector, and iii) members of cooperatives without compulsory SI. In the informal economic sector, the informal workers include: i) persons defined as establishment owners in the informal sector, ii) own-account workers in the informal sector, iii) wage workers (or employees) in the informal sector, and iv) family business workers in the establishment belonging to the informal sector. In the household economic sector, the informal workers include: i) own-account workers engaged in production of goods and

services for own consumption of their households, ii) wage workers in the household sector (i.e., workers hired by households).

As such, the informal workers could be generally categorised into four groups: i) self-employed, ii) informal wage workers, iii) family business workers, and iv) co-operative members. However, because of the changes in the Law on Social Insurance in 2014, members of cooperatives became participants of the compulsory social insurance scheme since 2016. In addition, this group accounted for a negligible proportion in the informal workers, and thus in this study, we removed this group. As such, there were three groups of workers for this study: self-employed, informal wage workers and family business workers.

As will be discussed below, this study uses the Heckman's two-step probit selection model, which requires a full set of observations for each variable of both informal and formal workers, and thus the full set of observations in the model was the total number of workers in LFS 2016. As a result, there were 400,583 observations in each model for the aforementioned groups of workers.

2.2 Method

The sample was conditionally selected according to the definition of informal workers and three sub-groups of informal workers by job position. Therefore, the OLS regression results in this case might be biased and inconsistent due to sample selection which was not at random. According to Wooldridge et al. (2016), some kind of sample selection could be ignored if the sample was determined solely by an exogenous explanatory variable. However, in this case, job position could be an endogenous explanatory variable of VSI participation since workers wanting to join the compulsory SI tended to choose formal jobs, while workers wanting not to participate in SI tended to choose informal jobs. This means that job position could depend on SI participation in general and VSI participation in particular. This is an endogenous sample selection problem.

Wooldridge et al. (2016) indicated that a regression with instrumental variables could solve two kinds of endogeneity problems (i.e., omitted variables and measurement error) while the two-step selection model could solve the endogenous sample selection problem. Therefore, in this study, with a binary dependent variable, the Heckman's two-step probit selection model was applied to examine and correct the sample selection problem. This model was developed by Van de Ven and Van Pragg (1981) and was applied in Morrissey et al. (2016).

In Heckman's two-step probit selection model, dependent variable Y_i is a binary variable, i.e., Y is equal to 1 if informal worker participated in VSI, and 0 otherwise. Heckman's two-step probit selection model is represented by two equations: i) the outcome equation, and ii) the selection equation.

The outcome probit equation is presented as below:

$$Y_i^* = \beta X_i + \varepsilon_i \quad (1)$$

where Y_i^* is an unobserved latent variable which reflects individual's utility level of participation in the VSI and determines probability of participation: $Y_i = 1$ if $Y_i^* > 0$ and $Y_i = 0$ if $Y_i^* \leq 0$, X_i is a vector of independent variables, and ε_i is random factors.

In addition, the decision to become self-employed, wage workers, or family business workers depends on explanatory variables in the following selection equation:

$$S_i^* = \alpha X_i + \phi Z_i + \mu_i, \text{ with } S_i = 1 \text{ if } S_i^* > 0; S_i = 0 \text{ if } S_i^* \leq 0 \text{ and } Y_i \text{ is observed in this equation if } S_i = 1 \tag{2}$$

where S_i^* is an unobserved latent variable which determines worker i to be a self-employed, informal wage worker, or family business worker, X_i is a vector of independent variables in the outcome equation (1) and may be other variables Z_i which are factors affecting S_i^* but not affecting Y_i^* and error term μ_i .

In this study, a variable about the location where the workers usually worked was included in the Z vector. This variable reflected a characteristic of informal job, but determined participation in VSI.

The error terms ϵ_i and μ_i are assumed to be normally distributed and their correlated coefficient is $Corr(\epsilon_i, \mu_i) = \rho$. If $S_i = 1$, Y_i is observed and a log likelihood equation of the selected equation is as below:

$$\begin{aligned} \ln L = & \sum_{\substack{i \in S \\ Y_i \neq 0}} w_i \ln \{ \Phi_2(\beta X_i + offset_i^\beta, \phi Z_i + offset_i^\phi, \rho) \} + \\ & \sum_{\substack{i \in S \\ Y_i = 0}} w_i \ln \{ \Phi_2(-\beta X_i + offset_i^\beta, \phi Z_i + offset_i^\phi, -\rho) \} \\ & \sum_{i \notin S} w_i \ln \{ 1 - \Phi(\phi Z_i + offset_i^\phi) \} \end{aligned} \tag{3}$$

where S is a set of observations where Y_i is observed, Φ_2 is the cumulative binary normal distribution function with expectation is equal to 0, and $\Phi(.)$ is a cumulative normal distribution function.

In the estimation results of the maximum likelihood equation, ρ (rho) is not directly estimated, but is estimated through a “ $\text{atanh } \rho$ ” as follows:

$$\text{Atanh } \rho = \frac{1}{2} \ln \left(\frac{1 + \rho}{1 - \rho} \right) \tag{4}$$

The value of ρ is in the range from -1 to 1 . If $\rho = 0$, the likelihood equation of the selection equation is equal to sum of the two equations including (1) and (2). It means that there is no problem of sample selection bias. The likelihood ratio (LR) test can be done by comparing the probability of the full model with the total probability of the outcome probit model and the selection model. According to Morrissey et al. (2016), a negative value of ρ reflects that the observed group of workers is less likely to participate in VSI than others. In contrast, a positive value of ρ reflects that the group of workers currently observed is more likely to participate in VSI than others. A smaller value of ρ indicates a less serious problem of sample selection bias in the model.

To make the results more reliable, this study also used a robust estimation option in the Heckman’s two-step probit selection model. Moreover, this study also estimated the model with bootstrap option which allowed an estimation of standard deviation and confidence intervals of the regression results according to the recovered sampling method from a finite sample. The bootstrap option usually gives the smallest standard deviation and more accurate confidence interval than conventional regression results.

2.3 Variables

The dependent variable Y in equation (1) is a binary variable (denoted by $d_voluntary_si$), in which $d_voluntary_si = 1$ if informal worker participated in VSI, and 0 otherwise.

The explanatory variables of X vector in equation (1) are summarised in Table 1, which represent seven factors of individual characteristics, three factors of household characteristics, and seven factors of employment characteristics.

The summary of statistical results for informal workers and their VSI coverage are shown in Tables 2–6. Table 2 shows that out of 329,520 informal workers, there were 50.63% self-employed, 23.36% wage workers, and 25.97% family business workers. VSI coverage rate was 1.02% for all informal workers, 0.23% for self-employed, 3.49% for wage workers and 0.32% for family business workers. Tables 3-6 summarise VSI coverage rates of informal workers along with their individual, household, employment characteristics and income quintiles, respectively.

Table 1. Expected sign of independent variables in the model

Characteristics	Variables	Description	Expected effect	References
Age group	<i>d_1524 (.ref)</i>	15-24 years old	Nonlinear effects	Auerbach et al. (2007) and Nivakoski (2014) showed positive effects of age on social insurance participation while Castel (2008) and Collins-Sowah et al. (2013) concluded the negative effects of age on social insurance participation.
	<i>d_2534</i>	25-34 years old		
	<i>d_3549</i>	35-49 years old		
	<i>d_5059</i>	50-59 years old		
	<i>d_60plus</i>	60 years old and above		
Gender	<i>d_female (.ref)</i>	Female workers	–	Auerbach et al. (2007); Collins-Sowah et al. (2013)
	<i>d_male</i>	Male workers		
Marital status	<i>d_single (.ref)</i>	Single workers	+	Collins-Sowah et al. (2013); Giles et al. (2013); Nivakoski et al. (2014)
	<i>d_married</i>	Married workers		
	<i>d_married_other</i>	Widowed, divorced and separated workers		
Education level	<i>d_noschooling (.ref)</i>	No schooling or incomplete primary education	+	Auerbach et al. (2007), Castle (2008), Carpio (2011), Giles et al. (2013), Roushdy and Selwaness (2019)
	<i>d_prim2sec</i>	Educational level from completed primary to completed high school		
	<i>d_professional</i>	Intermediate or professional college level		
	<i>d_tertiary</i>	University or higher degree		
Vocational training	<i>d_voc_training</i>	= 1 if workers participated in at least one vocational training program, = 0 if otherwise	+	

Table 1. Continued

Characteristics	Variables	Description	Expected effect	References
Duration of residency	<i>d_residence_lower1 (.ref)</i>	Residents for less than 1 year		Giles et al. (2013)
	<i>d_residence_1to5</i>	Residents for 1–5 years	+	
	<i>d_residence_above5</i>	Residents for more than 5 years	+	
Living area	<i>d_urban</i>	= 1 indicated urban, = 0 indicated rural	–	Castel (2008), Nivakoski (2014)
Head of household	<i>d_hh_head</i>	= 1 if the informal worker was a household head, and 0 otherwise.	+	Auerbach et al. (2007), Roushdy & Selwaness (2019)
No. of school-age children in household	<i>num_child_sa</i>	Total number of school-age (6–18 years old) children in household of each informal worker	+	Castel (2008), Heenkenda (2016)
Square of number of school-age children in household	<i>num_child_sa_sqr</i>	Square of number of school-age children in household	–	
No. of household members participating in social insurance system	<i>num_si_holding</i>	Total number of household members currently participating in the social insurance system	+	“Bandwagon effect” – Pindyck (2018)
State of economic activity	<i>d_working</i>	=1 indicated that worker had a job at the time of survey, and 0 otherwise	+	Mandigma, M.B.S. (2016)
Work experience	<i>d_workexp_less5 (.ref)</i>	Less than 5 years at current main job		
	<i>d_workexp_510</i>	From 5 to 10 years at current main job	+	Nivakoski (2014)
	<i>d_workexp_10plus</i>	More than 10 years at current main job	+	
Job position	<i>d_self_employed (.ref)</i>	Employers and self-employed workers		Auerbach et al. (2007)
	<i>d_wage_worker</i>	Wage earners	+	
	<i>d_fam_worker</i>	Unpaid family workers	–	
Type of establishment	<i>d_own (.ref)</i>	Agri-forestry-fishery households, freelancer workers and individual business establishment		Auerbach et al. (2007)
	<i>d_coop</i>	Cooperatives	+	
	<i>d_other</i>	State agencies, non-state business and organisations, foreign		

Table 1. Continued

Characteristics	Variables	Description	Expected effect	References
		organisations and other organisations and unions	+	
Labour contracts	<i>d_no_contract (.ref)</i>	No labour contract		Auerbach et al. (2007) showed the negative impact of part-time job.
	<i>d_under3mths_contract</i>	Less than 3 months labour contract	+	
	<i>d_verbal_contract</i>	Verbal agreement	+	
Business registration status	<i>firm_reg</i>	=1 implied that the firm had business registration, and 0 indicated otherwise	+	Auerbach et al. (2007)
Income of informal workers	<i>Group 1 (.ref)</i>	Poorest workers		Castel (2008), Carpio (2011), Heenkenda (2016)
	<i>Group 2</i>	Near poor workers	+	
	<i>Group 3</i>	Middle income workers	+	
	<i>Group 4</i>	Near rich workers	+	
	<i>Group 5</i>	Richest workers	+	

Source: Authors' summary.

Table 2. Voluntary social insurance coverage rate of informal workers with different job positions

Characteristics	VSI participation (%)	
	Yes	No
Informal workers	3,350 (1.02%)	326,170 (98.98%)
Self-employers	392 (0.23%)	166,454 (99.77%)
Wage workers	2,684 (3.49%)	74,276 (96.51%)
Family workers	270 (0.32%)	85,322 (99.68%)

Source: Authors' estimates, using *Labour Force Survey 2016*.

Table 3. Voluntary social insurance coverage rate of informal workers according to individual characteristics

Characteristics	VSI participation (%)	
	Yes	No
<i>Gender</i>		
Male	1,730 (1.02%)	168,434 (98.98%)
Female	1,620 (1.02%)	157,736 (98.98%)
<i>Age groups</i>		
15–24 years old	468 (1.07%)	43,195 (98.93%)
25–34 years old	1,285 (1.9%)	66,340 (98.1%)

Table 3. Continued

Characteristics	VSI participation (%)	
	Yes	No
35–49 years old	1,064 (0.91%)	115,971 (99.09%)
50–59 years old	415 (0.63%)	65,677 (99.37%)
60 years old and above	118 (0.34%)	34,987 (99.66%)
<i>Education level</i>		
No schooling or incomplete primary education	65 (0.10%)	65,017 (99.90%)
Primary to high school	1,885 (0.76%)	246,898 (99.24%)
Intermediate or professional college level	621 (5.76%)	10,152 (94.24%)
University or higher degree	779 (15.96%)	4,103 (84.04%)
<i>Marital status</i>		
Single	731 (1.51%)	47,729 (98.49%)
Married	2,485 (0.97%)	254,857 (99.03%)
Other statuses (widowed, divorced, and separated	134 (0.56%)	23,584 (99.44%)
<i>Duration of residency</i>		
Less than 1 year	56 (2.00%)	2,744 (98.00%)
1–5 years	172 (2.19%)	7,666 (97.81%)
More than 5 years	3,122 (0.98%)	315,760 (99.02%)

Source: Authors' estimates, using *Labour Force Survey 2016*.

Table 4. Voluntary social insurance coverage rate of informal workers according to household characteristics

Characteristics	VSI participation (%)	
	Yes	No
<i>Number of school-age children in household</i>		
0	2,195 (1.14%)	190,203 (98.86%)
1	753 (1.03%)	72,617 (98.97%)
2	375 (0.76%)	48,685 (99.24%)
3	25 (0.23%)	11,059 (99.77%)
4	2 (0.07%)	2,762 (99.93%)
From 5 to 8	0 (0.00%)	844 (100%)
<i>Number of household members participating in the social insurance system</i>		
0	7 (0.00%)	278,964 (100%)
1	1,523 (3.98%)	36,714 (96.02%)
2	1,361 (13.51%)	8,713 (86.49%)
3	328 (18.53%)	1,442 (81.47%)
4	118 (28.92%)	290 (71.08%)
5	13 (26.53%)	36 (73.47%)
6	0 (0.00%)	8 (100%)
7	0 (0.00%)	3 (100%)

Source: Authors' estimates, using *Labour Force Survey 2016*.

Table 5. Voluntary social insurance coverage rate of informal workers according to employment characteristics

Characteristics	VSI participation (%)	
	Yes	No
<i>State of economic activity</i>		
Worker had a job	3,317 (1.01%)	324,158 (98.38%)
Worker had no job	33 (1.61%)	2,012 (98.39%)
<i>Labour contracts</i>		
Labour contracts in term of under 3 months	0 (0%)	124 (100%)
Verbal contracts	154 (0.25%)	61,205 (99.75%)
Without any labour contracts	90 (0.69%)	12,962 (99.31%)
Labour contracts in term of over 3 months	3,106 (1.22%)	251,879 (98.78%)
<i>Job position</i>		
Cooperative members	4 (3.28%)	118 (96.72%)
Self-employed	392 (0.23%)	166,454 (99.77%)
Family workers	270 (0.32%)	85,322 (99.68%)
Wage workers	2,684 (3.49%)	74,276 (96.51%)
<i>Work experience</i>		
Less than 5 years	1,564 (1.66%)	92,543 (98.34%)
5–10 years	980 (1.31%)	73,853 (98.69%)
More than 10 years	805 (0.52%)	155,152 (99.48%)
<i>Type of establishment</i>		
Agri-forestry-fishery households, freelance workers and individual business establishment	805 (0.25%)	316,586 (99.75%)
Cooperatives	21 (5.61%)	353 (94.39%)
State agencies, non-state business and organisations, foreign organisations and other organisations and unions	2,524 (21.47%)	9,231 (78.53%)
<i>Living area</i>		
Rural	1,445 (0.62%)	230,319 (99.38%)
Urban	1,905 (1.95%)	95,851 (98.05%)

Source: Authors' estimates, using *Labour Force Survey 2016*.

Table 6. Voluntary social insurance coverage rate of informal workers according to income

Characteristics	VSI participation (%)	
	Yes	No
<i>Income quintiles</i>		
Poorest (Group 1)	243 (0.31%)	78,590 (99.69%)
Group 2	61 (0.26%)	23,067 (99.74%)
Group 3	676 (0.63%)	107,030 (99.37%)
Group 4	1,378 (2.01%)	67,157 (97.99%)
Richest (Group 5)	992 (1.93%)	50,326 (98.07%)

Source: Authors' estimates, using *Labour Force Survey 2016*.

3. Main Findings

The Chow-test was applied to assess the difference in impact structure between two sub-groups of informal workers, i.e., self-employed and the other groups. The test result showed that there was a difference in the impact of demographic and employment characteristics on the probability to participate in VSI between these two groups.¹ Accordingly, Heckman's two-step probit selection model could be applied for these sub-groups of informal workers. More particularly, four regressions were conducted for four samples, as follows:

Sample 1: all informal workers,

Sample 2: self-employed (the first sub-sample of informal workers),

Sample 3: wage workers (the second sub-sample of informal workers),

Sample 4: family business workers (the third sub-sample of informal workers).

The estimation results of Heckman's two-step probit selection model show that rho is statistically significant in the model for the three samples, 1, 2 and 3, but not for sample 4. These confirmed that there was an endogenous sample selection problem, and that Heckman's two-step probit selection model corrected this problem in the regression for samples 1–3, but not for sample 4.

Table 7 presents the estimated results at 1%, 5% and 10% significance levels. The results clearly show the significant impact of three main factors: i) the number of school-age children in household, ii) the number of household members who were participating in SI, and iii) job position of informal workers. In addition, Table 7 also illustrates differences in the marginal effect of worker's characteristics on the dependent variable for regressions with samples 2–4. These main findings can be summarised as below.

First, the impact of job position (presented in the model for Sample 1): wage workers had a 1.25 percentage-point higher probability to participate in VSI than the self-employed. This finding was similar to Auerbach et al. (2007) which showed that the self-employed were less likely to participate in SI than salaried workers.

Second, in all estimations for samples 1–4, the number of household members participating in SI (either compulsory or voluntary SI) had positively and statistically significant impact on the informal workers' probability to participate in VSI. The impact of this variable was the highest for wage workers (sample 3): one more member in household participating in SI could result in an increase of 21.1 percentage points in probability to participate in VSI of wage workers. The marginal effects of this variable were 0.124 for self-employed and 0.00964 for all informal workers. These results support the bandwagon effect in VSI participation behaviour in the consumer demand theory.

Third, the number of school-age children (from 6 to 18 years old) in a household had a statistically significant nonlinear effect on probability to participate in VSI for all samples of informal workers. The estimation results of marginal effect were positive

¹ Chow test results showed that at 1% significance level, $F_{value}=1354038.9 > F_{critical}=6.63$. Therefore, the null hypothesis was rejected. It meant that there were differences among coefficients in the model for the self-employed and the remaining group.

Table 7. Marginal effects of factors affecting probability of participating in VSI

	Informal workers	Self-employed	Wage workers	Family workers
	0.0042	0.3249	0.5132	0.0080
Probability (d_voluntary_si=1) (predict)				
Age group (Ref: 15–24)				
25–34	3.95E-05 (0.000582)	-0.2130*** (0.0449)	0.0780*** (0.0141)	0.00535 (0.0173)
35–49	-0.00121* (0.000630)	-0.286*** (0.0557)	0.0764*** (0.0168)	-0.00184 (0.00398)
50–59	-0.00331*** (0.000484)	-0.295*** (0.0683)	0.0806*** (0.0274)	-0.00756 (0.0180)
60+	-0.00322*** (0.000337)	-0.283*** (0.0780)	0.119** (0.0472)	-0.00141 (0.00530)
Gender (Ref: Female)				
Male	-0.00141*** (0.000368)	0.0384** (0.0154)	-0.208*** (0.00978)	-0.00644 (0.0147)
Marital status (Ref: Single)				
Married	-0.000299 (0.000452)	-0.00432 (0.0250)	-0.00737 (0.0120)	0.00347 (0.00878)
Widowed, divorced, and separated	0.00175 (0.00107)	0.0334 (0.0289)	-0.00432 (0.0261)	-0.00932 (0.0254)
Education level (Ref: Incomplete primary education)				
Primary to high school	0.00314*** (0.000750)	0.0333** (0.0170)	0.260*** (0.0235)	0.00846 (0.0219)
Intermediate or professional college level	0.00872*** (0.00290)	0.00475 (0.0269)	0.422*** (0.0462)	0.0121 (0.0337)
University or higher degree	0.00282 (0.00194)	-0.0565 (0.0344)	0.511*** (0.0547)	0.002 (0.0113)
Vocational training (Ref: No vocational training)	0.00103** (0.000432)	0.0719*** (0.0152)	0.0208 (0.0140)	0.00292 (0.0125)

Table 7. Continued

	Informal workers	Self-employed	Wage workers	Family workers
Duration of residency (Ref: less than 1 year)				
1–5 years	0.0038 (0.00403)	-0.0411 (0.0409)	0.0636* (0.0352)	-0.00422 (0.0115)
More than 5 years	0.00309*** (0.000876)	-0.0755* (0.0400)	0.121*** (0.0325)	-0.00055 (0.00651)
Living area (Ref: Rural)				
Urban	-0.00138*** (0.000336)	-0.0460*** (0.0123)	-0.0980*** (0.0148)	0.00751 (0.0182)
Head of household (Ref: Being head of household)				
	-0.000519 (0.000399)	-0.151*** (0.0320)	0.0265** (0.0121)	0.0126 (0.0437)
Number of school-age children in household				
	0.00241*** (0.000602)	0.0405*** (0.0131)	0.0415*** (0.0142)	0.0145 (0.0352)
Square of number of school-age children in household				
	-0.000709*** (0.000252)	-0.0140*** (0.00538)	-0.00903* (0.00536)	-0.00416 (0.0102)
Number of household members participating in the social insurance system				
	0.00964*** (0.000631)	0.124*** (0.00590)	0.211*** (0.0202)	0.0226 (0.0542)
State of economic activity (Ref: unemployment)				
	-0.000452 (0.00170)	-0.0377 (0.0448)	0.00181 (0.0554)	0.00646 (0.0172)
Work experience (Ref: Less than 5 years)				
5–10 years	0.00257*** (0.000505)	0.0376*** (0.0118)	0.0762*** (0.0140)	0.000327 (0.00266)
More than 10 years	0.00150*** (0.000374)	-0.0061 (0.0125)	0.0456*** (0.0153)	-0.00438 (0.0112)
Job position (Ref: Self-employed)				
Wage workers	0.0125*** (0.00115)			
Family workers	-0.000109 (0.00109)			

Table 7. Continued

	Informal workers	Self-employed	Wage workers	Family workers
Type of working establishment (Ref: Agri-forestry-fishery households, freelancer workers and individual business establishment)				
Cooperatives	0.0263*** (0.00891)		0.131*** (0.0466)	0.0697 (0.130)
State, non-state and foreign organisations	0.0203*** (0.00201)	-0.490*** (0.0387)	0.235*** (0.0323)	-0.000864 (0.00342)
Labour contract				
Less than 3 months	-0.00452*** (0.000310)		-0.521*** (0.0671)	
Verbal agreement	-0.00760*** (0.000497)	-0.427*** (0.0782)	-0.286*** (0.0204)	
Business registration status (Ref: No business registration)				
Group 2 (Near poor)	-0.000284 (0.000680)	0.848*** (0.0427)	0.138*** (0.0241)	0.0104 (0.0300)
Group 3 (Middle)	0.000359 (0.00106)	0.133*** (0.0457)	-0.0383 (0.0832)	-0.0104 (0.0283)
Group 4 (Near rich)	0.00348** (0.00137)	0.0708* (0.0414)	0.177*** (0.0640)	0.00249 (0.00866)
Group 5 (Richest)	0.00672*** (0.00175)	0.0811** (0.0410)	0.217*** (0.0653)	0.00775 (0.0184)
<i>Rho</i>	0.479*** (0.06579)	0.114*** (0.0414)	0.236*** (0.0651)	0.00548 (0.0135)
Observations	400,583	400,583	400,583	400,583

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Source: Authors' estimates, using *Labour Force Survey 2016*.

for variable “number of school-age children in household” but negative for variable “square of the number of school-age children in household”, meaning that one more school-age child increased informal workers’ probability to participate in VSI, but once the number of school-age children was higher, one more school-age child decreased informal workers’ probability to participate in VSI. These results are consistent with those in Heenkenda (2016), which showed that the number of dependents in a worker’s household was positively associated with his/her dropout from a voluntary pension scheme.

In addition to the aforementioned variables, Table 7 also shows different impacts of other explanatory variables depending on type of workers.

Age of workers had different effects on the probability of participating in VSI among groups of samples. In general, however, comparing marginal effect coefficients of variable “age” in all models shows that age had a nonlinear effect on probability to participate in VSI, and older workers were less likely to participate in VSI. For samples 1, 2 and 4, workers aged over 60 years old had a higher probability to participate in VSI than those aged 50–59. These results are not completely consistent with previous studies; for example, Auerbach et al. (2007), Carpio (2011) or Heenkenda (2016) showed that age had a positive impact on participation in SI, while Collins-Sowah et al. (2013) and Castel (2008) found a negative effect of age on SI participation.

Finding for the Vietnamese case in this study could be explained by the context of the social insurance system in Viet Nam. According to the Law on Social Insurance in 2014, the minimum number of years of contribution to get monthly pension is 20 years, while the retirement age is 55 years old for female workers and 60 years old for male workers. As such, workers aged 35 years and older were less likely to participate in VSI because they could not meet the requirement of the minimum number of years of contribution.

Gender had a negatively and statistically significant impact on probability to participate in VSI for the samples of all informal and wage workers, meaning that male workers are less likely to participate in VSI than their female counterparts. In contrast, for the self-employed, male workers had a 3.84 percentage point higher probability to participate in VSI than their female counterparts.

The highest education level had a positively and statistically significant effect on informal wage workers’ participation in VSI. This result was consistent with those in Raghupathi and Raghupathi (2020) and Przybytniowski (2017). For samples 1, 2 and 4, however, there were statistically significant effects of one or two out of three dummy variables of the highest education level. More specifically, self-employed with primary to high school level were 3.33 percentage points more likely to participate in VSI than those with incomplete primary school level. There was no statistically significant difference, however, between self-employed with primary to high school level and those with higher education level in participating in VSI. The results were similar for the sample of family business workers. Auerbach et al. (2007), Castel (2008) and Giles et al. (2013) found a positive impact, while Collins-Sowah et al. (2013) and Nivakoski (2014) found a negative impact of education level on probability to participate in SI.

For duration of residency, the regression results for samples 1 and 3 show that this variable had a positively and statistically significant effect on likelihood to participate

in VSI. In contrast, for sample 2, the self-employed with long-term residence was 7.55 percentage points less likely to participate in VSI than the other groups.

Regarding area of residence (urban vs. rural), in models for samples 1–3, urban workers had a 0.138 percentage points, 4.6 percentage points and 9.8 percentage points lower probability to participate in VSI than rural workers, respectively. In the model of sample 4, this variable was not statistically significant.

The variable “being the household head” had different impacts on the probability to participate in VSI, depending on the studied samples. For all informal workers (sample 1) and family business workers (sample 4), this variable had no statistically significant impact. It had a positive impact on the probability to participate in VSI for wage workers (sample 3), while a negative impact on the probability to participate in VSI for the self-employed (sample 2).

In regard to work experience, in the model of all four samples, workers having 5 to 10 years of experience in their current jobs were more likely than those with less than 5 years of experience to participate in VSI. It is interesting, however, that marginal effects of dummy variable “more than 10 years of experience” were lower than those for “5–10 years of experience” for samples 1 and 3. These results were consistent with those in Nivakoski (2014) and Roushdy and Selwaness (2019) which showed a non-linear effect of work experience on participation in VSI.

The variable showing type of working establishments had statistically significant effects on probability to participate in VSI for the self-employed (sample 2) and wage workers (sample 3). Similarly, in terms of business registration status of establishments, the self-employed and wage workers in registered establishments had 84.8 percentage points and 13.8 percentage points higher probability to participate in VSI than did those in non-registered establishments, respectively.

Results from four samples of informal workers show that a worker’s income level also had different effects on his/her probability to participate in VSI. For the sample of all informal workers, the near-rich group (the 4th income quintile) and the richest group (the 5th income quintile) were more likely to participate in VSI than the poorest group (the 1st income quintile). This finding was similar to those in Carpio (2011), Castel (2008) and Heenkenda (2016).

For the self-employed, there was a statistically significant nonlinear effect of income on their participation in VSI. The near-poor income group (the 2nd income quintile) and the richest group had the highest probabilities to participate in VSI (with the marginal effects at 0.133 and 0.114, respectively, relatively to the poorest group). The marginal effect for the middle group (the 3rd income quintile) and the near-rich group (the 4th income quintile) was 0.078 and 0.0811 relatively to the poorest group, respectively. For wage workers, income level had a positive significant effect on VSI participation for the middle to richest groups, but no significant effect for the poorest and the near-poor groups.

4. Policy Implications and Concluding Remarks

This study provided an empirical study about the impacts of job position and other characteristics of informal workers on their participation in VSI, using Heckman’s two-

step probit selection model. We found that: i) there was a nonlinear effect of the number of school-age children in household on VSI participation of informal workers in general and all sub-groups of informal workers by job position in particular, ii) the bandwagon effect was presented by positive significant effect of the number of household members participating in SI on VSI participation in all samples of informal workers, iii) job position had a statistically significant effect on VSI participation of informal workers, and iv) there were differences in effects of demographic and employment characteristics on VSI participation among different sub-groups of informal workers by job position.

These findings could help provide some policy implications to expand VSI coverage for informal workers in Vietnam. For example, a reduction of the minimum years of contribution would be helpful for workers at middle ages to be eligible for participating in VSI, and thus having monthly pension benefits later. Also, a matching contribution or child grant programs would facilitate workers with low income or having many children to be affordable with the required minimum contribution level.

Although this study could add new findings to empirical literatures, it had some limitations, mostly due to data. For instance, we did not know whether an informal worker and his/her spouse had private insurance since this might be an alternative for VSI.

References

- Auerbach, P., Genoni, M.E., & Pagés C. (2007). *Social security coverage and the labour market in developing countries* (IZA Discussion paper no. 2979). Institute of Labor Economics. <https://www.iza.org/publications/dp/2979/social-security-coverage-and-the-labor-market-in-developing-countries>
- Bhan, G., Goswami, A., & Revi, A. (2014), The intent to reside: Spatial illegality, inclusive planning, and urban social security. In Ministry of Housing and Urban Poverty Alleviation Government of India, *State of urban poor report 2013* (pp. 83–94). Oxford University Press.
- Carpio, M.A. (2011). *Do pension wealth, pension cost and the nature of pension system affect coverage? Evidence from a country where pay-as-you-go and funded systems coexist* (MPRA Paper No. 34926). University Library of Munich, Germany. <https://mpra.ub.uni-muenchen.de/34926/>
- Castel, P. (2008). *Voluntary defined benefit pension system willingness to participate: The case of Viet Nam* (SSRN Working Paper). <https://doi.org/10.2139/ssrn.1379607>
- Collins-Sowah, P.A., Kuwornu, J.K.M., & Tsegai, D. (2013). Willingness to participate in micro pension schemes: Evidence from the informal sector in Ghana. *Journal of Economics and International Finance*, 5(1), 21–34. <https://doi.org/10.5897/JEIF12.097>
- Dang, N.A., Tacoli, C., & Thanh, H.X. (2003). *Migration in Viet Nam: A review of information on current trends and patterns, and their policy implication*. Paper presented at 'Migration, Development and Pro-Poor Policies in Asia' Conference, Dhaka, 22–24 June. <https://iied.org/sites/default/files/pdf/migrate/G00479.pdf>
- Do, T.T. (2023). *Voluntary social insurance in Viet Nam*. Publishing House of the National Economics University, Viet Nam.
- Duong, L.B., Linh, T.G., & Thao, N.T.P. (2021). *Social protection for rural-urban migrants in Viet Nam: Current situation, challenges and opportunities* (CSP Research Report 08). Centre for Social Protection, Ha Noi, Viet Nam.

- Giles, J., Wang, D., & Park, A. (2013). Expanding social insurance coverage in urban China. In C. Giuliotti, K. Tatsiramos and K.F. Zimmermann (Eds.) *Labor Market Issues in China (Research in Labor Economics, Vol. 37)* (pp. 123–179). Emerald Group Publishing Limited. [https://doi.org/10.1108/S0147-9121\(2013\)0000037009](https://doi.org/10.1108/S0147-9121(2013)0000037009)
- GSO & ILO (2017). *2016 Report on informal employment in Viet Nam*. General Statistics Office and International Labour Organization. Hong Duc Publishing House.
- Heenkenda, S. (2016). *Determinants of dropouts from voluntary pension scheme: Evidence from Sri Lanka* (MPRA Paper No. 72481). Munich Personal RePEc Archive. <https://mpra.ub.uni-muenchen.de/72481/>
- ILO (2012). *Social security for all: Building social protection floor and comprehensive social security system*. International Labour Organization.
- Mandigma, M.B.S. (2016). Determinants of social insurance coverage in the Philippines. *International Journal of Social Science and Humanity, 6*(9), 660. <https://doi.org/10.18178/ijssh.2016.6.9.728>
- Morrissey, K., Kinderman, P., Pontin, E., Tai, S., & Schwannauer, M. (2016). Web based health surveys: Using a Two Step Heckman model to examine their potential for population health analysis. *Social Science & Medicine, 163*, 45–53. <https://doi.org/10.1016/j.socscimed.2016.06.053>
- Nivakoski, S. (2014). Determinants of pension coverage and retirement income replacement rates – Evidence from TILDA. *Economic and Social Review, 45*(3), 299–328.
- Pindyck, R.S., & Rubinfeld, D.L. (2014), *Microeconomics* (8th edition). Pearson Education.
- Posel, D.R. (2001). Who are the heads of household, what do they do, and is the concept of headship useful? An analysis of headship in South Africa. *Development Southern Africa, 18*(5), 651–670. <https://doi.org/10.1080/03768350120097487>
- Przybytniowski, W.P. (2017). The level of insurance knowledge of young people entering professional life – results of questionnaire surveys conducted among students of the Podkarpackie Voivodship. *Jagiellonian Journal of Management, 3*(2), 107–121. <https://doi.org/10.4467/2450114XJJM.17.008.7738>
- Raghupathi, V., & Raghupathi, W. (2020). The influence of education on health: An empirical assessment of OECD countries for the period 1995–2015. *Archives of Public Health, 78*(1), 1–18. <https://doi.org/10.1186/s13690-020-00402-5>
- Roushdy, R., & Selwaness, I. (2019). Who is covered and who under-reports: An empirical analysis of access to social insurance in Egypt. *Journal of International Development, 31*(8), 720–751. <https://doi.org/10.1002/jid.3434>
- Van de Ven, W.P., & Van Praag, B.M. (1981). The demand for deductibles in private health insurance: A probit model with sample selection. *Journal of Econometrics, 17*(2), 229–252. [https://doi.org/10.1016/0304-4076\(81\)90028-2](https://doi.org/10.1016/0304-4076(81)90028-2)
- Vietnam's National Assembly (2014). *The law on social insurance*. Hanoi, Viet Nam.
- Wooldridge, J.M., Wadud, M., & Lye, J. (2016). *Introductory econometrics* (Asia Pacific edition) Cengage AU.
- Yao, J., & Kim, B.C. (2015). Social insurance participation of rural migrant workers based on gender dimension: Evidence from four Chinese cities. *Asian Social Work and Policy Review, 9*(1), 57–69. <https://doi.org/10.1111/aswp.12045>