



PUEY UNGPHAKORN INSTITUTE
FOR ECONOMIC RESEARCH

Understanding a Less Developed Labor Market through the Lens of Social Security data

by

Nada Wasi, Chinnawat Devahastin Na Ayudhya, Pucktada Treeratpituk
and Chommanart Nittayo

January 2021

Discussion Paper

No. 147

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

Understanding a Less Developed Labor Market through the Lens of Social Security data

Nada Wasi, Chinnawat Devahastin Na Ayudhya

Pucktada Treeratpituk, and Chommanart Nittayo

December 2020

Abstract

While understanding labor market dynamics is crucial for designing the country's social protection programs, prohibitive longitudinal surveys are rarely available in less developed countries. We illustrate that employment history from Social Security records can provide several important insights by using data from a middle-income country, Thailand. First, in contrary to the traditional view, we find that the formal and informal sectors are quite connected. Our analysis of millions of individual histories by a machine learning technique shows that more than half of registered workers left the formal sector either seasonally or permanently long before their retirement age. This finding raises a question of whether the social protection schemes being separately designed for formal and informal workers are effective. Second, the semi-formal workers also had a much flatter wage-age profile compared to those always staying in the formal sector. This observation calls for effective redistributive tools to prevent earnings inequality to translate into disparities in old-age and transmit to the next generation. Lastly, on the employer size, we find that almost half of formally registered firms had fewer than five employees, the benchmark often used to define informal firms. This result suggests that the distributions of firm sizes differ across countries and the employer size alone is unlikely sufficient to define informal workers.

Keywords : Employment, Work history, Social Security, k-means clustering, Thailand

I. Introduction

Understanding the dynamics of individual employment patterns is important for several reasons. First, there exists an intertwining relationship between work patterns and social protections. On one hand, taxes and social programs can reduce work incentives, encourage early retirement, or reduce job search effort (see *e.g.*, Nicholson and Needels, 2006; Blundell *et al.*, 2016; Keane and Wasi, 2016). On the other hand, an appropriate design of social protections is contingent on the country's labor market structure. In developing countries, how supporting schemes for informal workers should be designed also depends on whether their workers choose to stay in either the formal or informal sector throughout their working lives or frequently switch between the two sectors (Winkler *et al.*, 2017).

Second, the trend of population aging brings several challenges including old-age poverty and fiscal sustainability of elderly supporting programs. Since pension is often based on beneficiaries' earnings-based contribution during their working life, individual earnings history can be used to assess whether pension is likely adequate. The work history can also be used to validate financing projection assumptions of several pay-as-you-go schemes.

Lastly, the longitudinal data of workers and firms can shed light on the connection between the formal and informal sectors, and the evolution of firm size distribution. These issues are of interest to development researchers. Traditionally, the classical development theory explains that the informal sector shrinks as a country develops. The efficient formal firms replace inefficient informal ones in the process of economic development and absorb labor from the informal sector (*e.g.*, Rauch, 1991; La Porta & Shleifer, 2014) The mean and dispersion of the distribution of firm size is found to increase with development (Poshke, 2018).

In this study, we illustrate that employment history from Social Security records can provide useful insights when longitudinal worker and firm surveys are not available. Administering longitudinal surveys are costly and require specific expertise, making it rarely

available in less developed countries. This article uses the data from a middle-income country, Thailand, where the Social Security Act came into force in 1990. The compulsory mandate originally applied to large firms only and was later extended to cover employers with at least one employee in April 2002. The monthly contributions from employers and employees enable tracking of both employers and employees over time.

Our contributions are two folds. First, we document dynamic employment patterns of workers and firms in a less developed labor market and discuss their policy implications. Due to data unavailability, there have not been many studies looking at employment dynamics in developing countries. Second, we illustrate that the k-means clustering, a machine learning technique, can be used to uncover the underlying key work patterns from millions of data points of individual job history.

Our results show that in Thailand, less than half of registered workers stayed in the formal sector for over eight years. Others left the formal sector temporarily, seasonally, or permanently long before their retirement age. These semi-formal workers also had a much flatter wage-age profile compared to those always staying in the formal sector. The implications of these work patterns on social protection schemes are discussed. On the employer side, we document that 48 percent of registered Thai firms are indeed micro businesses with fewer than five employees. This contrasts to a common belief that firms with fewer than five workers can be considered informal. In addition, we also document the increasing concentration in the labor market in terms of employments. The number of workers among the Top-6 firms quadrupled between 2002 and 2018 while more than half of medium-size businesses reduced their size.

The remaining of the paper is organized as follows. The next section presents the background on the labor market in developing countries and in Thailand including the country's existing social protection schemes. The data and methodology are discussed in

Section III. Section IV presents our results and the last section provides conclusions and discussions on policy implications.

II. Background

II.A The labor market in developing countries

The labor markets in developing countries are known to consist of both formal sector and sizable informal sector. Informality implies that informal firms avoid taxes and regulations, and informal workers are not protected from eminent risks such as sickness, unemployment, and old-age poverty. A range of alternatives have been used to define formal and informal economies. Common practices include defining “formal workers” as those registering with the official social security system; or working in a large firm. The threshold of number of employees for being considered large is arbitrarily and often depends on questionnaire designs. For instance, five, ten or 25 employees have been used (Feige, 1990; Henley *et al.*, 2009.)

Traditionally, the formal and informal sectors are viewed as disconnected parallel economies (*e.g.*, Rauch, 1991). Informal firms are typically small, have lower productivity and pay lower wage compared to the formal economy (Freeman, 2010; Rothenberg *et al.*, 2016). Informal workers are poorly educated and operate low-skill businesses (Banerjee and Duflo, 2007). When the country develops and grows, the informal sector shrinks as descendants of informal workers choose to work with formal firms having higher productivities and being able to offer higher wages (La Porta & Shleifer, 2014). The average firm sizes are also reported to increase with per capita income within and across countries (Poschke, 2018). Larger firms are often found to pay observationally equivalent workers higher wages (*e.g.*, Brown and Medoff, 1989) except some specific cases where few firms dominate (Benmelech *et al.*, 2018).

Recently, some of these views have been challenged. The formal and informal sectors appear quite connected in some countries. In Mexico and Brazil, workers shift from one sector

to the other in response to their economic conditions and transitions take place within a relatively short period of time (Bosch and Maloney, 2007; Meghir *et al.*, 2015). Moreover, despite the rapid economic growth, the informal sector in China, India and the Philippines does not seem to shrink (Freeman, 2010).¹ In addition, growing online labor platforms have led to a rising number of non-standard employees including the young and better educated who prefer flexible working schedule and/or location (OECD 2018).

The country's labor market structure also implies how its social protections should be designed. This issue has been particularly challenging for countries with a large informal economy. Extending social insurance programs available to formal workers to informal workers is not straightforward because most informal workers have irregular and low income. Several countries have adopted a non-contributory scheme for the informal sector, but the financial burden of such scheme is still a concern. The co-existence of the contributory scheme for the formal sector and non-contributory scheme for the informal sector could also imply an implicit tax on formal jobs and reduce incentives to enroll and contribute (Winkler *et al.*, 2017).

II.B The Thai labor market and its social protections

Evidence from various cross-sectional data sets suggest that the Thai formal sector has expanded over the past few decades. The estimated shares of informal workers, defined by those not having social protections from their employers, reduced from 62% in 2010 to 55% in 2019 (Thailand's National Statistical Office, 2019). The shares of non-farming households have continuously increased whereas the shares of households with farming income has largely declined (see Figure 1). Consistent with the development literature, the fall of the share of farming households is more apparent among the younger generation (left vs. right figures).

¹ Possible explanations include to labor-saving technology in the formal sector and manufacturing jobs being outsourced to the informal sector

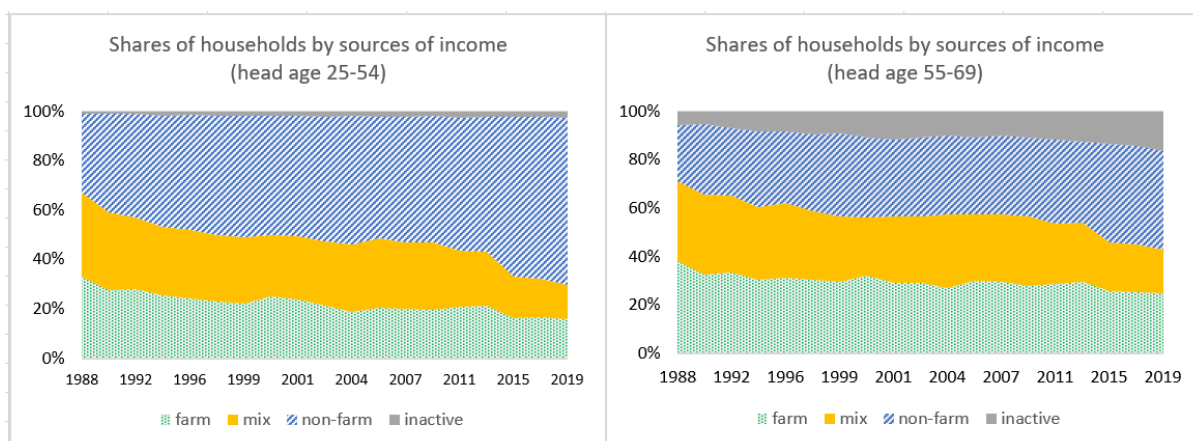


Figure 1. Shares of Households in Thailand by Their Income Sources

The decline in the agricultural sector share is offset by the risen shares of the manufacturing, trade and service sectors. While most of agricultural employment is informal, the informal sector also comprises other small-scale production and services such as street vendors, taxi driver, and household workers. Regarding the employer side, a previous study report that the total number of registered establishments rose from 44,744 to 92,095 firms between 1987 and 1996 (Wiboonchutikula, 2002).

We are not aware of any study documenting individual employment pattern of Thai workers or the evolution of firm size distribution.² Do the Thai workers choose whether to work in the formal or informal sector and stay there throughout their lives? or do they frequently move between the two sectors like Brazilian workers?

Regarding social protections, Thailand has several social insurance and social assistance programs for both formal and informal workers. Each scheme targets different sections of population and is regulated by different authorities. Table 1 summarizes key health insurance and old-age pension schemes.³ For formal workers, public and private employees

²There are studies looking at firms' financial status and their ownership (*e.g.*, Apaitan *et al.*, 2019; Bertrand *et al.*, 2008).

³See more details in Ratanabanchuen (2019) and Sriratanaban (2002).

are covered by distinct schemes. Civil servants are covered by the Civil Servant Medical Benefit Scheme (CSMBS) and Old Civil Servant pension.⁴ These schemes are non-contributory schemes and are viewed as a fringe benefit to help compensate for low public salary rates.

Table 1 : Summary of Key Social Protection Schemes in Thailand

Formal sector workers		Informal sector workers and the unemployed
Government	Private	
Health insurance		
Civil Servant Medical benefit scheme (non-contributory)	Social Security: Article 33 (contributory, mandatory)	Universal Healthcare (non-contributory, universal) Social Security : Article 39, 40 (contributory, voluntary)
Pension		
Old Civil Servant pension Government Pension fund (non-contributory)	Social Security: Article 33 (contributory, mandatory)	Social Security : Article 39, 40 (contributory, voluntary) National Savings fund (contributory, voluntary)
	Elderly allowance (non-contributory, universal)	

Formal private sector employees are covered by the Social Security scheme.⁵ The Thai Social Security Act was enacted in 1990 and originally required employers in non-agricultural sectors with 20 or more employees to register. It was later extended to cover employers with ten or more employees in 1993 and then employers with at least one employee in 2002. This compulsory mandate, also known as Article 33, requires contributions from employees, employers and the government. Current benefit provisions are health, disability, maternity, death, old-age pension and child allowance. The Thai Social Security regulation is relatively

⁴ Government employees employed after 1997 were enrolled into the Government Pension Fund. These young cohort still receive lifetime pensions from the Old Civil Service Pensions but at a lower rate.

⁵ State enterprise staff and private school teachers are excluded because each group is covered by their own insurance schemes.

generous in the sense that workers who voluntarily leave their jobs can claim unemployment benefit if they contribute for at least six months in one calendar year.

Regarding informal workers and the unemployed, there are two key universal non-contributory schemes: the Universal Healthcare Coverage Scheme (UCS) and elderly allowance program. The UCS, introduced in 2001, provides health care for those not covered by the social security and civil servant schemes. The elderly allowance, started in 2009, provides income assistance to all Thai citizens aged 60 years or older except for retired government officers. In addition, the Social Security Act provides voluntary contributory schemes for informal workers with a range of options where plans with higher coverage require higher contribution rates (Articles 39 and 40). The country also started the National Savings Fund (NSF) in 2015 where the government helps match the contribution (with a cap) and the fund provides retirement annuity when eligible members reached the age of 60.

Most of these public schemes are parallels in the sense that workers who are eligible for multiple schemes must choose only one. The exception is the elderly allowance where Social Security beneficiaries and NSF members aged 60 or over are also eligible as their pension are likely inadequate for livings. The maximum social security taxable income has been capped at 15,000 Baht for over thirty years and the pension benefits have never been adjusted for the rising cost of living.

III. Data & Methodology

Our data consist of employers and employees registering under the Social Security Article 33 during April 2002-October 2018. Although the Social Security Office (SSO) came into force in 1990, small firms were exempt until April 2002. During the studied period, the number of registered firms increased from 237,817 to 385,879, and the numbers of registered employees rose from 6.3 million to 11.5 million.

While SSO data lack details of some demographic information, such as education or occupation, its advantages over cross-sectional labor force surveys are that (i) we can track individual employees and firms over a long period of time; (ii) administrative data are not subject to the measurement error problem; and (iii) monthly data are more suitable for the purpose of understanding job entry and exit information.

We selected two cohorts of employees to study their employment dynamics: those aged 15-44 years in 2002 and those aged 15-44 years in 2010. These cohorts consisted of 5.2 million and 6.8 million workers, accounting for 82% and 77% of all SSO employees in 2002 and 2010, respectively. Each cohort was then followed for 96 months. We did not follow those 45 years old or older as they approached the SSO pensionable age of 55.

The amount of monthly information derived from over five million workers makes it impossible to analyze every possible path of work history. To uncover the key underlying work patterns, we use the k-means clustering technique to obtain natural segmentation. This machine learning technique helps classify individuals with a similar set of work history to be in the same segment (cluster). Nevertheless, the information being input to the machine is still based on researchers' judgment. We construct six variables to capture distinct information related to work history:

- (1) the number of jobs in the formal sector over 96 months;
- (2) job tenure;
- (3) the number of times an employee exited and returned to the formal sector;
- (4) out-of-formal sector duration;
- (5) the number of times an employee returned to the same employer after leaving; and
- (6) the last month observed as being employed in the formal sector.

The k-means clustering works as follows. To classify observations $i = 1, \dots, N$ with K characteristics $(x_k^i$ for $k = 1, \dots, K)$ into C clusters, the algorithm first randomly picks C cluster centroids. Let $\mu^c = (\mu_1^c, \dots, \mu_k^c, \dots, \mu_K^c)$ denotes the centroid or the mean characteristics for cluster c . Then, each observation is assigned to the cluster with the most similar centroid. The similarity is based on the Euclidean distance, where the distance between observation i and the centroid of a cluster c is calculated by:

$$d_E(\mathbf{x}^i, \mu^c) = \sqrt{\sum_{k=1}^K (x_k^i - \mu_k^c)^2}.$$

Once all N observations are assigned to clusters, the mean of the characteristics (μ^c) within each cluster is recomputed. Next, each observation is reassigned to the closest cluster based on the newly computed centroids. The process is repeated until all the centroids are stabilized. In our case, $K = 6$ and $N = 5.2$ million (6.8 million) workers for the first (second) cohort.

Note that our employment-related variables are at the individual level, when there are multiple observations for that individual (*e.g.*, job tenure), the median value is used. All variables except for the last month observed were transformed into a logarithmic form and standardized to have a mean of zero and a standard deviation equal to one.

To determine the appropriate number of clusters, we use the “elbow analysis” to find an optimal point that balances the trade-off between the within group homogeneity and the simplicity of the model. The total within-cluster distance is measured by $\sum_{c=1}^C \sum_{i \in c} d_E(\mathbf{x}^i, \mu^c)^2$ where a lower number indicates a higher degree of similarity. By increasing the number of clusters, the model can always raise the degree of similarity within group because the complex model has more freedom in grouping similar data points together. However, after a certain point the improvement in within group homogeneity becomes

marginal, resulting in an “elbow” shape plot. This kink point is typically selected as the appropriate number of clusters.⁶

Monthly wage data is deflated by the 2015 CPI. The wage data has been censored at the maximum social security taxable value, THB 15,000 (about USD 472) for the whole period. Not surprisingly, the number of censored observations increased over time from 10% in 2002 to 33% in 2018. Therefore, when discussing wages, we report median values, instead of mean.

IV. Results

IV.A What are the work patterns?

The k-means clustering technique suggests that there are four clusters underlying the work history of both cohorts as shown in Table 2. Figure A1 shows the “elbow plot” where the kink is at four clusters.

For the first cohort (aged 15-44 years in 2002), the largest cluster (38%) captures workers staying in the formal sector for the entire period (2002-2010). These employees had a stable job-- ninety percent had only one or two jobs within the span of eight years. Most of them never left the formal sector. Thus, we label this group *fully formal*. The second cluster (33%) typically had two to four jobs over eight years. When transitioning from one job to another job, they tended to exit the formal sector for two to fifteen months. This group is labeled *hoppers*.

The third cluster (14%) is labeled *seasonal* because each year these employees had two to eight months out of the formal sector, but then returned to the same employer. The months when they left the formal labor market were around May-October, which is the country’s

⁶ The concept is analogous to selecting the number of classes for the semi-parametric latent class model. The model selection criteria, such as BIC balances the in-sample fit with the model complexity (the number of parameters).

harvesting season. The last cluster (15%) consists of those staying in the formal sector for seven to 28 months and never returning (*shortly observed*).

Table 2: Employment characteristics in each cluster

Cohort 1: 2002-2010 Aged 15-44 years in 2002	cluster 1			cluster 2			cluster 3			cluster 4		
	fully formal: 38%			hoppers: 33%			seasonal: 14%			shortly observed: 15%		
	percentile			percentile			percentile			percentile		
Characteristics	25th	50th	75th	25th	50th	75th	25th	50th	75th	25th	50th	75th
1) number of jobs	1	1	2	2	3	4	5	6	8	1	1	1
2) job tenure	48	96	96	9	16	30.5	3	4	7	5	12	23
3) no. of times exiting the formal sector	0	0	0	1	1	2	2	3	4	0	0	0
4) out-of-formal sector duration	0	0	0	2	5	15.5	2	4	8	0	0	1
5) no. of times returning to same employer	0	0	0	0	0	0	1	1	2	0	0	0
6) last month observed in the formal sector	96	96	96	77	96	96	77	96	96	7	16	28
Cohort 2: 2010-2018 Aged 15-44 years in 2010	cluster 1			cluster 2			cluster 3			cluster 4		
Characteristics	fully formal: 42%			hoppers: 30%			seasonal: 13%			shortly observed: 15%		
	percentile			percentile			percentile			percentile		
Characteristics	25th	50th	75th	25th	50th	75th	25th	50th	75th	25th	50th	75th
1) number of jobs	1	1	2	2	3	4	5	6	8	1	1	1
2) job tenure	48	96	96	9	17	33	3	4	7	5	12	23
3) no. of times exiting the formal sector	0	0	0	1	1	2	2	3	4	0	0	0
4) out-of-formal sector duration	0	0	0	2	5	16	2	4	8	0	0	2
5) no. of times returning to same employer	0	0	0	0	0	0	1	1	2	0	0	0
6) last month observed in the formal sector	96	96	96	89	96	96	83	96	96	7	16	28

The results for the second cohort (aged 15-44 years in 2010) are surprisingly similar. The characteristics underlying the four clusters remain comparable, and hence are labeled with the same group names. Two exceptions are that the share of the *fully formal* cluster increases from 38 to 42 percent; and the *hoppers'* job tenure is slightly longer.

Figure 2 illustrates the employment patterns for typical workers in each of the four clusters. Each box represents one job for a period of approximately one year. Job A, Job B, Job C, Job D represent different employers, and grey boxes indicate periods of out-of-the formal

sector employment. While we do not observe whether the workers were unemployed or took informal jobs during their absence from the SSO records, the country’s persistent low employment rates (around 1%) suggests that most were unlikely to have been unemployed. The frequent transition could also be driven by Thai Social Security regulations which do not impose a heavy penalty for leaving the formal market temporarily or permanently.



Figure 2: Typical Work Patterns by Cluster

It is notable that three out of four clusters consist of workers who leave the formal sector temporarily or permanently. These kinds of dynamic patterns cannot be observed in cross-sectional data.

3.2 Are semi-formal workers high-skill or low-skill?

Previous studies have documented that informal workers were mostly poorly educated and received lower wages. Here we examine whether the semi-formal workers also earn lower wages, or they consist of some high-skill workers likely working through an online platform. As workers age, their wages are expected to grow because they accumulate more human capital through work experience (learning-by-doing). Figure 3 provides wage-age profiles for different groups of workers. The left figure shows that the median wages of *fully formal* are higher than

other groups at all ages. The gap is also widening as workers age. It is striking that the median wages for the three semi-formal groups are rather flat starting from about the age of 30.

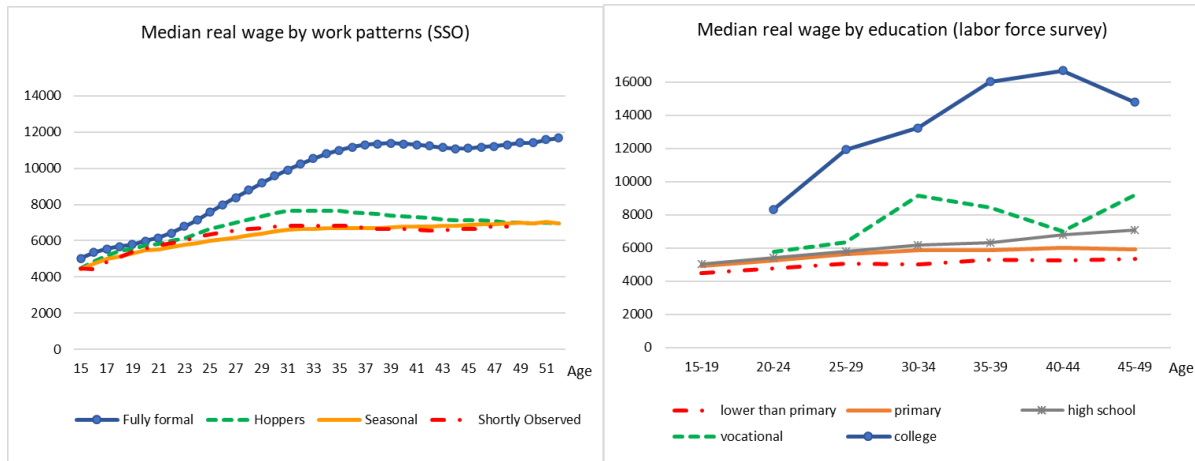


Figure 3: Median Real Wages by Work Patterns and Education Levels

In fact, the wage profiles of the *fully formal* vs. the three *semi-formal* groups resemble the wage profiles of *college* vs. *non-college* groups using cross-sectional labor force survey data for corresponding cohorts (Figure 3, right). Although this result suggests that the semi-formal workers in Thailand comprise low-wage low-skill workers, we cannot rule out a possibility that those who left and never returned (or never register) to the formal sector command higher earnings in the informal sector.

By further comparing workers who switched jobs with those staying in the same job, we find that changing jobs is associated with steeper wage growth for the *fully-formal* group, but not the *hoppers* (see Figure 4).⁷ This pattern confirms that the *fully-formal* workers likely possess higher skills and the reasons driving job switching are rather different. The *fully-formal* workers are offered higher wages at a new job as firms compete for talents. In contrast, the

⁷ This analysis uses a subset of workers in the two largest clusters who did not permanently leave the formal sector.

hoppers' short job tenures and flat wage profile likely reflect their relatively poor performance and/or the excess supply of their skill-level.

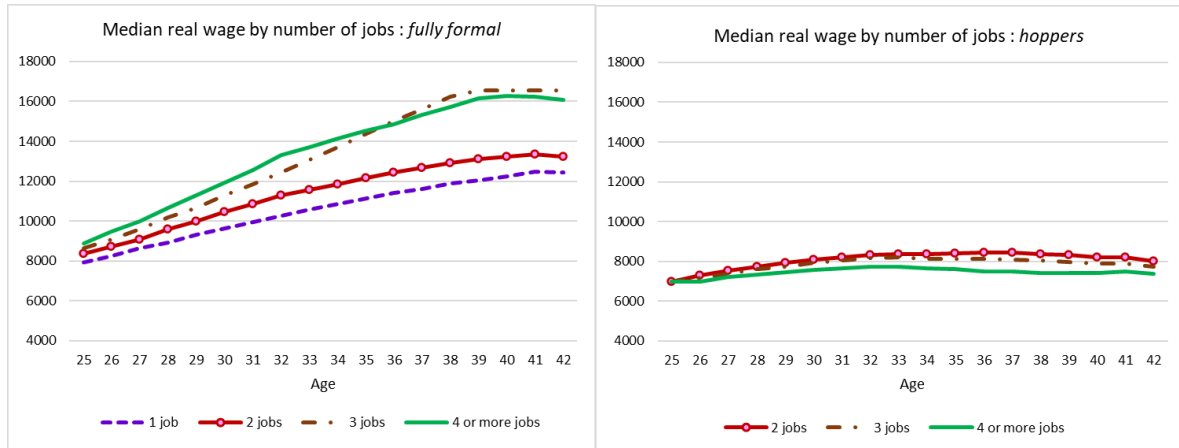


Figure 4: Median Real Wages of *fully formal* and *hoppers* by Number of Jobs

The wage-age profiles across the four clusters for the 2010 cohort are similar to those of the 2002 cohort, except that their slopes are steeper for certain years, likely driven by the 40% increase in minimum wage in April 2012. The results are available upon request. The employment patterns of different groups of workers and their associated wage profiles have implications for both policies aiming to reduce inequality and to extend social protections to informal workers. The issues will be discussed in the last section.

3.3 What are the stylized facts related to firm sizes?

Here we present three stylized facts related to firm sizes: the distribution, the evolution and wage premium.

First, during the studied period, 46-48% of formally registered Thai firms had less than five employees. This contrasts with the common practice, defining firms with five employees as informal firms. Table 3 presents the shares of number of firms and their employment by firm sizes in 2002 and 2018. Even though the total number of firms largely increased during this

period, the shares of number of firms by sizes remained stable. In both years, the shares of firms with 200 or fewer employees were 97.7%. The shares of firms with 2001-1000 and 1001 or more employees were 2% and .3%, respectively. However, the employment shares among firms with more than 1000 employees rose from 27% to 35% whereas the employment shares among firms with 51-1000 employees declined.

Table 3 : Shares of Firms and Their Employment by Firm Sizes

	No. of employees					
	less than 5	5-10	11-50	51-200	201-1000	1001 or more
Shares of number of firms						
2002	46%	26%	20%	5%	2%	0.3%
2018	48%	24%	21%	5%	2%	0.3%
Shares of employment						
2002	4%	7%	16%	18%	28%	27%
2018	4%	6%	16%	17%	24%	35%

Second, when further classifying large companies into four sizes, it turns out that the employment growth among large firms came from exceptionally large ones, *i.e.*, those with 5,000 or more employees.

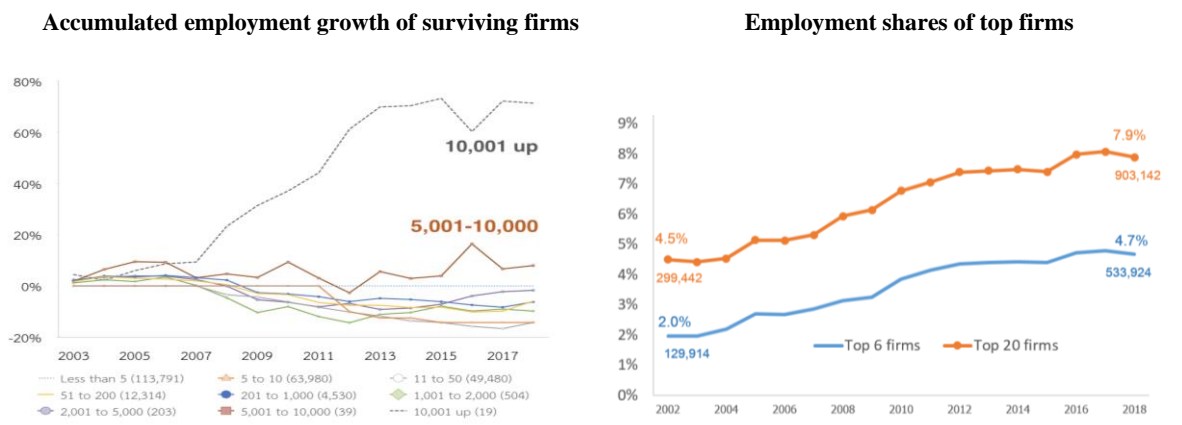


Figure 5: Employment Growth and Employment Shares by Firm Sizes

This can be seen in the left panel of Figure 5 where we plot the median accumulative employment growth since 2002 for surviving firms by their sizes. More than half of firms with less than 5000 employees were smaller compared to their size in 2002. The right panel depicts the rising concentration of employment among the Top-6 and Top-20 firms. The employment share of the Top-6 firms was more than doubled during this period, increasing from 2% (129,914 employees) in 2002 to 4.7% (533,924 employees) in 2018.

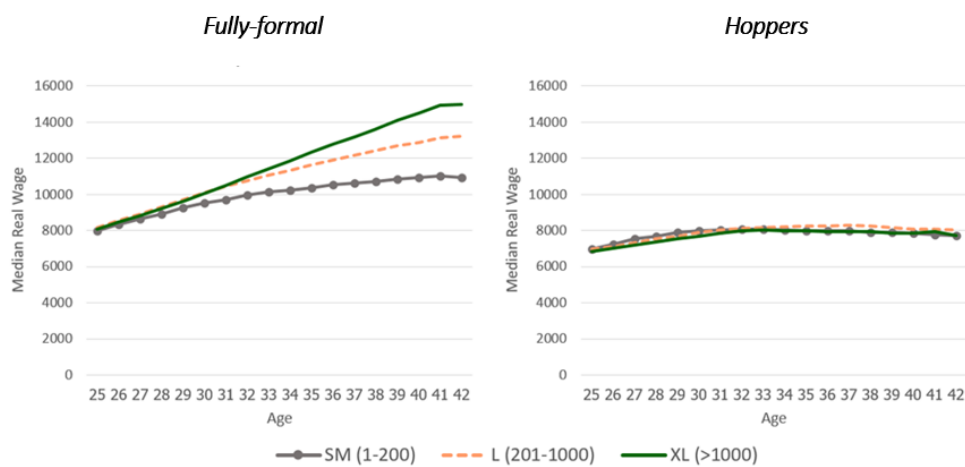


Figure 6: Median Real Wages by Firm Sizes

Lastly, we find that large firms paid higher wages for the likely high-skill workers, but not medium or low-skill workers. Figure 6 illustrates that among the *fully formal* of the same age, those working in a large firm earned more. However, for *hoppers*, who presumably possess lower skill levels than *fully formal*, their wages did not correlate with firm sizes. The results imply that in Thailand, larger firms pay higher wages only for workers who are more productive, but not necessary all the workers.

6. Conclusions and Policy Implications

We have illustrated that employment history from Social Security records can provide important labor market insights. These dynamics have been masked in annual cross-sectional surveys. Our key results and implications are as follows.

First, in contrary to the traditional view, our evidence supports an alternative view that the formal and informal sectors in developing countries are connected. More than half of the Thai employees who registered with the Social Security did not always stay in the formal sector. Some frequently transitioned in and out of the formal sector. This observation raises a question of whether the social protection schemes being independently designed for the formal and informal workers are effective. When low-wage workers often move between the two sectors, parallel schemes designed for each sector can be regressive and function improperly.

This is likely the case for Thailand where the country has the mandatory Social Security scheme and the National Savings Fund separately targeting formal and informal workers, respectively. Since both schemes use the numbers of contribution years to determine eligibility and pension benefits, workers frequently switching between the two sectors may not be eligible in either scheme. In addition, the unemployment provision has not functioned as an insurance, but rather served as an income assistance. Most of those claiming the unemployment benefits are those voluntarily leaving the formal sector, including those moving to the informal sector temporarily or permanently (Wasi *et al.*, 2018).

Second, the diverged wage-age profiles across different groups of workers point to a high degree of heterogeneity. The semi-formal workers had a much flatter wage-age profile compared to those always staying in the formal sector. This implies that without effective redistributive tools, earnings inequality can translate into disparities in old-age poverty and transmit to the next generation. Currently, Thailand has a progressive income tax system, but its Social Security pension does not have a redistributive feature. Conditional on having the

same number of service years, the pension-to-income replacement ratios are constant. The replacement rate is 20% of the average salary over the last 60 working months if the insured contributes for 15 years. For every additional year of contribution, the replacement rate increases by 1.5%. In other countries, such as the US, the replacement rate is higher for beneficiaries with lower career average income.

Furthermore, the heterogeneity of lifetime work patterns and earnings across individuals is often ignored in an actuarial valuation projection (for example, ILO, 2016). In such evaluation, everyone's wage grows at the same rate as the projected growth of the country's economy. Our result presents contradicting evidence that many workers do not reap benefits from the country's economic growth. Future research of this kind of projection should incorporate such heterogeneity into account, especially if researchers aim to evaluate impacts of a newly proposed regulation on various groups of insurers.

Lastly, on the firm size distribution, we find that almost half of the formally registered firms had fewer than five employees, the benchmark often used to define informal firms. This result suggests that distribution of firm sizes may differ across countries, and that using firm size alone may not be a sufficient criterion to define informal firms. In addition, we document the rising employment concentration among exceptionally large firms but declining in size among medium firms. Further examining reasons behind this pattern can be a fruitful future research. Possible driving forces include (i) large firms' ability to invest in advanced technology lead them to have higher productivities and able to expand; (ii) routine jobs in medium size firms are replaced with machines; and/or (iii) higher costs of hiring and firing employees due to labor protection laws create incentives for medium firms to stay small.

References

- Apaitan, T., Banernghansa, C., Paweenawat, A., & Samphantharak, K. (2019). Towards a Competitive Thailand: The Role of Market Power and Business Dynamism.
- Banerjee, A. V., & Duflo, E. (2007). The economic lives of the poor. *Journal of economic perspectives*, 21(1), 141-168.
- Benmelech, E., Bergman, N., & Kim, H. (2018). Strong employers and weak employees: How does employer concentration affect wages? Working Paper 24307, *National Bureau of Economic Research*.
- Bertrand, M., Johnson, S., Samphantharak, K., & Schoar, A. (2008). Mixing family with business: A study of Thai business groups and the families behind them. *Journal of financial Economics*, 88(3), 466-498.
- Blundell, R., French, E., & Tetlow, G. (2016). Retirement incentives and labor supply. In *Handbook of the economics of population aging, 1*, 457-566. North-Holland.
- Bosch, M. M. & Maloney, W. F., Gross Worker Flows in the Presence of Informal Labor Markets: Evidence from Mexico, 1987-2002 (June 2007). IZA Discussion Paper No. 2864, Available at SSRN: <https://ssrn.com/abstract=998484>
- Brown, C., & Medoff, J. L. (1989). The Employer Size-Wage Effect. *Journal of Political Economy*, 97, 1027–1059.
- Feige, E. L. (1990). Defining and estimating underground and informal economies: The new institutional economics approach. *World development*, 18(7), 989-1002.
- Freeman, R. B. (2010). Labor regulations, unions, and social protection in developing countries: Market distortions or efficient institutions?. In *Handbook of development economics* (Vol. 5, pp. 4657-4702). Elsevier.
- Henley, A., Arabsheibani, G. R., & Carneiro, F. G. (2009). On defining and measuring the informal sector: Evidence from Brazil. *World development*, 37(5), 992-1003.

- International Labour Organization (2016). Actuarial Valuation of Thailand Social Security Scheme administered by the Social Security Office as of 31 December 2013; <https://www.social-protection.org/gimi/gess/RessourcePDF.action?ressource.ressourceId=53893>
- Keane, M. P., & Wasi, N. (2016). Labour supply: the roles of human capital and the extensive margin. *The Economic Journal*, 126(592), 578-617.
- La Porta, R., & Shleifer, A. (2014). Informality and development. *Journal of Economic Perspectives*, 28(3), 109-26.
- Meghir, C., Narita, R., & Robin, J. M. (2015). Wages and informality in developing countries. *American Economic Review*, 105(4), 1509-46.
- National Statistical Office (2019). The informal employment survey 2019. Labour Statistic Groups. National Statistical Office, Thailand.
- Nicholson, W., & Needels, K. (2006). Unemployment insurance: Strengthening the relationship between theory and policy. *Journal of Economic Perspectives*, 20(3), 47-70.
- Organisation for Economic Co-operation and Development (OECD). (2018). The future of social protection: what works for non-standard workers?.
- Poschke, M. (2018). The firm size distribution across countries and skill-biased change in entrepreneurial technology. *American Economic Journal: Macroeconomics*, 10(3), 1-41.
- Rauch, J. E. (1991). Modelling the informal sector formally. *Journal of development Economics*, 35(1), 33-47.
- Ratanabanchuen, R. 2019. "The Pension System in Thailand." *Nomura Journal of Asian Capital Markets* 3 (2): 34.
- Rothenberg, A. D., Gaduh, A., Burger, N. E., Chazali, C., Tjandraningsih, I., Radikun, R., ... & Weiland, S. (2016). Rethinking Indonesia's informal sector. *World Development*, 80, 96-113.

- Sriratanaban, J. (2002). 'Civil Servant Medical Benefit Scheme: Unregulated fee-for-service and cost escalation,' In *Health Insurance Systems in Thailand*, Pramualratana P, Wibulpolprasert S (eds). Health Systems Research Institute; 41-52.
- Wasi, N., Chandrasakha, J. & Sessomboon, P. (2018). "No one want to be unemployed? : some evidence from the Thai Social Security data" *aBRIDGED article*, Puey Ungphakorn Institute for Economic (in Thai).
- Wiboonchutikula, P. (2002). Small and medium enterprises in Thailand: Recent trends. *Small Business Economics*, 18(1-3), 211-224.
- Winkler, H., Ruppert Bulmer, E., & Mote, H. (2017). *Expanding Social Insurance Coverage to Informal Workers*. World Bank.

Appendix

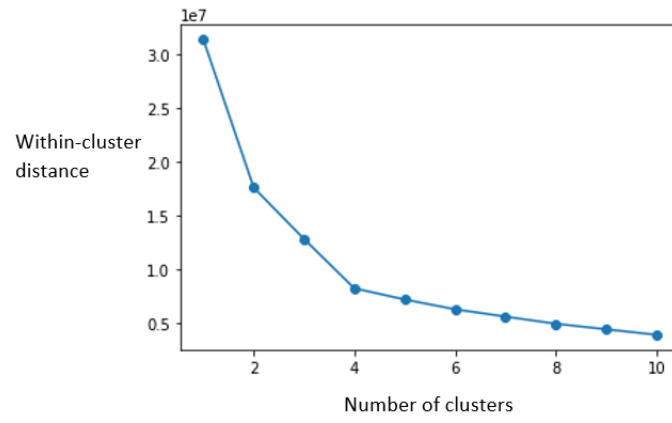


Figure A1: Total Within-Cluster Distance by Number of Clusters