Non-farm Occupational and Earnings Dynamics in Rural Thailand

Chayanee Chawanote and Christopher B. Barrett

Charles H. Dyson School of Applied Economics and Management,
Cornell University

March 2013 Revision

Acknowledgements: This work was supported by a Cornell Graduate School Research Travel Grant. We thank the National Statistical Office (NSO) of Thailand and the University of Chicago-UTCC Research Center for giving us permission to use the data and Garrick Blalock, Gary Fields and seminar participants at Cornell and Monash Universities for helpful comments. Any remaining errors are our responsibility alone.

Non-farm Occupational and Earnings Dynamics in Rural Thailand

Abstract

This study explores individual occupational and earnings dynamics in rural Thailand over the period 2005-2010. We find significant occupational transitions, mainly involving moving out of farming and into non-farm employment, rather than starting businesses, especially enterprises that employ others. Non-farm employers' and wage or salary employees' earnings stochastically dominate the distributions of the non-farm self-employed and all non-farm occupational categories' earnings distributions dominate that of farming, revealing an occupational ladder, with the most remunerative employment as a non-farm business owner/employer, and the worst as an agricultural worker or farmer. Occupational transitions into the rural non-farm economy are therefore associated with statistically significant earnings gains while transitions into farming are associated with earnings losses. These results are confirmed when we track individuals over time, and with a variety of methods to control for prospective unobserved heterogeneity. However, only a small number of individuals become and remain non-farm employers, reflecting the difficulty involved in starting, expanding or even maintaining a rural non-farm business that employs others. Furthermore, only a tiny share of household-based enterprises employ ten or more non-family members while roughly two-thirds of those with non-farm employment work in private firms of at least that size. Given that paid non-farm employment generates higher average returns than does non-farm self-employment, and the scant capacity of household-based enterprises to generate adequate non-farm labor demand to absorb migrants out of farming, our findings suggest that promoting rural non-farm employment by larger enterprises may be more important to rural income growth than promoting rural non-farm self-employment and household entrepreneurial activity.

JEL code: O1, J2, J6, I3
Keywords: income diversification, non-farm employment, rural livelihoods, self-employment, Thailand

1. Introduction

The rural non-farm economy (RNFE) is increasingly seen as a pathway out of poverty in low- and middle-income countries. As land becomes increasingly scarce, a transition to the rural non-farm sector becomes essential for many land-constrained rural households, a natural part of the 'agricultural transformation' intrinsic to economic development (Timmer 1988, 2002). Firms and activities in the RNFE provide essential linkages in the development process between agriculture and the macroeconomy, becoming a key contributor to increasing rural incomes, reducing poverty, and stimulating economic growth (Timmer 2002). Thus most rural households earn at least some income from non-farm sources, as non-farm workers, operating non-farm businesses, or both (Reardon 1997).

The growing number of empirical studies related to the RNFE can be divided into two general groups. The first group investigates the determinants of RNFE participation, either at household or individual levels (Gibson and Olivia, 2010; Jonasson and Helfand, 2010). The second group focuses on the impacts of RNFE participation on household income, rural poverty, and inequality (Reardon et al., 2000; Ferreira and Lanjouw, 2001; Cherdchuchai and Otsuka, 2006; Matsumoto et al., 2006; Hung et al., 2010). The literature relies overwhelmingly, however, on (repeated) cross-sectional evidence. The dynamic role of the RNFE and occupational transitions on rural household earnings has yet to be investigated intensively; Block and Webb (2001), Bezu et al. (2012) and Bezu and Barrett (2012) are notable exceptions that use longitudinal household-level data from Ethiopia. Past research suggests that those engaged in highly productive non-farm activities typically enjoy upward earnings mobility (Barrett, Reardon and Webb 2001; Block and Webb 2001; Lanjouw, 2001; Bezu et al., 2012; Bezu and Barrett, 2012). Of course, it is also likely that individuals with higher initial wealth and human capital are

more able to engage in high-return non-farm activities and benefit most from the RNFE, so there could be significant selection effects involved in this oft-found association (Barrett et al., 2005). As Banerjee and Newman (1993) theorize, in the presence of capital market imperfections, the ex ante poor tend to choose wage labor while the ex ante rich become entrepreneurs. Banerjee and Newman also emphasize the interplay between 'the distribution of income and wealth' and 'the dynamics of occupational choice'.

Evidence from many countries reveals considerable heterogeneity in the RNFE. But most household businesses consist either of self-employed enterprises without paid, non-family employees or small-sized firms with limited firm expansion (Fafchamps 1994; Haggblade et al., 2007). These businesses face several constraints, such as access to capital, skilled labor, entrepreneurial ability, and government registry requirements. Subsistence self-employment does not automatically transition into the enterprise growth that increases both the business owner's household income and employment within their region (Mondragon-Velez and Pena-Parga, 2008; de Mel et al., 2008; Schoar, 2010). It therefore seems important to differentiate between non-farm self-employment without hired workers and those household enterprises that hire non-family members, which we term entrepreneurs. Little is known empirically about the earnings transitions between farm work, rural non-farm employment, and rural non-farm self-employment – with or without employees, especially with adequate controls for prospective unobserved heterogeneity associated with selection into distinct occupational groups.

This study helps to fill these gaps by exploring rural non-farm occupational and earnings dynamics in rural Thailand, differentiating between non-farm employment, non-farm self-employment and non-farm entrepreneurship. More explicitly, the research questions this paper explores are as follows. First, what patterns of occupational transitions exist among farmers, non-

farm employees, the non-farm self-employed, and non-farm employers in rural Thailand? Second, how do rural non-farm employment and occupational transitions affect directional earnings mobility? Which occupational shifts – e.g., from farm to non-farm employee or from non-farm self-employed to non-farm employer – are associated with people increasing or decreasing earnings when controlling for other characteristics?

There have been a few previous studies on occupational mobility in developing countries (Fuwa, 1999; Quadrini, 2000; Mondragon-Velez and Pena-Parga, 2008). Mondragon-Velez and Pena-Parga (2008), in particular, explore the transitions between unemployed, wage-earner, self-employed and business owner status in seven main cities in Colombia. They mainly focus on the determinants of entry into and exit from urban self-employment and business ownership. They find that most newly self-employed and entrepreneurs transition from wage employment rather than from unemployment. However, they find extremely low transitions from self-employment to entrepreneurship (and vice versa). In studies of the determinants of income mobility, Cichello et al. (2005), Woolard and Klasen (2005), and Fields et al. (2005) found that the conditional effects of occupation and sector of employment were statistically significant in South Africa and Latin America.

This paper uses the nationally representative Thai Socio-Economic Survey (SES) panel data collected annually between 2005 and 2007, and the subsequent round in early 2010. This study thus uses more rounds of nationally representative panel data, with far more individual observations over a longer period, than any of the prior RNFE studies. This enables us to employ multiple empirical approaches, some of which would not be possible with simply two observations per individual or a much more modest number of observations, in order to more robustly identify the effects of occupational transitions on earnings dynamics in a nationally

representative sample. It also enables us to differentiate among alternative non-farm occupations in a way that matters fundamentally to rural development policy debates.

Little is known about earnings and occupational mobility in rural Thailand. Isvilanonda et al. (2000) indicate the growing importance of income from non-rice crops and non-farm activities, using survey data from six villages in 1987 and 1998. They find that while the number of the poor declined, income inequality has increased. Cherdchuchai and Otsuka (2006), using the same baseline survey data in 1987 and a new survey in 2004, investigate a structural shift of household income away from farm to non-farm income sources. Unlike Isvilanonda et al. (2000), they find that non-farm employment expansion reduces the income gap and the difference in poverty incidence between prosperous and poor regions. However, neither of these two papers investigates earnings mobility as it relates to occupational transitions.

We find significant occupational transitions over the course of just five years, mainly involving moves into farm and non-farm (salaried or wage) employee positions rather than into non-farm self-employment and employer positions. Transitions into non-farm employment result in statistically significant income gains, on average, while moves into farming or farm work are associated with reduced earnings. Non-farm employers' and employees' earnings distributions stochastically dominate those of both farmers and the self-employed (without employees), while that of farmers is stochastically dominated by each of the other three occupational groupings.

That core finding of an occupational ladder is reinforced by directional earnings mobility regression analysis when tracking the same individuals over time. Only a small number of individuals become non-farm employers, the most remunerative occupation group, reflecting the difficulty inherent to establishing and maintaining a business with employees. Moreover, less than one percent of these household enterprises employ ten or more family members (Chawanote

2013), indicating limited employment generation potential. Our findings suggest that promoting rural non-farm employment by attracting established businesses, government or not-for-profit agencies may be more important to rural poverty reduction than promoting rural non-farm self-employment in the hope that this leads to entrepreneurial rural non-farm job creation and higher rural household incomes.

2. Data and background

Thailand is a lower middle-income country by the World Bank's classification, with GDP per capita of $8,004 in 2009. The $2/day per person poverty headcount ratio was 11.5 percent of population in 2004, down from 16.7 percent after the 1997-98 financial crisis. The labor force participation rate was 73.2 percent of the total population ages 15 and above. Roughly 1.3 percent of the total labor force reported being unemployed between 2005 and 2009. Approximately 67 percent of the population from 2005 to 2009 lived in rural areas, with a steadily declining share employed in agriculture.

The Thai SES panel data were collected by the National Statistical Office (NSO) of Thailand in 2005 – 2007 and 2010. The repeated cross-sectional rounds of the well-respected SES have been used frequently by researchers (e.g., Schultz 1990, Paxson 1992, Mammen and Paxson 2000, Giné and Townsend 2004, Felkner and Townsend 2011). Beginning in 2005 NSO began tracking households and split-off individuals from sample households to create proper panel data, although these panel data appear not to have been exploited much, if at all.[1] We

[1] The SES panel data have been used for internal government reports and by some research institutes in Thailand. But we can find no English language publications that exploit these important data, likely due to the facts that printed information on the panel is only available in the Thai language, thereby sharply limiting awareness of their availability, and the data are not freely available to the public, but require a contract with NSO approving release and use, as well as purchase of the data.

therefore take particular care to explore attrition patterns and their implications (Appendix A), so as to enhance the usefulness of this rich longitudinal data set for future researchers.

For the first three rounds (2005-7), the survey was recorded in May, while the last (2010) round was surveyed in January. The survey has two main segments: i) household information on every member in the household, and ii) individual information on household members aged 15 years or older. Part one includes general information on household members, household characteristics and assets, and income from agriculture. Part two includes survey questions on education, health care, employment, incomes, expenditures, financial status (debt and savings), migration, and opinions on public policies. The survey covers every province in Thailand and randomly selects blocks of districts, sub-districts and villages, and finally selects ten households per village as in a two-stage stratified random sampling.[2] All statistics we report are adjusted for sampling weights.

Table 1 summarizes the Thai SES panel data.[3] The 2005 round surveyed 6,000 households with a total of 16,310 individuals and 9,897 individuals in rural areas. All individuals age 15 and over were tracked in the following years' surveys. Any adult who left the core household was tracked so long as they remained within Thailand and a new address for that split-off individual could be found by the survey team. Some individuals are missing from one round, but reappear in later rounds once they could be tracked again. We use only the balanced panel, in other words, only individuals present in all four rounds of the SES. Due to split-off households and attrition, the total number of individuals aged 15 years or older surveyed in all four rounds is 12,758, of whom 7,831 lived in rural areas. Given the substantial attrition, we take special care to control for the possible bias this might introduce (Appendix A).

[2] Each rural sub-district in the SES panel data has only one village.
[3] Further explanation of the rural sample across years and attrition issues are provided in Appendix A.

[Table 1 here]

Since the Thai SES panel only surveys at household and individual level, we match it with another village-level dataset. A rural development census, the National Rural Development (NRD) data set, was collected at the village level by the Community Development Department of Thailand.[4] NRD data that match the Thai SES panel data are only available for April to June 2005 and April to May 2007 and 2009. The data cover general conditions of the village and local economy, including the availability of public services and infrastructure, health and sanitation, village educational achievement, and agroecological conditions.

2.1 Definition of rural non-farm employment

We use only the Thai SES panel data on individuals who were employed in rural areas, including unpaid workers for household businesses, and those who were 15-70 years old. Over the five-year period of SES data collection, the unemployment rates in the rural areas included in the study ranged from 0.5 to 1.1 percent while employment rates ranged from 76.1 to 79.9 percent.[5] In the employment section of the SES, respondents were asked for their primary occupation, work status, and company size[6] for each of up to three jobs that they had worked in the past 12 months. The first job recorded in the dataset reflects the individual's current main job at the time of survey.[7] It should not be affected by seasonality for those who are in the farming

[4] The data are distributed by the University of Chicago-UTCC Research Center, Bangkok, Thailand.
[5] Other categories in the survey are waiting for seasonal work, looking for work, retired, long term illness and disabilities, caring for other household members, and going to school. We only focus on employment status since this better represents a group of earners and eliminates the possible variation in occupational transitions that would come with seasonal work.
[6] The company size, measured as the total number of workers including the owner, is categorized as 1 worker (i.e., no employees), 2-9 workers, 10-50 workers, 51-100 workers, 101-200 workers, 201-500 workers, or over 500 workers.
[7] Five percent of 1st jobs were reported as having ended in the 12 months prior to the survey while around 99 percent of the 2nd and 3rd jobs were reported as having ended in the previous year. The length of time

sector, given the survey timing. May is the beginning of the rice cultivation season in Northern and Central regions, rice harvesting season in Southern region, and other harvesting season for fruits in Eastern Thailand. Even though the 2010 survey round was in January, it is a main cultivation period for tapioca/cassava, cane, and other similar crops. The options for primary occupation in the survey are farmer/fisherman (crops, livestock, aquaculture, fishery, hunting and gathering), production (handicrafts and basic technology manufacturing), production (industry), merchandise/own business, government/state enterprise employee, company/business employee, and general worker/laborer. The work status question includes options for employer, self-employed without employees, working without pay for household business, government employee, state enterprise employee, private company employee, and cooperative group. These two questions – primary occupation and work status – are used to separate non-farm activities from farm activities at individual level.

Unfortunately, the household survey does not directly identify a respondent's employer, so we cannot match employees with employers. We do, however, know the size distribution of individual respondents' employers. As reflected in Table 2 for 2005 and 2010 (the 2006 and 2007 data exhibit qualitatively identical patterns), at most one-third of rural non-farm employees work for private businesses with ten or more employees (grey-shaded cells).[8] Parallel analysis of household enterprise data from the SES panel (Chawanote 2013) finds that less than one percent of household-owned enterprises in rural Thailand employ ten or more workers and very few of these enterprises exhibit any statistically significant employment growth over the 2005-10 period.

---

in 2nd jobs was consistently an order of magnitude shorter than in 1st jobs and time spent in 3rd jobs was only 16-25 % that in 2nd jobs. The 2nd and 3rd jobs reflect seasonal jobs or jobs that ended before the current, primary one. Dropping these short-term, temporary positions makes no qualitative difference to the analysis we report here.

[8] Furthermore, employees' reports of small-sized state owned enterprise or government employers almost certainly refers to respondents' immediate department/unit rather than to the entire agency. If similar underreporting of private firm size occurs, that reinforces our finding.

The striking mismatch between the jobs created by rural household enterprises and the jobs held by those who work for a salary or wages in the rural non-farm economy carries important policy implications. Donors' and governments' present emphasis on promoting rural household non-farm entrepreneurial activity might not offer an adequately broad platform to facilitate agrarian transformation and rural earnings growth.

[Table 2 here]

The rural non-farm sector includes all economic activities in rural areas except primary production in agriculture, livestock, fishing and hunting, and thus includes any employment in manufacturing, mining, trade, construction, transportation, communications, government and services (Lanjouw, 2001; Haggblade et al., 2002). Using this definition, those who reported their primary occupation as being anything other than farmer/fisherman are considered as working in the non-farm sector. Conversely, only those who reported their primary occupation as farmer/fisherman are considered as working in farm employment.

Previous studies that decompose rural non-farm employment have categorized it as either low-productivity wage labor or high-productivity salaried work or self-employment (e.g., Barrett et al., 2005; Jonasson and Helfand, 2009; Bezu et. al., 2012). In this study, however, entrepreneurship status is classified separately from self-employment without employees as this differentiates what are sometimes referred to as "subsistence" from "transformational" entrepreneurs, with the latter being the prospective source of new RNFE jobs (Schoar, 2010). Any respondent who employed non-family members in any non-farm activity is considered a 'non-farm employer' or 'entrepreneur', while anyone self-employed without employees, working without pay for household business, or working in a cooperative group is grouped into the category 'non-farm self-employment.' Both employers and the self-employed refer to those who

operate their own business and receive business profits as their primary earnings. Finally, 'non-farm employee' includes salaried and waged workers, i.e., those employed by the government, state enterprises, or private companies or not-for-profit agencies, and who have no claim to business profits.

Table 3 summarizes individuals' work status in rural Thailand. The percentage of workers in each occupation changed only slightly between 2005 and 2010. Although farmers and farm workers represent a plurality of rural Thai workers, more people are employed primarily in non-farm occupations. Non-farm employees account for the largest proportion of non-farm sector workers, while non-farm employers account for only one percent of the total employed population in rural Thailand. More than 90 percent of the rural non-farm self-employed do not create jobs outside the entrepreneur's household and of those who become employers, less than 1 percent create 10 or more jobs. This is an important point largely missed in the literature and in contemporary policy dialogues, which emphasize promoting entrepreneurship to ignite the rural non-farm economy.

[Table 3 here]

In the analysis that follows, we focus on the earnings and occupational dynamics of only those rural working age adults (15-70 years old) who were employed and surveyed in all four SES rounds, so as to avoid conflating transitions between unemployment and employment with transitions among occupations. This introduces the possibility of attrition bias, either due to exits from the sample – due to outmigration, death, unavailability, or another reason – or because of one or more periods of unemployment during the SES rounds. Appendix A explores the possibility of non-random attrition in detail, demonstrating that attrition indeed appears non-random, although the attrition-corrected regression results reported in the main body of the paper

are not statistically significantly different from the uncorrected results in Appendix A, Table A5-A7.

2.2 Earnings

Individual earnings are decomposed by source: farm earnings, non-farm business profits, and wages or salaries. Farm earnings and non-farm business profits are recorded at the household level. We use individual work hours per week in each enterprise to assign individual farm income to individual household members based on their share of total family labor time allocated to the farm enterprise. Similarly, non-farm business profits are allocated to all self-employed members in the household proportional to time self-employed members work in the household non-farm enterprise. Wage and salary earnings are already recorded at the individual worker level. All earnings are adjusted for the consumer price index for each region of Thailand to put them in real 2007 baht terms.[9]

We focus on structural occupational transitions and individual earnings mobility 2005-2007-2010. This allows us to use village-level controls available from the NRD (which, as indicated previously, was not fielded in 2006 and 2008). Plus, the longer spell length in the dynamics analysis minimizes the role of transitory shocks and measurement error, reducing the possibility of overstating structural economic mobility (Naschold and Barrett 2011).

Table 4 shows mean earnings by quartiles conditional on each occupation and with the lowest and highest one percent of earnings in each year cut off so as to eliminate extreme outliers likely to reflect measurement error. On average, non-farm employers enjoy the highest earnings while farmers receive the lowest earnings. Non-farm employees earn more on average than those

[9] Consumer price index data by region are reported by Thailand's Ministry of Commerce (www.moc.go.th).

engaged in non-farm self-employment in every quartile. Given that the rural poverty line in 2007 was 1,333 Baht[10] per capita per month, except for farmers and the non-farm self-employed, individuals in the first quartile on average earn above the poverty line. Twenty percent of the rural employed fall under the poverty line; almost eighty percent of the rural poor are farmers. However, this is based solely on these three earnings sources, excluding income from other sources such as remittances, incomes from house/land lending, or returns from financial assets. The poor seemed to be affected most by the country's 2008-9 economic downturn as the earnings averages in the first quartile in 2010 dropped from 2007, whereas the highest quartiles still enjoyed an increase in earnings on average. The economic slowdown also had an impact on non-farm businesses since earnings of both non-farm self-employed and employers in 2010 fell slightly from 2007.

[Table 4 here]

Figure 1 presents the cumulative frequency distributions of earnings by occupation in 2005 and 2010. Each non-farm occupation's earnings distribution stochastically dominates that of farming.[11] This ordering is consistent with the results presented in Table 4, where the most desirable remunerative is non-farm employer followed by non-farm employee, non-farm self-employment, and farming, in that order. Perhaps more interesting, the non-farm employers and non-farm employees distributions both first-order stochastically dominate the non-farm self-employed earnings distribution. There is no statistically significant stochastic dominance ordering between those two dominant distributions due to a small number of low earnings draws among non-farm employers. But mean earnings for employers are considerably higher, albeit with the gap closing over the 2005-10 period.

[10] In 2007, there were 33.72 baht per US dollar, referencing from the Office of the National Economic and Social Development Board (NESDB).
[11] We use Davidson and Duclos' (2000) test to confirm all the stochastic dominance results reported here.

[Figure 1 here]

3. Earnings changes and occupational transitions

Having already observed a clear earnings distribution ordering among occupations in each year, we would expect that transitions from farming into rural non-farm occupations should be associated with increased earnings, as should transitions from non-farm self-employment to non-farm employee or employer status. Conversely, transitions into farming, or into non-farm self-employment from the other two non-farm occupational categories, should be associated with reduced earnings. This intuition is confirmed by extending the repeated cross-sectional analysis to intertemporal transitions.

[Table 5 here]

The transition matrices presented in Table 5 show how movement across farm and non-farm employment categories relates to changes in income. The percentage change is calculated to show how occupational status in 2005 (row) changed by 2010 (column). We see that, other than for non-farm employers, work status primarily remains the same across the five years, with 67-79 percent of each group remaining in their original occupational sector. But almost 30 percent of 2005 non-farm employers had shed their employees and converted to merely self-employed status by 2010 while around 40 percent maintained their non-farm employer status for those five years, although the sub-sample size is small. As one would expect, transitions are more from farming into more remunerative non-farm employment rather than into farming. However, more people slip from non-farm self-employment into farming than graduate into the most remunerative non-farm employee or employer positions. Likewise, almost 12 times more non-farm employees slip back into non-farm self-employment than graduate into becoming non-farm

employers. The non-farm self-employed are more likely to transition into employer status than are those who did not previously run a non-farm business. But as in Mondragon-Velez and Pena-Parga (2008), we find an extremely low transition rate into being an employer, just 3.3 percent, 0.7 and 0.5 percent of the self-employed, non-farm employees and farmers or farm workers, respectively.

These results strongly suggest a 'gravity effect' on the occupation ladder: it is easier to move down into lower-return occupations than up into higher-return ones. These transition patterns indicate the difficulty of starting and expanding a business or even securing paid non-farm employment, since the earnings distributions for those two occupational groups first order stochastically dominate the earnings distributions of the other two categories. Constraints may include differences in physical and human asset endowments, access to finance, social connections, etc. We discuss these issues more in section 4 when we investigate the determinants of occupational transitions.

[Table 6 here]

Table 6 presents the median, mean and standard deviation percentage real earnings changes associated with each transition. None of the earnings changes are statistically significantly different from zero, reflecting the considerable dispersion observed in unconditional earnings transitions. The mean and median patterns are similar in their directional changes. Individuals who remained in their initial occupational categories enjoyed positive mean changes in earnings. Movement from farming into any non-farm employment generates earnings gains, on average, while moving into farm employment is associated with earnings losses, on average. But note that of the roughly 40 percent of non-farm employers who maintain their business and employees over the course of five years, most suffered a decline in earnings over the 2005-10

period. This underscores the considerable challenge of maintaining, much less growing employment through nonfarm household enterprises in rural Thailand.

[Figure 2 here]

We can generalize this analysis to explore the full distribution of earnings changes associated with each transition (Figure 2), dropping the lowest and highest one percent of earnings changes. The plots in Figure 2 show that some earnings changes distributions first order stochastically dominate others, although there is no clear and consistent ranking among the distributions of earnings changes of initial employer positions based on stochastic dominance tests. For those initially in farm work, the transition from farm work to non-farm employee status first order stochastically dominates staying in the farm sector. However, none of these earnings changes distributions reveal statistically significant second order stochastically dominant between transitions into non-farm self-employment, nor consistently significant transitions into non-farm employee or employer status.

4. Multivariate analysis of occupational shifts and earning mobility

4.1 Empirical Model

Especially given the absence of an explicit earnings change ordering among occupational transitions and the non-random nature of those transitions, multivariate regression analysis can help us better understand how changes in earnings associate with farm and non-farm occupational shifts. We emphasize that in these observational data, it is exceedingly difficult to control for all prospective sources of unobserved heterogeneity that might generate selection effects or spurious correlation between occupational transitions and earnings dynamics. We can convincingly establish associations only. But by employing a range of controls and estimation

techniques, each aimed at addressing a different source of prospective bias, we can check if the core qualitative results are robust to a range of statistical corrections that are each incomplete and imperfect but as a set offer a reasonably comprehensive approach to check the core results. The robustness of the findings and the quality of the data give us confidence that the strong and consistent statistical associations we find likely indicate a true causal relationship between occupational transitions and earnings dynamics in rural Thailand.

We employ a conditional mobility model in which change in earnings or change in log earnings are regressed on time-invariant and time-varying individual characteristics. In this class of model, changes in earnings are explained by initial earnings, gender, age, educational attainment, sector of employment, and geographic region, with occupation and sector of employment typically considered time-varying variables (Cichello et al., 2005; Fields, 2007). This framework allows us to explore how occupational shifts change earnings when controlling for other observable characteristics that are almost surely correlated with both earnings dynamics and occupational patterns. Following Fields (2007), the conditional micro mobility model is defined as:

$$\Delta \ln y_{it} = \alpha + \beta_1 \ln y_{i,t-1} + t \ln y_{i,t-1} \beta_2 + \Delta X_{it} \beta_3 + Z_i \beta_4 + \phi_i + \lambda_t + \varepsilon_{it} \qquad (1)$$

where $\Delta \ln y_{it}$ is the change in log reported real earnings from year t-1 to year t and $\ln y_{i,t-1}$ is the base year log reported real earnings, included as a control for autocorrelation.[12] Because the periodicity of the SES panel changed, from one year revisits between the 2005, 2006 and 2007 rounds, to a three year revisit in the 2010 round, we do not impose a single autocorrelation parameter. Instead, we add interaction terms between the base year earnings and year dummies for the 2006-7 and 2007-10 transitions. Z_i denotes a matrix of time-invariant individual and

[12] We use a logarithmic specification because it substantially improves goodness of fit relative to using earnings levels.

household characteristics, as observed in the initial year. Both age and age squared are included to control for life cycle effects that should be reflected in a positive (negative) sign on the linear (quadratic) term. Education is recorded as the highest level completed, with dummy variables for primary school, secondary school, high school/vocational school, and college degree and above, with less than primary school or none as a base level. Gender is described with a dummy variable taking value one for females, and marital status is described with a dummy taking value one for married persons. Since the observations are at the individual level, a dummy for household head is also included, as well as family size. An initial year asset index and household owned agricultural land separately from the asset index are also included to control for household capital endowments.[13] ΔX_{it} denotes employment transition experiences, which are represented by dummy variables for fifteen possible transitions, with staying in farm work as a base case. Finally, ϕ is a vector of sub-district fixed effects, λ is a vector of time fixed effects, and ε_{it} is a mean zero i.i.d error term, corrected for clustering and potential heteroskedasticity.

We hypothesize that the sectoral transitions' coefficient estimates in the log earnings equation follow the same ordering found in the unconditional analyses reported in section 3, even after controlling for individual and household characteristics. Moreover, we can also test the differences between occupational transitions' coefficients, given the initial or previous job, for earnings changes associated with those occupational shifts, similar to testing for stochastic dominance in Figure 2. That is, transitions into (out of) farming, or into (out of) non-farm self-employment from the other two non-farm occupational categories should be associated with reduced (increased) earnings.

[13] The estimation details of the asset index, constructed using factor analysis following Sahn and Stifel (2003), are reported in Appendix B. The index includes number of rooms, housing materials, electricity, cooking fuels, water supply, toilet, number of durable goods (e.g., microwave, refrigerator, air condition, fan, television, radio, VCD-DVD player, washing machine, cable television, cell phone, landline, computer and internet), number of vehicles (motorcycles, cars, trucks, tractors), and livestock.

4.2 Empirical results

Table 7 provides descriptive statistics of these variables for the whole sample and for each group. Given each group in 2005, the mean of the asset index is the highest for non-farm employers and lowest for farmers, although there is not much difference in means of the asset index between non-farm self-employment and non-farm employees. Non-farm employees have the highest proportion of college graduates as opposed to farmers that have the highest proportion of primary school graduates.

[Table 7 here]

The estimation results, using 2005-6, 2006-7, and 2007-10 transitions, are reported in Table 8.[14] Model (1) is estimated by OLS with bootstrapped standard errors and controlling for sub-district fixed effects. The occupational transition variables are jointly statistically significant in determining log earnings change. The occupational transitions' coefficient estimates show that individuals who were employed in non-farm activities and who remained in their initial positions all enjoyed a statistically significant gain in earnings relative to individuals who remained in farming. Conversely, those who transitioned into farming from non-farm occupations suffered statistically significant earnings losses compared to those individuals who stayed in farming. Meanwhile, all of the movements out of the farming sector result in statistically significantly positive log earnings changes. In every case, the highest point estimate for log earnings change is associated with movement into (or remaining) a non-farm employer, and is statistically significant. The conditional effects of age, education, marital status, gender, and household asset holdings all have the expected signs and are statistically significantly different from zero.

[14] The estimation results with absolute earnings, rather than log earnings, are qualitatively similar, as reported in Appendix Table C1.

[Table 8 here]

However, other unobserved characteristics may be confounding the OLS estimates in Model (1). The five-year, four-round panel data offers the opportunity, however, to control for individual-level fixed effects so as to control for time invariant unobservables. We present those estimates as model (2). Because one might be interested in the coefficient estimates on the time invariant individual and household characteristics, model (3) presents results using a Hausman-Taylor estimator, an instrumental variables approach that enables estimation of the coefficients of time-invariant regressors while still controlling for individual-level random effects.

In the individual fixed effects model, almost all of the occupational transitions still have statistically significantly positive estimated effects on log earnings changes with an ordering in magnitude that mirrors the unconditional earnings orderings apparent in Figure 1. Only transitions from non-farm self-employment and employer positions into non-farm employees have greater estimated expected percentage change than those transitions into or remaining non-farm employers. However, there is no statistically significant difference between these two pairs of coefficient estimates (Table 9).

Qualitatively similar results emerge from model (3)'s Hausman-Taylor (H-T) estimates. The biggest gains come from becoming a non-farm employer and all transitions out of farming are associated with gains relative to remaining in agriculture as a primary occupation. Although the sign and statistical significance of the H-T coefficient estimates of the time-invariant observed characteristics are similar to those in OLS estimation, the sign and significance of the H-T coefficient estimates on age and education are the opposite, but statistically insignificant. However, if one looks at the absolute earnings (rather than log earnings) H-T regressions (reported in Appendix Table C1), these signs on age, high school and college attainment are the

same as the OLS estimators and the coefficient estimates are statistically significant. In particular, there are noticeable life cycle, gender and family size effects, while both higher individual educational attainment and greater household assets strongly and statistically significantly increase earnings. [15]

[Table 9 here]

Table 9 presents the estimated differences in log earnings changes among occupational transitions, similar to the earnings dominance tests in Figure 2. The results confirm that moving to the farm sector from any non-farm occupation leads to statistically significantly lower earnings changes. On average, movements into the non-farm sector increase earnings relative to remaining in farming. By contrast, shifting between non-farm sectors results in mixed outcomes. In most cases, transitions from any of the non-farm occupations into another non-farm position leads to lower earnings growth than does staying. The lone exception is transitions from non-farm self-employment to being a non-farm employee, reinforcing the general impression that self-employment is less desirable than permanent salaried or wage employment. That result appears to hold even when controlling for characteristics and constraints.

4.3 Robustness checks

Although the previous regressions use individual fixed effects to control for unobserved time invariant characteristics in an attempt to disentangle the influence of occupational shifts on changes in earnings, time-varying unobservables could still drive both changes in earnings and in occupation, leading to spurious correlation that would undercut the argument that occupational transitions drive earnings gains. One important prospective class of time-varying factors

[15] As described in Appendix A, when we correct for non-random attrition using inverse probability weights, the coefficient estimates do not change significantly. See, in particular, Appendix Tables A5-A7.

unobserved in the SES data that could have such effects is village-level environmental and infrastructure variables. Improvements in village-scale infrastructure – roads, water, electricity, etc. – can change both the absolute and relative productivity of different occupations, thereby causing individual occupational transitions and hence earnings changes. Controlling for changes in infrastructure can therefore substantially obviate this prospective problem. Moreover, there might be costs associated with changing sectors and these costs (e.g, job search), are likely to decrease with the number of jobs and the rate of job growth in the local economy (Neal 1995). We therefore also control for total months spent working in the respondent's current job and changes in the ratio of total households working in particular occupations within the village. The ratios are calculated from the NRD data set to represent village employment conditions that could affect occupational switching in the village.

One approach to addressing the concern that time-varying unobservables might affect both earnings dynamics and occupational transitions is to predict the probability of these occupational movements in a first stage and then to use these predicted transition probabilities in two-stage estimation of equation (1). In order to do that, we have to first estimate the occupational transition probabilities using multinomial logit models, then use the predicted probabilities of occupational transition as explanatory variables in the second stage, log earnings regression. Our instruments are changes in village characteristics, reflecting changes in infrastructure and agricultural circumstances that affect the occupational choice decisions. This identification allows for more variation across villages while still allowing for variation in individual characteristics within villages. The first stage multinomial logit estimation details are discussed in Appendix D.

[Table 10 here]

Table 10 reports the results of both OLS and instrumental variables regressions, with control variables from the previous survey round and log earnings changes (2005-7 and 2007-10) as the dependent variable. Since we estimate each regression separately given the initial occupation in the first stage, the second stage must also be separately estimated for the three occupations besides non-farm employer. The average of the predicted probabilities in each initial group is the same as the percent share reported in the transition matrix given each original occupation (Table 5). In each equation, each sector transition is compared to staying in the original sector. The magnitudes of the control variables' coefficient estimates and their statistical significance are similar in both the OLS and IV estimations. But although the IV estimates on occupational transitions are generally consistent in sign and magnitude with the OLS estimates and with Table 8's pooled estimates (i.e., not conditional on initial occupation), they are not statistically significant. This likely reflects both the usual instrumental variables problem of lost efficiency and the problem of splitting the sample into smaller subsamples conditional on initial occupation, thus generating imprecise parameter estimates.

As another robustness check we estimate a multinomial logit model correcting for selection bias, following Dubin and McFadden (1984) and Bourguignon et al. (2007). Occupational changes might be subject to both selection bias and endogeneity. If each group of individuals that shifts occupation differs systematically in their unobservable characteristics (e.g., skills, motivation, ability), then regression results based on individuals' observed characteristics will be biased. This method has been implemented mostly in studies of wage determinants since individuals self-select into their industry of employment. It is likely that unobservable characteristics affecting wage rates also simultaneously determine selection into the sector in which individuals work. As described in Appendix E, we look at occupational changes that affect

earnings changes. Instead of only estimating coefficients, we calculate $E\left[\Delta \ln y \mid \hat{P}(j \text{ to } k), Z\right]$ for j,k = 0, 1, 2, 3. The estimated average earnings changes are presented in Table 11. Most of them are statistically significantly different from zero and show the expected signs, consistent with the earnings orderings manifest in the unconditional analyses described earlier. However, the coefficients of the second stage regression, especially the coefficients on the selection bias correction terms are statistically insignificant (Tables E1-E3). But the results of this robustness check are consistent with the previous, individual-level fixed effects and Hausman-Taylor estimates, as well as with the unconditional earnings orderings displayed in Figure 2. So the core story appears robust to any of a variety of different approaches that attempt to correct for prospective statistical weaknesses in any single estimation strategy we can apply.

[Table 11 here]

5. Conclusions

Economic growth almost always involves a transition from heavy dependence on farming to non-farm rural activity. This study reports on widespread occupational transitions in rural Thailand over a five year period, 2005-2010. Such transitions mainly involve moves into farm or non-farm employment more than into non-farm self-employment or employer positions. The non-farm employers' and employees' earnings distributions stochastically dominate the other categories' earnings distributions, while that of farmers is stochastically dominated by each of the three non-farm occupational groupings. As a result, transitions into the rural non-farm economy are associated with statistically significant earnings gains, while transitions into farming are associated with earnings losses. But not all non-farm occupations are equally lucrative. It is more common to move down into lower-return occupations, especially non-farm

self-employment and farming, than up into higher-return ones, reflecting a 'gravity effect' on the occupation ladder.

Multivariate regression results confirm that the biggest gains arise from becoming a non-farm employer and all transitions out of farming are associated with gains relative to remaining in agriculture as a primary occupation. Moreover, both higher individual educational attainment and greater household physical capital endowments strongly and statistically significantly increase earnings, indicating the joint importance of human and physical capital, as well as climbing the occupational ladder, to earnings mobility.

Although the most remunerative employment is as a RNFE business owner and employer, only a small number of individuals become non-farm employers. This result confirms similar findings from other developing countries that observe far more subsistence self-employment than business owners generating paid employment for others. The very small rate of transition into being a non-farm employer reflects the difficulty inherent to starting a business, while the fact that less than 40% of non-farm employers remain non-farm employers for five years indicates the challenges of even maintaining a rural non-farm business with employees. Moreover, there is a striking mismatch between at least two-thirds of rural non-farm employees working for an enterprise with ten or more employees, versus less than one percent of household enterprises employing ten or more people. Rural Thai households face considerable challenges in starting and maintaining, much less expanding, a non-farm business and those household enterprises create few jobs. So the greatest prospects for taking advantage of the earnings gains routinely associated with occupational transitions out of farming appear to come from finding salaried or wage employment with non-household enterprises. Rural development policy might therefore aim to increase remunerative non-farm employment opportunities by established employers and

rely less on trying to stimulate self-employment in the hopes that it will spark entrepreneurial activity and rural employment generation.

References

Banerjee, A. & A. Newman. (1993). ”Occupational Choice and the Process of Development.” *Journal of Political Economy, 101*, 274-298.

Barrett, C. B., M. Bezuneh, D. C. Clay, & T. Reardon. (2005). “Heterogeneous constraints, incentives and income diversification strategies in rural Africa.” *Quarterly Journal of International Agriculture, 44*(1), 37-60.

Barrett, C. B., T. Reardon, & P. Webb. (2001). “Non-farm income diversification and household livelihood strategies in rural Africa: concepts, dynamics, and policy implications.” *Food Policy, 26*, 315-331.

Baulch, B. & A. Quisumbing. (2011). “Testing and adjusting for attrition in household panel data,” Chronic Poverty Research Center Toolkit Note.

Becketti, S., W. Gould, L. Lillard, & F. Welch. (1998). “The Panel Study of Income Dynamics after Fourteen Years: An Evaluation.” *Journal of Labor Economics, 6*(3), 472-92.

Bezu, S. & C. B. Barrett (2012), “Employment dynamics in the rural nonfarm sector in Ethiopia: Do the poor have time on their side?” *Journal of Development Studies*, *48*(9), 1223-1240.

Bezu, S., C.B. Barrett & S. Holden. (2012)."Does Non-farm economy offer pathways for upward mobility? Evidence from a panel data study in Ethiopia," *World Development, 40*(8), 1634-1646.

Block, S. & P. Webb (2001). “The dynamics of livelihood diversification in post-famine Ethiopia,” *Food Policy, 26*(4), 333-350.

Bourguignon, F., M. Fournier, & M. Gurgand. (2007). "Selection bias corrections based on the multinomial logit model: Monte Carlo comparisons." *Journal of Economic Surveys, 2*(1), 174-205.

Chawanote, C. (2013). "Rural Household Non-farm Business: Wealth and Rainfall shocks on Startup, Expansion, or Exit?" Working paper, Cornell University.

Cherdchuchai, S., & K. Otsuka. (2006). "Rural income dynamics and poverty reduction in Thai villages from 1987 to 2004." *Agricultural Economics, 35*(3), 409-423.

Cichello, P. L., G. S. Fields, & M. Leibbrandt. (2005). "Earnings and Employment Dynamics for Africans in Post-apartheid South Africa: A Panel Study of KwaZulu-Natal." *Journal of African Economies, 14*(1), 143-190.

Davidson, R., & JY. Duclos. (2000) "Statistical Inference for Stochastic Dominance and for the Measurement of Poverty and Inequality." *Econometrica, 68*(6), 1435-1464

de Mel, S., D. McKenzie, & C. Woodruff (2008). "Who are the Microenterprise Owners?: Evidence from Sri Lanka", World Bank Working paper 4635.

Dubin, J. A. & McFadden, D. L. (1984). "An econometric analysis of residential electric appliance holdings and consumption.". *Econometrica, 52*(2), 345-362.

Fafchamps, M. (1994). "Industrial Structure and Microenterprises in Africa", *Journal of Developing Areas, 29*(1), 1-30.

Felkner, J.S. and R.M. Townsend (2011). "The Geographic Concentration of Enterprise in Developing Countries," *Quarterly Journal of Economics 126*(4), 2005-2061.

Ferreira, F. H.G., & P. Lanjouw. (2001). "Rural Non-farm Activities and Poverty in the Brazilian Northeast." *World Development, 29*(3), 509-528.

Fields, G. S. (2007). "What We Know (and Want to Know) About Earnings Mobility in Developing Countries." Working paper, Cornell University.

Fields, G. S. & M. L. Sanchez Puerta. (2005). "How Is Convergent Mobility Consistent with Rising Inequality? A Reconciliation in the Case of Argentina." Working paper, Cornell University.

Fitzgerald, J., P. Gottschalk, & R. Moffitt. (1998). "An analysis of sample attrition in panel data." *Journal of Human Resources, 33*(2), 251-299.

Fuwa, N. (1999). "An Analysis of Social Mobility in a Village Community: The Case of a Philippine Village." *Journal of Policy Modeling, 21*(1), 101-138.

Gibson, J., and S. Olivia (2010). "The Effect of Infrastructure Access and Quality on Non-Farm Enterprises in Rural Indonesia." *World Development, 38*(5), 717-726.

Giné, X. and R.M. Townsend (2004). "Evaluation of financial liberalization: a general equilibrium model with constrained occupation choice," *Journal of Development Economics 74*(2), 269–30.

Haggblade, S., P. Hazell, and T. Reardon. (2002). "Strategies for stimulating poverty-alleviating growth in the rural non-farm economy in developing countries." International Food and Policy Research Institute Environment and Production Technology Division Discussion Paper number 92.

Hazell, P. B.R., S. Haggblade, and Thomas Reardon (2007). "Structural Transformation of the Rural Non-farm Economy", in S.Haggblade, P.B.R. Hazell, and T. Reardon, eds., *Transforming the Rural Non-farm Economy* (Baltimore: Johns Hopkins University Press).

Hung, Pham Thai, Bui Anh Tuan, and Dao Le Thanh. (2010). "Is Non-farm Diversification a Way Out of Poverty for Rural Households ? Evidence from Vietnam in 1993-2006." PMMA working paper.

Isvilanonda, S., A. Ahmad, and M. Hossain. (2000). "Recent Changes in Thailand's Rural Economy: Evidence from six villages." *Economic and Political Weekly* December, 4644-4649.

Jonasson, E. and S. M. Helfand. (2010). "How Important are Locational Characteristics for Rural Non-agricultural Employment? Lessons from Brazil." *World Development, 38*(5), 727-741.

Lanjouw, P. (2001). "Non-farm Employment and Poverty in Rural El Salvador." *World Development, 29*(3), 529-547.

Liedholm, C. (2007). "Enterprise Dynamics in the Rural Non-farm Economy", in S.Haggblade, P.B.R. Hazell, and T. Reardon, eds., *Transforming the Rural Non-farm Economy* (Baltimore: Johns Hopkins University Press).

Mammen, K. and C. Paxson (2000). "Women's Work and Economic Development," *Journal of Economic Perspectives 14*(4), 141-164.

Matsumoto, T., Y. Kijima and T. Yamano. (2006). "The role of local non-farm activities and migration in reducing poverty: evidence from Ethiopia, Kenya, and Uganda", *Agricultural Economics*, *35*(3), 449-458.

Mondragon-Velez, C. and X. Pena-Parga. (2008). "Business Ownership and Self-Employment in Developing Economies: The Colombian Case" in J.Lerner and A. Schoar, eds., *International Differences in Entrepreneurship* (Cambridge, MA: National Bureau for Economic Research): 89-127.

Naschold, F. and C.B. Barrett. (2011). "Do Short-Term Observed Income Changes Overstate Structural Economic Mobility?" *Oxford Bulletin of Economics and Statistics, 73*(5), 705-717.

Neal, Derek. (1995). "Industry-Specific Human Capital: Evidence from Displaced Workers", *Journal of Labor Economics, 13*(4), 653-677.

Paxson, C. (1992). "Using Weather Variability to Estimate the Response of Savings to Transitory Income in Thailand," *American Economic Review 82*(1), 15-33.

Quadrini, V. (2000). "Entrepreneurship, Saving, and Social Mobility." *Review of Economic Dynamics, 3*(1), 1-40.

Reardon, T. (1997), "Using Evidence of Household Income Diversification to Inform Study of the Rural Non-farm Labor Market in Africa", *World Development, 25*(5), 735-747.

Reardon, T., J. Taylor, K. Stamoulis, P. Lanjouw and A. Balisacan. (2000). "Effect of Non-Farm Employment on Rural Income Inequality in Developing Countries: An Investment Perspective", *Journal of Agricultural Economics, 51*(2), 266-288.

Rural Development Information Center, Community Development Department, Ministry of Interior. The University of Chicago-UTCC Research Center [distributor], 2008.

Schoar, A. (2010). "The Divide Between Subsistence and Transformational Entrepreneurship", *Innovation Policy and the Economy, 10*(1), 57-81.

Sahn, D. E., and D. Stifel. (2003). "Exploring alternative measures of welfare in the absence of expenditure data." *Review of Income and Wealth, 49*(4), 463-489.

Schultz, T.P. (1990). "Testing the Neoclassical Model of Family Labor Supply and Fertility," *Journal of Human Resources 25*(4), 599-634.

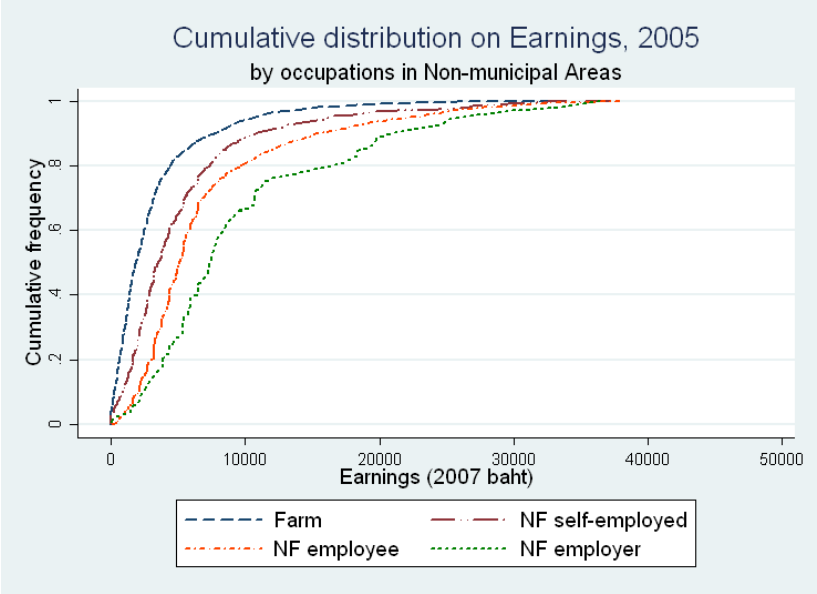
Timmer, C.P. (1988). "The Agricultural Transformation." In H. Chenery and T.N. Srinivasan eds., *Handbook of Development Economics,* Vol. 1. Amsterdam: North-Holland, pp. 275-331.

Timmer, C.P. (2002). "Agriculture and Economic Growth." In B. Gardner and G. Rausser, eds., *Handbook of Agricultural Economics, Vol. IIA*. Amsterdam: North-Holland. 1487-1546.

Woolard, I., and S. Klasen. (2005). "Determinants of Income Mobility and Household Poverty Dynamics in South Africa." *Journal of Development Studies, 41*(5), 865-897.

Wooldridge, J. M. (2002). *Economic Analysis of Cross-sectional and Panel Data*, Cambridge MA: MIT Press.

Figure 1: Cumulative distribution by occupation, 2005 and 2010



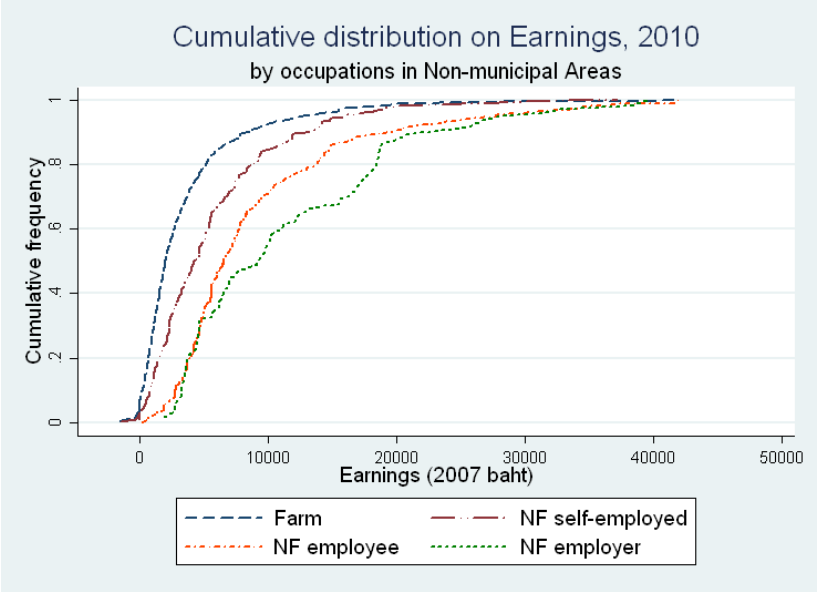
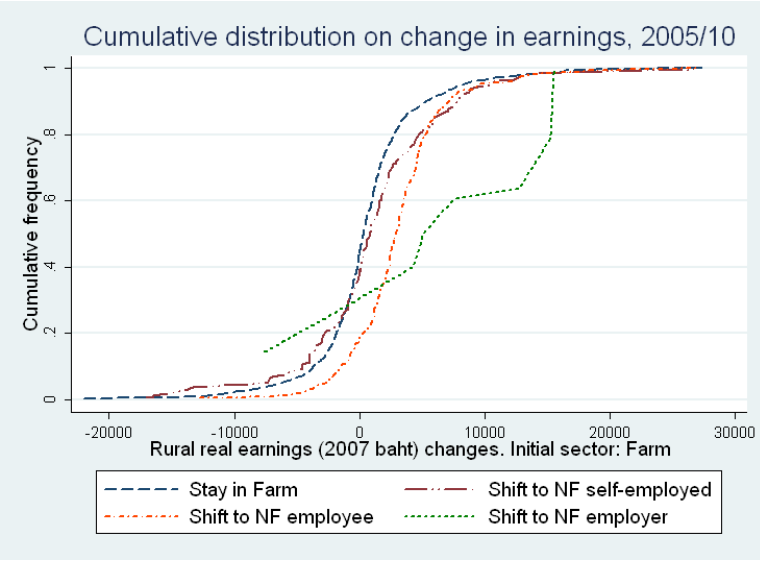
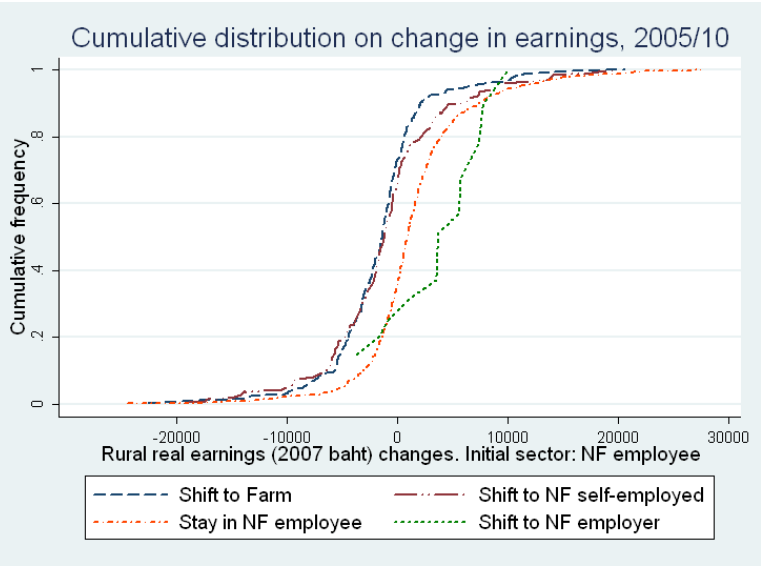


Figure 2: Cumulative distribution of change in earnings by occupational transition between 2005 and 2007

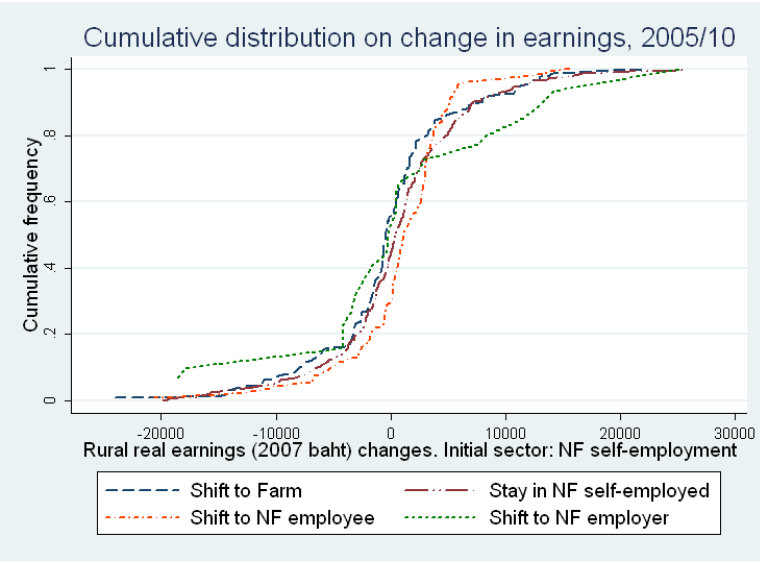
(a) Transitions from farm



(b)Transitions from non-farm employee



(c) Transitions from non-farm self-employed



(d) Transitions from non-farm employer

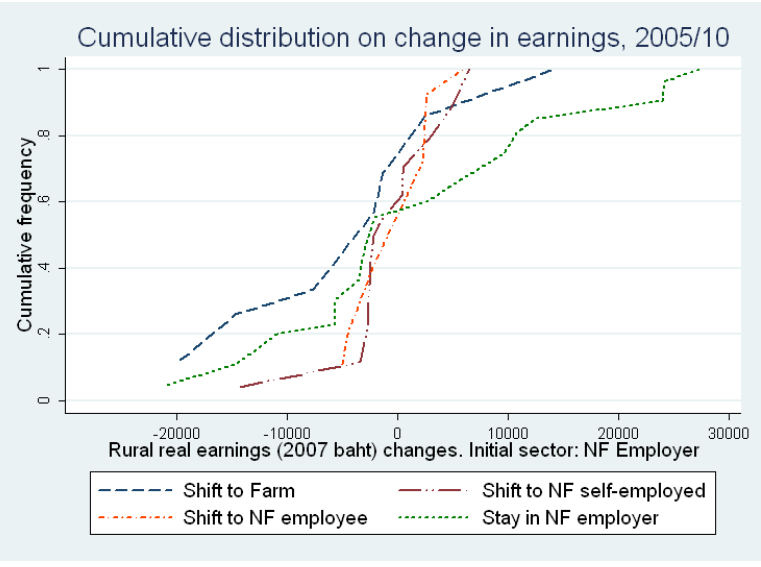


Table 1: Summary information on SES sample

Year	Individuals	Rural Individuals	Households	Rural Households
2005	16,310	9,897	6,000	3,680
2006	16,542	10,208	6,020	3,752
2007	16,490	10,350	5,955	3,783
2010	17,045	10,915	6,244	4,002
Observed all 4 years	12,758	7,831	5,229*	3,362*
Observed all 4 years (age ≤ 70)	11,484	7,000		
% Rural by 3 year total		61.4%		64.3%

* Based on the household ID that was recorded in 2005.

Table 2: Sector of individual non-farm employment, by employer size (percent)

2005	Government	State owned enterprise	Private sector	Total
2-9 workers	3.6	0.1	36.0	**39.8**
10-50 workers	8.4	0.4	19.1	**27.9**
51-100 workers	2.1	0.0	3.5	**5.7**
101-200 workers	1.4	0.2	3.8	**5.4**
201-500 workers	0.6	0.0	4.8	**5.4**
over 500 workers	6.6	0.8	8.5	**15.8**
Total (n=1,540)	**22.7**	**1.5**	**75.8**	**100.0**
2010				
2-9 workers	3.6	0.2	31.4	**35.2**
10-50 workers	12.3	0.6	19.5	**32.4**
51-100 workers	3.0	0.2	6.0	**9.2**
101-200 workers	1.5	0.3	4.9	**6.7**
201-500 workers	0.6	0.1	5.7	**6.4**
over 500 workers	3.7	0.2	6.3	**10.1**
Total (n=1,355)	**24.7**	**1.6**	**73.7**	**100.0**

Each cell reports a percentage of total non-farm employment.

Table 3: Work status in rural areas

Work Status (ages 15-70 years)	2005		2006		2007		2010	
	N	%	N	%	N	%	N	%
Unemployed	72	1.0	75	1.1	70	1.0	38	0.5
Other work status*	1,600	22.9	1,474	21.1	1,335	19.1	1,379	19.7
Employed status in each year	5,328	76.1	5,451	77.9	5,595	79.9	5,583	79.8
- Employed < 4 waves	1,238	17.7	1,361	19.4	1,505	21.5	1,493	21.3
- Employed status in all 4 waves	4,090	58.4	4,090	58.4	4,090	58.4	4,090	58.4
• Farm work	1,762	25.2	1,751	25.0	1,756	25.1	1,922	27.5
• NF self-employed	733	10.5	727	10.4	781	11.2	738	10.5
• NF employee	1,540	22.0	1,546	22.1	1,496	21.4	1,362	19.5
• NF employer	53	1.3	64	0.9	56	0.8	65	0.9
• Missing data on occupation	2	0.05	2	0.03	1	0.01	3	0.04
Total	7,000		7,000		7,000		7,000	

*Other work status includes waiting for seasonal work, unemployed, looking for work, retired, long term illness/disability, caring for other household members, student, and others.

Table 4: Mean individual earnings (per month) by quartile and occupation

2005	1st quartile	2nd quartile	3rd quartile	4th quartile	Overall
Farm work	309	1,234	2,513	7,647	3,397
	(242)	(258)	(478)	(4,773)	(4,001)
NF self-employed	1,051	2,741	4,866	12,203	5,503
	(649)	(439)	(822)	(6,618)	(5,484)
NF employee	2,220	4,181	6,143	15,683	7,684
	(800)	(517)	(749)	(7,045)	(6,593)
NF employer	3,295	6,450	9,554	22,287	10,471
	(1,379)	(752)	(1,383)	(6,817)	(8,197)

2007	1st quartile	2nd quartile	3rd quartile	4th quartile	Overall
Farm work	345	1,163	2,420	7,322	3,348
	(242)	(269)	(530)	(4,252)	(3,761)
NF self-employed	1,060	3,073	5,837	12,868	6,178
	(648)	(637)	(982)	(5,812)	(5,484)
NF employee	2,423	4,544	6,653	16,111	8,246
	(879)	(479)	(888)	(7,394)	(6,862)
NF employer	4,305	9,674	16,601	24,278	13,286
	(2,202)	(1,508)	(2,301)	(6,130)	(8,033)

2010	1st quartile	2nd quartile	3rd quartile	4th quartile	Overall
Farm work	260	1,348	2,819	9,407	4,004
	(424)	(307)	(651)	(6,660)	(5,231)
NF self-employed	943	3,003	5,676	13,348	5,808
	(594)	(687)	(863)	(5,484)	(5,528)
NF employee	2,839	5,309	8,362	21,676	9,951
	(1,123)	(628)	(1,360)	(8,941)	(8,900)
NF employer	3,656	6,537	14,038	25,549	11,407
	(806)	(1,127)	(3,295)	(7,971)	(8,934)

Units are 2007 Thai baht. The lowest and highest one percent of the rural income percentiles in each year have been omitted. Quartiles are based on employed status in each year. Standard deviation is shown in parentheses for each quartile by occupation.

Table 5: Transition matrix of occupational changes 2005/2010

Farm and Non-farm employment 2005	Farm work	Non-farm (NF) employment 2010			Total
		NF self-employed	NF employees	NF employers	
Farm work					
- *Number*	1,386	134	233	9	1,762
- *Row percentage*	78.66	7.6	13.22	0.51	100
NF self-employed					
- *Number*	153	462	93	24	732
- *Row percentage*	20.9	63.11	12.7	3.28	100
NF employees					
- *Number*	372	127	1,028	11	1,538
- *Row percentage*	24.19	8.26	66.84	0.72	100
NF employers					
- *Number*	9	15	8	21	53
- *Row percentage*	16.98	28.3	15.09	39.62	100
Total	1,920	738	1,362	65	4,085

The total number of individuals here differs from that reported in Table 2 because some employed individuals are missing data on occupation.

Table 6: Transition matrix of median and mean percentage change in earnings 2005/2010

Farm and Non-farm employment 2005	Farm work	Non-farm (NF) employment 2010		
		NF self-employed	NF employees	NF employers
Farm work				
- *Median of ΔY*	297	821	3,338	7,499
- *Mean of %ΔY*	1.17	2.63	5.07	5.12
- *(s.d.)*	(3.09)	(5.15)	(6.60)	(6.84)
NF self-employed				
- *Median of ΔY*	-479	156	2,223	481
- *Mean of %ΔY*	1.08	0.80	1.29	0.78
- *(s.d.)*	(3.69)	(2.69)	(2.23)	(2.43)
NF employees				
- *Median of ΔY*	-1,159	-1,239	975	3,722
- *Mean of %ΔY*	0.03	0.24	0.35	1.04
- *(s.d.)*	(1.70)	(1.75)	(1.03)	(1.23)
NF employers				
- *Median of ΔY*	-2,209	-1,754	2,279	-2,107
- *Mean of %ΔY*	-0.18	-0.12	-0.12	0.57
- *(s.d.)*	(0.99)	(0.56)	(0.85)	(1.36)

Earnings changes (*ΔY*) are in real Thai baht, adjusted by CPI for each region of Thailand with base year 2007. The reported statistics omit the top and the bottom one percent of the sample and correct for attrition weights as described in Appendix A.

Table 7: Summary statistics of variables used in the multivariate analysis

Variables	Mean	Std. Dev.	Description
Log earnings (t-1)	8.17	1.30	Previous year earnings
Individual characteristics			All in year 2005 from SES
Age	38.48	10.86	Individual's years of age
Age^2	1598.81	871.57	Square of individual's years of age
HH head	0.41	0.49	= 1 if individual is a household head; 0 otherwise
Married	0.80	0.40	= 1 if marital status is married; 0 otherwise
Female	0.42	0.49	= 1 if gender is female; 0 otherwise
Education (Base: None/less than primary school)			
- Primary school	0.61	0.49	= 1 if completed the primary school (grade 6)
- Secondary school	0.13	0.34	= 1 if completed the secondary school (grade 9)
- High/Vocational school	0.15	0.36	= 1 if completed the high/vocational school (grade 12)
- College and above	0.07	0.26	= 1 if completed college level or higher level
Total working months	72.89	100.55	How long has individual been working in this occupation
Household characteristics			
Family size	4.26	1.72	Number of members in the household
Owned agricultural land	38.48	108.07	Agricultural area owned by household (100 rai unit; 1 rai = 1,600 m^2)
Asset Index	0.24	1.05	Asset index for household wealth based on housing characteristics, durable goods, and agricultural lands

	Farm work		NF self-employed		NF employees		NF employers	
Variables	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Individual characteristics								
Age	40.31	10.92	39.93	11.32	36.06	10.14	39.26	10.21
HH head	0.43	0.50	0.37	0.48	0.39	0.49	0.78	0.42
Married	0.84	0.37	0.84	0.37	0.75	0.43	0.90	0.30
Female	0.40	0.49	0.57	0.50	0.37	0.48	0.14	0.35
Education								
- Primary school	0.75	0.43	0.62	0.49	0.47	0.50	0.42	0.50
- Secondary school	0.10	0.30	0.16	0.36	0.15	0.36	0.15	0.37
- High/Vocational school	0.09	0.29	0.13	0.34	0.20	0.40	0.35	0.48
- College and above	0.01	0.10	0.03	0.17	0.15	0.36	0.04	0.20
Total working months	78.06	111.58	63.11	86.65	71.90	94.52	76.42	94.91
Household characteristics								
Family size	4.28	1.64	4.20	1.74	4.27	1.75	4.31	2.49
Owned agricultural land	66.19	148.44	22.21	50.71	18.64	59.15	33.54	181.65
Asset Index	-0.03	0.89	0.45	0.85	0.39	1.15	1.14	1.91

Table 8: Multivariate regressions of log earnings change on selection variables

Dependent V: $\Delta \ln y_{it}$	(1) OLS (Bootstrap s.e.)		(2) Individual fixed effects		(3) Hausman-Taylor (random effects)	
	Coef.	Std. err.	Coef.	Std. err.	Coef.	Std. err.
Log(earning (t-1))	-0.69***	0.03	-1.24***	0.04	-1.03***	0.02
Log(earning (t-1))*t_06	-0.04	0.04	0.01**	0.004	0.01***	0.002
Log(earning (t-1))*t_07	-0.14***	0.04	0.03***	0.01	0.02***	0.002
Sector transitions						
Farm to (farm: base)						
NF self-employed	0.07	0.15	-0.26	0.36	0.18	0.15
NF employee	0.78***	0.06	0.90***	0.27	1.03***	0.07
NF employer	1.29***	0.26	1.19***	0.33	1.21***	0.30
NF self-employed to						
Farm	-0.15*	0.08	-0.05	0.15	0.03	0.10
NF self-employed	0.45***	0.06	0.46**	0.20	0.75***	0.13
NF employee	0.40***	0.07	0.75***	0.21	1.10***	0.11
NF employer	0.76***	0.22	0.58**	0.24	1.04***	0.25
NF employee to						
Farm	-0.54***	0.06	0.24	0.23	0.16**	0.08
NF self-employed	0.19**	0.09	0.76***	0.23	0.95***	0.13
NF employee	0.66***	0.05	1.30***	0.23	1.72***	0.09
NF employer	1.00***	0.22	1.51***	0.27	1.86***	0.27
NF employer to						
Farm	-0.27	0.21	-0.07	0.29	0.09	0.31
NF self-employed	0.53***	0.11	0.54**	0.23	0.94***	0.22
NF employee	0.26	0.21	0.94***	0.25	1.25***	0.22
NF employer	0.76***	0.10	0.77***	0.27	1.44***	0.27
Age	0.04***	0.01			-0.01	0.01
Age^2	-0.001***	0.0001			0.00	0.00
HH head	0.11***	0.02			0.16***	0.04
Married	0.05*	0.03			0.16***	0.04
Female	-0.06***	0.02			-0.03	0.03
Education (Less than primary school/none:base)						
Primary school	0.20**	0.08			-0.15	0.09
Secondary school	0.29***	0.09			-0.13	0.10
High/Vocational school	0.42***	0.09			-0.05	0.10
College and above	0.78***	0.09			0.08	0.11
Total working months	0.0002	0.0001			0.0003***	0.0001
Family size	-0.02*	0.01			-0.06***	0.01
Owned agricultural land	0.001*	0.0004			0.002***	0.001
Asset Index	0.21***	0.03			0.36***	0.03
Time effect (2006)	0.39	0.34				
Time effect (2007)	1.33***	0.32				
Individual fixed effects	No		Yes		Random effects	
Sub-district fixed effects	Yes		No		No	
Constant	4.30***	0.32	9.34***	0.37	8.03***	0.40
Adjusted R^2	0.41		0.65			
Wald test (bootstrapped)	2595.66***				4233.25***	

***P<0.01, **P<0.05, and *P<0.10. Variance-covariance matrices are clustered at sub-district level for Model (1) and (2), and are bootstrapped for Model (1) and (3). Model (2) uses attrition weights, as described in Appendix A. We omit the top and the bottom one percent of the sample used, yielding 11,342 observations used.

Table 9: Change in estimated log earnings differences by sectoral transition

Model from Table 7: Sectoral transitions	(1) differences	(2) differences	(3) differences
From farm:			
to NF self-employed vs. to NF worker	-0.71***	-1.16***	-0.84***
to NF self-employed vs. to NF employer	-1.22***	-1.45***	-1.03***
to NF worker vs. to NF employer	-0.51**	-0.29	-0.18
From NF self-employed:			
to farm vs. stay NF self-employed	-0.59***	-0.51***	-0.72***
to farm vs. to NF worker	-0.55***	-0.80***	-1.07***
to farm vs. to NF employer	-0.91***	-0.63***	-1.02***
stay NF self-employed vs. to NF worker	0.04	-0.29**	-0.35***
stay NF self-employed vs. to NF employer	-0.32	-0.13	-0.30
to NF worker vs. to NF employer	-0.36	0.17	0.06
From NF worker:			
to farm vs. to NF self-employed	-0.73***	-0.52***	-0.79***
to farm vs. stay NF worker	-1.20***	-1.06***	-1.56***
to farm vs. to NF employer	-1.54***	-1.27***	-1.70***
to NF self-employed vs. stay NF worker	-0.47***	-0.54***	-0.77***
to NF self-employed vs. to NF employer	-0.81***	-0.75***	-0.91***
stay NF worker vs. to NF employer	-0.34	-0.21	-0.15
From NF employer:			
to farm vs. to NF self-employed	-0.80***	-0.61**	-0.85***
to farm vs. to NF worker	-0.53*	-1.02***	-1.16***
to farm vs. stay NF employer	-1.03***	-0.85***	-1.34***
to NF self-employed vs. to NF worker	0.27	-0.41*	-0.31
to NF self-employed vs. stay NF employer	-0.23*	-0.24	-0.49**
to NF worker vs. stay NF employer	-0.50**	0.17	-0.18

Note: A negative entry implies that the latter transition (e.g., from farm to NF worker, in the first row) yields a higher expected log earnings change than does the former (e.g., from farm to NF self-employed, in the first row); and vice versa for positive entries. *, ** and *** indicate statistically significant differences at the 10, 5 and 1 % levels, respectively.

Table 10: Estimations of log earnings change in 2005-7 and 2007-10, given initial occupation in 2005

Dependent V: $\Delta \ln y_{it}$	(1) OLS Farm		(2) IV-OLS Farm		(3) OLS NF self-employed		(4) IV-OLS NF self-employed		(5) OLS NF employee		(6) IV-OLS NF employee	
	Coef.	Std. err.	Coef.	Std. err.	Coef.	Std. err.	Coef.	Std. err.	Coef.	Std. err.	Coef.	Std. err.
Log(earnings (t-1))	-0.87***	0.05	-0.93***	0.06	-0.94***	0.04	-0.98***	0.07	-0.62***	0.04	-0.54***	0.06
Log(earning (t-1))*t_07	-0.19**	0.08	-0.16**	0.08	-0.09	0.07	-0.06	0.09	0.02	0.06	0.09	0.06
Predicted probability of transitions to												
Farm	6.16***	0.86	5.53***	0.92	-0.94***	0.16	-0.94	0.76	-1.09***	0.10	-0.06	0.25
NF self-employed	-0.29	0.32	-5.62*	2.86	7.98***	0.69	8.28***	0.96	-0.37***	0.11	0.11	0.66
NF employee	0.78***	0.12	1.61*	0.88	-0.23	0.16	-0.40	0.63	5.19***	0.43	4.47***	0.53
NF employer	1.64***	0.58	1.86	1.43	0.33*	0.18	0.41	0.59	0.26*	0.14	2.07	1.55
Age	0.02	0.04	0.07*	0.04	0.03	0.03	0.03	0.03	0.0004	0.01	-0.0001	0.01
Age^2	-0.0003	0.0004	-0.001	0.0005	-0.0005	0.0003	-0.0006	0.0004	-0.0001	0.0002	-0.0001	0.0002
HH head	-0.02	0.12	-0.01	0.12	0.19**	0.09	0.17*	0.10	0.08*	0.05	0.06	0.05
Married	0.16	0.16	-0.11	0.17	-0.09	0.11	-0.11	0.13	0.03	0.04	-0.001	0.05
Female	-0.05	0.11	0.09	0.10	-0.16*	0.09	-0.15	0.11	-0.05	0.05	-0.03	0.05
Education (Less than primary school/none:base)												
Primary school	0.16	0.13	0.40**	0.17	-0.35	0.26	-0.30	0.26	0.10	0.08	0.03	0.09
Secondary school	0.20	0.17	0.58**	0.26	-0.47*	0.25	-0.45*	0.27	0.25***	0.09	0.18*	0.11
High/Vocational school	0.14	0.21	0.26	0.22	-0.16	0.28	-0.18	0.30	0.51***	0.10	0.51***	0.11
College and above	0.64*	0.33	1.33**	0.52	-0.40	0.35	-0.42	0.36	0.70***	0.10	0.74***	0.12
Working months	0.0002	0.0003	-0.0003	0.0005	0.00002	0.0004	0.0001	0.0004	0.001***	0.0003	0.001**	0.0003
Family size	-0.02	0.03	-0.01	0.03	-0.01	0.04	-0.01	0.04	-0.02	0.01	-0.01	0.02
Owned land	0.001**	0.0003	0.0003	0.001	0.001	0.001	0.001	0.002	0.0005***	0.0001	0.0002	0.0002
Asset Index	0.16*	0.09	0.43***	0.12	0.22***	0.07	0.24***	0.07	0.11***	0.03	0.09***	0.03
Time dummy	1.73***	0.63	1.60**	0.62	0.75	0.59	0.40	0.82	-0.18	0.51	-0.86	0.56
Adjusted R^2	0.60		0.57		0.74		0.71		0.42		0.28	
N. of obs	2773		2773		1174		1174		2417		2417	

***P<0.01, **P<0.05, and *P<0.10. We omit the top and the bottom one percent of the sample used. Variance-covariance matrices are clustered at village level and attrition weights applied.

Table 11: Average log earnings changes from the selection bias correction estimations

	$E[\Delta \ln y \mid \hat{P}(j \text{ to } k), Z]$	Std Dev.
Farm to		
Farm	0.11**	0.86
NF self-employed	0.34***	1.26
NF employee	1.43***	1.78
NF self-employed to		
Farm	-0.07	2.17
NF self-employed	0.13***	0.97
NF employee	0.10	1.23
NF employer	0.32*	0.89
NF employee to		
Farm	-0.93***	0.60
NF self-employed	-0.20***	0.70
NF employee	0.07***	0.24
NF employer	0.54***	0.78

***P<0.01, **P<0.05, and *P<0.10. We omit the top and the bottom one percent of the sample used. Estimation of transition from farm to NF employer is omitted since there are only nine observations.

Appendix A: Attrition analysis

Since only 7,000 individuals, ages 15-70 years old, appear in all four SES waves while total rural individuals in the same age range each year cover more than 9,000 individuals in the initial, 2005 round, there is a potential concern that non-random attrition might bias inferences based on the balanced panel. This appendix explores issues related to attrition in this panel and explains how we have addressed those issues.

Table A1 shows the changes in rural individuals across all four waves. Only 1-2 percent of rural residents moved to urban areas from one survey round to the next. So rural outmigration is not a significant potential source of selection problems.

Table A1: Sample individuals in rural areas, ages 15-70 years

Year	Total rural	Total 2005 rural remaining	Missing	Added in	Rural to urban	Urban to rural
2005	9,153	9,153	0	0		
2006	9,415	8,443	710	972	87	58
2007	9,475	7,892	551	1,583	72	67
2010	9,984	7,000	892	2,984	196	106

We restrict attention to just those rural individuals who were employed in all four rounds, so as to avoid conflating transitions between unemployment and employment with transitions among occupations. Roughly 10-15% of the sample transitioned between unemployment and employment between rounds. As a consequence, selection bias could be an issue if there are significant differences in characteristics between those individuals present and employed in all four survey waves and those who were rural residents in 2005 but not present and employed in all four waves, given attrition – due mainly to changing household composition (due to deaths or individuals aging out of the workforce as the survey progressed) – or transitions in and out of unemployment.

Given the possibility of selection problems from restricting our analysis to those rural residents who were surveyed and employed in all four waves, we test for differences in mean characteristics between individuals in the retained subsample and those who were initially surveyed in the 2005 round. Table A2 shows that most of each group's characteristics are statistically significantly different but the key dependent variable, changes in log earnings, do not show a significant difference between the groups. As expected, those who were employed in all four waves and remained in the subsample have higher earnings on average than the full rural sample surveyed in 2005.

Table A2: Mean characteristics tests across groups

Variables	All rural 2005	Subsample	t-stat
Age	39.25	38.56	-3.77***
HH head	0.34	0.41	12.30***
Married	0.66	0.80	17.63***
Family size	4.31	4.26	-1.58
Education			
- Primary school	0.57	0.61	5.57***
- Secondary school	0.17	0.13	-7.39***
- High/Vocational school	0.16	0.15	-1.94*
- College and above	0.05	0.07	5.35***
Total working months	62.36	72.51	5.00***
Owned agricultural land	3338.06	3762.43	3.11***
Asset Index	0.13	0.23	5.84***
Earnings 2005	5465.97	6301.63	7.59***
Earnings 2006	5560.43	6605.51	7.94***
Earnings 2007	5582.39	6435.02	10.20***
Earnings 2010	6310.29	7191.22	8.61***
Log earnings 2005	7.76	8.05	8.94***
Log earnings 2006	7.76	8.05	10.57***
Log earnings 2007	7.94	8.16	13.37***
Log earnings 2010	8.08	8.23	8.82***
Δ Log earnings 0506	-0.04	-0.01	1.45
Δ Log earnings 0607	0.16	0.10	-2.88***
Δ Log earnings 0710	0.10	0.07	-2.01**
Δ Log earnings 0507	0.13	0.10	-1.21
Δ Log earnings 0510	0.21	0.18	-1.09

***P<0.01, **P<0.05, and *P<0.10. Mean testing uses the survey individual weight.

Since tests of mean group characteristics show significant differences, we need to test for possible attrition bias. We follow the procedure suggested by Baulch and Quisumbing (2011), first estimating an attrition probit regression, then running the tests proposed by Becketti, Gould, Lillard and Welch (1998, hereafter BGLW), then adjusting using inverse probability weights, as suggested by Fitzgerald et al. (1998) and Wooldridge (2002).

Table A3: Attrition probit regressions

	(1)	(2)
Pseudo-R^2	0.1527	0.2547
Wald Tests (P-value)		
Rural village attrition rate	0.000	0.000
Log (earnings) 05	0.000	
Assets	0.213	0.051
Village variables[a]	0.750	
Demography and Education[b]	0.000	0.000
Other variables[c]	0.000	0.000
No. of observations[d]	7000	9897

[a] Village variables include dummy variables for electricity available to all households in the village, insufficient water for agriculture, year-round trafficable road, agricultural loss and stagnant flood, ratio of asphalt/concrete section to total of the most convenient route to the nearest major district, travel time to the nearest district, number of soil problems, and ratios of households in the village working in establishment, manufacturing, employment, and agriculture. [b] Demography and education variables include age, age^2, household head, gender, marital status, and number of household members. [c] Other variables include demographic and education variables, asset variables, and region dummies. Variance-covariance matrices are clustered at sub-district level. [d] In both columns, we use all observations in the initial rural survey. Since there are many missing values from log(earnings) and village variables, the number of observations in column (1) has dropped as reported.

We first run a probit regression where the dependent variable takes the value of one for an individual who dropped out of the sample after the first wave (from aging out of the workforce, dying, becoming unemployed, or moving and being untraceable), then regressing that dichotomous dependent variable on baseline variables that could affect both the likelihood of attrition and the outcome variable of interest (earnings). The attrition probit regressions reported in Table A3 column 1 include individual and household characteristics, household owned agricultural land and asset index, log earnings in the initial survey round, rural village attrition

rate,[16] dummy variables for each of the five main regions in Thailand, and village characteristics variables in 2005. The pseudo R^2 statistic can be interpreted as the proportion of attrition that is non-random. Wald tests are then performed to test whether observables jointly explain the predicted attrition probability. As shown in Table A3 column 1, the pseudo R^2 values are relatively low. Log earnings in 2005 and demographic variables (age, age^2, household head, female, and marital status) are statistically significantly different from zero at the 1% level, and jointly explain the attrition rate. On the other hand, village variables and asset variables insignificantly determine attrition probabilities. Many unemployed individuals have missing values of log earnings in 2005, however, as well as some missing values for village characteristics variables. If we exclude log earnings and village characteristics variables from attrition probit regressions, as shown in Table 3A column 2, we can use all the observations from the 2005 survey. This attrition probit yields a higher pseudo R^2. We can strongly reject the null hypothesis that attrition is unrelated to individual and household characteristics.

The BGLW test is based on an F-test of the joint significance of the attrition dummy and the interactions between attrition dummy and the explanatory variables when regressing the first wave outcome variable on the determinants of outcome variable, plus attrition dummy and its interactions. Hence, the model estimation is defined as

$$\Delta \ln Earnings_{0506,i} = \delta_0 + X_{05i}\delta_1 + \delta_2 Attrition_i + [X_{05i} \bullet Attrition_i]\delta_3 + v_i,$$

where X includes individual and household characteristics, household owned agricultural land and asset index, log earnings in 2005, rural village attrition rate, and regional dummy variables. The test determines whether there exist statistically significant differences between the coefficient estimates for retained and attritted individuals. The results presented in Table A4

[16] Village attrition rate is the ratio of total individuals dropped out of the survey after the first wave to total individuals in the village in 2005.

show that the attrition dummy and its interactions are jointly statistically significant, and hence reject the null hypothesis that sample attrition is random and has no effect on coefficient estimates.[17]

Table A4: BGLW Attrition pooling tests

	(1)
R-squared	0.3967
All Interactions	
(F-stat)	4.14***
(P-value)	0.000
Attrition dummy *(P-value)*	0.260
Log (earnings) 05 *(P-value)*	0.000
A*log(earnings) 05 *(P-value)*	0.957
No. of observations[a]	6507

Variance-covariance matrices are clustered at sub-district level.
[a] Observations are dropped due to missing log earnings changes variable.

These tests suggest that non-random attrition could bias the main estimations. We use the inverse probability weighted method to correct for possible biases. Following Baulch and Quisumbig (2011), we create the ratio of predicted values from the restricted regression and unrestricted regression of reversed attrition probit where the dependent variable, RA = 1 if non-attrition, where the unrestricted regression include the same explanatory variables as the attrition probits in Table A3 column (2), while the restricted regression excludes the auxiliary variables (demographic variables, household asset variables, rural village attrition rate, and region dummies) in the first period. The inverse probability weights vary from 0.40 to 26.17 with mean 0.92.

Tables A5 – A7 exhibit regression results that replicate the pooled OLS (with sub-district

[17] If we regress changes in log earnings between the second and the third wave on the first wave explanatory variables, we also have very low R^2 (around 0.016) and an insignificant F-test statistic on the null that all interaction terms equal zero. However, the log earnings interaction coefficient estimate remains statistically significant at the 10% level.

fixed effects), individual fixed effects, and Hausman-Taylor estimations, respectively, from the main body of the paper. Most of the coefficients and significance of individual coefficients are very similar. We therefore test whether coefficients with and without inverse probability weighted corrections differ statistically significantly for each of the three models. We cannot reject the null hypothesis that the parameters are equal with and without correction for attrition bias for both the pooled OLS and individual fixed effects models. A few variables' coefficient estimates are statistically significantly different for the Hausman-Taylor estimations. Therefore, although attrition bias indeed appears to exist, it seems that it does not matter much to our estimation results and resulting inferences.

Table A5: Pooled OLS regressions of log earnings change on selection variables

Dependent V: $\Delta \ln y_{it}$	(1) No Weight		(2) Attrition weight	
	Coef.	Std. err.	Coef.	Std. err.
Log(earning (t-1))	-0.69***	0.03	-0.68***	0.06
Log(earning (t-1))*t_06	-0.04	0.04	-0.13*	0.08
Log(earning (t-1))*t_07	-0.14***	0.04	-0.23***	0.06
Sector transitions				
Farm to (farm: base)				
NF self-employed	0.07	0.15	-0.25	0.32
NF employee	0.78***	0.06	0.97***	0.15
NF employer	1.29***	0.25	1.38***	0.26
NF self-employed to				
Farm	-0.15*	0.08	-0.01	0.11
NF self-employed	0.45***	0.06	0.56***	0.08
NF employee	0.40***	0.07	0.44***	0.12
NF employer	0.76***	0.22	1.02***	0.21
NF employee to				
Farm	-0.54***	0.07	-0.49***	0.15
NF self-employed	0.19*	0.10	0.32***	0.11
NF employee	0.66***	0.05	0.77***	0.06
NF employer	1.00***	0.22	1.08***	0.19
NF employer to				
Farm	-0.27	0.19	-0.15	0.21
NF self-employed	0.53***	0.11	0.68***	0.17
NF employee	0.26	0.21	0.27	0.25
NF employer	0.76***	0.10	0.89***	0.13
Age	0.04***	0.01	0.04***	0.01
Age^2	-0.001***	0.0001	-0.001***	0.0001
HH head	0.11***	0.02	0.08**	0.03
Married	0.05*	0.03	0.07*	0.04
Female	-0.06***	0.02	-0.06	0.04
Education (Less than primary school/none:base)				
Primary school	0.20**	0.09	0.26**	0.13
Secondary school	0.29***	0.09	0.35**	0.14
High/Vocational school	0.42***	0.09	0.53***	0.15
College and above	0.78***	0.10	0.88***	0.14
Total working months	0.0002	0.0001	0.0001	0.0002
Family size	-0.06***	0.02	-0.06	0.04
Owned agricultural land	0.001**	0.0003	0.001	0.0003
Asset Index	0.21***	0.03	0.22***	0.03
Time effect (2006)	0.39	0.35	1.23*	0.66
Time effect (2007)	1.33***	0.31	2.12***	0.54
Constant	4.30***	0.33	3.92***	0.55
Adjusted R^2	0.41		0.46	
N. of observations	11342		11342	

***P<0.01, **P<0.05, and *P<0.10. Variance-covariance matrices are clustered at sub-district level. We omit the top and the bottom one percent of the sample used.

Table A6: Individual fixed effects regressions of log earnings change on selection variables

Dependent V: $\Delta \ln y_{it}$	(1) No Weight		(2) Attrition weight	
	Coef.	Std. err.	Coef.	Std. err.
Log(earning (t-1))	-1.21***	0.02	-1.24***	0.04
Log(earning (t-1))*t_06	0.004	0.003	0.01**	0.004
Log(earning (t-1))*t_07	0.02***	0.003	0.03***	0.01
Sector transitions				
Farm to (farm: base)				
NF self-employed	0.06	0.17	-0.26	0.36
NF employee	0.77***	0.09	0.90***	0.27
NF employer	1.08***	0.30	1.19***	0.33
NF self-employed to				
Farm	-0.05	0.11	-0.05	0.15
NF self-employed	0.42***	0.14	0.46**	0.20
NF employee	0.74***	0.13	0.75***	0.21
NF employer	0.57**	0.23	0.58**	0.24
NF employee to				
Farm	0.02	0.08	0.24	0.23
NF self-employed	0.61***	0.14	0.76***	0.23
NF employee	1.16***	0.11	1.30***	0.23
NF employer	1.41***	0.22	1.51***	0.27
NF employer to				
Farm	-0.14	0.28	-0.07	0.29
NF self-employed	0.50***	0.19	0.54**	0.23
NF employee	0.86***	0.19	0.94***	0.25
NF employer	0.76***	0.23	0.77***	0.27
Constant	9.34***	0.21	9.34***	0.37
Adjusted R^2	0.62		0.65	
N. of observations	11342		11342	

***P<0.01, **P<0.05, and *P<0.10. Variance-covariance matrices are clustered at sub-district level. We omit the top and the bottom one percent of the sample used.

Table A7: Hausman-Taylor estimation of log earnings change on selection variables

Dependent V: $\Delta \ln y_{it}$	(1) No Weight		(2) Attrition weight	
	Coef.	Std. err.	Coef.	Std. err.
Log(earning (t-1))	-1.03***	0.02	-1.06***	0.01
Log(earning (t-1))*t_06	0.01***	0.002	0.01***	0.003
Log(earning (t-1))*t_07	0.02***	0.002	0.03***	0.003
Sector transitions				
Farm to (farm: base)				
NF self-employed	0.18	0.15	-0.10	0.08
NF employee	1.03***	0.07	1.24***	0.06
NF employer	1.21***	0.30	1.39***	0.31
NF self-employed to				
Farm	0.03	0.10	0.17**	0.08
NF self-employed	0.75***	0.13	1.02***	0.09
NF employee	1.10***	0.11	1.27***	0.11
NF employer	1.04***	0.25	1.28***	0.22
NF employee to				
Farm	0.16**	0.08	0.37***	0.07
NF self-employed	0.95***	0.13	1.29***	0.10
NF employee	1.72***	0.09	2.01***	0.07
NF employer	1.86***	0.27	1.97***	0.31
NF employer to				
Farm	0.09	0.31	0.23	0.38
NF self-employed	0.94***	0.22	1.18***	0.24
NF employee	1.25***	0.22	1.43***	0.33
NF employer	1.44***	0.27	1.64***	0.30
Age	-0.01	0.01	0.03***	0.01
Age^2	0.00	0.00	-0.001***	0.0001
HH head	0.16***	0.04	0.08*	0.05
Married	0.16***	0.04	0.24***	0.05
Female	-0.03	0.03	-0.09**	0.04
Education (Less than primary school/none:base)				
Primary school	-0.15	0.09	-0.10	0.09
Secondary school	-0.13	0.10	-0.03	0.10
High/Vocational school	-0.05	0.10	0.09	0.10
College and above	0.08	0.11	0.13	0.12
Total working months	0.0003***	0.0001	0.0003***	0.0001
Family size	-0.06***	0.01	-0.04***	0.01
Owned agricultural land	0.002***	0.001	0.00***	0.00
Asset Index	0.36***	0.03	0.35***	0.02
Constant	8.03***	0.40	6.98***	0.23
Wald test	4233.25		11371.17	
N. of observations	11342		11342	

***P<0.01, **P<0.05, and *P<0.10. Variance-covariance matrices are clustered at sub-district level. We omit the top and the bottom one percent of the sample used.

Appendix B: Asset Index

The asset index used as a proxy for household wealth was constructed using factor analysis based on the method of Sahn and Stifel (2003). Dummy variables for dwelling characteristics and number of durable goods owned are used in the analysis to obtain the first factor. All data come from the SES panel in 2005 (initial year). Table B1 reports summary statistics, factor loadings and scoring coefficients.

Table B1: Asset index summary statistics, factors and scoring coefficients

Variable	Mean	SD	Factor loading	Scoring coef.
Number of rooms	2.73	1.30	0.48	0.09
Housing materials (dummies)				
Cement	0.33	0.47	0.27	0.05
Wood	0.34	0.47	-0.31	-0.06
Others: local/reused materials	0.33	0.47	0.05	0.01
Having electricity	0.99	0.09	0.11	0.02
Cooking fuel (dummies)				
Gas	0.52	0.50	0.53	0.10
Others: Electricity, charcoal, wood	0.45	0.50	-0.52	-0.10
Water supply (dummies)				
Pipe	0.61	0.49	0.09	0.02
Underground water	0.35	0.48	-0.07	-0.01
Others: rain, open sources	0.04	0.19	-0.04	-0.01
Toilet: flush	0.99	0.11	0.11	0.02
Household items (number)				
Microwave	0.08	0.28	0.49	0.09
Refrigerator	0.86	0.47	0.53	0.10
Air condition	0.09	0.40	0.51	0.09
Fan	2.08	1.28	0.61	0.11
Radio	0.71	0.61	0.49	0.09
VCD-DVD player	0.66	0.57	0.56	0.10
Washing machine	0.37	0.50	0.62	0.11
Television	1.13	0.59	0.65	0.12
Cable television	0.01	0.09	0.08	0.01
Satellite dish (for TV)	0.01	0.10	0.21	0.04
Landline	0.17	0.39	0.56	0.10
Cell phone	0.77	0.85	0.65	0.12
Computer	0.09	0.31	0.62	0.11
Internet access	0.03	0.18	0.48	0.09

Variable	Mean	SD	Factor loading	Scoring coef.
Vehicles (number)				
Motorcycle	1.08	0.82	0.36	0.07
Car	0.07	0.28	0.46	0.08
Mini-truck/Van	0.19	0.46	0.45	0.08
2-wheel tractor	0.22	0.44	-0.09	-0.02
4-wheel tractor	0.01	0.11	0.10	0.02
6-wheel or higher	0.01	0.14	0.20	0.04
Owned livestock				
Buffalo	0.04	0.48	-0.07	-0.01
Cow	0.05	0.61	-0.06	-0.01

Appendix Table C1: Multivariate regressions of earnings change (level)

Dependent V: Δy_{it}	(1) OLS (Bootstrap s.e.)		(2) Individual fixed effects		(3) Hausman-Taylor (random effects)	
	Coef.	Std. err.	Coef.	Std. err.	Coef.	Std. err.
Earnings (t-1)	-0.42***	0.03	-1.20***	0.03	-0.88***	0.03
Earnings (t-1)*t_06	0.02	0.03	-0.01	0.02	-0.00	0.02
Earnings (t-1)*t_07	0.04	0.04	0.07***	0.02	0.07***	0.02
Sector transitions						
Farm to (farm: base)						
NF self-employed	1065.37***	307.73	628.74**	304.01	964.14***	336.78
NF employee	1989.65***	181.24	1653.30***	290.90	2583.88***	197.67
NF employer	4835.78***	1757.25	4480.41**	1855.96	4788.83**	2121.98
NF self-employed to						
Farm	-810.26***	269.86	-30.57	250.61	-188.32	289.07
NF self-employed	676.31***	156.64	1491.93***	331.03	1660.85***	428.65
NF employee	863.82***	313.75	1663.33***	394.25	2449.84***	348.19
NF employer	2439.56**	1151.30	2858.35**	1148.78	4252.66***	1192.15
NF employee to						
Farm	-1773.72***	188.06	246.61	240.27	139.43	187.65
NF self-employed	-376.87	356.18	1514.53***	490.37	1475.32***	418.37
NF employee	1504.56***	158.14	2757.15***	270.51	4134.05***	261.63
NF employer	2540.48**	1208.43	4878.30***	1163.42	5054.67***	1187.10
NF employer to						
Farm	-2736.03*	1403.42	-849.64	1477.14	-298.51	1602.25
NF self-employed	945.48	907.50	817.64	1109.49	2888.46**	1354.14
NF employee	-3259.43**	1410.57	263.01	1383.87	416.58	1434.27
NF employer	2175.61**	891.83	1135.96	2112.43	4613.94**	2147.88
Age	114.00***	22.02			66.26	56.32
Age^2	-1.44***	0.26			-0.98	0.63
HH head	160.92*	86.02			624.14***	175.51
Married	222.62**	111.28			689.83***	165.02
Female	-280.32***	72.31			-446.60***	147.73
Education (Less than primary school/none:base)						
Primary school	109.13	180.29			-597.44*	322.76
Secondary school	521.36**	220.94			-107.76	428.85
High/Vocational school	1099.60***	244.64			890.27*	476.54
College and above	4742.58***	467.58			7420.76***	714.96
Total working months	1.70***	0.43			1.38***	0.40
Family size	-113.13***	32.73			-339.10***	49.84
Owned agricultural land	2.01*	1.19			4.89**	2.18
Asset Index	754.34***	131.77			1972.12***	155.36
Time effect (2006)	125.86	175.98				
Time effect (2007)	498.90**	206.48				
Individual fixed effects	No		Yes		Yes	
Sub-district fixed effects	Yes		No		No	
Constant	-887.31*	518.39	5234.41***	178.71	2416.04	1478.51
Adjusted R^2	0.21		0.57			
Wald test (bootstrapped)	1033.32***				2205.58***	

***P<0.01, **P<0.05, and *P<0.10. Variance-covariance matrices are clustering at sub-district level for Model (1) and (2), and are bootstrapped for Model (1) and (3). Model (2) uses attrition weights. We omit the top and the bottom one percent of the sample used, yielding 11,634 observations used.

Appendix D: Determinants of occupational shifts

We estimate multinomial logit models to predict occupational transition probabilities and then use those predicted values in place of observed transitions in estimating equation (1). The multinomial logit is a form of random utility model, based on the premise that an individual compares her expected utility under different occupations and her constrained conditions:

$$U_t^e = U(O_{tj}, \ln y_{t-1}, Z, \Delta V \mid O_{t-1,j}) \tag{D1}$$

where O_{tj} is an indicator variable for occupation in period t and j = 0, 1, 2, 3 indicate farming occupation, non-farm self-employment, non-farm employee, and non-farm employer, respectively,. The variable ln y_{t-1} represents the log earnings reported in the previous year, indicating whether individuals might consider a change in occupation according to their past income draw in last period's occupation, $O_{t-1,j}$. Z denotes observed individual and household characteristics, just as in equation (1).

In these multinomial logit regressions we include ΔV in order to capture changes in community variables such as infrastructure and agricultural circumstances that reflect evolving environmental conditions. Table D1 enumerates these community level variables from the rural community census survey (NRD) and describes the changes in these variables for the 363 villages in the Thai SES data that match with the NRD data. Recall that the NRD data do not include 2006 and 2008. So, we study occupational transitions between 2005 and 2007 and between 2007 and 2010 based in part on changes in community variables between 2005 and 2007 and between 2007 and 2009, respectively, as a pooled panel to increase sample size.

Hence, given individuals' initial occupation, we estimate the multinomial logit:

$$\Pr(y = j \mid m) = \frac{\exp(X\Gamma(j))}{1 + \sum_{j=1}^{3} \exp(X\Gamma(j))} \tag{D2}$$

where $X\Gamma(j)_i = \gamma_0 + \gamma_1 y_{t-1,i} + Z_i\gamma_2 + \Delta V_c\gamma_3 + \nu_i$. To ensure model identification, $\Gamma(j)_i$is set to zero when individuals stay in their previous occupation, choice *m*. In particular, if the base case is 0 = staying as a farmer, then j = 1, 2, 3 refer to the shift from farm to non-farm self-employed, the shift from farm to non-farm employee, and the shift from farm to non-farm employer, respectively. The coefficients are then interpreted with respect to staying in one's initial occupation, the base category.

The multinomial logit models are estimated separately for each initial occupation. The small number of observations of non-farm employers precludes estimation from that base position. Hence, the average marginal effects reported in Tables D2-D4 reflect only the transitions from farming, non-farm self-employment, and non-farm employee positions, respectively. The changes in village level variables jointly statistically significantly determine the transition probabilities.

As one would expect, lower initial log earnings are strongly associated with occupational shifts. People leave lower-paying jobs in search of better-paying ones. Women exhibit lower occupational mobility than men do. The main pattern we see in transitions from farming into non-farm activities is that household asset endowments are strongly and statistically significantly positively associated with transitions into non-farm employer status but negatively associated with transitions into the farming and non-farm worker categories. Those with capital are clearly more likely to take their chances starting their own enterprise. On the other hand, owned agricultural land is positively correlated with transitions into farming, confirming that land is the main factor driving work in agriculture.

Village-level variables also matter considerably. In particular, individuals living in areas lacking sufficient water for agriculture are more likely to shift from farm into non-farm employer

sector. More year-round trafficable roads in the area fosters more shifts into rather than from farming. If road conditions are consistently reliable, farmers have less difficulty selling their products and hence remain on farm. Likewise, as access to electricity provided by the state increases to cover all households in a village, transitions from farm to non-farm self-employment become less likely, reflecting the importance of services to support economically viable farming.

Among those initially in the non-farm sector, fewer initial asset holdings increases the likelihood of moving into farming in the following year. Greater farm profitability in the area increases the likelihood of becoming a farmer and non-farm employer from self-employment, but increases the likelihood of becoming a non-farm worker. This reflects the linkage of agricultural prosperity to non-farm business growth.

Transportation infrastructure and employment conditions also play an important role in occupational transitions. Reduced travel time to the nearest district also induces a shift from non-farm worker to farming. More households working in establishments or manufacturing in the village also encourages transition into entrepreneurship, likely reflecting the expanded market for one's goods or services.

Table D1: Summary statistics of changes in village characteristics

Variable	Mean	Std. Dev.	Description
Village characteristics (changes)			Changes of village-level variables between 2005 and 2007 from NRD
All HH w/ electricity	0.09	0.45	*Change in access to electricity for all households:* -1 = all households had electricity in 2005, but not all households had electricity in 2007 (worse off), 0 = same condition, 1 = better off
Insuff water for agri. (100 Rai; 1Rai = 1600 m^2)	-0.66	9.39	*Change in areas having insufficient water for agriculture:* If increasing, there were more areas in the village reporting insufficient water for agriculture.
Travel time	0.15	12.85	*Change in travel time to the nearest district (minutes):* If decreasing, it exhibits more efficient travel time from the village to the nearest district by vehicle.
Year-round well usable road	0.03	0.58	*Change in main road being usable well throughout the year:* -1 = main road was used well in 2005, but not in 2007 (worse off), 0 = same condition, 1 = better off Being usable well also implies that there is no effect from severe flooding during monsoon/rainy season, not only usable in dry season
Ratio concrete of convenient route	0.01	0.13	*Change in ratio of asphalt/concrete sections to total of the most convenient route to the nearest major district (km):* Of the total length of the most convenient route to the nearest district, more asphalt/concrete length would imply better developed infrastructure
Agriculture loss	-0.04	0.54	*Change in land utilization problem of growing plants not meeting breakeven point:* -1 = no problem in 2007, but had this problem in 2005 (better off), 0 = same condition, 1 = had problem in 2007, but no problem in 2005 (worse off)
Stagnant flood	-0.01	0.45	*Change in Stagnant flood problem:* -1 = no problem in 2007, but had this problem in 2005 (better off), 0 = same condition, 1 = had problem in 2007, but no problem in 2005 (worse off)
Soil problems	-0.08	2.13	*Change in number of soil problem*: If increasing, there were more list of soil problems in the village
Ratio HH work in establishment	-0.001	0.18	*Change in ratio of HH working in businesses/services in the village*: If increasing, there were more HH work in that sector
Ratio HH work in manufacturing	-0.01	0.09	*Change in ratio of HH working in factories/manufacturing in the village*
Ratio HH member in employment	0.01	0.27	*Change in ratio of HH whose member working as employee in the village*
Ratio HH work in agriculture	-0.001	0.36	*Change in ratio of HH working in farming*

Note: All changes in village characteristics are calculated from the NRD data set.

Table D2: Multinomial logit estimation of determinants of transitions from farm work

	Farm to NF self-employed		Farm to NF worker		Farm to NF employer	
	AME	Std. Err.	AME	Std. Err.	AME	Std. Err.
Log(earnings t-1)	-0.0120***	0.0037	-0.0234**	0.0101	0.0009	0.0006
Age	0.0057*	0.0033	-0.0134*	0.0073	0.0050***	0.0017
Age^2	-0.0000	0.0000	0.0001	0.0001	-0.0001***	0.0000
HH head	0.0030	0.0238	0.0032	0.0312	0.0040**	0.0019
Married	-0.0503	0.0315	-0.0240	0.0229	-0.0074**	0.0034
Female	0.0269	0.0178	-0.0019	0.0196	-0.0083**	0.0034
Education						
Less than primary school/none (base)						
Primary school	0.0442	0.0308	-0.0537*	0.0320	-0.0011	0.0023
Secondary school	0.0750**	0.0335	-0.0529	0.0420	-0.0275***	0.0097
High/Vocational school	0.0300	0.0385	0.0069	0.0389	0.0006	0.0032
College and above	0.1176**	0.0588	0.0333	0.0605	-0.0231**	0.0099
Total working months	-0.0001**	0.0001	-0.0002*	0.0001	-0.0000	0.0000
Family size	0.0030	0.0041	0.0123*	0.0068	-0.0009	0.0007
Owned agricultural land	-0.0003***	0.0001	-0.0006*	0.0003	-0.00003**	0.0000
Asset Index	0.0435***	0.0096	-0.0402*	0.0244	0.0038***	0.0015
Change in village characteristics						
Insuff water for agri.	0.0002	0.0007	0.0005	0.0005	0.0001*	0.0001
All HH w/ electricity	-0.0344***	0.0125	0.0135	0.0170	-0.0017	0.0013
Year-round well usable road	0.0134	0.0114	-0.0004	0.0155	-0.0047**	0.0020
Ratio concrete of convenient route	0.0167	0.0469	-0.0672	0.0829	0.0005	0.0033
Travel time	-0.0000	0.0005	-0.0005	0.0008	-0.0000	0.0000
Agriculture loss	0.0125	0.0135	0.0062	0.0179	-0.0017	0.0011
Stagnant flood	-0.0003	0.0186	0.0115	0.0201	-0.0010	0.0013
Soil problems	-0.0030	0.0039	0.0069	0.0056	0.0001	0.0003
Ratio HH work in establishment	0.0358	0.0434	-0.0076	0.0555	-0.0094**	0.0038
Ratio HH work in manufacturing	-0.0112	0.0886	0.0065	0.0940	0.0059	0.0055
Ratio HH member in employment	-0.0165	0.0245	0.1217*	0.0636	0.0027	0.0041
Ratio HH work in agriculture	0.0005	0.0196	0.0316	0.0256	-0.0024	0.0020
Time dummy	0.0188	0.0187	0.0147	0.0261	0.0063**	0.0030
Number of observations	2773					
Pseudo R^2	0.1348					

***P<0.01, **P<0.05, and *P<0.10. We omit the top and the bottom one percent of the sample used. Inverse probability weight to correct attrition bias has been applied in the estimation. Variance-covariance matrices are clustered at village level. AME = average marginal effects.

Table D3: Multinomial logit estimation of determinants of transitions from non-farm self-employment

	NF self-employed to farm		NF self-employed to NF worker		NF self-employed to NF employer	
	AME	Std. Err.	AME	Std. Err.	AME	Std. Err.
Log(earnings t-1)	-0.0271***	0.0084	-0.0016	0.0080	0.0066*	0.0036
Age	0.0115*	0.0066	-0.0008	0.0071	0.0070***	0.0022
Age^2	-0.0001	0.0001	-0.0001	0.0001	-0.0001***	0.0000
HH head	-0.0124	0.0441	0.0035	0.0215	0.0180**	0.0077
Married	-0.0147	0.0551	-0.0416*	0.0250	-0.0159*	0.0095
Female	0.0017	0.0316	-0.0441**	0.0222	-0.0140*	0.0080
Education						
Less than primary school/none (base)						
Primary school	0.0737	0.0538	0.0209	0.0408	-0.0110	0.0111
Secondary school	0.0710	0.0677	-0.0158	0.0482	-0.0253*	0.0148
High/Vocational school	0.0563	0.0853	-0.0356	0.0510	-0.0021	0.0126
College and above	0.0860	0.1146	-0.0586	0.0703	0.0032	0.0213
Total working months	-0.0002*	0.0001	-0.0002	0.0001	0.0000	0.0000
Family size	0.0400***	0.0091	-0.0050	0.0059	0.0046**	0.0021
Owned agricultural land	0.0017***	0.0003	-0.0005	0.0003	-0.0000	0.0001
Asset Index	-0.0861***	0.0206	-0.0195*	0.0117	-0.0019	0.0044
Change in village characteristics						
Insuff water for agri.	-0.0017	0.0019	0.0010	0.0014	-0.0001	0.0002
All HH w/ electricity	-0.0183	0.0316	-0.0195	0.0234	0.0035	0.0067
Year-round well usable road	-0.0150	0.0226	0.0163	0.0149	0.0031	0.0059
Ratio concrete of convenient route	-0.1476	0.1009	0.0611	0.0535	0.0176	0.0191
Travel time	0.0002	0.0016	0.0010	0.0008	0.0001	0.0002
Agriculture loss	-0.0507*	0.0289	0.0377**	0.0158	-0.0126	0.0078
Stagnant flood	0.0348	0.0284	0.0046	0.0199	0.0044	0.0069
Soil problems	0.0035	0.0077	-0.0095**	0.0045	0.0017	0.0014
Ratio HH work in establishment	-0.1254	0.0821	0.0085	0.0380	-0.0159	0.0160
Ratio HH work in manufacturing	0.0303	0.1392	-0.1017	0.0982	0.0828***	0.0300
Ratio HH member in employment	-0.0518	0.0744	0.0934***	0.0322	-0.0091	0.0125
Ratio HH work in agriculture	-0.0030	0.0382	0.0223	0.0287	0.0038	0.0055
Time dummy	0.0507	0.0313	0.0088	0.0170	-0.0023	0.0065
Number of observations	1174					
Pseudo R^2	0.2154					

***P<0.01, **P<0.05, and *P<0.10. We omit the top and the bottom one percent of the sample used. Inverse probability weight to correct attrition bias has been applied in the estimation. Variance-covariance matrices are clustered at village level. AME = average marginal effects.

Table D4: Multinomial logit estimation of determinants of transitions from non-farm employee

	NF worker to farm		NF worker to NF self-employed		NF worker to NF employer	
	AME	Std. Err.	AME	Std. Err.	AME	Std. Err.
Log(earnings t-1)	-0.0776***	0.0141	-0.0275***	0.0097	-0.0021	0.0018
Age	0.0038	0.0052	0.0005	0.0040	0.0006	0.0011
Age^2	-0.0000	0.0001	-0.0000	0.0000	-0.0000	0.0000
HH head	-0.0127	0.0199	0.0328**	0.0140	0.0027	0.0043
Married	0.0279	0.0203	0.0154	0.0200	0.0035	0.0044
Female	-0.0149	0.0231	0.0193	0.0140	-0.0102*	0.0055
Education						
Less than primary school/none (base)						
Primary school	0.0948**	0.0459	0.0083	0.0486	0.0798***	0.0215
Secondary school	0.1407***	0.0510	-0.0100	0.0505	0.0758***	0.0209
High/Vocational school	0.0207	0.0547	-0.0182	0.0530	0.0828***	0.0220
College and above	0.0007	0.0721	-0.0773	0.0587	0.0036	0.0086
Total working months	-0.0001	0.0001	-0.0000	0.0001	-0.0000	0.0000
Family size	-0.0016	0.0066	0.0081	0.0056	0.0010	0.0009
Owned agricultural land	0.0005	0.0005	-0.0003*	0.0002	-0.0000	0.0000
Asset Index	-0.0643***	0.0134	0.0228**	0.0091	0.0032**	0.0016
Change in village characteristics						
Insuff water for agri.	0.0007	0.0016	-0.0016**	0.0007	-0.0000	0.0001
All HH w/ electricity	0.0356	0.0255	0.0088	0.0129	-0.0017	0.0024
Year-round well usable road	-0.0059	0.0185	0.0139	0.0135	0.0003	0.0027
Ratio concrete of convenient route	0.0035	0.0594	0.0127	0.0384	-0.0009	0.0096
Travel time	-0.0014***	0.0005	0.0000	0.0002	0.0000	0.0001
Agriculture loss	-0.0016	0.0173	-0.0170	0.0137	0.0012	0.0022
Stagnant flood	-0.0058	0.0214	0.0226**	0.0097	0.0010	0.0034
Soil problems	0.0007	0.0049	0.0034	0.0026	-0.0006	0.0009
Ratio HH work in establishment	-0.0820*	0.0480	0.0243	0.0256	0.0120**	0.0058
Ratio HH work in manufacturing	0.5429***	0.1445	0.0218	0.0507	-0.0160	0.0099
Ratio HH member in employment	-0.0128	0.0338	-0.0134	0.0261	-0.0110***	0.0038
Ratio HH work in agriculture	0.0675***	0.0217	0.0114	0.0143	-0.0051*	0.0029
Time dummy	0.0807***	0.0172	-0.0158	0.0159	0.0022	0.0027
Number of observations	2417					
Pseudo R^2	0.1814					

***P<0.01, **P<0.05, and *P<0.10. We omit the top and the bottom one percent of the sample used. Inverse probability weight to correct attrition bias has been applied in the estimation. Variance-covariance matrices are clustered at village level. AME = average marginal effects.

Appendix E: Selection bias corrections based on the multinomial logit model

Let the changes in log earnings in the s[th] alternative is given by

$$\Delta \ln y_s = \alpha_s + Z_s\beta_s + u_s \tag{E1}$$

where Z contains the same covariates as in equation (1), including log earnings in the previous period. Let $\Delta \ln y_s$ be observed only if alternative s – occupational shift to s given previous occupation – is chosen among four alternatives. Following the Heckman selection model, Dubin and McFadden (1984) and Bourguignon et al. (2007) include multiple correction terms to control for self-selection in the s[th] alternative instead of only an inverse Mills ratio term for self-selection correction. Hence, equation (E1) becomes

$$\Delta \ln y_s = \alpha_s + Z_s\beta_s + h(P_0,...,P_3) + e_s, \tag{E2}$$

where P_j is the probability that alternative j, j = 0, 1, 2, 3, will be chosen. P_j follows the multinomial logit model in equation (D2). Dubin and McFadden's assume that

$E(u_s) = \sigma_s \frac{\sqrt{6}}{\pi} \sum_{j \neq s} r_j \left[\eta_j - E(\eta_j)\right]$, where η_j is a disturbance term from equation (D1), conditional on the alternative s being chosen. Hence, (E2) becomes

$$\Delta \ln y_s = \alpha_s + Z_s\beta_s + \sigma_s \frac{\sqrt{6}}{\pi}\left[\sum_{j \neq s} r_j \left(\frac{P_j \ln(P_j)}{1-P_j}\right) + r_s \ln(P_s)\right] + e_s \tag{E3}$$

r_j is the correlation coefficient between disturbances u_s and η_j, and e_s is a residual with asymptotic mean zero.

Equation (E3) is estimated in two steps. The multinomial logit model for each of occupational shifts given previous occupation as in (D2) is first estimated and the predicted probabilities are substituted into the selectivity correction terms. Then, the predicted log earnings

changes of each occupational shift in the second stage are averaged and reported in Table 10. Tables E1 - E3 show estimation results in the second stage of each occupation transition.

Table E1: Second stage selection bias correction estimations of transitions from farm

	Remain in Farming (F)		F to NF Self-Employment (SE)		F to Non-Farm Worker (W)	
	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
Log(earnings$_{t-1}$)	-0.74***	0.03	-0.46**	0.19	-0.88***	0.05
Log(earning (t-1))*t_07	0.02	0.04	-0.30	0.19	-0.08	0.06
Age	0.04**	0.02	0.10	0.12	-0.02	0.03
Age^2	-0.001***	0.0002	-0.00	0.00	0.00	0.00
HH head	0.11*	0.06	-0.29	0.37	-0.02	0.10
Married	-0.06	0.10	0.10	0.69	0.01	0.15
Female	0.04	0.06	-0.23	0.37	-0.24**	0.10
Education						
Less than primary school/none (base)						
Primary school	0.04	0.11	-0.03	0.89	-0.24	0.15
Secondary school	0.18	0.15	0.58	1.02	-0.27	0.22
High/Vocational school	0.20	0.14	-0.19	1.10	-0.06	0.19
College and above	0.53	0.35	-0.06	1.66	0.25	0.47
Total working months	-0.00	0.00	0.00	0.00	0.001*	0.0005
Family size	-0.03	0.02	-0.14	0.12	0.05*	0.03
Owned agricultural land	0.00	0.00	-0.00	0.00	-0.00	0.00
Asset Index	0.34***	0.06	0.65*	0.34	0.09	0.09
Time dummy for 2007	0.18	0.29	1.91	1.45	0.62	0.44
Correction terms						
P(stay F)	0.87*	0.51	-3.27	2.85	-0.65	0.43
P(F to SE)	0.46	0.69	2.46	3.99	-0.69	0.99
P(F to Wkr)	-0.75	0.64	1.20	3.49	0.71	0.69
P(F to E)	-0.94	0.70	-5.76	6.73	0.87	1.10
Constant	5.15***	0.60	2.54	3.75	7.81***	0.84
Adjusted R^2	0.37		0.24		0.85	
N. of observations	2212		210		342	

***P<0.01, **P<0.05, and *P<0.10. P(.) is the coefficient of correction terms of $P_j lnP_j/(1-P_j)$, j=1,2,3 and lnP_0. Estimation of transition from farm to NF employer is omitted since there are only nine observations.

Table E2: Second stage selection bias correction estimations of transitions from non-farm self-employment

	SE to F		Remain SE		SE to W		SE to NF Employer (E)	
	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
Log(earnings t-1)	-0.84***	0.09	-0.82***	0.05	-0.92***	0.07	-0.26	12.63
Log(earning (t-1))*t_07	-0.05	0.10	0.03	0.08	-0.03	0.11	0.44	22.00
Age	0.09	0.09	-0.01	0.03	-0.06	0.05	-0.40	26.51
Age^2	-0.001	0.001	0.00003	0.0003	0.001*	0.001	0.003	0.34
HH head	0.12	0.28	0.09	0.10	0.10	0.22	1.03	6.79
Married	0.46	0.33	0.05	0.10	0.43	0.25	-0.13	46.88
Female	0.01	0.42	0.08	0.12	0.40	0.24	0.14	48.51
Education								
Less than primary school/none (base)								
Primary school	-0.68	0.51	0.04	0.21	-0.04	0.35	-2.62	18.11
Secondary school	-0.62	0.67	-0.12	0.25	0.43	0.37	-3.77	14.07
High/Vocational school	-0.27	0.52	0.08	0.24	0.22	0.42	-3.19	23.71
College and above	-1.38	1.95	0.08	0.32	0.96	0.76	-2.42	8.40
Total working months	0.001	0.001	0.001	0.0005	0.001	0.001	0.01	0.18
Family size	-0.03	0.07	-0.02	0.03	-0.01	0.07	-0.05	8.09
Owned agricultural land	0.00	0.00	0.00	0.00	0.00	0.00	-0.02	0.09
Asset Index	0.53**	0.24	0.32***	0.09	0.45***	0.15	1.21	14.47
Time dummy for 2007	0.46	0.87	-0.33	0.69	0.42	0.94	-4.85	231.88
Correction terms								
P(SE to F)	-0.08	0.28	0.05	0.80	-0.17	1.17	-9.76	125.12
P(stay SE)	2.16	1.54	0.06	0.37	-0.22	1.63	2.62	62.67
P(SE to Wkr)	0.86	1.78	-0.43	0.64	-0.64***	0.24	-3.92	179.15
P(SE to E)	-0.25	1.95	-1.26	0.81	-1.26	1.62	0.61	9.98
Constant	6.57**	2.53	7.11***	0.67	8.24***	1.38	10.50	578.35
Adjusted R^2	0.60		0.47		0.80		0.41	
N. of observations	215		802		128		29	

***P<0.01, **P<0.05, and *P<0.10. P(.) is the coefficient of correction terms of $P_j lnP_j/(1-P_j)$, j=0,2,3 and $\ln P_1$

Table E3: Second stage selection bias correction estimations of transitions from non-farm employee

	W to F		W to SE		Remain W		W to E	
	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
Log(earnings t-1)	-0.68***	0.25	-1.07***	0.31	-0.44***	0.05	-2.06	1.33
Log(earning (t-1))*t_07	0.20	0.23	0.06	0.37	0.001	0.03	-1.51***	0.38
Age	0.03	0.05	0.002	0.07	0.01	0.01	0.05	0.78
Age^2	-0.001	0.001	-0.00	0.00	-0.0001	0.0001	-0.0004	0.01
HH head	0.13	0.24	0.09	0.36	0.02	0.04	2.68	1.78
Married	-0.18	0.21	0.36	0.31	0.01	0.03	-3.75*	2.23
Female	0.08	0.20	0.11	0.34	-0.09***	0.03	1.09	1.43
Education								
Less than primary school/none (base)								
Primary school	-0.34	0.37	0.15	0.76	-0.03	0.07	1.03	2.09
Secondary school	-0.42	0.42	0.21	0.80	0.16	0.07	0.00	1.10
High/Vocational school	0.10	0.45	0.58	0.81	0.25***	0.07	-0.96	1.93
College and above	-0.26	0.48	1.37	1.18	0.44***	0.08	omitted	
Total working months	0.00	0.00	0.00	0.001	0.0006	0.0002	0.01	0.02
Family size	-0.05	0.05	0.02	0.07	-0.02**	0.01	0.30	0.61
Owned agricultural land	0.00	0.00	0.004	0.005	-0.0003	0.0004	0.003	0.04
Asset Index	0.33*	0.18	0.02	0.30	0.08***	0.02	0.89***	1.27
Time dummy for 2007	-1.72	1.96	-0.34	3.26	0.02	0.30	11.14***	2.76
Correction terms								
P(Wkr to F)	-0.08	0.29	0.28	3.02	0.25	0.41	-7.72	3.72
P(Wkr to SE)	1.09	1.70	-0.20	0.41	0.08	0.38	14.91***	4.03
P(stay Wkr)	1.68	1.25	-2.29	3.91	0.21	0.25	-21.09***	2.96
P(Wkr to E)	1.07	2.50	0.44	3.41	0.17	0.60	-1.08	0.82
Constant	7.03***	1.88	6.83*	3.50	3.71***	0.40	3.86	16.56
Adjusted R^2	0.14		0.22		0.22		0.80	
N. of observations	409		153		1834		21	

***P<0.01, **P<0.05, and *P<0.10. P(.) is the coefficient of correction terms of $P_j \ln P_j/(1-P_j)$, j=0,1,3 and $\ln P_2$