Barriers to Mobility or Sorting? Sources and Aggregate Implications of Income Gaps across Sectors and Locations in Indonesia

José Pulido* and Tomasz Święcki†

October 2018

Abstract

Existence of large income gaps between agricultural and non-agricultural workers in developing countries is well known, but the exact source of the gaps is debated. The two main hypotheses, barriers to labor mobility and sorting of workers based on unobserved comparative advantage, have distinct predictions for aggregate efficiency but are difficult to distinguish using only cross-sectional data typically available for developing countries. We use panel data from Indonesia Family Life Survey to document that workers who move out of agriculture see an income gain of over 20% while those who move into agriculture see a similar income loss, even if they stay in the same location. We then ask whether even such within-worker sector premia tell us anything about the presence of barriers to sectoral mobility. By themselves, they do not. However, taking into account a richer set of moments of the joint sector-income distribution over time allows us to identify the role of self-selection across sectors and of barriers to sectoral mobility. Our estimates indicate that while self-selection is important, there are also barriers that significantly misallocate workers across sectors. Removing such barriers would lead 35% of workers to reallocate and as a result would increase aggregate output by as much as 21%.

*Banco de la República - Colombia. E-mail: jpulidpe@banrep.gov.co

†Corresponding author, Vancouver School of Economics, University of British Columbia. Email: Tomasz.Swiecki@ubc.ca. We acknowledge financial support from Social Sciences and Humanities Research Council of Canada.

1 Introduction

Large and persistent gaps in average incomes of agricultural and non-agricultural workers in developing countries have been well documented. What exactly accounts for these gaps is still debated, however. A common view is that the gaps exist because workers cannot arbitrage them away due to broadly understood barriers to mobility across sectors and locations. Such barriers would suggest that labor is inefficiently allocated. An opposing view, recently gaining influence, is that the gaps simply reflect an efficient sorting of heterogeneous workers based on their observable and unobservable characteristics. The goal of this paper is to evaluate what the observed income gaps in Indonesia can tell us about the presence of mobility barriers and sorting and to quantify the aggregate losses from any uncovered worker misallocation.

The relative plight of agricultural workers compared to non-agricultural workers in developing countries is at first sight staggering. For example, wage workers outside of agriculture earn 80% more than workers in agriculture in a median of 13 countries studied by Herrendorf and Schoellman (2018). Similarly large gaps emerge when workers are split according to their place of residence rather than the sector of their occupation. For example, urban households in Vietnam in 1998 had real per capita consumption twice as high as rural households (Nguyen et al. (2007)). Such gaps could in principle merely reflect differences in the composition of workforce. For example, to the extent that urban workers are typically better educated than rural workers, the differences in their average wages could simply be picking up the return to additional human capital of urban workers. But studies for various developing countries find evidence that substantial rural-urban gaps remain after factoring out the effect due to differences in schooling achievement and other observable individual characteristics (see, e.g., Hnatkovska and Lahiri (2016) for India and Qu and Zhao (2008) for China). These residual gaps have been documented for wages, broader measures of income, expenditure, and consumption. Similar gaps have also been identified in a parallel literature using value added data to compare productivity of workers across sectors. Gollin, Lagakos and Waugh (2014) show using a wide sample of countries that workers in non-agriculture are twice as productive as workers in agriculture, after taking into consideration the differences in hours worked, schooling and quality of schooling between rural and urban areas.[1] Given the ubiquity of large residual gaps estimated using different countries, measures, and methodologies, they appear to be a real phenomenon rather than merely a measurement artifact.

It is therefore a puzzle why such gaps persist. Why do workers not switch to sectors and locations offering higher income to workers with their observable characteristics, eroding the premia? There are two main hypotheses in the literature. The first one is that the gaps are a manifestation of barriers to mobility. To the extent that these barriers are at least partially induced by policies, this view implies that labor is misallocated. Given the magnitude of the gaps, there are potentially

[1] Herrendorf and Schoellman (2015) caution that such gaps might overestimate true productivity differences due to potential measurement problems in agricultural value added.

large aggregate efficiency gains from mitigating the mobility frictions. In this spirit, Restuccia, Yang and Zhu (2008) calculate that distortions to the allocation of labor between agriculture and non-agriculture play an important role in explaining cross-country income differences.

An alternative explanation for the residual income gaps is that it is a result of sorting of workers across sectors or locations based on characteristics known to them but not observed by researchers. For example, a positive urban premium can be observed if workers choosing urban locations have on average more unobservable skills than rural workers conditional on their education attainment. This mechanism is the explanation of the urban premium recently proposed by Young (2013), who builds on an adaptation of the Roy (1951) model by Lagakos and Waugh (2013). Importantly, in this view residual gaps across sectors or locations can exist despite the allocation of labor being efficient.

Given the different implications of the two canonical explanations of income gaps for allocative efficiency, it is important to know which view is a better description of reality. A major shortcoming of the existing literature accounting for income gaps is that it relies on cross-sectional data. But as is well-known following Heckman and Honoré (1990), the estimation of selection models using only cross-sectional data faces identification challenges. Existing studies therefore need to rely on functional form assumptions (Bryan and Morten (2018)) or on indirect ways of detecting sorting (Young (2013)). In this paper, we argue that augmenting a standard model of sorting by including barriers to sectoral mobility requires longitudinal data to identify the parameters of interest, even when imposing functional forms assumptions. We exploit the panel dimension of a dataset collected in Indonesia, to provide more direct evidence of the extent barriers to sectoral mobility in a context of self-selection.

The Indonesia Family Life Survey (IFLS, Strauss, Witoelar and Sikoki (2016)) we use is uniquely well fitted for our goals. First, it is a longitudinal survey spanning a relatively long period of time, with five waves of the survey conducted between 1993 and 2014. Second, a feature of the survey design and implementation is that it exerts particular effort to track households and individuals even if they migrate, a critical feature for a country undergoing a process of urbanization. Third, IFLS records a rich set of socio-economic information on surveyed individuals. Fourth, with about 20000 surveyed individuals it is a large survey representative of more than 80% of the Indonesian population. Fifth, with roughly 35%/65% split between agricultural and non-agricultural workforce Indonesia is a relevant setting to investigate the gaps across boundaries traditionally used for developing countries. Finally, being the forth most populous country in the world Indonesia is an important country to study in its own right.

We begin our analysis in the next section by documenting some robust features of the Indonesian data. Just like in other developing countries, the ILFS data shows the existence of a large income gap across sectors in Indonesia. Controlling for observable worker characteristics, workers outside of agriculture earn 78% more than workers in agriculture. This is the non-agriculture premium we want to understand better.

Importantly, this premium already conditions on the rural vs. urban location. Much of the literature tends to associate rural employment with agriculture and urban employment with non-agriculture. This implicit isomorphism might lead to an intuition that non-agricultural premium is to be expected even in the absence of sorting or frictions because it compensates workers for the real cost of rural-to-urban migration. We find that logic to be misguided. In Indonesia 47% of rural workforce has primary employment outside of agriculture and 10% of urban workforce is employed primarily in agriculture, so we can meaningfully separate the non-agricultural and urban premia. In this paper we emphasize the sectoral dimension more because most of the rural-urban residual income gap can be accounted for by differences in sectoral composition of rural and urban areas combined with the large non-agriculture premium. The direct urban premium estimated at 23% in the cross-section of workers, while not trivial, is substantially smaller than the 78% non-agriculture premium.

Moving beyond these cross-sectional premia, we exploit the panel structure of our data by relying on within-worker variation in income across sectors and locations. This approach follows Katz and Summers (1989) and the subsequent long tradition of estimating inter-industry wage differentials in developed countries. In Indonesia, it reduces the residual gaps roughly by half, to 39% for non-agriculture and 9% for urban locations. Digging even deeper, we compare the income growth of workers moving out of agriculture relative to those staying in agriculture, and of workers moving out of non-agriculture relative to those who stay employed in non-agriculture. Because we also have detailed migration data, we can do this calculation even conditional on staying in the same very narrowly defined geographical areas (village level). Perhaps our most surprising finding is that the non-agriculture premium exists even within such local markets and that it is approximately symmetric for switches in both directions. Workers who move out of agriculture see an income gain of 22% while those who move into agriculture see a loss of 23%, even if they stay in the same village. The reported premia are robust to a host of concerns about sample selection, estimation method, and measurement issues.

The fact that half of the non-agriculture premium disappears after controlling for time-invariant unobserved heterogeneity informally suggests that sorting does indeed occur and is important. The question is if the 22% average excess income gain received by a worker who switches from agriculture to non-agriculture can be reconciled with an efficient sorting based on comparative advantage alone. In principle, it can. This is because the industry premia, even estimated using within-worker variation, have by themselves little empirical content. We show this by extending a standard model of self-selection based on both permanent and transitory components of comparative advantage to include different types of barriers to sectoral mobility. In particular, we consider utility costs of switching sectors (Dixit and Rob (1994); Cameron, Chaudhuri and McLaren (2007); Artuç, Chaudhuri and McLaren (2010); Dix-Carneiro (2014)) and frictions preventing individuals from working in their preferred sectors (akin to search costs as in Taber and Vejlin (2016)). We demonstrate that the same cross-sectional and within-worker non-agriculture premia can be rationalized by different

combinations of comparative advantage shock processes and barriers to mobility. In particular, by picking the right covariance matrix for the transitory component of comparative advantage we can generate large within-worker premia in the absence of any barriers. Similarly, we can have large barriers to mobility despite observing zero non-agricultural premia. The premia alone cannot tell us if there is any worker misallocation or not.

The comparative advantage process and barriers to mobility can be separately identified once we impose some parametric structure and exploit a richer set of moments of the joint sector-income distribution over time. We use indirect inference (Gourieroux, Monfort and Renault, 1993) for the structural estimation of our model, where the selected auxiliary models are the main reduced-form regressions that characterize the data features we are interested in and that allow us to identify the full set of structural parameters, including the mobility barriers.

Our findings suggest that while self-selection clearly occurs, both types of barriers - utility switching costs and inability to select the preferred sector - significantly improve the overall fit of the model compared to the frictionless specification. They are both able to qualitatively match simultaneously the sectoral premia and the patterns of the moments of the joint distribution of income. For the switching costs specification, we estimate opposite signs for the switching costs away from and towards agriculture. This pattern is observationally similar to receiving a positive compensating differential for working in agriculture. If we assume that the choice of a sector is always voluntary (but switching is costly), then the model uses utility compensation for moving to agriculture to rationalize why so many workers make the move despite taking an income cut.

A considerably better fit to the data, however, is offered by the model which recognizes that not all sectoral transitions are voluntary. In fact, our central estimate implies that half of the transitions between non-agriculture and agriculture we see happen for random reasons (these can be interpreted as life events forcing an individual to switch the sector of employment) rather than in response to shocks to the comparative advantage. Once a worker lands in her sub-optimal sector, moving to the preferred sector is difficult. Given its superior empirical performance, our preferred model relies on this type of mobility friction.

The barriers to mobility are quantitatively important. To make this point, we conduct a counterfactual exercise in our baseline model in which the frictions are removed entirely. This thought experiment is standard in the misallocation literature, though of course extreme because we do not know how the frictions could be completely eliminated in practice. With that caveat in mind, removing all barriers to intersectoral mobility would result in a large reallocation of workers. Overall, 35% of workforce would be in a different sector than in the baseline equilibrium. Since the initially misallocated workers reap large income gains from the reallocation (their income doubles on average), the adjustment has a sizable effect on aggregate output, raising it by 21.5%. Agricultural employment contracts by 8 p.p., but output and productivity increase by double digits in both sectors.

Among the large literature on the income gap between agriculture and non-agriculture, the

most closely related work consists of a handful of papers that also exploit individual-level panel information for developing countries. Beegle, De Weerdt and Dercon (2011) offer early evidence of large within-individual gains in Kenia, but their focus is on consumption gains from migration rather than the more puzzling income gains from sector switching conditional on not migrating. Perhaps the closest, in concurrent work Hicks et al. (2017) also use the IFLS and find smaller within-individual non-agricultural premium. As we explain in section 3.2.5, the divergence in our findings stems from differences in the data selection and focusing on different measures of interest. More importantly, we argue that the non-agricultural premium by itself is not necessarily an informative statistic, and we estimate a structural model that allows us to quantitatively evaluate the importance of barriers to sectoral mobility.

While our results are non-experimental and the magnitudes we report depend on the structural assumptions we make, we believe our key finding of barriers to mobility is also broadly consistent with the limited existing experimental evidence. In a randomized small-scale setting, Bryan, Chowdhury and Mobarak (2014) find substantial gains from inducing workers in Bangladesh to work outside their village, though again the focus is on consumption gains from migration making direct comparison difficult. More closely, Sarvimaki, Uusitalo and Jantti (2018) using a natural experiment in Finland find large income gains for workers who abandoned farming as a result of forced migration.

2 Data

In this section we describe the data, only highlighting the features of the dataset most relevant for our analysis. Comprehensive details about the design and implementation of the IFLS are reported in Strauss, Witoelar and Sikoki (2016).

Our primary source of data is the Indonesia Family Life Survey. The first IFLS was conducted in 1993, with subsequent waves in 1997, 2000, 2007, and 2014. From the outset the IFLS was designed as a long-term panel survey, which allows us to compare life trajectories of individuals making different occupational and locational choices. Furthermore, the IFLS puts considerable effort into tracking individuals over time. This feature is rare among longitudinal household surveys in developing countries, which typically lose respondents who move out of an original survey area. As a measure of tracking success, Thomas et al. (2012) report that the 2007 IFLS managed to interview 87% of individuals who were eligible to be tracked. Tracking movers is crucial for drawing conclusions from a comparison of migrants and stayers when the decision to migrate is not random.

The IFLS is a large-scale survey, conducted in 13 of the 27 Indonesian provinces. Because the ones excluded are mostly outlying provinces, the sample is representative of 83% of Indonesian population. The first wave interviewed 22019 individuals and the number of respondents grew to 58337 in the fifth wave. In our analysis we restrict attention to adults (15 years or older) who are employed and therefore answer the detailed work module of the survey. The definition of employed

is expansive and comprises all persons who answered affirmatively to any of the following categories: i) their primary activity during the past week was working, trying to work or helping to earn income; ii) had worked for pay at least 1 hour during the past week ; iii) had a job or business, but were temporarily not working during the past week; iv) had worked at a family-owned (farm or non-farm) business during the past week.

For those individuals, the dataset we construct records their annual income, the sector where they worked according to the job that consumed the most time, years of schooling, work experience by sector and standard demographic characteristics such as age and gender. In addition, we use information on the household location in each survey wave and the movements recorded in the migration module of the survey to construct individual location histories at various levels of administrative detail.

Our main outcome variable of interest is annual income. The annual income can be derived from wages, from net profits of a business (such as a farm), or from other sources such as government transfers. We believe that total income is the appropriate measure in a setting where work on a family farm is pervasive and where half of the workforce does not report any wage work.

Following a standard distinction for developing countries, we split locations according to whether they are rural or urban. The rural-urban status of each survey location is determined by the Indonesian Central Bureau of Statistics (BPS) based on multiple criteria. Along a sectoral dimension, we classify workers as employed either in agriculture or in non-agriculture comprising all other sectors.[2]

Table 1 reports descriptive statistics for the constructed dataset. In our analysis below we focus on the 22829 individuals whom we observe in at least two waves of the survey, for a total of 70586 observations.[3]

3 Income Gaps across Sectors and Locations

3.1 Baseline Results

In this section we present the key patterns of income gaps across sectors and locations in Indonesia. The gaps are estimated using Mincerian regressions with the following general form

$$\ln y_{islt} = X_{it}\beta + D_N + D_U + D_i + \varepsilon_{islt}, \tag{1}$$

where y_{islt} denotes income of an individual i working in sector s (agriculture or non-agriculture), living in location type l (rural or urban) in year t. X_{it} collects standard individual covariates such as sex, years of education, experience and experience squared, as well as year and province

[2]This two-sector partition is common in macro-development literature and is sufficient to illustrate the puzzle of low agricultural incomes. We have also divided non-agriculture further into manufacturing and services. The income gaps between manufacturing and services are small relative to the gaps between those two sectors and agriculture.

[3]Depending on the specification the effective sample size can be smaller as we do not observe all variables for all individuals.

dummies. D_N and D_U capture the non-agriculture and urban premia of interest, while D_i captures the time-invariant component individual heterogeneity.

The baseline specification is a reduced form relationship between income and certain observable and unobservable worker characteristics. If workers switch between sectors randomly, then the D_N premium has a simple interpretation of an average gain that a worker can get by moving from agriculture to non-agriculture. If, on the other hand, workers sort across sectors (and locations) based on their unobserved comparative advantage as in Roy (1951) then the premia estimated using equation (1) need not have a simple interpretation and a structural model is needed for an exhaustive analysis. While our argument in this paper is that sorting is indeed important and we therefore estimate a structural model later, we begin by discussing the reduced form OLS estimates as they have a long tradition and they will be used as auxiliary models in our structural estimation.

As a starting point we estimate equation (1) without any controls except for the sector dummies.[4] This specification simply compares average incomes across sectors and, as can be seen in the first column of Table 2, these incomes vary greatly. Compared to agriculture, incomes in non-agriculture are on average 84 log points [lp] (or 131%) higher.[5][6] The second column compares urban and rural incomes. The urban premium stands at a similarly dramatic 65 lp (or 91%). A natural question is whether the urban and sectoral premia capture the same variation in the data.

Many studies take a dichotomous view of economic activity in developing countries. A classical divide in development literature goes along the rural vs. urban dimension. Macroeconomists tend to work with sectoral data and hence use the agriculture vs. non-agriculture split. But both literatures often implicitly consider both partitions as interchangeable, for example by associating structural transformation (decline of agricultural employment share) with urbanization (increase in urban share). The joint distribution of workers across sectors and locations shown in Table 1 suggests that such interchangeability is too crude in Indonesia. In 2000 (around the middle of our sample period) the share of rural workers at 59% was quite a bit higher than the 37% share of agricultural workers. Among rural workers 45% had primary employment outside of agriculture, while 11% of urban workforce was employed in agriculture.

So can the raw urban premium be explained by different composition of sectors in rural and urban locations or are urban workers paid more in the same sectors? Column 3 of Table 2 estimates the urban and sectoral premia jointly. Controlling for sectors reduces the urban premium almost by half, yet it is still high at 41 lp. Controlling for type of location has a smaller impact on sectoral premia, still at 69 lp. These numbers are the first indication that sector of employment might have a stronger effect on income than place of residence directly.

This point is further strengthened by controlling for individual worker characteristics in the

[4] In all specifications we control for year and province fixed effects. Observations are weighted by their longitudinal survey weights and standard errors are clustered at the level of primary sampling units of the survey.

[5] Because the coefficients of interest are often large in magnitude we report them directly in log points and only occasionally translate them to exact percentage differences.

[6] Reported coefficients are statistically significant at 5% level or lower unless mentioned otherwise.

Mincer regression. Column 4 shows the urban premium of 21 lp and non-agriculture premium of 57 lp. Controlling for observables reduces the urban premium by half once again, while the sectoral premium again changes much less. These residual (controlling for observables) income gaps are also about as much as what can be calculated with cross-sectional data. They therefore correspond most directly to the gaps calculated in most other studies.

Using the panel structure of our data we are in a position to begin addressing the issue of sorting on unobservables. The specification in column 5 adds worker fixed effects to the set of controls. If workers sort themselves into their initial careers based on their unobservable skill but subsequently switch sectors randomly, then cross-sectional non-agriculture premium would suffer from a selction bias but the worker fixed-effects premium could be interpreted as an average gain from switching to agriculture. Using only within-worker variation to identify the gaps reduces the urban premium by more than half to 8 lp. While not trivial, a 9% additional income gain associated with moving from rural to urban location while keeping the same sector of employment is not shocking either. In contrast, the non-agriculture premium is still surprisingly large. The same worker switching from agriculture to non-agriculture without changing the rural-urban status sees on average an additional income gain of 33 lp (or 39%). Column 6 paints a similar picture using sightly more flexible specification with a full set of interactions between sector and urban dummies. Staying in a rural area and switching away from agriculture gives an income boost of 33 lp.

Because the large premia estimated on switchers are the most novel and surprising reduced-from finding of this paper, we now explore the mobility pattern in our data more carefully. The first panel of Table 3 presents the count of wave-to-wave transitions between sectors and the third panel shows the associated transition matrix.[7] About 20% of workers in agriculture transition to non-agriculture between survey waves, and 12% on workers in non-agriculture switch to agriculture. Overall, 24% of workers change the sector at least once while in our sample. The fact that there are almost as many cases of workers moving into agriculture as cases of workers moving out of agriculture is puzzling in light of the large negative premium associated with working in agriculture. The second and fourth panels of Table 3 records analogous transitions along the rural-urban dimension. There is less mobility between rural and urban areas, with a change in location status in 9% of cases. About 17% of workers move between rural and urban locations at least once while in our sample. As expected in a developing country, there are more than twice as many transitions from rural to urban than in the opposite direction, resulting in net migration to urban areas.

The sectoral and urban premia reported so far show an average effect of moving in and out of the sector or location. We now reevaluate the income gaps while taking the direction of transitions into account. The estimating equation now takes the form

$$\Delta \ln y_{islt} = \Delta X_{it}\beta + \Delta D_{ss'} + \Delta D_{ll'} + \Delta \varepsilon_{islt}, \tag{2}$$

[7]These are transitions between two consecutive observations for each worker rather than year-to-year transitions. The time between the waves of the survey varies form two to seven years.

where $\Delta D_{ss'}$ and $\Delta D_{ll'}$ capture the direction of sectoral and locational transition. Results are reported in the first column of Table 4. Along the locational dimension, workers who move from rural to urban areas see an income increase of 9 lp relative to those who stay in rural areas. Workers who move into rural areas have an income shortfall of 16 lp relative to those who stay in urban areas. Results for the non-agriculture premium are once again stronger. Relative to workers who remain in agriculture, workers switching out of agriculture see an additional income growth of 22 lp. Workers who switch from non-agriculture to agriculture see an income loss of 33 lp relative to workers remaining in non-agriculture.

So far we have established existence of a significant income premium for working in non-agriculture controlling for movements between rural and urban locations. But it is still conceivable that movements within rural and urban locations, if correlated with sector switching and having an independent effect on income, might bias the estimates of the sectoral premium. Now we isolate geographic mobility completely using the detailed migration information provided by our dataset. We interact the direction of sectoral transition variable $\Delta D_{ss'}$ with an indicator for whether a worker migrated across village boundary (or correspondingly fine location for cities). The second column of Table 4 displays the results, with workers staying in agriculture and staying within a village as a reference category. Workers who migrate and move out of agriculture have the largest income gains; workers who migrate and move into agriculture suffer the largest relative income losses. But perhaps the most striking results are for workers who do not migrate: those who switch out of agriculture gain additional 20 lp in income relative to those who remain in agriculture. Those switching into agriculture see an income loss of 26 lp relative to non-movers who remain employed in non-agriculture.

That such large non-agricultural premium can be identified from within-worker sector switches within very narrow geographical areas is truly surprising. Moreover, it is not easily reconciled with the workhorse models of labor markets in developing countries. If workers are sorting across sectors according to comparative advantage that is fixed over time and switching is costless and happens because of changes in sectoral prices (as in the standard Roy model), then we should not expect to see a large premium for switchers, and we should expect flows to be in one direction only. If switching is costly and occurs only if the income gain justifies incurring the mobility cost then we should see a positive premium regardless of the direction of the voluntary switch. In contrast, we see workers switching to agriculture taking systematic cuts to their income that is of similar magnitude as gain for workers switching in the opposite direction. Thus it seems that there is a pure premium associated with working in non-agriculture. There are several possible rationalizations for why the premium might exist in an equilibrium. First, workers might demand utility compensation for switching to agriculture. We explore both unrestricted (i.e., possibly negative) switching costs and compensating differentials as a possible rationalization of the observed choices. In the case of compensating differentials, workers simply attach higher non-monetary value to working on a farm than for other jobs. Our concern is that given the harsh realities of farm work in developing

countries this explanation is not quite compelling. For that reason, we also consider a friction that results in some workers switching sectors randomly rather than voluntarily. Finally, as we argue later, the premium can arise even in the frictionless setting when switching happens because of idiosyncratic shocks to the comparative advantage over time. In order to quantify the relevance of these explanations, in the next two sections we develop and estimate a structural model capturing them all. But before we get to the model, we show the robustness of our motivating reduced-form findings.

3.2 Robustness

In this subsection we illustrate that the existence of a non-agriculture premium is robust to a number of concerns about measurement, interpretation and estimation. Our baseline point of reference is the 57 lp cross-sectional premium and 33 lp within-worker premium reported earlier in column 4 and 5 of Table 2.

3.2.1 Job Type

The first exercise incorporates information on a type of job workers engage in as this helps to illuminate the nature of labor markets in Indonesia. Workers in IFLS can be consistently classified into 4 categories: self-employed, private workers, government workers and unpaid family workers. As Table 5 reports, self-employment is the most common work status, accounting for almost half of employment. Private sector workers earning wages and salaries - a category that would usually be the focus in studies based on developed countries - constitutes less that a third of the workforce. Almost 15% of workers who typically help in household work or in a family business or farm are classified as unpaid family workers. These workers nevertheless can report income and are included in the analysis, but our results are robust to dropping this category altogether. The second panel of Table 5 also reports the 10 most common occupations. The point of this table is to show what non-agriculture typically means in Indonesia. It is more about being a self-employed street vendor rather than having a formal factory job in manufacturing.

Controlling for the job type has a small impact on the non-agriculture premium, e.g., reducing it from 33 lp to 29 lp in the worker fixed effects regression. More interestingly, Table 6 reports the results of interacting job type with a direction of switch. For the two main categories, self-employed and private workers, there is about 25 lp premium for switching to non-agriculture relative to staying in agriculture. Workers switching away from non-agriculture suffer a loss of similar magnitude relative to workers remaining in non-agriculture. The similarity of results for self-employed and wage workers can come as a surprise. The non-agriculture premium for wage workers could be in principle rationalized along similar lines as intersectoral or even inter-firm wage differentials documented for developed countries. There might be good non-agricultural jobs that pay more than bad agricultural jobs because employers in non-agriculture for some reason share rents with

their employees. But such rent-sharing explanation would be silent as to why we see a similar premium for self-employed workers switching sectors since they are the residual claimants of their effort. The sectoral premium for the self-employed is thus perhaps our most surprising finding.

Going back to earlier discussion, the non-agriculture premium could reflect compensating differentials, if, e.g, workers value flexible schedule associated with farm work. But the fact that the premium exists for self-employed in both sectors makes compensating differentials less compelling as an explanation. Furthermore, going one step further we can show that the premium of the same magnitude exists even for self-employed workers switching sectors while staying in the same narrow location. We do not find it plausible that workers willingly give up 25% of their income because they prefer to run a farm than a non-farm business in the same village.

3.2.2 Wages and Consumption

While our preferred outcome variable is annual income, there can be concerns about the quality of that self-reported measure. The problem could be particularly stark for self-employed who often have to allocate family business income to individuals. As a robustness check we now restrict attention to annual wage income that is less likely to suffer from measurement problems. Doing so comes at the expense of restricting the sample by more then half to individuals who work for wages in the private or government sector. Table 7 illustrates that the same pattern of premia can be observed using data for wages as for total income, though the magnitudes are a little smaller. Controlling for worker fixed effects, the non-agriculture premium is 23 lp, while the urban premium 11 lp. Despite the sample size being significantly reduced the premia are still precisely estimated.

Since the IFLS records consumption expenditure, it offers an additional way of verifying that working in non-agriculture allows a higher standard of living. One drawback of consumption data in the present context is that it is recorded at a household level, whereas the focus of the paper is on individual decisions. This requires some adjustments to make the results comparable. The first column of Table 8 reports results of a household-level cross-sectional regression of log per capita expenditure (Log PCE) on a continuous variable measuring the share of household income derived from non-agriculture and an urban dummy. Column 4 reports a corresponding calculation for per capita household income. Households that derive higher share of income from non-agriculture have a higher per capita consumption, though the elasticity is not as large as for income. The rest of Table 8 reverts to individual level regressions, but with dependent variables still at the household level. Column 3 results indicate that if a member of a household moves from agriculture to non-agriculture than the average consumption in the household increases by over 7 lp. This might appear as a modest number compared to the baseline income premium so two comments are in order. First, since a survey worker typically accounts for less than 60% of income in his household, the coefficient should be scaled by the inverse of that share to be interpretable as an increase in consumption associated with all household workers switching to non-agriculture. This transformation would increase the

non-agriculture consumption premium to about 13 lp. To illustrate that this transformation is reasonable column 6 performs it on per capita income variable. The transformed coefficient of 35 lp is very close to the baseline non-agricultural premium. Second, in similar specifications the consumption premium is still only 1/3-1/2 as large as income premium. In light of permanent income logic perhaps it should not be surprising that an income shock associated with switching sectors has only partial pass-through to consumption.

3.2.3 Heterogeneity in Mincerian Returns

The baseline regressions control for standard Mincerian determinants of income such as education and experience. The coefficients on these determinants do not vary between sectors and rural/urban locations, however. A recent paper by Herrendorf and Schoellman (2018) argues that this might lead to an overstatement of the residual income gaps, if, e.g, non-agriculture offers higher returns to education and experience. To address this concern, we now allow the Mincerian returns to vary be sector and location. Table 9 reports the associated premia, calculated as the average marginal effects of switching for the population.[8] While some underlying returns do indeed differ by sector, this has no significant effect on the estimated premia of interest.

3.2.4 Additional Jobs and Home Production

Workers are assigned to a sector according to whether their main job is in agriculture or non-agriculture. Correspondingly, the annual income is constructed using the income from the main job. Some workers, however, have more than one job. If having a secondary job is more common for agricultural and rural workers then we might overestimate the non-agriculture and urban premia. Columns 3 and 4 of Table 10 show the premia estimated when instead we take into account income from worker's both primary and secondary jobs. This adjustment reduces the premia by about a fifth.[9]

Another concern is that by focusing on income we are not taking into account home production which is not trivial in developing countries. If agricultural households do not include food produced and consumed in-house in their income then this could lead to an overstatement of the non-agriculture premium. The IFLS data allows us to assess how important this consideration is because it asks households to report the value of goods and services produced for own consumption. The average share of self-produced consumption is about 10%, but is predictably higher in rural areas (13%) than in urban areas (7%). As a robustness check we therefore scale up individual incomes (from both main and secondary job) by the inverse of the share of self-produced consumption in a household the individual belongs too. This effectively increases incomes of workers in rural

[8] Results are similar if we calculate the average marginal effects for switchers instead.

[9] Using a continuous measure of the share of income a worker derives from non-agriculture instead of a dummy for the primary job leads to similar results, with a cross-sectional premium of 47 lp and 26 lp in the specification with worker fixed effects.

and predominantly agricultural households. As columns 5 and 6 report, this has little effect on the estimated premia. Columns 7 and 8 consider an adjustment even more favorable for agriculture - scaling incomes by the inverse of the share of home-produced food in total food consumption. This again does not affect the estimated non-agriculture premium much, though the urban premium becomes insignificant.

3.2.5 Hours Worked

All the results so far show that workers in agriculture have lower annual income than workers in non-agriculture. One natural question is to what degree this income difference is driven by systematic differences in labor supply across sectors. To investigate this issue, Table 11 adds hours worked per year to the set of individual controls. Controlling for hours worked reduces the non-agriculture premium by about a fifth. In particular, comparison of columns 2 and 4 shows that the premium identified from switchers falls from the baseline level of 33 lp to 27 lp. This reflects the fact that workers in non-agriculture work more hours, as illustrated in Table A.2 in the Appendix. Column 2 of that table shows that the same workers supply on average 15% more hours when they switch to non-agriculture.

Whether one actually should condition on hours work in calculating the sectoral premia can be debated. The answer depends on the interpretation one wants to give to the premia and on the reason hours differ across sectors. In this paper, the non-agricultural premium is meant to capture an increase in the annual income that can be expected by a worker switching away from agriculture. To the extent that the switch is associated with higher labor supply, this increase in hours should be included as part of the benefit of switching. Our baseline measure therefore does not control for hours. In our view, thus calculated premium is a more interesting object than a premium netting out the effect of hours. The reason is that a sector of employment and supply of hours are best seen as a package. Our conjecture is that lower hours worked in agriculture observed for the same individuals are an indication that these individuals are frequently underutilized in agriculture, perhaps because of intrinsic seasonality of farm work.[10] If workers are forced to be idle for stretches of time in agriculture, then their low average utilization should be considered as a part of the productivity gap between agriculture and non-agriculture.

Another interesting feature seen in columns 3 and 4 in Table 11 is that the elasticity of annual income with respect to annual hours worked is only about one half. This means that income per hour is declining in hours worked, consistent with diminishing returns to labor. Combining this observation with higher hours in non-agriculture explains why the non-agricultural premium in terms of income per hour (columns 5 and 6) is smaller than the premium controlling for hours (columns 3 and 4).[11] However, even when identified off switching workers income per hour is

[10] Table A.1 and column 3-4 of Table A.2 show that the results are robust to including the secondary job. This alleviates a concern that lower hours in the main job for agricultural workers are offset by having a second job.

[11] If we control for hours worked in the income per hour specifications in columns 5 and 6 then the premia would

still significantly (19 lp) higher in non-agriculture. We report these numbers mainly because some of the literature interprets measures of income per hour as "wages" and uses them to calculate sectoral wage premia. In particular, in a concurrent paper also using the IFLS data Hicks et al. (2017) argue that non-agricultural premium in Indonesia largely disappears when they use their preferred regression of income per hour with worker fixed effects. There are two main reasons why our substantive findings are different. First, in our implementation we only rely on information on income and hours reported contemporaneously by the survey respondents. In contrast, Hicks et al. (2017) also rely on recall information for several years prior to the survey. As discussed in more detail in Appendix B, the recall information is likely subject to non-classical measurement error which can bias the estimated non-agricultural premium downwards. Second, even though our results are robust to controlling for hours and looking at hourly income, as argued earlier our conceptually preferred specification does not take hours into account. Comparing income per hour could indeed be preferable in a setting in which workers are offered constant hourly wages and freely choose the sector to which to allocate their marginal hour of work. But if hours are largely dictated by the nature of work in a sector then sector is the relevant "marginal" choice. Since we find the second case to be more plausible in the context of Indonesian labor markets we do not adjust our preferred non-agricultural premia for differences in hours.

3.2.6 Long-Run Income Growth

One of our most surprising findings is that workers who switch from non-agriculture to agriculture suffer an income loss of around 30%. To be more precise, an interpretation of coefficients in the first column of Table 4 is that a worker who switches away from non-agriculture between two survey waves has an income growth over that period 33 lp lower relative to what he would be expected to get if he remained in non-agriculture. Taking a large income cut could nevertheless be a rational decision for a worker maximizing his lifetime discounted income if he expects that the current loss of income will be compensated by higher future income growth in agriculture. This argument potentially has some merit because over our sample period average income growth was indeed higher in agriculture. To illustrate these differential trends, Table A.3 in the appendix shows the evolution of the non-agricultural premium over time. While it is strong and statistically significant throughout, it does decline over our sample period, especially in the cross-section, consistent with agricultural incomes partially converging to those in non-agriculture.

However, if switching workers could accurately predict the future income path then we would expect that over a long period of time those who took a cut switching to agriculture are not worse off than workers who remained in non-agriculture. As a first test of this hypothesis we look at income growth over the entire 21-year period spanned by IFLS 1-5. Column 1 of Table 12 shows that workers who started in non-agriculture in 1993 but switched to agriculture by 2014 had income growth over

be identical to those in columns 3 and 4.

that period lower by 37 lp compared to those who began and finished in non-agriculture. This result suggests that switchers to agriculture do not make up their initial loss even after a prolonged time.

By using a single long time difference, the previous exercise identifies an average effect of switching among workers with diverse interim sectoral employment histories. Our second exercise exploits this interim information. For this purpose, we consider employment histories spanned by three observations at equal 7-year intervals (i.e. those individuals with data for 1993, 2000, 2007 or 2000, 2007, 2014). We are interested in comparing the change in income over the 14-year span for workers who made different sectoral decisions during that period. Figure 1 shows the mean log wages for a few key histories. In particular, compare income of NAA-history workers (i.e. those who switched from non-agriculture to agriculture during the first 7-year period and stayed in agriculture during the second 7-year period) to income of NNN-history workers (who remained in non-agriculture throughout). Before the switch, NAA-workers had on average lower incomes, consistent with idea that those who switch are negatively selected from non-agricultural workers. More importantly, after the switch their incomes decline relative to NNN-workers. This is another reflection of the loss from switching emphasized in this paper. But the gap between NAA- and NNN-workers does not significantly narrow over the subsequent 7-year period. So crucially, over the entire 14-year period incomes of non-agricultural workers who permanently switched in the first half of the period fall back relative to those who stayed in non-agriculture.

Column 2 of Table 12 casts this analysis into a regression framework with the usual controls. We find that workers who switched from non-agriculture to agriculture during the first 7-year period and were still in agriculture at the end of the second 7-year period had a cumulative growth over 14 years lower by 19 lp (significant at 0.05 level) than if they had remained in non-agriculture over this period. Similarly, workers who switched into non-agriculture in the first period and remained there had long-run income higher by 15 lp than if they had remained in agriculture, though that effect is less precisely estimated (significant at 0.10 level).[12] Overall we take these results as evidence that workers who chose agriculture have lower incomes even in the long run.

4 Model of Sorting across Sectors with Barriers to Sectoral Mobility

In the previous section we establish that income differences between agriculture and non-agriculture in Indonesia are evident not only in the cross-section, but also among workers who switch sectors. One of the main goals of this paper is to investigate what these reduced-form findings tell us about the extent of self-selection and barriers to sectoral mobility. To this end, in this section we introduce a simple discrete-time model of the labor supply in which heterogenous workers self-

[12] In principle we could construct even longer histories which would allow us to control for pre- and post-trends of various groups. Unfortunately, between the number of possible histories increasing and the number of individuals with required data decreasing with history lengths, these longer histories would have limited statistical power.

select into sectors in each period based on the value of their human capital. Workers switch across sectors due to exogenous variation in the prices of human capital over time and due to the presence of an idiosyncratic time-varying component in their sector-specific human capital that resembles transitory productivity shocks. As we argue in the next section, the latter component is able to generate by itself a within-individual sectoral premium, with the sign depending on the magnitude of its relative dispersion across sectors. However, our structural estimation suggests that in order to simultaneously fit the magnitudes of the premia, the allocation and transition of workers across sectors over time, and the moments of the joint income distribution, some frictions to sectoral mobility are needed in the model.

For that reason, we evaluate different types of barriers to sectoral mobility that misallocate workers across sectors. We first consider switching costs across sectors (Dixit and Rob (1994); Cameron, Chaudhuri and McLaren (2007); Artuç, Chaudhuri and McLaren (2010); Dix-Carneiro (2014)) and, relatedly, compensating differentials (Rosen (1986); Taber and Vejlin (2016)). Switching costs act as utility burdens that constrain voluntary switches, inducing misallocation across sectors. Since we estimate opposite signs for the switching costs away from and towards agriculture, the model with switching costs performs similarly to a specification in which workers receive a positive compensating differential for working in agriculture. Next, we consider a specification with imperfect self-selection, where we allow for frictions that prevent individuals from working in their preferred sector. These frictions could be rationalized by on-the-job searching frictions (Gautier, Teulings and Van Vuuren (2010); Gautier and Teulings (2015)), for example. In contrast to the case of utility costs, where mobility barriers bind only for workers with relatively small differences in comparative advantage, in the alternative specification even workers with a strong comparative advantage in one sector can be affected by the frictions. Because frictions affect the inframarginal workers, the induced allocative inefficiency in this case generates a larger impact on the aggregate income.

4.1 Frictionless Economy

Suppose workers choose the sector at each time t to maximize their contemporaneous utility.[13] Let Ω_{it} be a vector of state variables for an individual i at time t. The income an individual receives in sector s is a product of the exogenous price of human capital in sector s at time t, R_t^s, and the amount of human capital the worker can supply to that sector:

$$y_t^s\left(\Omega_{it}\right) = R_t^s h^s\left(\Omega_{it}\right).$$

The supply of human capital depends on both observable and unobservable components. The former are gathered in a vector of covariates X_{it}, which in our estimation includes gender, the

[13] Given our empirical findings in section 3.2.6 we abstract from modeling inter-temporal choices. Turning the model into a dynamic one would greatly complicate the analysis, without necessarily affecting the insights from our simpler framework.

urban-rural location, years of schooling, years of working experience and the square of working experience. Notice that since we emphasize in this paper the sectoral dimension of the residual wage premia, we abstract from the choice of location, and treat the urban-rural choice just as another covariate. Since the sectoral premia are robust to heterogeneous Mincerian returns across sectors, we assume for simplicity homogenous returns on covariates, and hence we focus our attention on self-selection based on the unobservable components. The set of unobservables includes a time invariant component θ_i^s, representing the permanent comparative advantage of worker i in sector s, and an idiosyncratic time-varying term ε_{it}^s, resembling a transitory productivity shock that affects the comparative advantage of the same worker i in sector s at time t:

$$h^s\left(\Omega_{it}\right) = \exp\left(X_{it}'\beta + \theta_i^s + \varepsilon_{it}^s\right).$$

As in standard selection models, the functional form assumptions on the distribution of the components of comparative advantage are key for identification. We assume that the permanent component θ_i^s is i.i.d. across individuals, drawn from a normal distribution $N\left(\mu_\theta,\, \Sigma_\theta\right)$. Productivity shocks are also normal i.i.d. across individuals and time, $\varepsilon_{it}^s \sim N\left(\mu_\varepsilon,\, \Sigma_\varepsilon\right)$ and for identification purposes, orthogonal across sectors. We impose the normalization $\mu_\theta = \mu_\varepsilon = \mathbf{0}$ in order to identify the evolution of prices of human capital over time.

Let us now describe the worker's problem. The worker is choosing at the beginning of period t where to work. At the time of the decision she knows the value of the comparative advantage components and the human capital prices. Her problem is:

$$V\left(\Omega_{it}\right) = \max_s \left\{V^s\left(\Omega_{it}\right)\right\}, \tag{3}$$

where the value of working in sector s in the frictionless case is simply the logarithm of the income, $V^s\left(\Omega_{it}\right) = V_{ef}^s\left(\Omega_{it}\right) = \ln y_t^s\left(\Omega_{it}\right)$.

Finally, we assume the researcher observes individual income $\hat{y}_{it}^s$ subject to a pure idiosyncratic measurement error ν_{it}:

$$\ln \hat{y}_{it}^s = \ln y_t^s\left(\Omega_{it}\right) + \nu_{it}.$$

We assume the measurement error has mean zero and is normal i.i.d. across individuals and time, $\nu_{it} \sim N\left(\mathbf{0}, \sigma_\nu^2\right)$. Notice that an alternative interpretation of this error is as an ex-post productivity shock that affect the worker's observable income, but not her sectoral choice. Since in this case the ex-post shock does not affect workers' self-selection into sectors, the model delivers the same predictions under both specifications.

Our model abstracts from the possibility that workers drop out from the labor market, to focus attention on the role of sorting and the barriers to sectoral mobility introduced in the next section in explaining non-agriculture premia among active workers. For this reason, in the structural estimation we use the balanced panel of workers with income recorded in the five available waves of

IFLS. Denote by Θ the set of all structural parameters. The elements of Θ are listed and described in the first column of Table 13.

4.2 Economies with Barriers to Sectoral Mobility

The first type of barrier to sectoral mobility that we consider is a utility cost of switching across sectors. This cost could reflect tangible expenditures such as training or transportation costs, or intangibles such as social adjustment costs. Denote by $\phi^{s's}$ the utility cost for switching from sector s' (which was chosen in $t-1$) to sector s, common to all individuals. The value of working in sector s is then:

$$V_{sc}^{s}\left(\Omega_{it}\right)=\ln y_{t}^{s}\left(\Omega_{it}\right)-\ln C^{s_{t-1}s_{t}}\left(\Omega_{it}\right),$$

where

$$C^{s_{t-1}s_{t}}\left(\Omega_{it}\right)=C^{s's}=\begin{cases}\phi^{s's} & \text{if } s\neq s'\\ 1 & \text{if } s=s'\end{cases}.$$

The problem of the worker is the same as in (3) but with $V^{s}\left(\Omega_{it}\right)=V_{sc}^{s}\left(\Omega_{it}\right)$. In the general case we do not restrict $\ln\phi^{s's}$ to be positive, so in principle switching costs could also measure a utility compensation for $\ln\phi^{s's}<0$.

As we show in the next section, we estimate opposite signs for $\ln\phi^{AN}$ and $\ln\phi^{NA}$ (where A and N denote agriculture and non-agriculture, respectively), a pattern that is observationally similar to receiving a positive compensating differential for working in agriculture.[14] In the case of a compensating differential, the value of working in a sector can be defined as:

$$V_{cd}^{s}\left(\Omega_{it}\right)=\ln y_{t}^{s}\left(\Omega_{it}\right)+\ln C^{s},$$

where

$$C^{s}=\begin{cases}cd & \text{if } s=A\\ 1 & \text{if } s=N\end{cases}$$

and thus the differential cd measures the additional utility that a worker obtains by working in agriculture (relative to working in non-agriculture). This differential can be related to any attribute of the agricultural work that is valued by individuals: less exposure to pollution, crime, or crowding, more flexible work schedules, etc. Note that both switching costs and compensating differentials act as if proportionally scaling down or up income, so they can be interpreted in terms of annual earnings. Further, notice that the specification with switching costs adds two parameters (ϕ^{AN},ϕ^{NA}) to Θ, whereas the model with a compensating differential only adds one (cd).

Finally, we also consider a different kind of barrier to the allocation of workers across sectors.

[14] The model with a compensating differential cd is observationally equivalent to a model with switching costs $\phi^{AN}=cd$, $\phi^{NA}=1/cd$.

We want to capture an idea that workers do not always get to work in a sector that they would like, even if they have a strong comparative advantage in that sector. These frictions can be interpreted as life events forcing an individual to switch the sector of employment, and are meant to capture in a simple way the underlying search frictions. Specifically, we assume that at the beginning of each period an individual gets a random draw such that she will be able to choose the sector she desires with probability $1-p\left(\Omega_{it}\right)$ and she will be forced to work in the other sector with probability $p\left(\Omega_{it}\right)$. The probability p of being forced to accept a job in a sector other than desired can depend on the worker's state. In particular, we want to allow for the possibility that it might be more difficult to switch a sector than to keep working in the same sector, by letting the probability differ between those who desire to switch and those who desire to stay:

$$p^{s_{t-1}s_t}\left(\Omega_{it}\right) = p^{s's} = \begin{cases} p^T & \text{if } s \neq s' \\ p^S & \text{if } s = s' \end{cases}.$$

Similarly as with the switching costs, this specification adds two parameters (p^T,p^S) to Θ.

5 Structural Estimation

In this section, we describe the estimation procedure and the identification of the parameters of the structural model. The estimation method is Indirect Inference (Gourieroux, Monfort and Renault (1993)). We rely on the functional form assumptions to deliver a proof for identification in a simplified version of the model.

5.1 Estimation Procedure

The first step in the estimation procedure is to choose a set of auxiliary regression models that summarize the main features of the data we want to capture: the sectoral premia, the moments of the joint distribution of income and the workers' sectoral decisions over time. Those auxiliary models are used to compute the Indirect Inference loss function to be minimized, and hence they must be simple to estimate multiple times. As we explain below, for identification this method does not require that those auxiliary regressions are well specified (i.e. models which are exact reduced forms of the structural model, in which case Indirect Inference is equal to MLE). However, we do need that the selected models provide us enough information about the moments in the data that allow us to identify the set of structural parameters Θ.

We first reduce the dimensionality of Θ by estimating the Mincerian returns β in a linear regression of log-income on observables controlling for the interaction between sectoral choice and year. With the estimated $\hat{\beta}$ we construct the residual income net of observables that is used for estimating the rest of the model. We can estimate β separately from the rest of Θ because the observables do not affect the self-selection of workers given our assumption of homogeneity in the

Mincerian returns across sectors.[15]

To estimate the rest of the structural parameters in Θ, we select the following seven auxiliary models: i) a log-residual income linear regression on the sector choice, controlling for time fixed effects; ii) a log-residual income linear regression on the sector choice, controlling for time and individual fixed effects; iii) a log-residual income linear regression on the direction of sector switching between waves, controlling for time fixed effects; iv) a log-residual income linear regression in first differences on the direction of sector switching between waves, controlling for the first differences in years of the waves; v) a log-residual income linear regression on the interaction between sectoral choice and year; vi) a sectoral choice linear probability model on time dummy variables; and vii) a sectoral choice linear probability model on the previous sectoral choice. The role of each of these models in identifying the structural parameters is explained in the next subsection.

For efficiency reasons, we only use the coefficients of interest of the selected auxiliary regressions in the Indirect Inference loss function. Hence, we use the following 29 coefficients of the seven auxiliary models: 1-2) the non-agriculture premia in models i) and ii); 3-6) the sector-specific premia for switching workers from models iii) and iv); 7-23) the full set of estimated coefficients from models v) to vii); 24-25) sector-specific residual variance in model v); and 26-29) sector-specific residual variance for non-switching and switching workers in model iv). Table 14 summarizes the auxiliary models as well as the selected coefficients.

Arrange the values of the selected coefficients estimated in the actual data in the vector $\hat{\delta}$. The elements of vector $\hat{\delta}$ are displayed in the third column of Table 15 and remain fixed during the estimation procedure. The Indirect Inference loss function is computed as the weighted sum of the squared differences between the values in $\hat{\delta}$ and the values for the same set of coefficients obtained from simulations of the structural model. For weights, we use factors that represent the importance of the estimated coefficient in the identification of the structural parameters of the model, assigned after extensive experimentation. Appendix C describes their magnitudes and presents technical aspects of the estimation procedure in more detail.

Finally, in the models with switching costs or involuntary choices there is an issue of endogeneity of observing workers' initial sector allocation in the panel. We address it by introducing a pre-sample period zero with sectoral choice free of switching costs in the first case, and with a probability of being forced to work in the undesired sector independent of the worker's state in the second case.[16] We use pre-sample information on covariates when available to construct the distribution of the initial conditions. This way, although the auxiliary regressions are computed only for the five years in the sample, the data generating process of the model produces draws also for period zero.

[15] A conceptually identical but numerically less efficient estimation could be performed including β and using log-income to compute all auxiliary regressions, controlling for observables. This is why we can safely drop the effects of observables from our identification proof in Appendix D.

[16] We make this probability equal to p^S.

5.2 Identification

In this section we discuss how from the selected coefficients of the auxiliary regressions we obtain the set of moments that allows us to identify the parameters in Θ. Those moments are enumerated in Appendix D, where we demonstrate how Θ is identified in a simplified version of the model with two periods. In the proof, we take advantage of the functional form assumptions to extend the standard cross-sectional moments by including moments of the income distribution of the switching workers across waves, available only thanks to the panel dimension of the data, with the aim to set up a system of equations to solve for all parameters in Θ. We fully expect our reasoning to generalize to the same setting with a larger number of years. To verify this hypothesis, we generate multiple samples from the model with simulated covariates over the number of years as observed in the data, using different sets of parameters values. We find that the chosen auxiliary regressions allow the estimation procedure to obtain the values of parameters used to generate each sample.

Let us first comment on the main insights from the demonstration in Appendix D. In the frictionless economy sectoral decisions do not depend on workers' histories, so the model behaves in each period t as the standard log-normal Roy model with comparative advantage $u_{it}^s = \theta_i^s + \varepsilon_{it}^s$. In this case, we can use standard arguments of Heckman and Honoré (1990) to identify from repeated cross-sectional moments the prices of human capital (which, given our normalizations, act as the means of the distribution comparative advantage) and the variance matrix of u_{it}^s in each period augmented by the variance of measurement error. Only with panel data we can separately identify the variances of the permanent and transitory components of comparative advantage and the variance of measurement error, inferred from the moments of the growth in income of switchers. The intuition is that the amount of additional information that switchers provide about the joint distribution of income in response to changes in relative human capital prices is similar to the information obtained from exclusion restrictions and support conditions in the process of non-parametric identification of cross-sectional non-normal Roy models.[17]

We are able to find the analytical expressions for the moments of the income distribution of the employment transition groups across waves exploiting the property that draws of u_{it}^s in different periods of time are joint normally distributed, since each one is the sum of two normally distributed random variables. In this way, we can express the transition probabilities across waves and the observed moments of the growth in income for switchers using upper truncated multivariate normal distributions, where the prices of human capital in the two periods affect the truncation values. We verify that by adding this information from the switchers to the standard cross-sectional moments we can set up a system of equations with an unique solution for all parameters in Θ.

For the case of switching costs and frictions, sectoral choices depend on workers' histories and with only repeated cross-sectional data we can no longer identify either the prices of human capital

[17] See French and Taber (2011) for a detailed discussion about parametric identification of selection models through distributional assumptions and nonparametric identification using exclusion restrictions and support conditions.

or the variance matrix of u_{it}^s augmented by the variance of measurement error. That is, we obtain the non-identification result that even with the log-normality assumptions, the standard Roy model is not identified in the presence of barriers to sectoral mobility. We can generate two combinations of Θ, with different values for at least one parameter other than the corresponding barriers, that produce exactly the same set of cross-sectional moments. This is due to the fact that the cross-sectional moments depend on the distribution of the previous sectoral choices. Therefore, in order to identify the full set of parameters in Θ we need panel data even under log-normality assumptions.

In a similar way as with the moments of the employment transition groups in the frictionless economy, for the models with barriers to sectoral mobility we can derive the closed-form solutions of the cross-sectional moments, the transition probabilities and the moments of the growth in income for switchers, expressed all of them in terms of moments of upper truncated multivariate normal distributions, but with a dimensionality that grows with the number of time periods in the panel. Switching costs affect the truncation values of the distributions, similar to the human capital prices, whereas the probabilities of forced switches shift the entire distribution. We verify again that adding the moments of the growth in income for switchers and the transition probabilities to the standard cross-sectional moments we obtain a system of equations with an unique solution for all parameters in Θ, including switching costs.

Now we discuss in detail how the selected coefficients of the auxiliary regressions in our Indirect Inference loss function capture the set of required moments for identification. First, linear probability models vi) and vii) describe the distribution of sectoral choice in each cross-section and the average transition probabilities between waves, respectively. Combined, these models characterize the evolution of the joint distribution of sectoral choice over time, and hence they deliver the probabilities of sectoral transition across all waves. Second, for the moments of income growth, we use for the first moments both the within-individual premium from model ii) and the premia for switching workers relative to stayers in the model in first differences iv). The difference between the two is that the latter model takes into account the direction of transition, so it can actually inform the estimation procedure with the observed gains of switchers to non-agriculture and the losses of workers switching to agriculture separately, unlike the fixed-effects premium. For the second moments we use the residual variances for workers switching to each sector from model iv).

For the cross sectional moments, model v) informs us about the conditional expected incomes in each combination sector-year, since it includes a full set of interactions for sector and year. Those coefficients, taking together with the cross-sectional premium in model i), characterize the first cross-sectional moments. We collect the residual variances for the pool of workers in each sector from model v) and the residual variances for non-switching workers to each sector from model iv), to account for the second cross-sectional moments. That is, since we have an identical distribution of the productivity shocks across years, we only need those four variances to characterize the second cross-sectional moments, instead of a full set of variances from each combination of sector and year.

Our set of cross-sectional moments in Appendix D also includes third central moments. The

intuition of including those moments is that self-selection produces right skewness in a sector's earnings if the variance of the comparative advantage draws in that sector is larger than the covariance (Heckman and Honoré (1990)). Hence, those moments contribute to inform us about the relative magnitudes of the elements of the variance-covariance matrix of comparative advantage. However, in practice, estimating those third moments can be problematic since they can be very sensitive to outliers. Likely for that reason, and up to the best our knowledge, no estimation of selection models by Indirect Inference resorts to their use. We substitute the information of these moments with model iii), which compares the average performance of a switching worker to each sector with their peer group after the switch. This comparison, which is only possible with panel data, provides us with evidence regarding the nature of sorting in the data, in particular, whether there is positive or negative hierarchical sorting into each sector. Hence, as in standard selection models, those premia are informative about the relation between the covariance and the ratio of variances of the comparative advantage draws. Thus model iii) offers the same amount of information that the third central moments provide in our identification strategy.

It is worth to emphasize that Indirect Inference does not require that all auxiliary models are well specified for identification and consistent estimation of the parameters. As Sauer and Taber (2017) argue, the only requirement is that each structural parameter has to have an independent effect on at least one coefficient of the auxiliary models. Thus, a necessary condition for identification is to check that each parameter monotonically affects at least one coefficient of the auxiliary regressions, and that it produces a unique combination of responses on all the selected coefficients. We verify this requirement holds by visualizing the effects of changes in each structural parameter in the selected coefficients, keeping the remaining parameters constant, over the domain where those parameters are expected to lie on.[18]

6 Results

In this section we present the structural estimation results and use them to quantify the importance of barriers to mobility and of self-selection. We begin with a frictionless model and show that it fails to explain some salient features of the data. Models featuring frictions provide a much better fit to the data and imply a large extent of misallocation in Indonesia. Finally, we discuss the empirical content of the reduced form non-agriculture premia when viewed through the lens of the model.

6.1 Estimation Results

Column (1) of Table 13 shows the values of the Indirect Inference point estimates for the 16 structural parameters in the frictionless economy, and the standard errors from 100 bootrstraps. Given those estimated parameters the model generates the values for the 29 coefficients of the auxiliary

[18]The grids (one per each structural parameter) of 29 plots (one per each selected coefficient of the auxiliary regressions) are available upon request.

regressions displayed in column (4) of Table 15. The last row of this table shows the value of the loss function, indicating the overall fit of the model (with smaller values indicating a better fit).

Perhaps surprisingly, the model without any frictions is not only able to replicate the cross-sectional non-agriculture premium, but also to generate a sizable within-individual premium. In the estimated frictionless economy, workers who switch from non-agriculture to agriculture see their incomes decline by 21 lp on average. This striking result can be explained by a selection effect generated by the transitory productivity shocks. As a result of this mechanism the fixed effect premium is shaped largely by the variance of transitory shocks across sectors. We formally state this result for a simplified version of the model in the following proposition.

Proposition 1. *Consider the frictionless model with two periods and human capital prices equal across sectors and over time. Then the average growth of log income of workers switching from agriculture to non-agriculture is positive if and only if* $\sigma^2_{\varepsilon N} > \sigma^2_{\varepsilon A}$. *Furthermore, the average growth of log income of workers switching from non-agriculture to agriculture has the same magnitude but is of the opposite sign.*

Proof. See Appendix E. □

Since with two periods the fixed effects premium is simply equal to to the average growth of log income of switchers (taken with appropriate signs), we immediately have the following implication.

Corollary 1. *Under the same conditions as in Proposition 1, the non-agriculture premium identified from a regression with worker fixed effects is positive if and only if* $\sigma^2_{\varepsilon N} > \sigma^2_{\varepsilon A}$.

To understand these results, observe that after workers sort themselves into sectors in the first period, the only reason a worker would switch to a different sector next period is a change in the balance of productivity shocks, $\varepsilon^N_{it} - \varepsilon^A_{it}$. With equal variances of shocks across sectors, the average growth in income is the same for switchers in both directions, so the within-individual premium is null. But in the case of asymmetric variances, the shocks with a larger dispersion have a higher chance to take extreme values, resulting in larger average increase in income of workers shifting to the sector with the larger variance. Thus, the sign of the non-agriculture premium after controlling for worker fixed effects depends only on the relative size of the variance of the productivity shocks: it is positive when the variance is larger in non-agriculture, and negative otherwise. This reasoning carries over quantitatively to the estimated general model with multiple periods and evolving human capital prices.

The main message from this discussion is that finding a large non-agricultural income premium after controlling for worker fixed effects, as we find for Indonesia, *by itself* does not indicate that workers face any frictions in choosing their sector of employment. In principle, the premium can be explained simply by larger dispersion of productivity shocks faced by non-agricultural workers. But the pattern of variances have observable implications for moments other than sectoral premia.

In particular, the frictionless model struggles to simultaneously account for non-agricultural premia and the pattern of the residual variances of workers' earnings in the data (the variance is larger in the agriculture sector). To generate the cross-sectional and fixed effects non-agriculture premia, the frictionless model forces the relative magnitudes of the variances for both the permanent and transitory components of comparative advantage to be opposite to the pattern observed in the residual variances. This enables it to display a relatively good fit for the premia (0.56 lp and 0.21 lp in the model versus 0.57 lp and 0.40 lp in the data for the cross-sectional and the fixed-effects premia, respectively), but at the expense of generating residual variances that are completely reversed relative to the data (compare coefficients δ_{24}, δ_{25} and δ_{26}, δ_{27} in columns (2) and (4) of Table 15). To explain jointly the premia and the patterns of the residual variances, we need to introduce some frictions to the sectoral allocation in the model.

The first type of friction is represented by utility costs of switching sectors. When we restrict the switching costs to be positive, which is a standard and perhaps natural case, we find that they have effectively no impact on the estimates. The reason is that the estimated costs are small in magnitude, and in particular, the zero bound for the cost of switching from non-agriculture to agriculture is binding. This result might seem surprising, given that the literature estimating utility costs of switching sectors typically finds them to be large, often equivalent to multiples of a worker's annual income (e.g. Artuc, Lederman and Porto (2015)). But the magnitudes might not be easily comparable across studies, as they depend on what other mechanisms of sector determination are built into the respective models. In our case, when we allow for self-selection according to comparative advantage then positive switching costs do not have much additional explanatory power. In particular, if switching to agriculture was costly then it would be even more puzzling why so many workers make the move.

The situation is different if we remove the restriction on the sign of the switching costs. Column (2) of Table 13 shows the estimates for the model with unrestricted switching costs, and column (5) of Table 15 the corresponding coefficients of the auxiliary models. The switching costs are of opposite signs, approximately symmetric in magnitude, and of a large magnitude. A worker switching from agriculture to non-agriculture faces a cost of 64 lp of annual income equivalent (i.e. roughly equivalent to her annual income). That is, a worker who actually moves from agriculture to non-agriculture, must have a value of her human capital in non-agriculture at least 90% largeer than in agriculture. For smaller differences, the worker remains in agriculture. A worker switching towards agriculture receives a utility *compensation* equivalent to almost doubling her new agricultural income. That is, a worker who actually switches from non-agriculture to agriculture, could have a value of her human capital in agriculture as much as 47% smaller than in non-agriculture.

Because of this implied compensation, the model now has an easier time justifying why workers switch to agriculture. It can rationalize the income cuts of workers switching to agriculture in terms of negative switching cost so it does not need to rely on the counterfactual pattern of residual income variances. It can therefore generate both a within-individual premium that is close to the

one observed in the data (0.35 lp in the model versus 0.40 lp in the data) and deliver the correct qualitative patterns for the residual variances (larger variances in agriculture, see coefficients δ_{24} to δ_{27} of Table 15). In summary, the overall fit of the model with switching costs is substantially better (last row of Table 15).

Since the estimated switching costs are nearly symmetric (i.e. $\phi^{AN}\phi^{NA}$ is close to 1), the model with switching is similar to a specification with a single positive compensating differential for working in agriculture.[19] Columns (3) in Table 13 and (6) in Table 15 show, respectively, the estimated parameters and the obtained auxiliary coefficients for the latter model. In this case, individuals are willing to be paid less to work in agriculture simply because it is a sector they enjoy more. This estimated preference is strong, as it is equivalent to increasing a worker's agricultural income by 61 lp (or 89%). Comparing columns (5) and (6) in Table 15 shows that the compensating differential model fits the data nearly as well as the more flexible model with switching costs.

These estimates demonstrate that in order to be consistent with the salient features of worker-level panel data on sectoral employment and income, a model built on revealed preferences (i.e. voluntary choices) needs to make switching to agriculture attractive in some non-pecuniary terms. While estimating compensating differentials has a long history, we recognize that in our context they are not a particularly satisfying explanation. Ultimately, such utility-based compensation is a residual force that allows the model to rationalize choices otherwise difficult to explain. We therefore explore an alternative conceptual approach to think about barriers to sectoral mobility. Instead of treating all observed sectoral transitions as a result of voluntary choices, the alternative is to recognize that sometimes workers switch sectors for reasons independent of their productivity.

First we consider a specification with a single probability p of a worker being forced to a different sector than she would desire. This probability is estimated to be 0.06 (see column (4) in Table 13), which might not seem large, but in fact implies that most of the observed switches are of this random nature. This parsimonious explanation fits the data noticeably better (see column (7) in Table 15) than the models with utility switching costs.

Next, we increase the model's flexibility by allowing the probability of the involuntary sector allocation to depend on whether the workers wants to switch or to remain in the same sector as in the previous period. This specification captures the notion that switching a sector might be more difficult than staying put. This is indeed the case: as reported in column (5) in Table 13, the probability that a worker who wants to remain in a sector has to switch anyway is $p^S = 0.11$, whereas a worker wanting to switch most likely will not get the chance to do so ($p^T = 0.81$). These numbers imply that 63% of the observed transitions from non-agriculture to agriculture are driven by chance rather than in response to productivity shocks. The effect is not symmetric, in that only 32% of switches to non-agriculture are forced by randomness.

The explanation offered by this model for the prevalence of income-reducing transitions to agriculture is thus that these transitions are largely random events. Furthermore, once a worker

[19] Recall the result in footnote 14.

finds herself in a non-desired sector she can be "trapped" there for a while, because it is difficult to transition to the other sector. The model with these features provides a considerably better fit to the data than all the alternatives presented above, as can be seen from column (8) in Table 15.[20] In particular, it can match closely not only the qualitative pattern of non-agriculture premia and residual variances but also their magnitudes. It is also the only specification that can replicate the asymmetry in the magnitude of income growth of switchers to agriculture and switchers to non-agriculture (coefficients δ_5 and δ_6) that is observed in the estimation sample. Since this model offers superior empirical performance and what we believe is a compelling underlying mechanism, it is our preferred specification and the basis for further analysis.[21]

6.2 Counterfactual Exercises

We now proceed to quantify the importance of mobility barriers across sectors by computing the counterfactual equilibrium in which the barriers are removed. While this counterfactual is intended to illustrate the response of labor supply to the removal of such frictions, it is worth pointing out that the exercise lacks a general equilibrium adjustments of factor prices. Such adjustments can dampen the reallocation of workers, so our results should be regarded as an upper limit of the full impact.

We simulate counterfactual data setting $p^S = p^T = 0$ while keeping the remaining elements of $\hat{\Theta}$ and the values of covariates as in our baseline model. We first discuss the implications of eliminating the frictions for aggregate income and then present sectoral outcomes. Denoting by N the total number of individuals in the panel, we compute the number of individuals reallocated after the barriers are removed, equal to M, and the fraction of the population that is reallocated, $m = \frac{M}{N}$. To decompose the impact of workers' misallocation on total income Y into its different margins, denote by Y_m the sum of earnings of the misallocated individuals. Further, denote by ψ_m the ratio of the average income of the misallocated individuals to the average income in the population, $\psi_m \equiv \frac{NY_m}{MY}$. Thus, the percentage growth rate of total income after removing mobility frictions can be expressed as the product of three terms:[22]

$$\Delta\%Y = m\psi_m\Delta\%Y_m. \tag{4}$$

The first term represents the fraction of the population that is reallocated, the second term how

[20] This is a fair comparison since the model has the same number of free parameters as the model with switching costs.

[21] Extending the model to also include compensating differentials or positive switching costs has very little effect on the estimated probabilities of involuntary switches and the model fit.

[22] To prove this, for any variable x in the observed data let x' denote its counterfactual value in the frictionless economy, and $\hat{x} \equiv \frac{x'}{x}$ the proportional change. Since individuals who remain in the same sector do not observe any adjustment in their income, we can express: $\hat{Y} = \hat{Y}_m \frac{Y_m}{Y} + \frac{(Y-Y_m)}{Y}$. After some manipulation we can rewrite the latter expression as: $\hat{Y} = m\psi_m(\hat{Y}_m - 1) + 1$, and hence the percentage growth rate of total income after removing switching costs, $\Delta\%Y = 100(\hat{Y} - 1)$, as in equation (4).

important on average is the income of those individuals relative to the whole population in the data, and the third term the growth rate in the total income of all misallocated individuals.

Table 16 presents the results of the calculation. The main finding is that removing workers' mobility barriers across sectors leads to a significant reallocation of workers across sectors (35% of the total labor force) and to a large increase in income of misallocated workers (which doubles on average). As a result, it produces a sizable impact in aggregate terms: an increase of around 21.5% in total income (pooled across all years). It is worth noting that the effect would have been even larger if the misallocated workers were average earners. However, in our estimated model, the representative misallocated worker earns 57% of what the average worker earns in the whole panel (largely because the misallocated workers cannot realize their full earning potential when they are in the wrong sector). This fact moderates the effect of the reallocation of those workers on the adjustment in the aggregate income. It is also worth noting that in our baseline specification income is the only determinant of utility, so increases in income result in identical increases in welfare.[23]

Table 17 breaks down the results further by sector. Removing barriers to mobility would result in an agricultural employment shrinking by 8.1 p.p. as a share of total workforce. While this net change is not small, it is significantly smaller than the 35 p.p. gross flows of workers between sectors. Gross flows exceed net flows because there are workers wrongly allocated in both sectors. Furthermore, because the misallocated workers have on average lower productivity than the average worker in their sector, removing the misallocation increases (labor) productivity in both sectors, by 10.1% in non-agriculture and a whooping 44.4% in agriculture.[24] Consequently, output increases in both sectors. In particular, it increases by 14.2% in agriculture despite the sector contracting in terms of employment.

In summary, our results indicate that labor is misallocated to a significant degree in Indonesia because of barriers to mobility across sectors. Eliminating such barriers would potentially lead to large aggregate productivity gains. Our work does not offer a practical guide to how the barriers can be eliminated in practice, but it highlights that policies easing frictions workers face in making sectoral choices could have a large positive impact on the economy.

6.3 Industry Premia Revisited

With the structural model at our disposal, we now use it to shed more light on the empirical content of the reduced-form sectoral premia of the kind we estimated in section 3.

There is a strand in the literature (e.g., Hicks et al. (2017), Herrendorf and Schoellman (2018)) arguing that if substantial cross-sectional non-agriculture premium largely disappears after controlling for worker fixed effects, then the data can be explained by an efficient sorting of workers. In

[23] In contrast, in a model with barriers to mobility modeled as utility costs of switching, income and welfare would diverge, with the average growth in utility smaller than in income, but positive.

[24] The estimated processes of permanent and transitory components of comparative advantage draws imply that both sectors are "standard" in the Roy model terminology of Heckman and Honoré (1990).

section 6.1 we explained that frictionless sorting does not imply that there should be zero premium identified from within-worker variation. The flip-side of this argument is that once we allow for barriers to sectoral mobility, the absence of the within-worker premium does not imply that the allocation is efficient. There can be many combinations of processes for permanent and transitory components of comparative advantage draws and barriers to mobility that result in the same cross-sectional and (possibly zero) within-worker premia. To separately identify the role of frictions and of sorting we have to look beyond industry premia at a rich set of moments observable in a panel of workers.

To illustrate this discussion, column (2) in Table 18 reports the cross-sectional and within-worker non-agriculture premia obtained from data simulated in a counterfactual removing frictions in our baseline model (discussed in the previous subsection). Even though the allocation is perfectly efficient in this case, the non-agriculture premium from a regression with worker fixed effects is not zero, but in fact strongly negative at -31 lp. The negative premium is a natural consequence of larger variance of productivity shocks faced by workers in agriculture.

The level of the fixed-effect premium by itself therefore does not have clear implications for the strength of barriers to mobility if sorting is also present. But the difference between the fixed effect and cross-sectional premia does indeed indicate the presence of sorting. To illustrate this point, we consider an alternative counterfactual scenario in which self-selection is eliminated. Specifically, we set $\sigma^2_{\theta^A}$, $\sigma^2_{\theta^N}$, $\sigma^2_{\varepsilon^A}$, $\sigma^2_{\varepsilon^N}$ all to zero. In this case all workers are identical and would prefer non-agriculture as it offers higher prices for human capital. There is no sorting, and both sectors employ workers because of the frictions restricting workers from selecting their preferred sector. As column (3) in Table 18 confirms, when transitions between sectors are purely random the fixed effect premium takes effectively the same value as the cross-sectional premium.

To summarize, comparing the cross-sectional and within-worker sector premia can be a useful diagnostic for detecting self-selection. But detecting barriers to sectoral mobility in observational data requires imposing sufficient structure and using data beyond the sectoral premia.

7 Conclusions

We present extensive reduced-form evidence of a substantial premium for working outside of agriculture in Indonesia. The same individual switching to work in non-agriculture gains about 25-30% income, while an individual switching in the opposite direction faces an income loss of a similar magnitude. We argue that in order to generate simultaneously those premia and the main moments of the joint distribution of income, we need to extend the models that attribute income gaps across sectors only to sorting of workers by including barriers to sectoral mobility that misallocate workers across sectors.

Our preferred way of thinking about barriers to mobility is that they restrict the ability of workers to work in their desired sectors. Such frictions misallocate a large fraction of workers across

sectors (35% in our baseline specification), and imply large income gains (of around 100%) for the misallocated workers when they reallocate. As a result, output in Indonesia could increase by as much as 21% if barriers to mobility across sectors were removed.

In this paper we are agnostic about the root causes of the barriers to sectoral mobility. Investigating what constitutes such barriers, why they persist, and what policies can be used as a remedy would be fruitful avenue for future research.

References

Artuc, Erhan, Daniel Lederman, and Guido Porto. 2015. "A mapping of labor mobility costs in the developing world." *Journal of International Economics*, 95(1): 28 – 41.

Artuç, Erhan, Shubham Chaudhuri, and John McLaren. 2010. "Trade shocks and labor adjustment: A structural empirical approach." *The American Economic Review*, 100(3): 1008–1045.

Beegle, Kathleen, Joachim De Weerdt, and Stefan Dercon. 2011. "Migration and economic mobility in Tanzania: Evidence from a tracking survey." *Review of Economics and Statistics*, 93(3): 1010–1033.

Bound, John, Charles Brown, and Nancy Mathiowetz. 2001. "Measurement error in survey data." In . Vol. 5 of *Handbook of Econometrics*, , ed. James J. Heckman and Edward Leamer, 3705 – 3843. Elsevier.

Bruins, Marianne, James A. Duffy, Michael P. Keane, and Anthony A. Smith Jr. 2018. "Generalized indirect inference for discrete choice models." Unpublished (Forthcoming Journal of Econometrics).

Bryan, Gharad, and Melanie Morten. 2018. "The aggregate productivity effects of internal migration: Evidence from Indonesia." Unpublished.

Bryan, Gharad, Shyamal Chowdhury, and Ahmed Mushfiq Mobarak. 2014. "Underinvestment in a profitable technology: The case of seasonal migration in Bangladesh." *Econometrica*, 82(5): 1671–1748.

Cameron, Stephen, Shubham Chaudhuri, and John McLaren. 2007. "Trade shocks and labor adjustment: Theory." National Bureau of Economic Research Working Paper 13463.

de Nicola, Francesca, and Xavier Gine. 2014. "How accurate are recall data? Evidence from coastal India." *Journal of Development Economics*, 106: 52 – 65.

Dix-Carneiro, Rafael. 2014. "Trade liberalization and labor market dynamics." *Econometrica*, 82(3): 825–885.

Dixit, Avinash, and Rafael Rob. 1994. "Switching Costs and Sectoral Adjustments in General Equilibrium with Uninsured Risk." *Journal of Economic Theory*, 62(1): 48 – 69.

Duncan, Greg J., and Daniel H. Hill. 1985. "An investigation of the extent and consequences of measurement error in labor-economic survey data." *Journal of Labor Economics*, 3(4): 508–532.

French, Eric, and Christopher Taber. 2011. "Identification of models of the labor market." In . *Handbook of Labor Economics*, , ed. Orley Ashenfelter and David Card, Vol. 4, 537 – 617. Elsevier.

Gautier, Pieter A., and Coen N. Teulings. 2015. "Sorting and the output loss due to search frictions." *Journal of the European Economic Association*, 13(6): 1136–1166.

Gautier, Pieter A., Coen N. Teulings, and Aico Van Vuuren. 2010. "On-the-job search, mismatch and efficiency." *The Review of Economic Studies*, 77(1): 245–272.

Gibson, John, and Bonggeun Kim. 2010. "Non-classical measurement error in long-term retrospective recall surveys." *Oxford Bulletin of Economics and Statistics*, 72(5): 687–695.

Godlonton, Susan, Manuel A. Hernandez, and Michael Murphy. 2016. "Anchoring bias in recall data: Evidence from Central America." International Food Policy Research Institute (IFPRI) IFPRI discussion papers 1534.

Gollin, Douglas, David Lagakos, and Michael E. Waugh. 2014. "The agricultural productivity gap." *The Quarterly Journal of Economics*, 129(2): 939–993.

Gourieroux, C., A. Monfort, and E. Renault. 1993. "Indirect inference." *Journal of Applied Econometrics*, 8: S85–S118.

Heckman, James J., and Bo E. Honoré. 1990. "The empirical content of the Roy model." *Econometrica*, 58(5): 1121–1149.

Herrendorf, Berthold, and Todd Schoellman. 2015. "Why is measured productivity so low in agriculture?" *Review of Economic Dynamics*, 18(4): 1003 – 1022.

Herrendorf, Berthold, and Todd Schoellman. 2018. "Wages, human capital and barriers to structural transformation." *American Economic Journal: Macroeconomics*, 10(2): 1–23.

Hicks, Joan Hamory, Marieke Kleemans, Nicholas Y. Li, and Edward Miguel. 2017. "Reevaluating agricultural productivity gaps with longitudinal microdata." National Bureau of Economic Research Working Paper 23253.

Hnatkovska, Viktoria, and Amartya Lahiri. 2016. "Urbanization, structural transformation and rural-urban disparities."

Kan, Raymond, and Cesare Robotti. 2017. "On moments of folded and truncated multivariate normal distributions." *Journal of Computational and Graphical Statistics*, 26(4): 930–934.

Katz, Lawrence F., and Lawrence H. Summers. 1989. "Industry rents: Evidence and implications." *Brookings Papers on Economic Activity*, 209. Copyright - Copyright Brookings Institution 1989; Last updated - 2014-05-07; CODEN - BPEAD5.

Lagakos, David, and Michael E. Waugh. 2013. "Selection, agriculture, and cross-country productivity differences." *The American Economic Review*, 103(2): 948–980.

Nguyen, Binh T., James W. Albrecht, Susan B. Vroman, and M. Daniel Westbrook. 2007. "A quantile regression decomposition of urban–rural inequality in Vietnam." *Journal of Development Economics*, 83(2): 466 – 490.

Qu, Zhaopeng (Frank), and Zhong Zhao. 2008. "Urban-rural consumption Inequality in China from 1988 to 2002: Evidence from quantile regression decomposition." Institute for the Study of Labor (IZA) IZA Discussion Papers 3659.

Restuccia, Diego, Dennis Tao Yang, and Xiaodong Zhu. 2008. "Agriculture and aggregate productivity: A quantitative cross-country analysis." *Journal of Monetary Economics*, 55(2): 234 – 250.

Rosen, Sherwin. 1986. "The theory of equalizing differences." In . Vol. 1 of *Handbook of Labor Economics*, 641 – 692. Elsevier.

Roy, Andrew Donald. 1951. "Some thoughts on the distribution of earnings." *Oxford Economic Papers*, 3(2): 135–146.

Sarvimaki, Matti, Roope Uusitalo, and Markus Jantti. 2018. "Habit Formation and the Misallocation of Labor: Evidence from Forced Migrations." mimeo.

Sauer, Robert M., and Christopher R. Taber. 2017. "Indirect inference with importance sampling: An application to women's wage growth." IZA Discussion Papers 11004.

Strauss, J., F. Witoelar, and B. Sikoki. 2016. "The fifth wave of the Indonesia Family Life Survey (IFLS5): Overview and field report." RAND working papers WR-1143/1-NIA/NICHD.

Taber, Christopher, and Rune Vejlin. 2016. "Estimation of a Roy/search/compensating differential model of the labor market." National Bureau of Economic Research Working Paper 22439.

Tallis, G. M. 1961. "The moment generating function of the truncated multi-normal distribution." *Journal of the Royal Statistical Society. Series B (Methodological)*, 23(1): 223–229.

Thomas, Duncan, Firman Witoelar, Elizabeth Frankenberg, Bondan Sikoki, John Strauss, Cecep Sumantri, and Wayan Suriastini. 2012. "Cutting the costs of attrition: Results from the Indonesia Family Life Survey." *Journal of Development Economics*, 98(1): 108 – 123. Symposium on Measurement and Survey Design.

Young, Alwyn. 2013. "Inequality, the urban-rural gap, and migration." *The Quarterly Journal of Economics*, 128(4): 1727–1785.

Tables

Table 1: Descriptive Statistics

	IFLS 1: 1993	IFLS 2: 1997	IFLS 3: 2000	IFLS 4: 2007	IFLS 5: 2014
Share of male	0.60	0.62	0.59	0.58	0.57
Mean age	41.4	38.1	39.0	40.7	41.2
Mean years of schooling	5.4	6.1	7.1	7.8	8.7
Joint distribution over sectors and locations					
Total Agriculture	0.45	0.35	0.36	0.36	0.29
Rural Agriculture	0.42	0.31	0.32	0.31	0.24
Urban Agriculture	0.03	0.03	0.04	0.05	0.05
Total Non-Agriculture	0.55	0.65	0.64	0.64	0.71
Rural Non-Agriculture	0.27	0.30	0.27	0.25	0.27
Urban Non-Agriculture	0.28	0.35	0.37	0.39	0.44
Total Rural	0.69	0.62	0.59	0.56	0.50
Total Urban	0.31	0.38	0.41	0.44	0.50
No. observations	9714	12875	17931	20874	24475
Main sample: panel of workers with 2+ observations					
No. observations			70586		
No. individuals			22829		

Table 2: Sectoral and Urban Income Premia

	(1) Log Income	(2) Log Income	(3) Log Income	(4) Log Income	(5) Log Income	(6) Log Income
Non-Agriculture	0.839***		0.686***	0.574***	0.332***	
	(0.041)		(0.040)	(0.036)	(0.033)	
Urban		0.647***	0.405***	0.207***	0.084**	
		(0.045)	(0.042)	(0.036)	(0.032)	
Agr.×Urban						0.062
						(0.055)
Non-Agr.×Urban						0.416***
						(0.046)
Non-Agr.×Rural						0.326***
						(0.039)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes
Indiv. cont.				Yes	Yes	Yes
Individual FE					Yes	Yes
Observations	48299	48308	48299	44494	44497	44497
R^2	0.412	0.394	0.424	0.503	0.518	0.518

Notes: Individual controls: education, experience, experience sq., and sex. Observations weighted by longitudinal survey weights. Standard errors clustered by enumeration areas (primary sampling units of the survey) in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table 3: Transitions across Sectors and Locations

Sector transitions	No. of cases	Share of total
AA	13214	27.68
AN	3886	8.14
NA	3546	7.43
NN	27098	56.76
Total	47744	100.00
Indiv. who switch at least once		23.89

Location transitions	No. of cases	Share of total
RR	23299	48.79
RU	3171	6.64
UR	1166	2.44
UU	20121	42.13
Total	47757	100.00
Indiv. who switch at least once		16.91

		Sector in T+1	
		Agricult.	Non-Agr.
Sector in T	Agricult.	0.78	0.22
	Non-Agr.	0.12	0.88

		Location in T+1	
		Rural	Urban
Location in T	Rural	0.90	0.10
	Urban	0.05	0.95

Notes: XY indicates a transition from sector (or location type) X to Y between two consecutive observations for an individual. A - Agriculture, N - Non-agriculture, R - Rural, U - Urban.

Table 4: Premia for Switchers and Stayers

	(1) Δ Log Income	(2) Δ Log Income
Sector transitions		
AN	0.220***	
	(0.050)	
NA	-0.392***	
	(0.049)	
NN	-0.066***	
	(0.023)	
Location transitions		
RU	0.091*	
	(0.047)	
UR	-0.199***	
	(0.058)	
UU	-0.040*	
	(0.023)	
Sector trans. × Migration		
AA × Migrate		-0.108
		(0.092)
AN × Stay		0.196***
		(0.053)
AN × Migrate		0.275**
		(0.108)
NA × Stay		-0.379***
		(0.054)
NA × Migrate		-0.472***
		(0.110)
NN × Stay		-0.117***
		(0.021)
NN × Migrate		-0.008
		(0.039)
Δ Year FE	Yes	Yes
Δ Province FE	Yes	Yes
Δ Indiv. cont.	Yes	Yes
Observations	27697	24858
R^2	0.075	0.075

Notes: XY indicates a transition from sector (or location type) X to Y between two consecutive observations for an individual. A - Agriculture, N - Non-Agriculture, R - Rural, U - Urban. Migrate indicates movement outside of the village boundary. Omitted categories: staying in agriculture (AA) and staying in rural area (RR) in column 1; staying in agriculture within the same village (AA×Stay) in column 2. Individual controls: education, experience, experience sq., and sex. Observations weighted by longitudinal survey weights. Standard errors clustered by enumeration areas (primary sampling units of the survey) in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table 5: Job Types and Top Occupations

Job Type	Empl. share
Self-employed	0.471
Private worker	0.318
Government worker	0.068
Unpaid family worker	0.142

Top 10 Occupations	Empl. share
Agricultural and animal husbandry workers	0.352
Salesmen, shop assistants and related workers	0.136
Bricklayers, carpenters and other construction workers	0.038
Maids and related housekeeping service workers NEC	0.038
Working proprietors (catering and lodging services)	0.034
Transport equipment operators	0.032
Teachers	0.031
Food and beverage processors	0.027
Working proprietors (wholesale and retail trade)	0.026
Service workers NEC	0.025
Cumulative	0.739

Notes: Employment shares reported for IFLS 4 (2007).

Table 6: Premia for Switchers and Stayers by Job Type

	(1) Self-employed	(2) Private Worker	(3) Government	(4) Unpaid Family
AN-AA	0.259***	0.245***	0.111	0.335
	18.31	11.98	0.43	1.21
NA-NN	-0.309***	-0.274***	-0.225	-0.871*
	33.61	17.89	1.02	3.79

Notes: Table presents tests based on results of a first-difference regression (2) (c.f. column 1 in Table 4) with direction of sectoral switch interacted with job type. Reported are the difference in coefficients of interest and the value of an *F(1,296)* test that the difference is zero. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table 7: Wage Premia

	(1) Log Income	(2) Log Income	(3) Log Wage	(4) Log Wage
Non-Agriculture	0.574***	0.332***	0.490***	0.231***
	(0.036)	(0.033)	(0.051)	(0.050)
Urban	0.207***	0.084**	0.193***	0.119***
	(0.036)	(0.032)	(0.042)	(0.035)
Year FE	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes
Indiv. cont.	Yes	Yes	Yes	Yes
Individual FE		Yes		Yes
Observations	44494	44497	23139	23140
R^2	0.503	0.518	0.556	0.601

Notes: Individual controls: education, experience, experience sq., and sex. Observations weighted by longitudinal survey weights. Standard errors clustered by enumeration areas (primary sampling units of the survey) in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table 8: Consumption Premia

	(1) Log PCE	(2) Log PCE	(3) Log PCE	(4) Log PCI	(5) Log PCI	(6) Log PCI
NA sh. in HH income	0.305***			0.702***		
	(0.017)			(0.040)		
Non-Agr.		0.214***	0.075***		0.492***	0.197***
		(0.014)	(0.013)		(0.030)	(0.024)
Urban	0.315***	0.161***	0.095***	0.416***	0.225***	0.063*
	(0.029)	(0.024)	(0.026)	(0.043)	(0.034)	(0.037)
Non-Agr./$\overline{Y_{ih}/Y_h}$		0.382	0.134		0.884	0.352
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes
Indiv. cont.		Yes	Yes		Yes	Yes
Individual FE			Yes			Yes
Observations	40168	53546	53550	38365	51690	51693
R^2	0.707	0.742	0.784	0.504	0.520	0.541

Notes: Specifications (1) and (4) estimated at a household level with observations weighted by longitudinal household survey weights. (1) also includes the number of household members (level and squared) as controls. *NA sh. in HH Income* is a continuous variable measuring the share of non-agriculture in household's income. Specifications (2)-(3) and (5)-(6) estimated at an individual level. Individual controls: education, experience, experience sq., and sex. Observations weighted by longitudinal survey weights. Standard errors clustered by enumeration areas (primary sampling units of the survey) in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table 9: Premia with Heterogeneity in Mincerian Returns

	(1)	(2)	(3)	(4)
	Log Income	Log Income	Log Income	Log Income
Non-Agriculture	0.574***	0.332***	0.625***	0.314***
	(0.036)	(0.033)	(0.039)	(0.034)
Urban	0.207***	0.084**	0.200***	0.074**
	(0.036)	(0.032)	(0.034)	(0.032)
Year FE	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes
Indiv. controls	Yes	Yes	Yes	Yes
Individual FE		Yes		Yes
Het. in Mincer			Yes	Yes
Observations	44494	44497	44494	44497
R^2	0.503	0.518	0.506	0.520

Notes: Columns (3) and (4) allow for differences in Mincerian returns across sectors and locations. Average marginal effect for the population reported. Average effects for switchers are similar. Individual Mincerian controls: education, experience, experience sq., and sex. Observations weighted by longitudinal survey weights. Standard errors clustered by enumeration areas (primary sampling units of the survey) in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table 10: Premia with Additional Jobs and Home Production

	Base	Base	Add. Job	Add. Job	Add+HH TC	Add+HH TC	Add+HH FC	Add+HH FC
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Log Income	Log Income	Log Income	Log Income	Log Income	Log Income	Log Income	Log Income
Non-Agr.	0.574***	0.332***	0.501***	0.264***	0.462***	0.251***	0.447***	0.245***
	(0.036)	(0.033)	(0.034)	(0.032)	(0.033)	(0.032)	(0.032)	(0.032)
Urban	0.207***	0.084**	0.171***	0.063*	0.141***	0.057*	0.124***	0.051
	(0.036)	(0.032)	(0.034)	(0.034)	(0.033)	(0.034)	(0.033)	(0.034)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Indiv. cont.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual FE		Yes		Yes		Yes		Yes
Observations	44494	44497	44489	44492	44489	44492	44489	44492
R^2	0.503	0.518	0.514	0.538	0.513	0.540	0.515	0.545

Notes: *Base* is the baseline specification involving primary job only. *Add. Job* also includes secondary job. *HH TC* scales income by the inverse of the share of self-produced consumption in household's overall consumption. *HH FC* scales income by the inverse of the share of self-produced food in household's food consumption. Individual controls: education, experience, experience sq., and sex. Observations weighted by longitudinal survey weights. Standard errors clustered by enumeration areas (primary sampling units of the survey) in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table 11: Premia with Hours Worked

	(1) Log Income	(2) Log Income	(3) Log Income	(4) Log Income	(5) Log Inc./Hour	(6) Log Inc./Hour
Non-Agriculture	0.574***	0.332***	0.441***	0.271***	0.297***	0.185***
	(0.036)	(0.033)	(0.034)	(0.032)	(0.036)	(0.036)
Urban	0.207***	0.084**	0.160***	0.084***	0.109***	0.076***
	(0.036)	(0.032)	(0.031)	(0.026)	(0.029)	(0.028)
Log Hours/Year			0.496***	0.432***		
			(0.011)	(0.011)		
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes
Indiv. cont.	Yes	Yes	Yes	Yes	Yes	Yes
Individual FE		Yes		Yes		Yes
Observations	44494	44497	43841	43843	43841	43843
R^2	0.503	0.518	0.592	0.595	0.478	0.493

Notes: Individual controls: education, experience, experience sq., and sex. Observations weighted by longitudinal survey weights. Standard errors clustered by enumeration areas (primary sampling units of the survey) in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table 12: Long Run Premia

	1993-2014 (1) Δ Log Income	93-07/00-14 (2) Δ Log Income
AA-AN	0.172	
	1.38	
NA-NN	-0.369***	
	9.10	
ANN-AAA		0.147*
		2.79
NAA-NNN		-0.186**
		4.62
Observations	2567	7857
R^2	0.105	0.098

Notes: Column 1 presents tests based on results of a first-difference regression (2), where the difference is over the period 1993-2014. Reported are the difference in coefficients of interest and the value of an *F(1,288)* test that the difference is zero. Column 2 presents tests based on a first-difference specification over 14 years (1993-2007 or 2000-2014) controlling for direction of switch during the first and second 7-year period. Reported are the difference in coefficients of interest and the value of an *F(1,292)* test that the difference is zero. Other controls and weights are as in column 1 in Table 4. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table 13: Parameter Estimates

Parameter	(1) Frictionless	(2) Unrestricted switching costs	(3) Compensating differential	(4) Single probability of involuntary choices	(5) Heterogeneous probabilities of involuntary choices
Variance of permanent comparative advantage in sector s ($\sigma^2_{\theta^s}$) and covariance ($\sigma_{\theta^{AN}}$)					
$\sigma^2_{\theta^A}$	0.29	0.50	0.52	0.36	0.41
	(0.03)	(0.05)	(0.05)	(0.02)	(0.02)
$\sigma^2_{\theta^N}$	0.63	0.45	0.48	0.72	0.64
	(0.04)	(0.04)	(0.04)	(0.02)	(0.03)
$\sigma_{\theta^{AN}}$	0.26	0.16	0.18	0.32	0.26
	(0.04)	(0.04)	(0.05)	(0.02)	(0.02)
Variance of transitory productivity shocks in sector s ($\sigma^2_{\varepsilon^s}$)					
$\sigma^2_{\varepsilon^A}$	0.00	0.12	0.12	0.00	0.25
	(0.00)	(0.03)	(0.03)	(0.00)	(0.02)
$\sigma^2_{\varepsilon^N}$	0.06	0.00	0.01	0.00	0.03
	(0.01)	(0.01)	(0.01)	(0.00)	(0.02)
Variance of measurement error (σ^2_ν)					
σ^2_ν	0.73	0.72	0.71	0.65	0.47
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
Price of human capital in sector s at time t (R^s_t)					
R^A_1	0.80	0.45	0.47	0.79	0.77
R^A_2	1.29	0.73	0.75	1.22	1.15
R^A_3	1.18	0.57	0.62	1.07	1.10
R^A_4	1.41	0.81	0.88	1.35	1.51
R^A_5	1.74	1.07	1.12	1.62	2.00
R^N_1	1.08	1.33	1.31	1.21	1.48
R^N_2	1.74	1.95	1.94	1.89	2.20
R^N_3	1.36	1.69	1.66	1.53	1.79
R^N_4	1.77	2.21	2.16	1.99	2.15
R^N_5	2.16	2.42	2.50	2.54	2.52
Switching cost of moving from sector s to sector s' ($\phi^{ss'}$)					
$\ln \phi^{AN}$	–	0.64	–	–	–
		(0.04)			
$\ln \phi^{NA}$	–	-0.63	–	–	–
		(0.03)			
Compensating differential					
$\ln cd$	–	–	0.61	–	–
			(0.04)		
Probabilities of involuntary choices					
p	–	–	–	0.06	–
				(0.00)	
p^S	–	–	–	–	0.11
					(0.01)
p^T	–	–	–	–	0.81
					(0.02)

Notes: Standard errors from 100 bootstraps in parentheses. All standard errors for R^s_t are smaller than 0.1 times the point estimate value.

Table 14: Auxiliary Models and Selected Coefficients

Auxiliary model	Selected coefficients	Coefficient description
i) Log-residual income linear regression on the sector choice: $\ln \tilde{y}_{its} = c + 1\{d_{it} = N\}\delta_1 + D_t + \varepsilon_{ist}$	δ_1	Non-agriculture premium (cross-sectional)
ii) Log-residual income linear regression on the sector choice: $\ln \tilde{y}_{its} = c + 1\{d_{it} = N\}\delta_2 + D_t + D_i + \varepsilon_{ist}$	δ_2	Non-agriculture premium (within-individual)
iii) Log-residual income linear regression on the direction of sector switching: $\ln \tilde{y}_{its} = c + 1\{d_{it-1} = s, d_{it} = s'\}\gamma_{ss'} + D_t + \varepsilon_{ist}$	$\delta_3 = \gamma_{NA}$ $\delta_4 = \gamma_{AN} - \gamma_{NN}$	Premia for switchers to each sector relative to their peers post-switch
iv) Log-residual income linear regression in first differences on the direction of sector switching: $\Delta \ln \tilde{y}_{its} = 1\{d_{it-1} = s, d_{it} = s'\}\gamma_{ss'} + \Delta D_t + \varepsilon_{ist}$	$\delta_5 = \delta_{AN}$ $\delta_6 = \delta_{NA} - \delta_{NN}$	Premia for switchers to each sector relative to non-switching workers
v) Log-residual income linear regression on the interaction between sector choice and year: $\ln \tilde{y}_{its} = \delta_7 + \{1\{d_{it} = N\} \times 1\{d_{it} = t\}\}\gamma_{s\times t} + \varepsilon_{ist}$	δ_7 $\delta_8 = \gamma_{A\times 2} \ldots \delta_{16} = \gamma_{N\times 5}$	Constant Interactions sector and year
vi) LPM of sector choice on time dummy variables: $1\{d_{it} = N\} = \delta_{22} + 1\{d_{it} = t\}\gamma_t + \varepsilon_{ist}$	δ_{17} $\delta_{18} = \gamma_2 \ldots \delta_{21} = \gamma_5$	Constant Year dummies
vii) LPM of sector choice on previous sector choice: $1\{d_{it} = N\} = \delta_{27} + 1\{d_{it-1} = N\}\delta_{28} + \varepsilon_{ist}$	δ_{22},δ_{23}	Constant and lagged sector choice
viii) Residual variances:	δ_{24}, δ_{25} δ_{26}, δ_{27} δ_{28}, δ_{29}	For workers in each sector from model v) For non-switching workers in each sector from model iv) For switching workers to each sector from model iv)

Notes: LPM stands for linear probability model. $\tilde{y}_{its}$ is the residual income of individual i in time t working in sector s, that satisfies $\ln \tilde{y}_{its} = \ln y_{its} - X'_{it}\hat{\beta}$, where y_{its} is the observed income, X'_{it} is the set of observables that includes gender, urban-rural location, years of schooling, years of working experience and square of years of working experience, and $\hat{\beta}$ is the vector of estimated coefficients on observables in the log-income linear regression on the interaction between sector choice and year conditional on observables: $\ln y_{its} = \delta + X'_{it}\beta + \{1\{d_{it} = N\} \times 1\{d_{it} = t\}\}\gamma_{s\times t} + \varepsilon_{ist}$. D_t corresponds to year fixed-effects and D_i to individual fixed-effects. Δx is the first difference of variable x. $1\{d_{it} = N\}$ is a dummy indicating whether individual i works in non-agriculture in period t, $1\{d_{it-1} = s, d_{it} = s'\}$ is a set of dummies indicating whether individual i in period $t-1$ worked in sector s and in period $t$worked in sector s', and $1\{d_{it} = t\}$ is a set of dummies indicating whether the observation of worker i corresponds to period t. The omitted category in models iii) and iv) is AA, in model v) is $A \times 1$ and in model vi) is $t = 1$.

Table 15: Coefficients of Auxiliary Regression Models

(1) Coefficient δ_i (weight Ω_i)	(2) Data ($\hat{\delta}_i$)	(3) Standard error in the data	(4) Frictionless	(5) Unrestricted switching costs	(6) Compensating differential	(7) Single probability of involuntary choices	(8) Heterogenous probabilities of involuntary choices
Non-agriculture premia: cross-sectional (δ_1) and within-individual (δ_2)							
δ_1 (1)	0.57	(0.03)	0.56	0.60	0.60	0.58	0.48
δ_2 (1)	0.40	(0.05)	0.21	0.35	0.35	0.34	0.40
Premia for switchers to agriculture (δ_3, δ_6) and to non-agriculture. (δ_4, δ_5). The first element in (a, b) is relative to peers post-switch; the second to non-switching workers							
δ_3 (5)	-0.05	(0.06)	-0.05	-0.10	-0.10	-0.07	-0.04
δ_4 (5)	-0.31	(0.05)	-0.41	-0.36	-0.37	-0.35	-0.24
δ_5 (5)	0.15	(0.07)	0.21	0.29	0.31	0.28	0.24
δ_6 (5)	-0.42	(0.06)	-0.21	-0.34	-0.33	-0.36	-0.40
Constant (δ_7) and coefficients on interaction sector and year ($\delta_8 : A \times 2,\ \delta_9 : A \times 3, \ldots \delta_{16} : N \times 5$)							
δ_7 (5)	-0.17	(0.10)	-0.18	-0.20	-0.18	-0.19	-0.18
δ_8 (1)	0.38	(0.07)	0.47	0.45	0.45	0.47	0.41
δ_9 (1)	0.34	(0.07)	0.38	0.27	0.27	0.32	0.38
δ_{10} (1)	0.63	(0.07)	0.56	0.55	0.55	0.56	0.67
δ_{11} (1)	0.85	(0.08)	0.78	0.77	0.78	0.76	0.94
δ_{12} (5)	0.76	(0.06)	0.60	0.66	0.64	0.63	0.70
δ_{13} (1)	1.10	(0.06)	1.06	1.04	1.03	1.04	1.07
δ_{14} (1)	0.89	(0.06)	0.91	0.89	0.88	0.88	0.85
δ_{15} (1)	1.05	(0.06)	1.12	1.17	1.16	1.12	1.03
δ_{16} (1)	1.27	(0.07)	1.33	1.30	1.33	1.34	1.19
Constant (δ_{17}) and coefficients on year dummies ($\delta_{18} : t = 2,\ \delta_{19} : t = 3...$)							
δ_{17} (10)	0.70	(0.01)	0.67	0.70	0.68	0.71	0.67
δ_{18} (10)	0.01	(0.02)	0.00	-0.04	-0.03	0.00	-0.03
δ_{19} (10)	-0.02	(0.02)	-0.09	0.00	-0.02	-0.03	-0.05
δ_{20} (10)	-0.03	(0.02)	-0.04	-0.03	-0.05	-0.02	-0.07
δ_{21} (10)	-0.04	(0.02)	-0.05	-0.12	-0.09	0.01	-0.09
Constant (δ_{22}) and lagged sector choice (δ_{23})							
δ_{22} (10)	0.21	(0.01)	0.20	0.22	0.22	0.23	0.16
δ_{23} (10)	0.68	(0.01)	0.66	0.63	0.62	0.68	0.71
Residual variance of workers in agriculture (δ_{24}) and non-agriculture (δ_{25})							
δ_{24} (3)	1.24	(0.04)	1.01	1.13	1.14	0.99	1.13
δ_{25} (3)	0.95	(0.03)	1.19	1.10	1.12	1.21	1.09
Residual variance of non-switching/switching workers in/to agriculture (δ_{26}, δ_{29}) and in/to non-agriculture (δ_{27}, δ_{28})							
δ_{26} (3)	1.43	(0.06)	1.44	1.59	1.57	1.27	1.44
δ_{27} (3)	1.08	(0.04)	1.56	1.45	1.44	1.31	1.01
δ_{28} (3)	1.73	(0.14)	1.58	1.52	1.54	1.74	1.80
δ_{29} (3)	1.86	(0.14)	1.51	1.52	1.51	1.77	1.83
Overall fit			2.013	1.439	1.462	0.939	0.414

Notes: A description of the auxiliary regressions is done in Table 14. Ω_i refers to the $i-$th element of the diagonal of the matrix Ω.

Table 16: Counterfactual: Aggregate Income

Variable	Notation	Counterfactual
Growth rate (%) in total income: $(1) * (2) * (3)$	$\Delta\%Y_i$	21.5
		(2.3)
(1) Fraction of the population reallocated	m	0.35
		(0.02)
(2) Ratio of average income of reallocated workers to average income	ψ_m	0.57
		(0.02)
(3) Growth rate (%) in total income of reallocated workers	$\Delta\%Y_m$	106.5
		(8.5)

Notes: Results correspond to the counterfactual exercise of eliminating involuntary switches. Standard errors from 100 bootstraps in parentheses.

Table 17: Counterfactual: Sectoral Allocation and Productivity

Variable	Agriculture	Non-Agriculture
Baseline employment share	0.39	0.61
Counterfactual employment share	0.30	0.70
Counterfactual employment growth (%)	-21.0	13.1
Counterfactual output growth (%)	14.2	24.6
Counterfactual productivity growth (%)	44.4	10.1

Notes: Results correspond to the counterfactual exercise of eliminating involuntary switches.

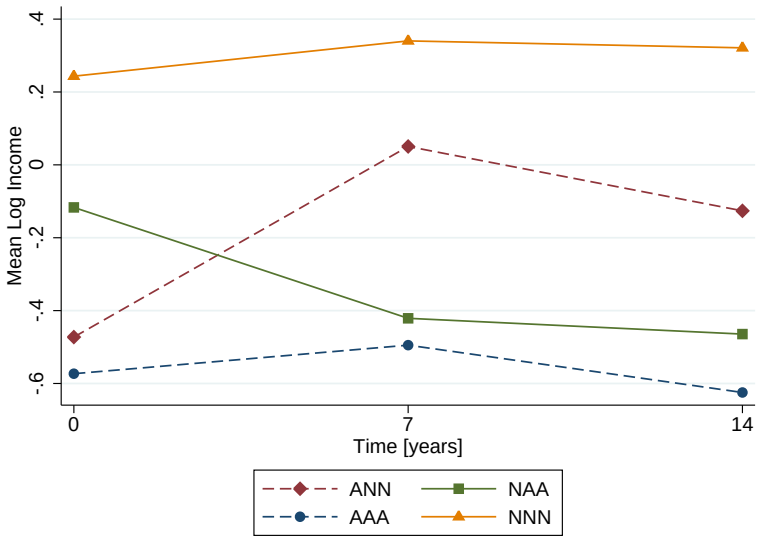
Table 18: Sectoral Premia in Counterfactuals

Coef.	(1) Baseline model	(2) No frictions	(3) No sorting
Non-agriculture premia: cross-sectional (δ_1) and within-individual (δ_2)			
δ_1	0.48	0.18	0.46
δ_2	0.40	-0.31	0.44

Notes: Baseline model is from column (8) of Table 15. No frictions imposes $p^T = p^S = 0$. No sorting imposes $\sigma^2_{\theta A}$, $\sigma^2_{\theta N}$, $\sigma^2_{\varepsilon A}$, $\sigma^2_{\varepsilon N}$ all equal to zero.

Figures

Figure 1: Mean Log Income by Employment History



Notes: Figure plots mean log income (after controlling for year and province fixed effects) by employment history spanned by three observations at 7-year intervals. XYZ indicates that worker was in sector X during the first observation (in 1993 or 2000), in sector Y during the second observation 7 years later (in 2000 or 2007), and in sector Z during the third observation 14 years later (in 2007 or 2014). A - Agriculture, N - Non-Agriculture. For clarity only histories of switchers who stick to their new sector and of always stayers are reported.

Appendix

A Additional Tables

Table A.1: Premia with Hours Worked: Additional Jobs

	(1) Log Income	(2) Log Income	(3) Log Income	(4) Log Income	(5) Log Inc./Hour	(6) Log Inc./Hour
Non-Agriculture	0.501***	0.264***	0.390***	0.216***	0.275***	0.150***
	(0.034)	(0.032)	(0.032)	(0.031)	(0.034)	(0.035)
Urban	0.171***	0.063*	0.143***	0.063**	0.112***	0.057**
	(0.034)	(0.034)	(0.030)	(0.026)	(0.028)	(0.028)
Log Hours/Year			0.509***	0.445***		
			(0.011)	(0.011)		
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes
Indiv. cont.	Yes	Yes	Yes	Yes	Yes	Yes
Individual FE		Yes		Yes		Yes
Observations	44489	44492	43819	43821	43819	43821
R^2	0.514	0.538	0.603	0.615	0.495	0.514

Notes: Income and hours from both the main job and the secondary job. Individual controls: education, experience, experience sq., and sex. Observations weighted by longitudinal survey weights. Standard errors clustered by enumeration areas (primary sampling units of the survey) in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table A.2: Hours Worked

	Base (1) Log Hours	Base (2) Log Hours	Add. Job (3) Log Hours	Add. Job (4) Log Hours
Non-Agriculture	0.286***	0.152***	0.234***	0.119***
	(0.024)	(0.029)	(0.022)	(0.027)
Urban	0.101***	0.014	0.062***	0.012
	(0.020)	(0.031)	(0.018)	(0.032)
Year FE	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes
Indiv. cont.	Yes	Yes	Yes	Yes
Individual FE		Yes		Yes
Observations	43841	43843	43819	43821
R^2	0.053	0.023	0.052	0.026

Notes: *Base* is the baseline specification involving primary job only. *Add. Job* also includes secondary job. Individual controls: education, experience, experience sq., and sex. Observations weighted by longitudinal survey weights. Standard errors clustered by enumeration areas (primary sampling units of the survey) in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table A.3: Premia over Time

(A) Cross-Sectional Premia

	Pooled	1993	1997	2000	2007	2014
	(1)	(2)	(3)	(4)	(5)	(6)
	Log Income	Log Income	Log Income	Log Income	Log Income	Log Income
Non-Agriculture	0.574***	0.792***	0.721***	0.547***	0.461***	0.449***
	(0.036)	(0.070)	(0.052)	(0.051)	(0.048)	(0.058)
Urban	0.207***	0.388***	0.271***	0.227***	0.204***	0.097
	(0.036)	(0.057)	(0.051)	(0.051)	(0.049)	(0.062)
Year FE	Yes					
Province FE	Yes	Yes	Yes	Yes	Yes	Yes
Indiv. cont.	Yes	Yes	Yes	Yes	Yes	Yes
Individual FE						
Observations	44494	5296	8548	10293	10619	9738
R^2	0.503	0.382	0.333	0.244	0.267	0.249

(B) Premia with Worker Fixed Effect

	Pooled	1993-97	1997-00	2000-07	2007-14
	(1)	(2)	(3)	(4)	(5)
	Log Income	Log Income	Log Income	Log Income	Log Income
Non-Agriculture	0.332***	0.339***	0.292***	0.303***	0.217***
	(0.033)	(0.071)	(0.052)	(0.056)	(0.059)
Urban	0.084**	0.210***	0.097	0.156***	0.144**
	(0.032)	(0.068)	(0.087)	(0.058)	(0.058)
Year FE	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes
Indiv. cont.	Yes	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes	Yes
Observations	44497	13844	18841	20912	20360
R^2	0.518	0.242	0.205	0.396	0.282

Notes: *Pooled* is the baseline sample with observations from IFLS 1-5. Panel A: cross-sectional regressions run separately for each survey wave. Panel B: panel regressions run separately for each two consecutive survey waves. Individual controls: education, experience, experience sq., and sex. Observations weighted by longitudinal survey weights. Standard errors clustered by enumeration areas (primary sampling units of the survey) in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

B Recall Bias

Each wave of the IFLS asks respondents about the income they earned over the past year. Throughout the paper we use this contemporaneously recorded income as our main dependent variable. In addition, the survey asks respondents to retrospectively recall employment information for several years prior to the survey. While this recall information can in principle be used to supplement the contemporaneous data and increase the sample size, retrospective survey data is known to raise serious quality concerns (cf. Bound, Brown and Mathiowetz (2001)). For this reason we do not use retrospective income information in our analysis. In this appendix we explain this choice in more detail and argue that it can largely explain why our results differ from those concurrently obtained by Hicks et al. (2017).

The first three columns of the first panel of Table B.1 show the non-agricultural premia estimated on the contemporaneous data recorded by the IFLS. These numbers are similar to those reported in Table 11 (columns 3, 4, and 6) in the main text, but not identical because the specifications and sample are modified to ease comparison with Hicks et al. (2017). In particular, we discard the information from the most recent wave of IFLS as it has not been incorporated by these authors. Columns 4-6 show the corresponding premia estimated on data from retrospective recall. Compared to the contemporaneous estimates, the cross-sectional premium (controlling for hours) drops from 71 lp to 53 lp, premium with worker fixed effects (controlling for hours) drops from 25 lp to 11 lp, and the 19 lp premium in terms of income per hour (with worker FE) disappears entirely.

These patterns are not surprising in light of research on biases arising in recall surveys. One such well documented bias is that past income reported by workers is biased towards their usual income.[25] For example, Gibson and Kim (2010) show the extent of this bias for US wage workers by comparing their self-reported retrospective earnings with administrative records. They also demonstrate that underreporting transitory income changes generates non-classical measurement error that biases the regression coefficients towards zero if the mismeasured variable is the dependent variable. This result is consistent with the reduced non-agricultural premia we find using recall data if workers cannot accurately recall how much higher their income was in years in which they worked in non-agriculture. Furthermore, the problem is likely to be exacerbated when the identifying variation comes from changes in income of individual workers over time. This would explain why the fall in the premium is proportionately much larger in the specification with worker fixed effects. Finally, the problems with measurement error are likely to be compounded when the dependent variable is constructed by dividing reported income by reported hours. That retrospectively recalled hours are unreliable is suggested by comparing coefficients on hours in columns 5 and 2. The elasticity of income with respect to hours implied by column 5 is less than 0.15, only 1/3 of the 0.44 elasticity implied by the corresponding column 2 for contemporaneous data. The implausibly low elasticity for recalled hours indicates that their relationship to income should be treated with great caution in recall data. In a rare validation study observing both hours worked and earnings, Duncan and Hill (1985) find that *"interview reports of average hourly earnings, obtained by dividing the interview reports of annual earnings by reports of annual work hours, appeared to be exceedingly unreliable"*

[25] Another bias with similar implications in this context is an anchoring bias, where respondents use an answer to a previously answered question as a mental anchor for subsequent answers. Godlonton, Hernandez and Murphy (2016) find strong evidence of this behavior in a survey of Central American farmers: retrospectively recalled income correlates more highly with current income (about which the respondents are asked first) than with income over the recall period that had been reported contemporaneously in the past. This type of cognitive bias is likely to be present in IFLS too, since IFLS also first asks about contemporaneous income and then asks respondents to retrospectively recall past income.

and caution against their use. The particularly low signal-to-noise ratio in income per hour derived from retrospective data can explain why the results become insignificant in column 6.

The take-away message from this discussion is that using data from retrospective recall in our application would introduces biases in our key results. These recall biases can be strong in IFLS since the respondents are asked retrospective questions about multiple years prior to the survey (up to a maximum of 10 years), and the quality of recall information deteriorates with time elapsed from the pertaining event (see, e.g. de Nicola and Gine (2014) in a developing country context). There are no obvious offsetting benefits to including the retrospective data. Statistical power, in particular, is not an issue, since the baseline sample of contemporaneous responses is large enough to allow us to estimate the key non-agricultural premium precisely.

We conclude this appendix by showing that the inclusion of retrospective data is likely the main reason why the substantive results on the strength of the non-agricultural premium reported by Hicks et al. (2017) are different than ours. In contrast to our results, they argue that the non-agricultural premium in Indonesia mostly disappears once individual fixed effects are allowed for. To aid comparison, columns 1-3 in the second panel of Table B.1 repeat the same exercise as columns 1-3 and 4-6 in panel A, but now on a sample pooling the contemporaneous and retrospective responses. The estimates lie roughly half way between the two corresponding numbers reported in the first panel. This means that the pooled-sample estimates are significantly attenuated relative to those based on better-measured contemporaneous data that we favor. For comparison, columns 4-6 copy the corresponding estimates from Hicks et al. (2017) (columns 2, 6, and 7 of their Table 5A), who use pooled contemporaneous and retrospective data. While we cannot replicate their results exactly without detailed knowledge of their data processing protocol, the estimates in columns 1-3 come close. Based on this exercise, we expect that their results would much have been much more in line with ours had they not used the retrospective data.[26]

C Estimation Procedure

This Appendix presents some technical aspects about the estimation procedure. The vector of structural parameters in the frictionless economy, denoted by Θ, is constituted by the following set of 21 elements: $\{R_t^s, \beta, \sigma^2_{\theta^s}, \sigma_{\theta^{AN}}, \sigma^2_{\varepsilon^s}, \sigma^2_\nu\}$ for $t = 1, ..5$ and $s = A, N$ (denoting agriculture and non-agriculture, respectively). In this set, $\sigma^2_{\theta^s}$ and $\sigma^2_{\varepsilon^s}$ denote the variances in Σ_θ and Σ_ε respectively, $\sigma_{\theta^{AN}}$ the covariance in Σ_θ, and β is comprised of the Mincerian returns on the five mentioned covariates, denoted β_{sex}, β_{loc}, β_{edu}, β_{exp} and β_{exp2}, respectively. For the model with switching costs, Θ is augmented by $\left\{\phi^{AN}, \phi^{NA}\right\}$, whereas for the model with compensating differentials Θ is augmented by *cd*. The Indirect Inference loss function, denoted by $Q(\Theta)$, is computed as the weighted sum of the squared differences between the values in $\hat{\delta}$ and the values for those obtained from simulations of the structural model, that is:

$$Q(\Theta) = \left(\hat{\delta} - \hat{\delta}^s(\Theta)\right)' \Omega \left(\hat{\delta} - \hat{\delta}^s(\Theta)\right)$$

where $\hat{\delta}^s(\Theta)$ corresponds to the same vector of selected coefficients of the auxiliary models estimated with data simulated from the structural model with parameters Θ, and Ω is a diagonal weighting

[26] Furthermore, their headline result depends on using income per hour as their preferred measure. We do not use hours data in our preferred specifications, both because of measurement issues for hours described in this appendix and conceptual issues discussed in section 3.2.5.

Table B.1: Retrospective Recall

(A) Contemporaneous vs. Recall Data

	Contemporaneous			Retrospective		
	(1) Log Inc.	(2) Log Inc.	(3) Log Inc./Hr	(4) Log Inc.	(5) Log Inc.	(6) Log Inc./Hr
Non-Agriculture	0.707***	0.245***	0.192***	0.525***	0.110***	-0.038
	(0.013)	(0.022)	(0.024)	(0.020)	(0.039)	(0.052)
Log Hours	0.604***	0.462***		0.140***	-0.012	
	(0.039)	(0.046)		(0.051)	(0.045)	
Log Hours Squared	0.000	-0.002		0.018***	0.016***	
	(0.005)	(0.005)		(0.006)	(0.005)	
Age squared		-0.000***	-0.000***		-0.001***	-0.000***
		(0.000)	(0.000)		(0.000)	(0.000)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Individual FE		Yes	Yes		Yes	Yes
Observations	48626	48626	48626	63498	63498	63498
R-sq	0.423	0.540	0.433	0.161	0.192	0.158

(B) Pooled Data vs. Hicks et al. (2017)

	Pooled Data			Hicks et al. (2017)		
	(1) Log Inc.	(2) Log Inc.	(3) Log Inc./Hr	(4) Log Inc.	(5) Log Inc.	(6) Log Inc./Hr
Non-Agriculture	0.588***	0.173***	0.076***	0.514***	0.171***	0.047
	(0.015)	(0.019)	(0.021)	(0.016)	(0.025)	(0.031)
Log Hours	0.385***	0.206***		0.531**	0.323***	
	(0.040)	(0.037)		(0.025)	(0.034)	
Log Hours Squared	0.006	0.009**		-0.021***	-0.014**	
	(0.005)	(0.004)		(0.005)	(0.006)	
Age squared		-0.000***	-0.000***		-0.001***	-0.000***
		(0.000)	(0.000)		(0.000)	(0.000)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Individual FE		Yes	Yes		Yes	Yes
Observations	107933	107933	107933	115897	115897	115897
R-sq	0.303	0.353	0.263			

Notes: *Contemporaneous* measures based on values reported for last year. *Retrospective* measures obtained from recall part of the survey. *Pooled Data* combines contemporaneous and retrospective observations. Sample restricted to IFLS 1-4. Sample includes all individuals with at least one observation of income and hours worked. *Income* is average monthly labor income from primary and secondary job. Contemporaneous income obtained by dividing annual income by 12. *Hours* are average monthly hours from primary and secondary job obtained as (weeks worked per year)*(normal hours per week)/12. Observations are not weighted. Standard errors clustered at the individual level in parentheses. Significance levels: * p<0.10, ** p<0.05, *** p<0.01. Columns 4-7 in Panel B are columns 2, 6, 7, respectively, from Table 5A in Hicks et al. (2017).

matrix. For weights, we use factors that represent the importance of the estimated coefficient in the identification of the structural parameters of the model. The values of those factors, displayed in the second column of Table 15, were assigned after extensive experimentation with simulations of the model. We proceed next to comment on their magnitudes, particularly for those weights that differ from one.

As we argue in the main text, the within-individual variation is key for identification. Our priority in the estimation procedure is to make the structural model able to deliver the observed premia for switchers in the data. Procuring the right amount of switchers in each year is crucial because of their size depends the precision of the obtained premia. For this reason, the coefficients of the linear probability models have the largest weights in the loss function, by a factor of 10. Moreover, we also impose larger weights (by a factor of 5) to the two sets of switchers' sectoral premia in which we are interested in. First, to the premia in the model in differences iv), since this regression actually forces the structural model to deliver both the gains of switchers to non-agriculture and the cuts in income of workers switching to agriculture, allowing the estimation procedure to identify any possible asymmetry in the switching costs across sectors. Second, to the premia in model iv), which compares the average performance of a switching worker to each sector with their peer group after the switch, informing about the nature of sorting. In addition, we know that the estimated coefficients of the interactions in model iv) help to identify the growth in relative human capital prices. Thus, the information about the full path of these prices can be recovered once the regression pin down the conditional expected income in each sector in the first year. For this reason, we force a greater accuracy in the coefficients of the constant and the interaction term of non-agriculture and the first year through larger weights, by a factor of 5. Finally, given the importance of the residual variances to identify the joint distribution of comparative advantage, they also have larger weights in the loss function, in this case by a factor of 3.

We minimize $Q(\Theta)$ using in each evaluation H different simulated samples each with size equal to the number of observations in the balanced panel ($N \times T = 8760$)[27]. For observables, we take in each simulated sample the same values we observe in the data. To choose H, we numerically explore how $Q(\Theta)$ varies in a fixed number of simulations only due to changes in the seed of the random numbers, as a function of the number of individuals. We found that the range of variation of $Q(\Theta)$ starts to stabilize after we include 80000 individuals. Thus, we choose $H = 40 \approx 80000/1752$.

Optimizing $Q(\Theta)$ is challenging since the discrete choices in the selection model create a non-smooth function, that behaves as a step function in some regions of the parameters' space. To deal with this problem, we use an algorithm with repeated iterations of an evolutionary method, particularly particles-swarm optimization, to find the solution[28]. We start with 16 implementations of particles-swarm optimization in a wide range of the feasible parameter space. In each of these 16 implementations we work with 96 particles, that initially are randomly and uniformly distributed. In a second stage, we perform 8 implementations of particles-swarm optimization in a range of the parameter space bounded by the smallest and largest solution for each parameter in the first stage, plus a parameter-dependent margin error. In this stage we use the same number and distribution

[27]We use the approach to compute the solution on the sample generated by $H \times N \times T$ observations, a method that is equivalent to compute the average of H times the solution of each sample of $N \times T$ observations, although computationally it is more efficient.

[28]Other possibilities recently developed in the literature include the use of a logistic-kernel of simulated latent utilities instead of endogenous variables (Bruins et al., 2018) or Monte Carlo importance sampling (Sauer and Taber, 2017). However, the possibility to use those techniques is model-dependent. As we argue next, our algorithm does not face problems to find an accurate solution.

for the initial particles. Finally we perform an additional optimization in which we initialize 8 of the 96 particles in the solutions found in the second stage. The estimate $\hat{\Theta}$ is the solution that minimizes $Q(\Theta)$ in the 25 implementations described of particles-swarm optimization. Using several numerical simulations we test that our algorithm ensures two-decimal accuracy in the solutions, as opposed to alternative optimization techniques.

D Identification

In this Appendix we demonstrate how parameters in Θ are identified for the same model as in our baseline specification, but with only two periods and abstracting from the effect of observables in income. For illustrative purposes, we first show how identification is achieved in the frictionless economy, and next we proceed to the model with switching costs. We refer to the Agriculture sector as A and to the Non-Agriculture sector as N. We denote $r_t = r_r^N - r_t^A$, $u_{it}^s = \theta_i^s + \varepsilon_{it}^s$ for $s = A, N$ and $\tilde{\sigma}_{ks}^2 = \sigma_{ks}^2 - \sigma_{kAN}$ for $k = u, \theta$ and $s = A, N$. Notice that $\sigma_{us}^2 = \sigma_{\theta s}^2 + \sigma_{\varepsilon s}^2$ for $s = A, N, AN$. Further, we denote the st. dev. of $u_{it} \equiv \left(u_{it}^A - u_{it}^N\right)$ as $\sigma_u^* = \sqrt{\tilde{\sigma}_{uA}^2 + \tilde{\sigma}_{uN}^2}$ and the st. dev. of $\theta_i \equiv \left(\theta_i^A - \theta_i^N\right)$ as $\sigma_\theta^* = \sqrt{\tilde{\sigma}_{\theta A}^2 + \tilde{\sigma}_{\theta N}^2}$.

Model for the Frictionless Economy

Without switching costs, sectoral decisions do not depend on workers' histories, so the model behaves in each period t as the standard Roy model with comparative advantage u_{it}^s, where, excluding the variance of measurement error, we can identify the variance matrix Σ_u and the prices of human capital r_t^s from cross-sectional data. This is consequence of the normality assumptions on the distribution of both $\left(\theta_i^A, \theta_i^N\right)$ and $\left(\varepsilon_{it}^A, \varepsilon_{it}^N\right)$, which imply that in each period $\left(u_{it}^A, u_{it}^N\right)$ is joint normally distributed with variance Σ_u, and hence standard arguments of Heckman and Honoré (1990) for identification in the normal case can be applied. However, only with panel data we can decompose Σ_u into Σ_θ and Σ_ε, the variances of the permanent and transitory components respectively, and identify σ_ν^2, the variance of measurement error, using the information obtained from the switching workers in the panel.

Letting $\lambda(\cdot) = \frac{\phi(\cdot)}{\Phi(\cdot)}$ with ϕ and Φ the PDF and CDF of a standard normal, and using properties of normal random variables following Heckman and Honoré (1990), we can obtain the following derivations for the first three observed moments of the income distribution in each period t:

$$P(t = N) = \Phi\left(\frac{r_t}{\sigma_u^*}\right) \tag{A.1}$$

$$E\left(y_{it}^N \mid t = N\right) = r_t^N + \frac{\tilde{\sigma}_{uN}^2}{\sigma_u^*}\lambda\left(\frac{r_t}{\sigma_u^*}\right) \tag{A.2}$$

$$E\left(y_{it}^A \mid t = A\right) = r_t^A + \frac{\tilde{\sigma}_{uA}^2}{\sigma_u^*}\lambda\left(\frac{-r_t}{\sigma_u^*}\right) \tag{A.3}$$

$$Var\left(y_{it}^N \mid t = N\right) = \sigma_{uN}^2 + \left(\frac{\tilde{\sigma}_{uN}^2}{\sigma_u^*}\right)^2\left[-\lambda\left(\frac{r_t}{\sigma_u^*}\right)\frac{r_t}{\sigma_u^*} - \lambda^2\left(\frac{r_t}{\sigma_u^*}\right)\right] + \sigma_\nu^2 \tag{A.4}$$

$$Var\left(y_{it}^A \mid t = A\right) = \sigma_{uA}^2 + \left(\frac{\tilde{\sigma}_{uA}^2}{\sigma_u^*}\right)^2\left[\lambda\left(\frac{-r_t}{\sigma_u^*}\right)\frac{r_t}{\sigma_u^*} - \lambda^2\left(\frac{-r_t}{\sigma_u^*}\right)\right] + \sigma_\nu^2 \tag{A.5}$$

$$E\left(\left[y_{it}^{N}-E\left(y_{it}^{N}\,|\,t=N\right)\right]^{3}\,|\,t=N\right)=\left(\frac{\tilde{\sigma}_{uN}^{2}}{\sigma_{u}^{*}}\right)^{3}\lambda\left(\frac{r_{t}}{\sigma_{u}^{*}}\right)\left[2\lambda^{2}\left(\frac{r_{t}}{\sigma_{u}^{*}}\right)+3\lambda\left(\frac{r_{t}}{\sigma_{u}^{*}}\right)\frac{r_{t}}{\sigma_{u}^{*}}+\left(\frac{r_{t}}{\sigma_{u}^{*}}\right)^{2}-1\right] \quad \text{(A.6)}$$

$$E\left(\left[y_{it}^{A}-E\left(y_{it}^{A}\,|\,t=A\right)\right]^{3}\,|\,t=A\right)=\left(\frac{\tilde{\sigma}_{uA}^{2}}{\sigma_{u}^{*}}\right)^{3}\lambda\left(\frac{-r_{t}}{\sigma_{u}^{*}}\right)\left[2\lambda^{2}\left(\frac{-r_{t}}{\sigma_{u}^{*}}\right)-3\lambda\left(\frac{-r_{t}}{\sigma_{u}^{*}}\right)\frac{r_{t}}{\sigma_{u}^{*}}+\left(\frac{r_{t}}{\sigma_{u}^{*}}\right)^{2}-1\right] \quad \text{(A.7)}$$

With information of $T=2$ repeated cross sections to compute the LHS of this system of 14 equations, we can identify r_t^N, r_t^A and the combination $\sigma_\nu^2\mathbb{I}_2+\Sigma_u$ (8 parameters). Let us now show the additional information we can obtain from panel data. We exploit the property that $(u_{it},u_{it'})$ for $t'\neq t$ and (u_{it},u_{it}^s) for $s=A,N$ are joint normally distributed, since each element is the sum of two normally distributed random variables. Denoting the CDF of a bivariate normal distribution with mean $\mathbf{0}'$ and variance Σ evaluated at the vector A as $\mathbf{\Phi}\left(A,\Sigma\right)$, the probability of transition from N to N is given by:

$$\begin{aligned}P(2=N,1=N)=&P\left(\left\{r_2^A+u_{i2}^A<r_2^N+u_{i2}^N\right\},\left\{r_1^A+u_{i1}^A<r_1^N+u_{i1}^N\right\}\right)\\=&P\left(\left\{u_{i2}<r_2\right\},\left\{u_{i1}<r_1\right\}\right)\\=&\mathbf{\Phi}\left(\overrightarrow{r}_{NN},\Sigma_T\right)\end{aligned} \quad \text{(A.8)}$$

with $\overrightarrow{r}_{NN}=[r_2,r_1]'$ and $\Sigma_T=\begin{bmatrix}\sigma_u^{*2} & \sigma_\theta^{*2}\\ \sigma_\theta^{*2} & \sigma_u^{*2}\end{bmatrix}$. The probability of transition from N to A is given by:

$$\begin{aligned}P\left(2=A,1=N\right)=&P\left(\left\{-u_{i2}<-r_2\right\},\left\{u_{i1}<r_1\right\}\right)\\=&\mathbf{\Phi}\left(\overrightarrow{r}_{NA},\Sigma_W\right)\end{aligned} \quad \text{(A.9)}$$

with $\overrightarrow{r}_{NA}=[-r_2,r_1]'$ and $\Sigma_W=\begin{bmatrix}\sigma_u^{*2} & -\sigma_\theta^{*2}\\ -\sigma_\theta^{*2} & \sigma_u^{*2}\end{bmatrix}$. Similarly:

$$P\left(2=N,1=A\right)=\mathbf{\Phi}\left(\overrightarrow{r}_{AN},\Sigma_W\right) \quad \text{(A.10)}$$

$$P\left(2=N,2=A\right)=\mathbf{\Phi}\left(\overrightarrow{r}_{AA},\Sigma_T\right) \quad \text{(A.11)}$$

with $\overrightarrow{r}_{AN}=[r_2,-r_1]'$ and $\overrightarrow{r}_{AA}=[-r_2,-r_1]'$.

Now consider the values of the expected income in the second period for each transition group of workers. In the frictionless economy we do not need directly those expected values, but we illustrate here how to compute them to introduce some notation that we use hereafter. The income of stayers in N in period 2 is given by:

$$\begin{aligned}E\left(y_{i2}^N|\,2=N,1=N\right)=&r_2^N+E\left(u_{i2}^N|\,2=N,1=N\right)\\=&r_2^N+E\left(u_{i2}^N|\,\left\{u_{i2}<r_2\right\},\left\{u_{i1}<r_1\right\}\right)\end{aligned}$$

Notice that the expected value in the second term of the RHS can be expressed as:

$$E\left(X_1^{k_1}X_2^{k_2}X_3^{k_3}|\,-\infty<X_i<b_i,i=1,2,3\right) \quad \text{(A.12)}$$

with $X_1 = u_{i2}^N$, $X_2 = u_{i2}$, $X_3 = u_{i1}$, $k_1 = 1, k_2 = k_3 = 0$, $b_1 = \infty$, $b_2 = r_2$ and $b_3 = r_1$. This expected value is the moment of the upper truncated multivariate normal distribution with mean $\mathbf{0}$ and variance:

$$\Sigma_{NN} = \begin{bmatrix} \sigma_{uN}^2 & -\tilde{\sigma}_{uN}^2 & -\tilde{\sigma}_{\theta N}^2 \\ -\tilde{\sigma}_{uN}^2 & \sigma_u^{*2} & \sigma_\theta^{*2} \\ -\tilde{\sigma}_{\theta N}^2 & \sigma_\theta^{*2} & \sigma_u^{*2} \end{bmatrix} = \begin{bmatrix} \sigma_{uN}^2 & \Lambda_{NN} \\ \Lambda_{NN}' & \Sigma_T \end{bmatrix}$$

where the vector Λ_{NN} is defined as $\Lambda_{NN} = [-\tilde{\sigma}_{uN}^2, -\tilde{\sigma}_{\theta N}^2]$. In general terms, we can denote the expected value in (A.12) for the particular case $k_2 = k_3 = 0$ and $b_1 = \infty$ as the function $\mathbf{M}_3(\cdot)$ of the variance matrix Σ, the elements b_2 and b_3 stacked up in a vector B and the coefficient $k = k_1$, that is:

$$\mathbf{M}_3(B, \Sigma, k) \equiv E\left(X_1^k | \ -\infty < X_1 < \infty, -\infty < X_2 < B_1, -\infty < X_3 < B_2\right)$$

with $\{X_1, X_2, X_3\} \sim \mathbf{N}(\mathbf{0}, \Sigma)$. Then we can rewrite:

$$E\left(y_{i2}^N \,|\, 2 = N, 1 = N\right) = r_2^N + \mathbf{M}_3(\overrightarrow{r}_{NN}, \Sigma_{NN}, 1)$$

To evaluate $\mathbf{M}_3(\cdot)$ we can use for example the recurrence relations developed by Kan and Robotti (2017) to compute numerically the moment generating function of the truncated multivariate normal distribution (first obtained by Tallis (1961))[29]. Following similar arguments, we can show that the income of each transition group in period 2 is given by:

$$\begin{aligned} E\left(y_{i2}^A | \, 2 = N, 1 = N\right) =& r_2^A + \mathbf{M}_3(\overrightarrow{r}_{NA}, \Sigma_{NA}, 1) \\ E\left(y_{i2}^N | \, 2 = N, 1 = A\right) =& r_2^N + \mathbf{M}_3(\overrightarrow{r}_{AN}, \Sigma_{AN}, 1) \\ E\left(y_{i2}^A | \, 2 = A, 1 = A\right) =& r_2^A + \mathbf{M}_3(\overrightarrow{r}_{AA}, \Sigma_{AA}, 1) \end{aligned}$$

with $\Sigma_{NA} = \begin{bmatrix} \sigma_{uA}^2 & \Lambda_{NA} \\ \Lambda_{NA}' & \Sigma_W \end{bmatrix}$, $\Sigma_{AN} = \begin{bmatrix} \sigma_{uN}^2 & \Lambda_{AN} \\ \Lambda_{AN}' & \Sigma_W \end{bmatrix}$ and $\Sigma_{AA} = \begin{bmatrix} \sigma_{uA}^2 & \Lambda_{AA} \\ \Lambda_{AA}' & \Sigma_T \end{bmatrix}$ where $\Lambda_{NA} = [-\tilde{\sigma}_{uA}^2, \tilde{\sigma}_{\theta A}^2]$, $\Lambda_{AN} = [-\tilde{\sigma}_{uN}^2, \tilde{\sigma}_{\theta N}^2]$ and $\Lambda_{AA} = [-\tilde{\sigma}_{uA}^2, -\tilde{\sigma}_{\theta A}^2]$. Notice that the second and third moments of each transition group can be computed as functions of $\mathbf{M}_3(\cdot, \cdot, 2)$ and $\mathbf{M}_3(\cdot, \cdot, 3)$ respectively. We do not need those expressions here, so we deduce those moments only for the model with switching costs.

Now let us compute the moments of the growth in income for switchers. For switching workers from A to N, the first moment is:

$$\begin{aligned} E\left(y_{i2}^N - y_{i1}^A | \, 2 = N, 1 = A\right) &= r_2^N - r_1^A + E\left(u_{i2}^N - u_{i1}^A | \, 2 = N, 1 = A\right) \\ &= r_2^N - r_1^A + E\left(u_{i2}^N - u_{i1}^A | \ \{u_{i2} < r_2\}, \{-u_{i1} < -r_1\}\right) \\ &= r_2^N - r_1^A + \mathbf{M}_3(\overrightarrow{r}_{AN}, \tilde{\Sigma}_{AN}, 1) \end{aligned} \tag{A.13}$$

with $\tilde{\Sigma}_{AN} = \begin{bmatrix} \sigma_{uA}^2 + \sigma_{uN}^2 - 2\sigma_{\theta AN} & \tilde{\Lambda}_{AN} \\ \tilde{\Lambda}_{AN}' & \Sigma_W \end{bmatrix}$ and $\tilde{\Lambda}_{AN} = [-\tilde{\sigma}_{uN}^2 - \tilde{\sigma}_{\theta A}^2, \tilde{\sigma}_{uA}^2 + \tilde{\sigma}_{\theta N}^2]$. Similarly, the

[29] Particularly, we can use the function *multivutmom* developed by Kan and Robotti (2017) in the Matlab package ftnorm. The instruction to compute $\mathbf{M}_3(B, \Sigma, k)$ is simply *multivutmom*$([k\ 0\ 0], [\text{inf}\ \ B_1\ \ B_2], [0\ 0\ 0], \Sigma)$.

expected value of the growth in income for switchers from N to A is:

$$E\left(y_{i2}^{A}-y_{i1}^{N}|\,2=A,1=N\right)=r_{2}^{A}-r_{1}^{N}+\mathbf{M}_{3}(\overrightarrow{r}_{NA},\tilde{\Sigma}_{NA},1) \tag{A.14}$$

where $\tilde{\Sigma}_{NA}=\begin{bmatrix}\sigma_{uA}^{2}+\sigma_{uN}^{2}-2\sigma_{\theta AN} & \tilde{\Lambda}_{NA}\\ \tilde{\Lambda}_{NA}' & \Sigma_{W}\end{bmatrix}$ and $\tilde{\Lambda}_{NA}=[-\tilde{\sigma}_{uA}^{2}-\tilde{\sigma}_{\theta N}^{2},\tilde{\sigma}_{uN}^{2}+\tilde{\sigma}_{\theta A}^{2}]$. The variances of the growth in income for switchers are defined as:

$$\begin{aligned}
&Var\left(y_{i2}^{N}-y_{i1}^{A}|\,2=N,1=A\right)\\
&=E\left(\left(u_{i2}^{N}-u_{i1}^{A}\right)^{2}|\,\{u_{i2}<r_{2}\},\{-u_{i1}<-r_{1}\}\right)-E\left(u_{i2}^{N}-u_{i1}^{A}|\,2=N,1=A\right)^{2}+2\sigma_{\nu}^{2}\\
&=\mathbf{M}_{3}(\overrightarrow{r}_{AN},\tilde{\Sigma}_{AN},2)+\left(\mathbf{M}_{3}(\overrightarrow{r}_{AN},\tilde{\Sigma}_{AN},1)\right)^{2}+2\sigma_{\nu}^{2}
\end{aligned} \tag{A.15}$$

And similarly:

$$Var\left(y_{i2}^{A}-y_{i1}^{N}|\,2=A,1=N\right)=\mathbf{M}_{3}(\overrightarrow{r}_{NA},\tilde{\Sigma}_{NA},2)+\left(\mathbf{M}_{3}(\overrightarrow{r}_{NA},\tilde{\Sigma}_{NA},1)\right)^{2}+2\sigma_{\nu}^{2} \tag{A.16}$$

The system of 22 equations (A.1)-(A.11) and (A.13)-(A.16) has a unique solution for the 10 elements of Θ. We verified this after extensive experimentation using global solvers over a broad range of feasible values for Θ. This shows that the cross-sectional moments, the transition probabilities across waves for each group of workers and the two first moments of the income growth for switchers are enough moments to identify the full set of parameters.

Model with Switching Costs

In the model with switching costs, we will require exactly the same set of 22 moments computed above to identify the 12 elements of Θ. The difficulty to obtain expressions for those moments relies on the fact that sectoral decisions depend now on workers' histories, and hence all moments, including the cross-sectional ones, depend on the income distributions of the previous periods. We deduce here the general rules to deduce expressions for those moments. Denote the CDF of a multivariate normal distribution with mean $\mathbf{0}'$ and variance Σ evaluated at the vector A as $\Phi(A,\Sigma)$. To compute the cross sectional moments in period 1, we need first the distribution of sectoral choices, that depend on the frictionless decisions in period zero. The probability of choosing non-agriculture in period 1 is:

$$\begin{aligned}
P(1=N)=&P(1=N,0=A)+P(1=N,0=N)\\
=&P\left(\left\{r_{1}^{A}+u_{i1}^{A}<r_{1}^{N}+u_{i1}^{N}-\ln\phi^{AN}\right\},\left\{r_{0}^{N}+u_{i0}^{N}<r_{0}^{A}+u_{i0}^{A}\right\}\right)\\
&+P\left(\left\{r_{1}^{A}+u_{i1}^{A}-\ln\phi^{NA}<r_{1}^{N}+u_{i1}^{N}\right\},\left\{r_{0}^{A}+u_{i0}^{A}<r_{0}^{N}+u_{i0}^{N}\right\}\right)\\
=&P\left(\left\{u_{i1}<r_{1}-\ln\phi^{AN}\right\},\{-u_{i0}<-r_{0}\}\right)+P\left(\left\{u_{i1}<r_{1}+\ln\phi^{NA}\right\},\{u_{i0}<r_{0}\}\right)\\
=&\Phi(\overrightarrow{r}_{AN},\Sigma_{W})+\Phi(\overrightarrow{r}_{NN},\Sigma_{T})
\end{aligned}$$

where now : $\overrightarrow{r}_{AN} = \left[r_1 - \ln \phi^{AN}, -r_0\right]', \overrightarrow{r}_{NN} = \left[r_1 + \ln \phi^{NA}, r_0\right]'$ and Σ_W and Σ_T as in the model without switching costs. The values of the expected income in N the first period are:

$$\begin{aligned}
&E\left(y_{i1}^N | \, 1 = N\right) \\
=&r_1^N + E\left(u_{i1}^N | \, 1 = N\right) \\
=&r_1^N + \frac{E\left(u_{i1}^N | \, 1 = N, 0 = A\right) P\left(1 = N, 0 = A\right) + E\left(u_{i1}^N | \, 1 = N, 0 = N\right) P\left(1 = N, 0 = N\right)}{P(1 = N)} \\
=&r_1^N + \frac{\mathbf{M_3}(\overrightarrow{r}_{AN}, \Sigma_{AN}, 1)\Phi\left(\overrightarrow{r}_{AN}, \Sigma_W\right) + \mathbf{M_3}(\overrightarrow{r}_{NN}, \Sigma_{NN}, 1)\Phi\left(\overrightarrow{r}_{NN}, \Sigma_T\right)}{\Phi\left(\overrightarrow{r}_{AN}, \Sigma_W\right) + \Phi\left(\overrightarrow{r}_{NN}, \Sigma_T\right)}
\end{aligned}$$

with $\mathbf{M}_3(\cdot)$, Σ_{AN}, Σ_{NN} as in the model without switching costs. Similarly:

$$E\left(y_{i1}^A | \, 1 = A\right) = r_1^A + \frac{\mathbf{M_3}(\overrightarrow{r}_{AA}, \Sigma_{AA}, 1)\Phi\left(\overrightarrow{r}_{AA}, \Sigma_T\right) + \mathbf{M_3}(\overrightarrow{r}_{NA}, \Sigma_{NA}, 1)\Phi\left(\overrightarrow{r}_{NN}, \Sigma_T\right)}{\Phi\left(\overrightarrow{r}_{AA}, \Sigma_T\right) + \Phi\left(\overrightarrow{r}_{NA}, \Sigma_W\right)}$$

where $\overrightarrow{r}_{AA} = \left[-r_1 + \ln \phi^{AN}, -r_0\right]', \overrightarrow{r}_{NA} = \left[-r_1 - \ln \phi^{NA}, r_0\right]'$ and Σ_{AA}, Σ_{NA} as in the model without switching costs.

The variances can be computed simply by:

$$\begin{aligned}
&Var\left(y_{i1}^N | \, 1 = N\right) \\
=&\frac{\mathbf{M_3}(\overrightarrow{r}_{AN}, \Sigma_{AN}, 2)\Phi\left(\overrightarrow{r}_{AN}, \Sigma_W\right) + \mathbf{M_3}(\overrightarrow{r}_{NN}, \Sigma_{NN}, 2)\Phi\left(\overrightarrow{r}_{NN}, \Sigma_T\right)}{\Phi\left(\overrightarrow{r}_{AN}, \Sigma_W\right) + \Phi\left(\overrightarrow{r}_{NN}, \Sigma_T\right)} - E\left(u_{i1}^N | \, 1 = N\right)^2 + \sigma_\mu^2 \\
&Var\left(y_{i1}^A | \, 1 = A\right) \\
=&\frac{\mathbf{M_3}(\overrightarrow{r}_{AA}, \Sigma_{AA}, 2)\Phi\left(\overrightarrow{r}_{AA}, \Sigma_T\right) + \mathbf{M_3}(\overrightarrow{r}_{NA}, \Sigma_{NA}, 2)\Phi\left(\overrightarrow{r}_{NN}, \Sigma_T\right)}{\Phi\left(\overrightarrow{r}_{AA}, \Sigma_T\right) + \Phi\left(\overrightarrow{r}_{NA}, \Sigma_W\right)} - E\left(u_{i1}^A | \, 1 = A\right)^2 + \sigma_\mu^2
\end{aligned}$$

The third central moments are computed as:

$$\begin{aligned}
&E\left(\left[y_{i1}^N - E\left(y_{i1}^N | \, 1 = N\right)\right]^3 | \, 1 = N\right) \\
=&\frac{\mathbf{M_3}(\overrightarrow{r}_{AN}, \Sigma_{AN}, 3)\Phi\left(\overrightarrow{r}_{AN}, \Sigma_W\right) + \mathbf{M_3}(\overrightarrow{r}_{NN}, \Sigma_{NN}, 3)\Phi\left(\overrightarrow{r}_{NN}, \Sigma_T\right)}{\Phi\left(\overrightarrow{r}_{AN}, \Sigma_W\right) + \Phi\left(\overrightarrow{r}_{NN}, \Sigma_T\right)} \\
&- 3E\left(u_{i1}^N | \, 1 = N\right)\left[Var\left(y_{i1}^N | \, 1 = N\right) - \sigma_\mu^2\right] - \left[E\left(u_{i1}^N | \, 1 = N\right)\right]^3 \\
&E\left(\left[y_{i1}^N - E\left(y_{i1}^N | \, 1 = N\right)\right]^3 | \, 1 = N\right) \\
=&\frac{\mathbf{M_3}(\overrightarrow{r}_{AA}, \Sigma_{AA}, 3)\Phi\left(\overrightarrow{r}_{AA}, \Sigma_T\right) + \mathbf{M_3}(\overrightarrow{r}_{NA}, \Sigma_{NA}, 3)\Phi\left(\overrightarrow{r}_{NN}, \Sigma_T\right)}{\Phi\left(\overrightarrow{r}_{AA}, \Sigma_T\right) + \Phi\left(\overrightarrow{r}_{NA}, \Sigma_W\right)} \\
&- 3E\left(u_{i1}^A | \, 1 = A\right)\left[Var\left(y_{i1}^N | \, 1 = A\right) - \sigma_\mu^2\right] - \left[E\left(u_{i1}^N | \, 1 = A\right)\right]^3
\end{aligned}$$

Now let examine the cross sectional moments for period 2. To compute the probability of being in a sector, we need to know the probability of occurrence of all possible paths that an individual can

exhibit before choosing a given sector[30]. This is, $P(2=N)=P(2=N,1=N)+P(2=N,1=N)$ where in turn the probability of transition from N to N is given by:

$$\begin{aligned}
&P(2=N,1=N)\\
=&P(2=N,1=N,0=A)+P(2=N,1=N,0=N)\\
=&P\left(\left\{u_{i2}<r_2+\ln\phi^{NA}\right\},\left\{u_{i1}<r_1-\ln\phi^{AN}\right\},\{-u_{i0}<-r_0\}\right)\\
+&P\left(\left\{u_{i2}<r_2+\ln\phi^{NA}\right\},\left\{u_{i1}<r_1+\ln\phi^{NA}\right\},\{u_{i0}<r_0\}\right)\\
=&\Phi\left(\overrightarrow{r}_{ANN},\Sigma_{WT}\right)+\Phi\left(\overrightarrow{r}_{NNN},\Sigma_{TT}\right)
\end{aligned}$$

with: $\overrightarrow{r}_{ANN}=\left[r_2+\ln\phi^{NA},\overrightarrow{r}'_{AN}\right]',\overrightarrow{r}_{NNN}=\left[r_2+\ln\phi^{NA},\overrightarrow{r}'_{NN}\right]'$ and:

$$\Sigma_{WT}=\begin{bmatrix}\sigma_u^{*2}&\sigma_\theta^{*2}&-\sigma_\theta^{*2}\\ \sigma_\theta^{*2}&\sigma_u^{*2}&-\sigma_\theta^{*2}\\ -\sigma_\theta^{*2}&-\sigma_\theta^{*2}&\sigma_u^{*2}\end{bmatrix},\Sigma_{TT}=\begin{bmatrix}\sigma_u^{*2}&\sigma_\theta^{*2}&\sigma_\theta^{*2}\\ \sigma_\theta^{*2}&\sigma_u^{*2}&\sigma_\theta^{*2}\\ \sigma_\theta^{*2}&\sigma_\theta^{*2}&\sigma_u^{*2}\end{bmatrix}$$

Following similar arguments, we can show that the remaining probabilities of transition can be expressed as:

$$\begin{aligned}
P(2=A,1=N)=&\Phi\left(\overrightarrow{r}_{ANA},\Sigma_{WW}\right)+\Phi\left(\overrightarrow{r}_{NNA},\Sigma_{TW}\right)\\
P(2=A,1=A)=&\Phi\left(\overrightarrow{r}_{AAA},\Sigma_{TT}\right)+\Phi\left(\overrightarrow{r}_{NAA},\Sigma_{WT}\right)\\
P(2=N,1=A)=&\Phi\left(\overrightarrow{r}_{AAN},\Sigma_{TW}\right)+\Phi\left(\overrightarrow{r}_{NAN},\Sigma_{WW}\right)
\end{aligned}$$

with: $\overrightarrow{r}_{ANA}=\left[-r_2-\ln\phi^{NA},\ \overrightarrow{r}'_{AN}\right]',\ \overrightarrow{r}_{NNA}=\left[-r_2-\ln\phi^{NA},\ \overrightarrow{r}'_{NN}\right]'$,
$\overrightarrow{r}_{AAA}=\left[-r_2+\ln\phi^{AN},\ \overrightarrow{r}'_{AA}\right]'\overrightarrow{r}_{NAA}=\left[-r_2+\ln\phi^{AN},\ \overrightarrow{r}'_{NA}\right]',\overrightarrow{r}_{NAN}=\left[\ r_2-\ln\phi^{AN},\ \overrightarrow{r}'_{AN}\right]'$ and:

$$\Sigma_{WW}=\begin{bmatrix}\sigma_u^{*2}&-\sigma_\theta^{*2}&\sigma_\theta^{*2}\\ -\sigma_\theta^{*2}&\sigma_u^{*2}&-\sigma_\theta^{*2}\\ \sigma_\theta^{*2}&-\sigma_\theta^{*2}&\sigma_u^{*2}\end{bmatrix},\Sigma_{TW}=\begin{bmatrix}\sigma_u^{*2}&-\sigma_\theta^{*2}&-\sigma_\theta^{*2}\\ -\sigma_\theta^{*2}&\sigma_u^{*2}&\sigma_\theta^{*2}\\ -\sigma_\theta^{*2}&\sigma_\theta^{*2}&\sigma_u^{*2}\end{bmatrix}$$

Now consider the values of the expected income in the second period. Again, we need to know the expected income for each transition group. The income of stayers in N in period 2 is given by:

$$\begin{aligned}
&E\left(y_{i2}^N|\,2=N,1=N\right)\\
=&r_2^N+E\left(u_{i2}^N|\,2=N,1=N\right)\\
=&r_2^N+\left[E\left(u_{i2}^N|\,2=N,1=N,0=A\right)P\left(2=N,1=N,0=A\right)\right.\\
&\left.+E\left(u_{i2}^N|\,2=N,1=N,0=N\right)P\left(2=N,1=N,0=N\right)\right]/P\left(2=N,1=N\right)
\end{aligned}$$

Similarly as above, consider the moment $E\left(X_1^{k_1}X_2^{k_2}X_3^{k_3}X_4^{k_4}|\ -\infty<X_i<b_i,i=1,2,3,4\right)$ of the upper truncated multivariate normal distribution $\mathbf{N}\left(\mathbf{0},\Sigma\right)$ with $k_2=k_3=k_4=0$ and $b_1=\infty$ as

[30]Unfortunately, we cannot use Bayes' rule to derive the expressions of the joint probability from the marginals, since for the latter ones there is no closed form solution.

a function $\mathbf{M}_4(B, \Sigma, k)$ of the variance Σ, the elements b_2, b_3 and b_4 stacked up in a vector B and the coefficient $k = k_1$ that is:

$$\mathbf{M_4}(B, \Sigma, k) \equiv E\left(X_1^k | \ -\infty < X_1 < \infty, -\infty < X_2 < B_1, -\infty < X_3 < B_2, -\infty < X_4 < B_3\right)$$

So we can express:

$$E\left(y_{i2}^N | \, 2 = N, 1 = N\right)$$
$$= r_2^N + \frac{\mathbf{M_4}(\overrightarrow{r}_{ANN}, \Sigma_{ANN}, 1)\Phi\left(\overrightarrow{r}_{ANN}, \Sigma_{WT}\right) + \mathbf{M_4}(\overrightarrow{r}_{NNN}, \Sigma_{NNN}, 1)\Phi\left(\overrightarrow{r}_{NNN}, \Sigma_{TT}\right)}{\Phi\left(\overrightarrow{r}_{ANN}, \Sigma_{WT}\right) + \Phi\left(\overrightarrow{r}_{NNN}, \Sigma_{TT}\right)}$$

with: $\Sigma_{NNN} = \begin{bmatrix} \sigma_{uN}^2 & \Lambda_{NNN} \\ \Lambda_{NNN}' & \Sigma_{TT} \end{bmatrix}$ and $\Sigma_{ANN} = \begin{bmatrix} \sigma_{uN}^2 & \Lambda_{ANN} \\ \Lambda_{ANN}' & \Sigma_{WT} \end{bmatrix}$, where $\Lambda_{NNN} = [-\tilde{\sigma}_{uN}^2, -\tilde{\sigma}_{\theta N}^2, -\tilde{\sigma}_{\theta N}^2]$ and $\Lambda_{ANN} = [-\tilde{\sigma}_{uN}^2, -\tilde{\sigma}_{\theta N}^2, \tilde{\sigma}_{\theta N}^2]$. Similarly, the expected incomes in period 2 for the remaining groups are:

$$E\left(y_{i2}^A | \, 2 = A, 1 = N\right)$$
$$= r_2^A + \frac{\mathbf{M_4}(\overrightarrow{r}_{ANA}, \Sigma_{ANA}, 1)\Phi\left(\overrightarrow{r}_{ANA}, \Sigma_{WW}\right) + \mathbf{M_4}(\overrightarrow{r}_{NNA}, \Sigma_{NNA}, 1)\Phi\left(\overrightarrow{r}_{NNA}, \Sigma_{TW}\right)}{\Phi\left(\overrightarrow{r}_{ANA}, \Sigma_{WW}\right) + \Phi\left(\overrightarrow{r}_{NNA}, \Sigma_{TW}\right)}$$
$$E\left(y_{i2}^A | \, 2 = A, 1 = A\right)$$
$$= r_2^A + \frac{\mathbf{M_4}(\overrightarrow{r}_{AAA}, \Sigma_{AAA}, 1)\Phi\left(\overrightarrow{r}_{AAA}, \Sigma_{TT}\right) + \mathbf{M_4}(\overrightarrow{r}_{NAA}, \Sigma_{NAA}, 1)\Phi\left(\overrightarrow{r}_{NAA}, \Sigma_{WT}\right)}{\Phi\left(\overrightarrow{r}_{AAA}, \Sigma_{TT}\right) + \Phi\left(\overrightarrow{r}_{NAA}, \Sigma_{WT}\right)}$$

$$E\left(y_{i2}^N | \, 2 = N, 1 = A\right)$$
$$= r_2^N + \frac{\mathbf{M_4}(\overrightarrow{r}_{AAN}, \Sigma_{AAN}, 1)\Phi\left(\overrightarrow{r}_{AAN}, \Sigma_{TW}\right) + \mathbf{M_4}(\overrightarrow{r}_{NAN}, \Sigma_{NAN}, 1)\Phi\left(\overrightarrow{r}_{NAN}, \Sigma_{WW}\right)}{\Phi\left(\overrightarrow{r}_{AAN}, \Sigma_{TW}\right) + \Phi\left(\overrightarrow{r}_{NAN}, \Sigma_{WW}\right)}$$

with:

$$\Sigma_{ANA} = \begin{bmatrix} \sigma_{uA}^2 & \Lambda_{ANA} \\ \Lambda_{ANA}' & \Sigma_{WW} \end{bmatrix}, \ \Sigma_{NNA} = \begin{bmatrix} \sigma_{uA}^2 & \Lambda_{NNA} \\ \Lambda_{NNA}' & \Sigma_{TW} \end{bmatrix}, \ \Sigma_{AAA} = \begin{bmatrix} \sigma_{uA}^2 & \Lambda_{AAA} \\ \Lambda_{AAA}' & \Sigma_{TT} \end{bmatrix}$$
$$\Sigma_{NAA} = \begin{bmatrix} \sigma_{uA}^2 & \Lambda_{NAA} \\ \Lambda_{NAA}' & \Sigma_{WT} \end{bmatrix}, \ \Sigma_{AAN} = \begin{bmatrix} \sigma_{uN}^2 & \Lambda_{AAN} \\ \Lambda_{AAN}' & \Sigma_{TW} \end{bmatrix}, \ \Sigma_{NAN} = \begin{bmatrix} \sigma_{uN}^2 & \Lambda_{NAN} \\ \Lambda_{NAN}' & \Sigma_{WW} \end{bmatrix}$$

where $\Lambda_{ANA} = [-\tilde{\sigma}_{uA}^2, \tilde{\sigma}_{\theta A}^2, -\tilde{\sigma}_{\theta A}^2]$, $\Lambda_{NNA} = [-\tilde{\sigma}_{uA}^2, \tilde{\sigma}_{\theta A}^2, \tilde{\sigma}_{\theta A}^2]$, $\Lambda_{AAA} = [-\tilde{\sigma}_{uA}^2, -\tilde{\sigma}_{\theta A}^2, -\tilde{\sigma}_{\theta A}^2]$, $\Lambda_{NAA} = [-\tilde{\sigma}_{uA}^2, -\tilde{\sigma}_{\theta A}^2, \tilde{\sigma}_{\theta A}^2]$, $\Lambda_{AAN} = [-\tilde{\sigma}_{uN}^2, \tilde{\sigma}_{\theta N}^2, \tilde{\sigma}_{\theta N}^2]$, $\Lambda_{NAN} = [-\tilde{\sigma}_{uN}^2, \tilde{\sigma}_{\theta N}^2, -\tilde{\sigma}_{\theta N}^2]$. Combining the last three equations with the probabilities of transition into each sector, it is straightforward to derive the first moments of the cross-sectional distribution of earnings in period 2. The second and third moments can be derived as in period 1, as functions of $\mathbf{M}_4(\cdot, \cdot, 2)$ and $\mathbf{M}_4(\cdot, \cdot, 3)$ respectively.

Finally consider the growth in income for switchers from A to N, that is:

$$\begin{aligned}
&E\left(y_{i2}^{N}-y_{i1}^{A}|\,2=N,1=A\right)\\
=&r_{2}^{N}-r_{1}^{A}+E\left(u_{i2}^{N}-u_{i1}^{A}|\,2=N,1=A\right)\\
=&r_{2}^{N}-r_{1}^{A}+[E\left(u_{i2}^{N}-u_{i1}^{A}|\,2=N,1=A,0=A\right)P\left(2=N,1=A,0=A\right)\\
&+E\left(u_{i2}^{N}-u_{i1}^{A}|\,2=N,1=A,0=N\right)P\left(2=N,1=A,0=N\right)]/P\left(2=N,1=A\right)
\end{aligned}$$

The unknown terms are those that involved expected values, that can be obtained from $\mathbf{M_4}(\overrightarrow{r}_{AAN},\tilde{\Sigma}_{AAN},1)$ and $\mathbf{M_4}(\overrightarrow{r}_{NAN},\tilde{\Sigma}_{NAN},1)$ respectively, with:

$$\tilde{\Sigma}_{AAN}=\left[\begin{array}{cc}\sigma_{uA}^{2}+\sigma_{uN}^{2}-2\sigma_{\theta AN} & \tilde{\Lambda}_{AAN}\\ \tilde{\Lambda}_{AAN}^{\prime} & \Sigma_{TW}\end{array}\right],\ \tilde{\Sigma}_{NAN}=\left[\begin{array}{cc}\sigma_{uA}^{2}+\sigma_{uN}^{2}-2\sigma_{\theta AN} & \tilde{\Lambda}_{NAN}\\ \tilde{\Lambda}_{NAN}^{\prime} & \Sigma_{WW}\end{array}\right]$$

where:

$$\tilde{\Lambda}_{AAN}=\left[-\tilde{\sigma}_{uN}^{2}-\tilde{\sigma}_{\theta A}^{2},\tilde{\sigma}_{uA}^{2}+\tilde{\sigma}_{\theta N}^{2},\tilde{\sigma}_{\theta A}^{2}+\tilde{\sigma}_{\theta N}^{2}\right],\ \tilde{\Lambda}_{NAN}=\left[-\tilde{\sigma}_{uN}^{2}-\tilde{\sigma}_{\theta A}^{2},\tilde{\sigma}_{uA}^{2}+\tilde{\sigma}_{\theta N}^{2},-\tilde{\sigma}_{\theta A}^{2}-\tilde{\sigma}_{\theta N}^{2}\right]$$

The variance can be expressed in terms of $\mathbf{M_4}(\overrightarrow{r}_{AAN},\tilde{\Sigma}_{AAN},2)$ and $\mathbf{M_4}(\overrightarrow{r}_{AAN},\tilde{\Sigma}_{AAN},2)$, as in the frictionless case.

Similarly, for the growth in income for switchers from N to A we need expressions for:

$$E\left(u_{i2}^{A}-u_{i1}^{N}|\,2=A,1=N,0=A\right),E\left(u_{i2}^{A}-u_{i1}^{N}|\,2=A,1=N,0=N\right)$$

that can be obtained from $\mathbf{M_4}(\overrightarrow{r}_{ANA},\tilde{\Sigma}_{ANA},1)$ and $\mathbf{M_4}(\overrightarrow{r}_{NNA},\tilde{\Sigma}_{NNA},1)$ respectively, with:

$$\tilde{\Sigma}_{ANA}=\left[\begin{array}{cc}\sigma_{uA}^{2}+\sigma_{uN}^{2}-2\sigma_{\theta AN} & \tilde{\Lambda}_{ANA}\\ \tilde{\Lambda}_{ANA}^{\prime} & \Sigma_{WW}\end{array}\right],\ \tilde{\Sigma}_{NNA}=\left[\begin{array}{cc}\sigma_{uA}^{2}+\sigma_{uN}^{2}-2\sigma_{\theta AN} & \tilde{\Lambda}_{NNA}\\ \tilde{\Lambda}_{NNA}^{\prime} & \Sigma_{TW}\end{array}\right]$$

where:

$$\tilde{\Lambda}_{ANA}=\left[-\tilde{\sigma}_{uA}^{2}-\tilde{\sigma}_{\theta N}^{2},\tilde{\sigma}_{uN}^{2}+\tilde{\sigma}_{\theta A}^{2},-\tilde{\sigma}_{\theta N}^{2}-\tilde{\sigma}_{\theta A}^{2}\right];\ \tilde{\Lambda}_{NNA}=\left[-\tilde{\sigma}_{uA}^{2}-\tilde{\sigma}_{\theta N}^{2},\tilde{\sigma}_{uN}^{2}+\tilde{\sigma}_{\theta A}^{2},\tilde{\sigma}_{\theta N}^{2}+\tilde{\sigma}_{\theta A}^{2}\right]$$

The variance can be expressed in terms of $\mathbf{M_4}(\overrightarrow{r}_{ANA},\tilde{\Sigma}_{ANA},2)$ and $\mathbf{M_4}(\overrightarrow{r}_{NNA},\tilde{\Sigma}_{NNA},2)$. As in the frictionless case, we verified the system of 22 moments has a unique solution for the 12 elements of Θ.

Model with Involuntary Switches

Now consider the model with involuntary switches. Once again, the objective is to find expressions for the same set of 22 moments to identify the 12 elements of Θ (the same 10 parameters as in the frictionless economy plus the two probabilities of involuntary choices). Denote the probability of being forced to accept a job in a sector other than desired for workers who want to stay in the same sector as p^S, and for workers who want to switch sector as p^T. Further, define $p_n^S=1-p^S$ and $p_n^T=1-p^T$. As in the structural estimation, for the allocation of sectors in period 0 we impose that a fraction p^S of workers are forced to be out of their desired sector.

First, consider the cross sectional moments in period 1. The observed probabilities of transition across sectors between period 0 and 1 can be computed as the sum of the products of the probabilities of occurrence of all four possible transition paths in the frictionless equilibrium (desired transitions) and the proportions of workers of each desired transition group that end up in each observed transition group, shares that depend on the staying/switching status of the corresponding desired transitions. These proportions are displayed in Table D.1. Each cell can be read as the fraction of workers of the desired transition group (displayed in columns) that ends up in the observed transition group (displayed in rows).

Table D.1: Probabilities of ending up in a given transition group for period 1

		Desired transition*			
		NN	NA	AN	AA
Observed transition*	NN	$\left(p_n^S\right)^2$	$p_n^S p^T$	$p^S p_n^S$	$p^S p^T$
	NA	$p_n^S p^S$	$p_n^S p_n^T$	$\left(p^S\right)^2$	$p^S p_n^T$
	AN	$p^S p_n^T$	$\left(p^S\right)^2$	$p_n^S p_n^T$	$p_n^S p^S$
	AA	$p^S p^T$	$p^S p_n^S$	$p_n^S p^T$	$\left(p_n^S\right)^2$

*The first and second letters correspond to the sector in period 0 and in period 1, respectively.

Thus, denoting as P^d the desired transitions, the probability of observing a transition from non-agriculture in period 0 to non-agriculture (NN) in period 1 is:

$$\begin{aligned} P\left(0=N,1=N\right) &= \left(p_n^S\right)^2 P^d\left(0=N,1=N\right) + p_n^S p^T P^d\left(0=N,1=A\right) \\ &\quad + p^S p_n^T P^d\left(0=A,1=N\right) + p^S p^T P^d\left(0=A,1=A\right) \\ &= \left(p_n^S\right)^2 \Phi\left(\overrightarrow{r}_{NN},\Sigma_T\right) + p_n^S p^T \Phi\left(\overrightarrow{r}_{NA},\Sigma_W\right) + p^S p_n^T \Phi\left(\overrightarrow{r}_{AN},\Sigma_W\right) + p^S p^T \Phi\left(\overrightarrow{r}_{AA},\Sigma_T\right) \end{aligned}$$

where $\overrightarrow{r}_{NN}, \overrightarrow{r}_{NA}$, $\overrightarrow{r}_{AN}, \overrightarrow{r}_{AA}$, Σ_W and Σ_T are as in the model with switching costs, but with $\phi_{AN} = \phi_{NA} = 1$. Arrange in a column vector $\overrightarrow{P}_1^d$ the probabilities of all four possible transition paths in the frictionless economy, i.e.:

$$\overrightarrow{P}_1^d = \left[\Phi\left(\overrightarrow{r}_{NN},\Sigma_T\right)\ \Phi\left(\overrightarrow{r}_{NA},\Sigma_W\right)\ \Phi\left(\overrightarrow{r}_{AN},\Sigma_W\right)\ \Phi\left(\overrightarrow{r}_{AA},\Sigma_T\right)\right]'$$

The probability of observing a transition NN from period 0 to period 1 can be written in a compact form as the cross product:

$$P\left(0=N,1=N\right) = \overrightarrow{F}_{NN} \times \overrightarrow{P}_1^d$$

where $\overrightarrow{F}_{NN}$ is a (row) vector with the shares of workers corresponding to the observed transition group NN, displayed in the first row of Table D.1. Similarly,

$$P\left(0=s,1=s'\right) = \overrightarrow{F}_{ss'} \times \overrightarrow{P}_1^d$$

for $s, s' = A, N$, where $\overrightarrow{F}_{NA}$, $\overrightarrow{F}_{AN}$, $\overrightarrow{F}_{AA}$ are row vectors with the shares of workers displayed in

the second, third and fourth rows of Table D.1, respectively. Hence:

$$P\left(1=s\right)=P\left(0=N,1=s\right)+P\left(0=A,1=s\right)=\left(\overrightarrow{F}_{Ns}+\overrightarrow{F}_{As}\right)\times\overrightarrow{P}_{1}^{d}$$

for $s=A,N$. Now consider the expressions for the values of the expected income in the first period. For sector N we have:

$$\begin{aligned}E\left(y_{i1}^{N}|\,1=N\right)=&r_{1}^{N}+E\left(u_{i1}^{N}|\,1=N\right)\\=&r_{1}^{N}+\frac{E\left(u_{i1}^{N}|\,1=N,0=A\right)\overrightarrow{F}_{AN}\times\overrightarrow{P}_{1}^{d}+E\left(u_{i1}^{N}|\,1=N,0=N\right)\overrightarrow{F}_{NN}\times\overrightarrow{P}_{1}^{d}}{\left(\overrightarrow{F}_{NN}+\overrightarrow{F}_{AN}\right)\times\overrightarrow{P}_{1}^{d}}\end{aligned}\tag{A.17}$$

To find expressions for $E\left(u_{i1}^{N}|\,1=N,0=A\right)$ and $E\left(u_{i1}^{N}|\,1=N,0=N\right)$ we need to compute the expected value of u_{i1}^{N} for the individuals in all possible desired transition groups. Such expected values are the elements of the column vector:

$$\overrightarrow{E}_{11}^{N}=\left[\mathbf{M_{3}}(\overrightarrow{r}_{NN},\Sigma_{NN},1)\ \mathbf{M_{3}}(\overrightarrow{r}_{NA},\Sigma_{NA|N},1)\ \mathbf{M_{3}}(\overrightarrow{r}_{AN},\Sigma_{AN},1)\ \mathbf{M_{3}}(\overrightarrow{r}_{AA},\Sigma_{AA|N},1)\right]'$$

with $\mathbf{M}_{3}(\cdot)$, Σ_{NN} and Σ_{AN} as in the model without switching costs (with $\phi_{AN}=\phi_{NA}=1$), and now $\Sigma_{NA|N}=\begin{bmatrix}\sigma_{uN}^{2} & -\Lambda_{AN}\\ -\Lambda_{AN}' & \Sigma_{W}\end{bmatrix}$ and $\Sigma_{AA|N}=\begin{bmatrix}\sigma_{uN}^{2} & -\Lambda_{NN}\\ -\Lambda_{NN}' & \Sigma_{T}\end{bmatrix}$. Thus, the conditional expected values of u_{i1}^{N} can be computed through the following weighted average:

$$E\left(u_{i1}^{N}|\,1=N,0=s\right)=\frac{\left[\overrightarrow{F}_{sN}\cdot\left(\overrightarrow{P}_{1}^{d}\right)'\right]\times\overrightarrow{E}_{11}^{N}}{\overrightarrow{F}_{sN}\times\overrightarrow{P}_{1}^{d}}\tag{A.18}$$

for $s=A,N$, where $\cdot$ denotes the dot product. Substituting (A.18) for $s=A,N$ in (A.17) we get:

$$E\left(y_{i1}^{N}|\,1=N\right)=r_{1}^{N}+\frac{\left[\left(\overrightarrow{F}_{NN}+\overrightarrow{F}_{AN}\right)\cdot\left(\overrightarrow{P}_{1}^{d}\right)'\right]\times\overrightarrow{E}_{11}^{N}}{\left(\overrightarrow{F}_{NN}+\overrightarrow{F}_{AN}\right)\times\overrightarrow{P}_{1}^{d}}$$

Using a similar reasoning, we can obtain:

$$E\left(y_{i1}^{A}|\,1=A\right)=r_{1}^{A}+\frac{\left[\left(\overrightarrow{F}_{NA}+\overrightarrow{F}_{AA}\right)\cdot\left(\overrightarrow{P}_{1}^{d}\right)'\right]\times\overrightarrow{E}_{11}^{A}}{\left(\overrightarrow{F}_{NA}+\overrightarrow{F}_{AA}\right)\times\overrightarrow{P}_{1}^{d}}$$

where $\overrightarrow{E}_{11}^{A}$ is the column vector with elements:

$$\overrightarrow{E}_{11}^{A}=\left[\mathbf{M_{3}}(\overrightarrow{r}_{NN},\Sigma_{NN|A},1)\ \mathbf{M_{3}}(\overrightarrow{r}_{NA},\Sigma_{NA},1)\ \mathbf{M_{3}}(\overrightarrow{r}_{AN},\Sigma_{AN|A},1)\ \mathbf{M_{3}}(\overrightarrow{r}_{AA},\Sigma_{AA},1)\right]'$$

with Σ_{AA} and Σ_{NA} as in the model with switching costs (with $\phi_{AN}=\phi_{NA}=1$), and now $\Sigma_{AN|A}=$

$\begin{bmatrix} \sigma_{uA}^2 & -\Lambda_{NA} \\ -\Lambda_{NA}' & \Sigma_W \end{bmatrix}$ and $\Sigma_{NN|A} = \begin{bmatrix} \sigma_{uA}^2 & -\Lambda_{AA} \\ -\Lambda_{AA}' & \Sigma_T \end{bmatrix}$. By the same token, the formulas for the second and third moments of the income distribution in period 1 can be obtained by substituting the expressions of the conditional expected values of u_{i1}^{s2} and u_{i1}^{s3} (for $s = A, N$), respectively, into the formulas of the corresponding moments derived for the model with switching costs. We obtain:

$$Var\left(y_{i1}^s | \, 1 = s\right) = \frac{\left[\left(\overrightarrow{F}_{Ns} + \overrightarrow{F}_{As}\right) \cdot \left(\overrightarrow{P}_1^d\right)'\right] \times \overrightarrow{E}_{12}^s}{\left(\overrightarrow{F}_{Ns} + \overrightarrow{F}_{As}\right) \times \overrightarrow{P}_1^d} - E\left(u_{i1}^s | \, 1 = s\right)^2 + \sigma_\mu^2$$

$$E\left(\left[y_{i1}^s - E\left(y_{i1}^s | \, 1 = s\right)\right]^3 | \, 1 = s\right) = \frac{\left[\left(\overrightarrow{F}_{Ns} + \overrightarrow{F}_{As}\right) \cdot \left(\overrightarrow{P}_1^d\right)'\right] \times \overrightarrow{E}_{13}^s}{\left(\overrightarrow{F}_{Ns} + \overrightarrow{F}_{As}\right) \times \overrightarrow{P}_1^d} - 3E\left(u_{i1}^s | \, 1 = s\right)\left[Var\left(y_{i1}^s | \, 1 = s\right) - \sigma_\mu^2\right] - \left[E\left(u_{i1}^s | \, 1 = s\right)\right]^3$$

for $s = A, N$, where:

$$\overrightarrow{E}_{1j}^A = \left[\mathbf{M_3}(\overrightarrow{r}_{NN}, \Sigma_{NN|A}, j)\ \mathbf{M_3}(\overrightarrow{r}_{NA}, \Sigma_{NA}, j)\ \mathbf{M_3}(\overrightarrow{r}_{AN}, \Sigma_{AN|A}, j)\ \mathbf{M_3}(\overrightarrow{r}_{AA}, \Sigma_{AA}, j)\right]'$$
$$\overrightarrow{E}_{1j}^N = \left[\mathbf{M_3}(\overrightarrow{r}_{NN}, \Sigma_{NN}, j)\ \mathbf{M_3}(\overrightarrow{r}_{NA}, \Sigma_{NA|N}, j)\ \mathbf{M_3}(\overrightarrow{r}_{AN}, \Sigma_{AN}, j)\ \mathbf{M_3}(\overrightarrow{r}_{AA}, \Sigma_{AA|N}, j)\right]'$$

for $j = 2, 3$.

Now consider the cross sectional moments for period 2. Once again, we need the proportion of workers of each desired transition group that end up in the corresponding observed transition groups. These shares are displayed in Table D.2, which can be read in the same way as Table D.1.

Table D.2: Probabilities of ending up in a given transition group for period 2

Observed transition* \ Desired transition*	NNN	NNA	NAN	NAA	ANN	ANA	AAN	AAA
NNN	$\left(p_n^S\right)^3$	$\left(p_n^S\right)^2 p^T$	$\left(p_n^S\right)^2 p^T$	$p_n^S \left(p^T\right)^2$	$p^S \left(p_n^S\right)^2$	$p^S p_n^S p^T$	$p^S p^T p_n^S$	$p^S \left(p^T\right)^2$
NNA	$\left(p_n^S\right)^2 p^S$	$\left(p_n^S\right)^2 p_n^T$	$p_n^S p^T p^S$	$p_n^S p^T p_n^T$	$\left(p^S\right)^2 p_n^S$	$p^S p_n^S p_n^T$	$\left(p^S\right)^2 p^T$	$p^S p^T p_n^T$
NAN	$p_n^S p^S p_n^T$	$p_n^S \left(p^S\right)^2$	$p_n^S \left(p_n^T\right)^2$	$p_n^S p_n^T p^S$	$\left(p^S\right)^2 p_n^T$	$\left(p^S\right)^3$	$p^S \left(p_n^T\right)^2$	$\left(p^S\right)^2 p_n^T$
NAA	$p_n^S p^S p^T$	$\left(p_n^S\right)^2 p^S$	$p_n^S p_n^T p^T$	$\left(p_n^S\right)^2 p_n^T$	$\left(p^S\right)^2 p^T$	$\left(p^S\right)^2 p_n^S$	$p^S p_n^T p^T$	$p^S p_n^T p_n^S$
ANN	$p^S p_n^T p_n^S$	$p^S p_n^T p^T$	$\left(p^S\right)^2 p_n^S$	$\left(p^S\right)^2 p^T$	$\left(p_n^S\right)^2 p_n^T$	$p_n^S p_n^T p^T$	$\left(p_n^S\right)^2 p^S$	$p_n^S p^S p^T$
ANA	$\left(p^S\right)^2 p_n^T$	$p^S \left(p_n^T\right)^2$	$\left(p^S\right)^3$	$\left(p^S\right)^2 p_n^T$	$p_n^S p_n^T p^S$	$p_n^S \left(p_n^T\right)^2$	$p_n^S \left(p^S\right)^2$	$p_n^S p^S p_n^T$
AAN	$p^S p^T p_n^T$	$\left(p^S\right)^2 p^T$	$p^S p_n^S p_n^T$	$\left(p^S\right)^2 p_n^S$	$p_n^S p^T p_n^T$	$p_n^S p^T p^S$	$\left(p_n^S\right)^2 p_n^T$	$\left(p_n^S\right)^2 p^S$
AAA	$p^S \left(p^T\right)^2$	$p^S p^T p_n^S$	$p^S p_n^S p^T$	$p^S \left(p_n^S\right)^2$	$p_n^S \left(p^T\right)^2$	$\left(p_n^S\right)^2 p^T$	$\left(p_n^S\right)^2 p^T$	$\left(p_n^S\right)^3$

*The first, second and third letters correspond to the sector in period 0, period 1 and period 2, respectively.

Arrange in a column vector $\overrightarrow{P}_2^d$ the probabilities of all eight possible transition paths in the

frictionless economy, i.e.:

$$\overrightarrow{P}_2^d = \left[\begin{array}{cccc} \Phi\left(\overrightarrow{r}_{NNN}, \Sigma_{TT}\right) & \Phi\left(\overrightarrow{r}_{NNA}, \Sigma_{TW}\right) & \Phi\left(\overrightarrow{r}_{NAN}, \Sigma_{WW}\right) & \Phi\left(\overrightarrow{r}_{NAA}, \Sigma_{WT}\right) \dots \\ \Phi\left(\overrightarrow{r}_{ANN}, \Sigma_{WT}\right) & \Phi\left(\overrightarrow{r}_{ANA}, \Sigma_{WW}\right) & \Phi\left(\overrightarrow{r}_{AAN}, \Sigma_{TW}\right) & \Phi\left(\overrightarrow{r}_{AAA}, \Sigma_{TT}\right) \end{array} \right]'$$

where $\overrightarrow{r}_{ss's''}\ \forall\ s, s', s'' = N, A$ and Σ_{TT}, Σ_{TW}, Σ_{TW} and Σ_{WW} are as in the model with switching costs, but with $\phi_{AN} = \phi_{NA} = 1$. Thus, the probabilities of observing any of the eight possible transition paths can be written in a compact form as:

$$P\left(0 = s, 1 = s', 2 = s''\right) = \overrightarrow{F}_{ss's''} \times \overrightarrow{P}_2^d\ \forall\ s, s', s'' = N, A$$

where $\overrightarrow{F}_{ss's''}$ is a row vector with the shares of workers displayed in the corresponding row of Table D.2 of the observed transition from sector s in period 0 to sector s' in period 1 and to sector s'' in period 2. Hence:

$$P\left(2 = s\right) = \left(\overrightarrow{F}_{NNs} + \overrightarrow{F}_{ANs} + \overrightarrow{F}_{NAs} + \overrightarrow{F}_{AAs}\right) \times \overrightarrow{P}_2^d$$

for $s = A, N$. In a similar way to the cross sectional moments in period 1, the expressions for all moments of the income distribution in period 2 can be obtained after substituting the probabilities of transition derived above and the expressions for the conditional expectations of u_{i2}^s, u_{i2}^{s2} and u_{i2}^{s3} into the corresponding formulas for the cross sectional moments in period 2 of the model with switching costs. Thus, define the column vectors:

$$\overrightarrow{E}_{2j}^A = \left[\begin{array}{cccc} \mathbf{M_4}(\overrightarrow{r}_{NNN}, \Sigma_{NNN|A}, j) & \mathbf{M_4}(\overrightarrow{r}_{NNA}, \Sigma_{NNA}, j) & \mathbf{M_4}(\overrightarrow{r}_{NAN}, \Sigma_{NAN|A}, j) & \mathbf{M_4}(\overrightarrow{r}_{NAA}, \Sigma_{NAA}, j) \dots \\ \mathbf{M_4}(\overrightarrow{r}_{ANN}, \Sigma_{ANN|A}, j) & \mathbf{M_4}(\overrightarrow{r}_{ANA}, \Sigma_{ANA}, j) & \mathbf{M_4}(\overrightarrow{r}_{AAN}, \Sigma_{AAN|A}, j) & \mathbf{M_4}(\overrightarrow{r}_{AAA}, \Sigma_{AAA}, j) \end{array} \right]'$$

$$\overrightarrow{E}_{2j}^N = \left[\begin{array}{cccc} \mathbf{M_4}(\overrightarrow{r}_{NNN}, \Sigma_{NNN}, j) & \mathbf{M_4}(\overrightarrow{r}_{NNA}, \Sigma_{NNA|N}, j) & \mathbf{M_4}(\overrightarrow{r}_{NAN}, \Sigma_{NAN}, j) & \mathbf{M_4}(\overrightarrow{r}_{NAA}, \Sigma_{NAA|N}, j) \dots \\ \mathbf{M_4}(\overrightarrow{r}_{ANN}, \Sigma_{ANN}, j) & \mathbf{M_4}(\overrightarrow{r}_{ANA}, \Sigma_{ANA|N}, j) & \mathbf{M_4}(\overrightarrow{r}_{AAN}, \Sigma_{AAN}, j) & \mathbf{M_4}(\overrightarrow{r}_{AAA}, \Sigma_{AAA|N}, j) \end{array} \right]'$$

for $j = 1, 2, 3$, where $\mathbf{M}_4(\cdot)$ and $\Sigma_{ss's''}\ \forall\ s, s', s'' = N, A$ are as in the model with switching costs (with $\phi_{AN} = \phi_{NA} = 1$), but now:

$$\Sigma_{NNN|A} = \left[\begin{array}{cc} \sigma_{uA}^2 & -\Lambda_{AAA} \\ -\Lambda_{AAA}' & \Sigma_{TT} \end{array} \right] \quad \Sigma_{NAN|A} = \left[\begin{array}{cc} \sigma_{uA}^2 & -\Lambda_{ANA} \\ -\Lambda_{ANA}' & \Sigma_{WW} \end{array} \right] \Sigma_{ANN|A} = \left[\begin{array}{cc} \sigma_{uA}^2 & -\Lambda_{NAA} \\ -\Lambda_{NAA}' & \Sigma_{WT} \end{array} \right]$$

$$\Sigma_{AAN|A} = \left[\begin{array}{cc} \sigma_{uA}^2 & -\Lambda_{NNA} \\ -\Lambda_{NNA}' & \Sigma_{TW} \end{array} \right] \quad \Sigma_{AAA|N} = \left[\begin{array}{cc} \sigma_{uN}^2 & -\Lambda_{NNN} \\ -\Lambda_{NNN}' & \Sigma_{TT} \end{array} \right] \Sigma_{NNA|N} = \left[\begin{array}{cc} \sigma_{uN}^2 & -\Lambda_{AAN} \\ -\Lambda_{AAN}' & \Sigma_{TW} \end{array} \right]$$

$$\Sigma_{NAA|N} = \left[\begin{array}{cc} \sigma_{uN}^2 & -\Lambda_{ANN} \\ -\Lambda_{ANN}' & \Sigma_{WT} \end{array} \right] \quad \Sigma_{ANA|N} = \left[\begin{array}{cc} \sigma_{uN}^2 & -\Lambda_{NAN} \\ -\Lambda_{NAN}' & \Sigma_{WW} \end{array} \right]$$

Additionally, define the (transition) event $\Omega \equiv \{2 = s', 1 = s\}$ for any $s, s' = A, N$. Thus, the formulas for the first, second and third moments of the income of individuals in each transition Ω are given by:

$$E\left(y_{i2}^{s'} | \Omega\right) = r_2^{s'} + E\left(u_{i2}^{s'} | \Omega\right) = r_2^{s'} + \frac{\left[\left(\overrightarrow{F}_{Nss'} + \overrightarrow{F}_{Ass'}\right) \cdot \left(\overrightarrow{P}_2^d\right)'\right] \times \overrightarrow{E}_{21}^{s'}}{\left(\overrightarrow{F}_{Nss'} + \overrightarrow{F}_{Ass'}\right) \times \overrightarrow{P}_2^d}$$

$$Var\left(y_{i2}^{s'}|\,\Omega\right)=\frac{\left[\left(\overrightarrow{F}_{Nss'}+\overrightarrow{F}_{Ass'}\right)\cdot\left(\overrightarrow{P}_2^d\right)'\right]\times\overrightarrow{E}_{22}^{s'}}{\left(\overrightarrow{F}_{Nss'}+\overrightarrow{F}_{Ass'}\right)\times\overrightarrow{P}_2^d}-E\left(u_{i2}^{s'}|\,\Omega\right)^2+\sigma_\mu^2$$

$$E\left(\left[y_{i2}^{s'}-E\left(y_{i2}^{s'}|\,\Omega\right)\right]^3|\,\Omega\right)=\frac{\left[\left(\overrightarrow{F}_{Nss'}+\overrightarrow{F}_{Ass'}\right)\cdot\left(\overrightarrow{P}_2^d\right)'\right]\times\overrightarrow{E}_{23}^{s'}}{\left(\overrightarrow{F}_{Nss'}+\overrightarrow{F}_{Ass'}\right)\times\overrightarrow{P}_2^d}$$
$$-3E\left(u_{i2}^{s'}|\,\Omega\right)\left[Var\left(y_{i2}^{s'}|\,\Omega\right)-\sigma_\mu^2\right]-\left[E\left(u_{i2}^{s}|\,\Omega\right)\right]^3$$

Combining those formulas with the probabilities of each transition group, it is straightforward to derive the expressions for the expected income of the pool of workers in each sector in period 2.

Finally, we need to derive the expressions for the first and second moments of the growth of income for switchers. Since we already have the mapping of probabilities between the actual and the desired transitions, we only need to find the expected values of the differences $u_{i2}^N-u_{i1}^A$ and $u_{i2}^A-u_{i1}^N$ for individuals in all desired transitions groups. For this, in a similar way as we did above, define the (column) vectors:

$$\overrightarrow{E}_{2j}^{AN}=\left[\begin{array}{c}\mathbf{M_4}(\overrightarrow{r}_{NNN},\tilde{\Sigma}_{NNN|AN},j)\ \mathbf{M_4}(\overrightarrow{r}_{NNA},\tilde{\Sigma}_{NNA|AN},j)\ \mathbf{M_4}(\overrightarrow{r}_{NAN},\tilde{\Sigma}_{NAN},j)\ \mathbf{M_4}(\overrightarrow{r}_{NAA},\tilde{\Sigma}_{NAA|AN},j)...\\ \mathbf{M_4}(\overrightarrow{r}_{ANN},\tilde{\Sigma}_{ANN|AN},j)\ \mathbf{M_4}(\overrightarrow{r}_{ANA},\tilde{\Sigma}_{ANA|AN},j)\ \mathbf{M_4}(\overrightarrow{r}_{AAN},\tilde{\Sigma}_{AAN},j)\ \mathbf{M_4}(\overrightarrow{r}_{AAA},\tilde{\Sigma}_{AAA|AN},j)\end{array}\right]'$$

$$\overrightarrow{E}_{2j}^{NA}=\left[\begin{array}{c}\mathbf{M_4}(\overrightarrow{r}_{NNN},\tilde{\Sigma}_{NNN|NA},j)\ \mathbf{M_4}(\overrightarrow{r}_{NNA},\tilde{\Sigma}_{NNA},j)\ \mathbf{M_4}(\overrightarrow{r}_{NAN},\tilde{\Sigma}_{NAN|NA},j)\ \mathbf{M_4}(\overrightarrow{r}_{NAA},\tilde{\Sigma}_{NAA|NA},j)...\\ \mathbf{M_4}(\overrightarrow{r}_{ANN},\tilde{\Sigma}_{ANN|NA},j)\ \mathbf{M_4}(\overrightarrow{r}_{ANA},\tilde{\Sigma}_{ANA},j)\ \mathbf{M_4}(\overrightarrow{r}_{AAN},\tilde{\Sigma}_{AAN|NA},j)\ \mathbf{M_4}(\overrightarrow{r}_{AAA},\tilde{\Sigma}_{AAA|NA},j)\end{array}\right]'$$

for $j=1,2$, where $\mathbf{M}_4(\cdot)$ and $\tilde{\Sigma}_{NAN}$, $\tilde{\Sigma}_{AAN}$, $\tilde{\Sigma}_{NNA}$, $\tilde{\Sigma}_{ANA}$, are as in the model with switching costs (with $\phi_{AN}=\phi_{NA}=1$), but now:

$$\tilde{\Sigma}_{ss's''|AN}=\left[\begin{array}{cc}\sigma_{uA}^2+\sigma_{uN}^2-2\sigma_{\theta AN} & \tilde{\Lambda}_{ss's''|AN}\\ \tilde{\Lambda}_{ss's''|AN}^{'} & \tilde{\Sigma}_{ss's''}(2,2)\end{array}\right],\ \tilde{\Sigma}_{ss's''|NA}=\left[\begin{array}{cc}\sigma_{uA}^2+\sigma_{uN}^2-2\sigma_{\theta AN} & \tilde{\Lambda}_{ss's''|NA}\\ \tilde{\Lambda}_{ss's''|NA} & \tilde{\Sigma}_{ss's''}(2,2)\end{array}\right]$$

$\forall\ s,s',s''=N,A$ (that is, the matrix in the position (2,2) of $\tilde{\Sigma}_{ss's''|AN}$ and $\tilde{\Sigma}_{ss's''|NA}$ is the same matrix of the position (2,2) of $\tilde{\Sigma}_{ss's''}$ in the model with switching costs), where:

$$\begin{array}{ll}\tilde{\Lambda}_{AAA|AN}=\left[\ \tilde{\sigma}_{uN}^2+\tilde{\sigma}_{\theta A}^2,\ \ \tilde{\sigma}_{uA}^2+\tilde{\sigma}_{\theta N}^2,\ \ \tilde{\sigma}_{\theta N}^2+\tilde{\sigma}_{\theta A}^2\right] & \tilde{\Lambda}_{AAA|NA}=\left[-\tilde{\sigma}_{uA}^2-\tilde{\sigma}_{\theta N}^2,-\tilde{\sigma}_{uN}^2-\tilde{\sigma}_{\theta A}^2,-\tilde{\sigma}_{\theta N}^2-\tilde{\sigma}_{\theta A}^2\right]\\ \tilde{\Lambda}_{ANA|AN}=\left[\ \tilde{\sigma}_{uN}^2+\tilde{\sigma}_{\theta A}^2,-\tilde{\sigma}_{uA}^2-\tilde{\sigma}_{\theta N}^2,\ \ \tilde{\sigma}_{\theta N}^2+\tilde{\sigma}_{\theta A}^2\right] & \tilde{\Lambda}_{AAN|NA}=\left[\ \tilde{\sigma}_{uA}^2+\tilde{\sigma}_{\theta N}^2,-\tilde{\sigma}_{uN}^2-\tilde{\sigma}_{\theta A}^2,-\tilde{\sigma}_{\theta N}^2-\tilde{\sigma}_{\theta A}^2\right]\\ \tilde{\Lambda}_{ANN|AN}=\left[-\tilde{\sigma}_{uN}^2-\tilde{\sigma}_{\theta A}^2,-\tilde{\sigma}_{uA}^2-\tilde{\sigma}_{\theta N}^2,\ \ \tilde{\sigma}_{\theta N}^2+\tilde{\sigma}_{\theta A}^2\right] & \tilde{\Lambda}_{ANN|NA}=\left[\ \tilde{\sigma}_{uA}^2+\tilde{\sigma}_{\theta N}^2,\ \ \tilde{\sigma}_{uN}^2+\tilde{\sigma}_{\theta A}^2,-\tilde{\sigma}_{\theta N}^2-\tilde{\sigma}_{\theta A}^2\right]\\ \tilde{\Lambda}_{NAA|AN}=\left[\ \tilde{\sigma}_{uN}^2+\tilde{\sigma}_{\theta A}^2,\ \ \tilde{\sigma}_{uA}^2+\tilde{\sigma}_{\theta N}^2,-\tilde{\sigma}_{\theta N}^2-\tilde{\sigma}_{\theta A}^2\right] & \tilde{\Lambda}_{NAA|NA}=\left[-\tilde{\sigma}_{uA}^2-\tilde{\sigma}_{\theta N}^2,-\tilde{\sigma}_{uN}^2-\tilde{\sigma}_{\theta A}^2,\ \ \tilde{\sigma}_{\theta N}^2+\tilde{\sigma}_{\theta A}^2\right]\\ \tilde{\Lambda}_{NNA|AN}=\left[\ \tilde{\sigma}_{uN}^2+\tilde{\sigma}_{\theta A}^2,-\tilde{\sigma}_{uA}^2-\tilde{\sigma}_{\theta N}^2,-\tilde{\sigma}_{\theta N}^2-\tilde{\sigma}_{\theta A}^2\right] & \tilde{\Lambda}_{NAN|NA}=\left[\ \tilde{\sigma}_{uA}^2+\tilde{\sigma}_{\theta N}^2,-\tilde{\sigma}_{uN}^2-\tilde{\sigma}_{\theta A}^2,\ \ \tilde{\sigma}_{\theta N}^2+\tilde{\sigma}_{\theta A}^2\right]\\ \tilde{\Lambda}_{NNN|AN}=\left[-\tilde{\sigma}_{uN}^2-\tilde{\sigma}_{\theta A}^2,-\tilde{\sigma}_{uA}^2-\tilde{\sigma}_{\theta N}^2,-\tilde{\sigma}_{\theta N}^2-\tilde{\sigma}_{\theta A}^2\right] & \tilde{\Lambda}_{NNN|NA}=\left[\ \tilde{\sigma}_{uA}^2+\tilde{\sigma}_{\theta N}^2,\ \ \tilde{\sigma}_{uN}^2+\tilde{\sigma}_{\theta A}^2,\ \ \tilde{\sigma}_{\theta N}^2+\tilde{\sigma}_{\theta A}^2\right]\end{array}$$

Using the definitions of $\overrightarrow{E}_{2j}^{AN}$ and $\overrightarrow{E}_{2j}^{NA}$ for $j=1,2$ is straightforward to derive the expressions for the first and second moments of the growth in income for switchers. For example, defining the (transition) event $\Omega^{NA}\equiv\{2=A,1=N\}$, the corresponding expressions for switchers from N to A are:

$$E\left(y_{i2}^A-y_{i1}^N|\,\Omega^{NA}\right)=r_2^A-r_1^N+E\left(u_{i2}^A-u_{i1}^N|\,\Omega^{NA}\right)=r_2^A-r_1^N+\frac{\left[\left(\overrightarrow{F}_{NNA}+\overrightarrow{F}_{ANA}\right)\cdot\left(\overrightarrow{P}_2^d\right)'\right]\times\overrightarrow{E}_{21}^{NA}}{\left(\overrightarrow{F}_{NNA}+\overrightarrow{F}_{ANA}\right)\times\overrightarrow{P}_2^d}$$

$$Var\left(y_{i2}^{A}-y_{i1}^{N}|\,\Omega^{NA}\right)=\frac{\left[\left(\overrightarrow{F}_{NNA}+\overrightarrow{F}_{ANA}\right)\cdot\left(\overrightarrow{P}_{2}^{d}\right)'\right]\times\overrightarrow{E}_{22}^{NA}}{\left(\overrightarrow{F}_{NNA}+\overrightarrow{F}_{ANA}\right)\times\overrightarrow{P}_{2}^{d}}-E\left(u_{i2}^{A}-u_{i1}^{N}|\,\Omega^{NA}\right)^{2}+2\sigma_{\mu}^{2}$$

Similar expressions apply for switchers from N to A. As in the cases of the frictionless economy and the economy with switching costs, with the explicit formulas for the set of 22 moments of interest we numerically verified that the system has an unique solution for the 12 elements of Θ.

E Proofs

Proof of Proposition 1

Under the assumptions of Proposition 1, the expression for expected log income growth of switchers to from agriculture to non-agriculture given in (A.13) simplifies to:

$$E\left(y_{i2}^{N}-y_{i1}^{A}|\,2=N,1=A\right)=\mathbf{M}_{3}(\overrightarrow{0},\tilde{\Sigma}_{AN},1),$$

and where $\tilde{\Sigma}_{AN}$ can be simplified to

$$\tilde{\Sigma}_{AN}=\begin{bmatrix}\sigma_{\theta}^{*2}+\sigma_{\varepsilon A}^{2}+\sigma_{\varepsilon N}^{2} & -\left(\sigma_{\theta}^{*2}+\sigma_{\varepsilon N}^{2}\right) & \sigma_{\theta}^{*2}+\sigma_{\varepsilon A}^{2}\\ -\left(\sigma_{\theta}^{*2}+\sigma_{\varepsilon N}^{2}\right) & \sigma_{\theta}^{*2}+\sigma_{\varepsilon A}^{2}+\sigma_{\varepsilon N}^{2} & -\sigma_{\theta}^{*2}\\ \sigma_{\theta}^{*2}+\sigma_{\varepsilon A}^{2} & -\sigma_{\theta}^{*2} & \sigma_{\theta}^{*2}+\sigma_{\varepsilon A}^{2}+\sigma_{\varepsilon N}^{2}\end{bmatrix}.$$

Re-write $\tilde{\Sigma}_{AN}$ in terms of the correlation matrix C_{AN} :

$$C_{AN}=\begin{bmatrix}1 & \frac{-\left(\sigma_{\theta}^{*2}+\sigma_{\varepsilon N}^{2}\right)}{\sigma_{\theta}^{*2}+\sigma_{\varepsilon A}^{2}+\sigma_{\varepsilon N}^{2}} & \frac{\sigma_{\theta}^{*2}+\sigma_{\varepsilon A}^{2}}{\sigma_{\theta}^{*2}+\sigma_{\varepsilon A}^{2}+\sigma_{\varepsilon N}^{2}}\\ \frac{-\left(\sigma_{\theta}^{*2}+\sigma_{\varepsilon N}^{2}\right)}{\sigma_{\theta}^{*2}+\sigma_{\varepsilon A}^{2}+\sigma_{\varepsilon N}^{2}} & 1 & \frac{-\sigma_{\theta}^{*2}}{\sigma_{\theta}^{*2}+\sigma_{\varepsilon A}^{2}+\sigma_{\varepsilon N}^{2}}\\ \frac{\sigma_{\theta}^{*2}+\sigma_{\varepsilon A}^{2}}{\sigma_{\theta}^{*2}+\sigma_{\varepsilon A}^{2}+\sigma_{\varepsilon N}^{2}} & \frac{-\sigma_{\theta}^{*2}}{\sigma_{\theta}^{*2}+\sigma_{\varepsilon A}^{2}+\sigma_{\varepsilon N}^{2}} & 1\end{bmatrix}$$

and denote ρ_{ij}^{AN} the (i,j) element of C_{AN}. Using our definition of $\mathbf{M}_{3}(\cdot)$ and explicit formulas for the moments of the upper-truncated multivariate normal distribution in the trivariate case (derived from recurrence relations) from Kan and Robotti (2017), $\mathbf{M}_{3}(\overrightarrow{0},\tilde{\Sigma}_{AN},1)$ can be re-written as:

$$\mathbf{M}_{3}(\overrightarrow{0},\tilde{\Sigma}_{AN},1)=-\sqrt{\sigma_{\theta}^{*2}+\sigma_{\varepsilon A}^{2}+\sigma_{\varepsilon N}^{2}}\phi(0)\left[\frac{\rho_{12}^{AN}\Phi_{2}\left([\infty,0];\rho_{13\cdot2}^{AN}\right)}{\Phi_{3}\left([\infty,0,0];C_{AN}\right)}+\frac{\rho_{13}^{AN}\Phi_{2}\left([\infty,0];\rho_{12\cdot3}^{AN}\right)}{\Phi_{3}\left([\infty,0,0];C_{AN}\right)}\right]$$

with $\rho_{ij\cdot k}^{AN}=\frac{\rho_{ij}^{AN}-\rho_{ik}^{AN}\rho_{jk}^{AN}}{\sqrt{\left(1-\left(\rho_{ik}^{AN}\right)^{2}\right)\left(1-\left(\rho_{jk}^{AN}\right)^{2}\right)}}$. Noticing that in our case $\Phi_{2}\left([\infty,0];\rho_{ij\cdot k}^{AN}\right)=\frac{1}{2}\ \forall\ i,j,k$, we have

$$\mathbf{M}_{3}(\overrightarrow{0},\tilde{\Sigma}_{AN},1)=\frac{\phi(0)\left[\sigma_{\varepsilon N}^{2}-\sigma_{\varepsilon A}^{2}\right]}{2\sqrt{\sigma_{\theta}^{*2}+\sigma_{\varepsilon A}^{2}+\sigma_{\varepsilon N}^{2}}\Phi_{3}\left([\infty,0,0];C_{AN}\right)},$$

which is positive if and only if $\sigma^2_{\varepsilon N} > \sigma^2_{\varepsilon A}$. Following the same steps we find the expected log income growth of switchers to from non–agriculture to agriculture as:

$$\mathbf{M}_3(\overrightarrow{0}, \tilde{\Sigma}_{NA}, 1) = -\frac{\phi(0)\left[\sigma^2_{\varepsilon N} - \sigma^2_{\varepsilon A}\right]}{2\sqrt{\sigma^{*2}_{\theta} + \sigma^2_{\varepsilon A} + \sigma^2_{\varepsilon N}}\Phi_3\left([\infty, 0, 0]\,; C_{NA}\right)}.$$

Furthermore, it can be verified that $\Phi_3\left([\infty, 0, 0]\,; C_{NA}\right) = \Phi_3\left([\infty, 0, 0]\,; C_{AN}\right)$, which implies that $\mathbf{M}_3(\overrightarrow{0}, \tilde{\Sigma}_{AN}, 1)$ and $\mathbf{M}_3(\overrightarrow{0}, \tilde{\Sigma}_{NA}, 1)$ have the opposite sign but the same magnitude. QED.