

Informal work in sub-Saharan Africa

Dead-end or stepping stone?

Michael Danquah, Simone Schotte^{**}, and Kunal Sen

December 2019

Abstract: Despite rapid economic growth in recent decades, informality remains a persistent phenomenon in the labour markets of many low- and middle-income countries. A key issue in this regard concerns the extent to which informality itself is a persistent state. Using panel data from Ghana, South Africa, Tanzania, and Uganda, this paper presents one of the very few analyses providing evidence on this question in the context of sub-Saharan Africa. Our results reveal an important extent of heterogeneity in the transition patterns observed for workers in upper-tier versus lower-tier informality. Given the limited alternative job opportunities available particularly to those in lower-tier informal self-employment, who often remain locked in a situation of inferior pay and conditions, specific policies that seek to enhance the livelihoods of workers in this most disadvantaged segment may be more relevant in the Sub-Saharan context, than policies that aim to reduce the regulatory barriers to formalisation.

Key words: Informality, segmentation, labour market dynamics, sub-Saharan Africa

JEL classification: D31, J46, J62, O12

Acknowledgements: This paper was prepared as part of the work under the UNU-WIDER project ‘Transforming informal work and livelihoods’. The authors are thankful for valuable feedback and comments on the paper from participants at UNU-WIDER Summer School and the fourth annual conference on ‘Inclusive Growth in Mozambique’ (IGM). The authors thankfully acknowledge Risto Rönkkö for excellent research assistance.

^{*} United Nations University World Institute for Development Economics Research (UNU-WIDER), Katajanokanlaituri 6 B, FI-00160 Helsinki, Finland; schotte@wider.unu.edu.

1 Introduction

In classical accounts of economic development, economic growth is seen to be accompanied by a decline in informal employment. Yet, in many low- and middle-income countries informal forms of economic activity remain a persistent phenomenon despite rapid economic growth in recent decades (Kanbur, 2017). Given these trends, the informal economy now comprises more than 60 per cent of total global employment and more than 90 per cent of all micro and small enterprises (MSEs) worldwide – with the highest shares being observed in sub-Saharan Africa (SSA) and South Asia (ILO, 2018). With premature deindustrialisation and the growth of the informal service sector across these regions, it seems likely that the trajectory towards informalisation may further be intensified in the future (Rodrik 2016).

Informal employment has primarily been portrayed as providing a source of livelihood for the poor (La Porta and Shleifer 2014). At the same time, there is an increasing consensus in recent theoretical and empirical studies that recognise the extent of heterogeneity in informal work (see, *inter alia*, Basu et al., 2018; Chen, 2012; De Mel et al., 2010; De Vreyer and Roubaud, 2013; Grimm et al., 2012; Ngumkeu, 2014; Kanbur, 2017; Paulson and Townsend, 2005; Perry et al., 2007). For example, within informal wage employment, one may observe workers employed as casual labourers in poorly paid unskilled jobs, at the lower end, along with skilled workers employed in better paid jobs that are not covered by labour legislation or social protection provisions but nevertheless require some professional training to obtain these jobs. Similarly, within informal self-employment, one may observe subsistence own-account or household entrepreneurs (often referred to as “penniless entrepreneurs”, see Banerjee and Duflo, 2007, or the “reluctant self-employed”, see Basu et al., 2018) along with more dynamic entrepreneurs with higher productivity and growth potential (sometimes referred to as “constrained gazelles” and “top performers”, see Grimm et al., 2012). Informal workers thus range from multi-dimensionally deprived individuals in subsistence activities which exhibit low returns, are easily accessible, and undesirable relative to formal sector employment, to workers in activities which are better paid, exhibit barriers to entry, and may even be preferred to formal sector employment (Fields, 2019). This internal duality between a primarily exclusion-driven “lower-tier” and a more exit-driven “upper-tier” in informality can be observed both in wage employment and self-employment.

A key issue concerning the persistence of informality in the labour markets of many low- and middle-income countries is whether informality itself – especially in lower-tier work – is a persistent state, such that the most disadvantaged workers are locked in a situation of inferior pay and conditions, or if instead informality is a transient state that all workers are roughly equally likely to experience at some point throughout their working life. Closely related to this is the question whether informal employment provides a “stepping stone” toward formal positions or, on the contrary, presents a “dead end” without better job perspectives, with the result that informal workers either stay in this position or drop out of the labour force (Slonimczyk and Gimpelson, 2015). In this paper, we examine the likelihood of workers to move from lower-tier to upper-tier informal work, and to formal work, as well as the earning implications of such transitions, using comparable panel data for four countries in Sub-Saharan Africa – Ghana, South Africa, Tanzania and Uganda.

From a theoretical standpoint, informal work can be seen as both a “dead end” and a “stepping stone.” On the one hand, informal jobs may be thought to increase workers’ human capital through the accumulation of work experience, add to their social capital through the expansion of professional networks, and have positive signalling effects (in terms of motivation, autonomy, and willingness to work, for example). Moreover, some employers may use informal positions (that can easily be terminated at any time) as a screening device to overcome information asymmetries before offering regular positions (Slonimczyk and Gimpelson, 2015). On the other hand, as

informal employment has often been thought of as a condition of “underemployment,” it may also be associated with human capital depreciation and stigma effects similar to those linked to unemployment. In addition, standard search theory would suggest that workers who take up informal jobs – while continue searching for better formal jobs in the evenings, on weekends, or where possible during working hours – have a lower probability of success than those in open unemployment (Fields, 1990). Finally, if labour markets are segmented, entry-barriers to formal employment may impede easy transitions from informality (Slonimczyk and Gimpelson, 2015; Ulyssea, 2010). Importantly, employment transitions are not unidirectional and existing research particularly on Latin America has highlighted the partly voluntary nature of informality. From this exit perspective, particularly formal wage employment is often presented as an entry point to upper-tier informal self-employment, where workers choose to leave wage jobs to set up their own self-employment activities (Basu et al., 2019; Maloney, 2003, 2004).

A growing literature has examined the extent of heterogeneity in informal work, and patterns of mobility within and across informal and formal employment in the Latin American context (Perry et al., 2007) and, more recently, a small set of transition countries (Commander et al., 2013; Slonimczyk and Gimpelson, 2015). We contribute to this literature by examining the nature, magnitude and direction of employment transition patterns in the informal economy in sub-Saharan Africa, where our knowledge of such transitions is limited. A strength of our analysis is the comparative nature of our study, which allows us to assess whether the patterns of transitions that we observe is specific to one country context or holds true for other countries in our sample. The location of the countries in our study – in Western Africa (Ghana), Eastern Africa (Tanzania and Uganda) and Southern Africa (South Africa) – provides a basis for making generalisable claims on the patterns of mobility within and across the informal economy in Sub-Saharan Africa. The four are among the few sub-Saharan African countries for which at least two waves of household panel data are available, and where variable definitions could be harmonised across countries.

Specifically, in this paper we offer a comparative perspective on the composition of employment and document transition patterns across different formality states, separating between wage and self-employment. Importantly, we not only distinguish between formal and informal employment, understood as the set of economic activities that are not regulated or protected by the state, but offer a more nuanced view that additionally differentiates between “upper-tier” and “lower-tier” informality. We then examine the individual and household-level characteristics that can be associated with these transition patterns and analyse the link between employment and income dynamics. The differentiation between formal, upper-tier informal and lower tier informal employment proves consequential both in terms of employment dynamics and earnings differentials, with earnings being significantly higher in upper-tier informal activities compared to the lower-tier, and highest in formal work. Our results show high persistence in the lower-tier segment of informality, where especially self-employed workers tend to remain locked in a situation of inferior pay and conditions. Informal wage jobs, by contrast, can present a steppingstone into formal employment relationships, especially for those in the more dynamic upper-tier segment. Last, in contrast to the findings from Latin America, we find a relatively strong segmentation between wage and self-employment in the sub-Saharan African case, with few workers exiting formal or upper-informal wage employment for self-employment.

The paper is organized as follows. Section 2 discusses the issue of heterogeneity in informal employment in reference to existing theoretical and empirical studies. Section 3 introduces the data, defines the different employment statuses distinguished in this paper, and discusses the empirical methods used to analyse employment and income dynamics. Section 4 provides a descriptive cross-country analysis of the composition of employment and investigates the patterns of employment mobility using transition matrices. Section 5 presents the results from the econometric analysis. Section 6 summarises our findings and concludes.

2 Heterogeneity in informal employment

Economic theory has provided two competing causal explanations for the existence of informality. One view puts labour market rigidities prevalent in the urban labour market into focus. It considers the informal economy as comprising marginal activities that (in absence of unemployment benefits) provide a source of livelihood for those in society who have been excluded from modern economic opportunities. This *exclusion* may be attributed to above-market-clearing wages set institutionally in the formal sector and/or a mismatch between workers' skills and the skills demanded in modern formal sector jobs. An alternative perspective, on the contrary, assumes a perfectly competitive labour market and considers informality as the result of individual utility-maximisation where individuals *exit* the formal sector after evaluating the costs and benefits of informality relative to formality (Hart, 1973; Maloney, 1999, 2003, 2004). In this view, individuals are endowed with heterogeneous skills, which are remunerated differently in different sectors, and have different preferences, for example, with respect to the willingness to take risks and the value attached to autonomy and independent work.

While both these approaches assume homogeneity within the informal sector, Fields was among the first scholars to describe the informal sector as “having *its own internal duality*” (Fields, 2005: 25). He suggests a bifurcation of the informal economy, where the “upper-tier” is characterised by relatively high earnings and attractive employment conditions while the remaining workers, who are rationed out of both formal sector jobs and upper-tier informal work, are absorbed into the lower informal segment. This synthesized approach reconciles the formerly competing *exit* versus *exclusion* perspectives, in the sense that each explains a different sub-component of informal employment (Kanbur, 2017; Perry et al., 2007; Tonin, 2013). In this regard, workers choose upper-tier informal activities in preference to formal sector work, while lower-tier informal activities present an opportunity of as a last resort.

Although using different methodologies, a wide range of recent studies endorse this two-tier view. This includes both conceptually driven and data-driven approaches. The former assume a dichotomous division of informality according to a-priori defined criteria. Among these, studies emphasising exclusion as the main mechanism typically classify workers into “limited entry” and “free-entry” activities, using entry requirements (skills/capital), permanence (location of the workplace/hours of operation), and firm-size (employment of non-family labour) as key criteria (Fields 1990, 2005; Rogerson, 1996; Ngumkeu, 2014). Those studies emphasising the choice-element (voluntary vs. involuntary exits) rather rely on workers' self-reported reasons for leaving the last (formal) job (Maloney, 2003, 2004). Despite adapting a different methodological angle, more data-driven approaches that derive the number of segments and their characteristics inductively from the data, often arrive at a similar dual structure. Prominent among these are the studies by Cunningham and Maloney (2001), who apply factor and cluster analysis to micro-enterprise data from Mexico, and Günther and Launov (2012), who fit a finite mixture model to household survey data from Côte d'Ivoire in order to detect unobserved heterogeneity in the composition of the informal sector.

An interesting debate between these alternative approaches has sparked with regard to the relative size of the upper-tier and lower-tier informal segments. On the one hand, scholars working on economically more dynamic regions of Latin America and South East Asia, argue that most people are in informal activities by choice (Cunningham and Maloney, 2001). On the other hand, for economically less dynamic regions particularly in Africa, most studies emphasise job rationing as the main determinant dividing a small group of successful entrepreneurs from a much larger group of unsuccessful ones who queue for preferred jobs while subsisting in informality (Fields, 1990; Rogerson, 1996). These conflicting views tend to reflect context-specific differences, while methodological differences may also play a role.

3 Data and methodology

3.1 Data sources and definitions

This study is based on the employment modules of living standard household surveys from four Sub-Saharan African countries: Ghana (Ghana Socio-Economic Panel Survey, GSPS), South Africa (National Income Dynamics Study, NIDS), Tanzania (Tanzania National Panel Study, TZNPS), and Uganda (Uganda National Panel Study, UNPS). In addition to the geographic focus on Sub-Saharan Africa, the choice of countries is based on the share of employment outside of smallholder agriculture and the availability of at least two recent waves of a nationally representative panel data with individual-level information on demographic characteristics, labour earnings, and employment, including direct or indirect information concerning the individual's formality status in employment. For reasons of data availability and cross-country comparability, we focus the analysis on the two most recent waves of panel data available in each of the four countries under study. The data have been collected between 2010 and 2017 with a two- to four-year time gap between panel waves (see Appendix A). The sample is restricted to workers in prime working age (15 to 65 years old).

In defining the activity status of a worker, our study considers the status of the worker in the major job – defined as the economic activity on which the worker has spent most of his or her working time in the case of reporting multiple jobs.¹ In line with the dual structure of informal employment introduced in Section 2, we distinguish between formal employment, upper-tier informal employment, lower-tier informal employment, and unemployment. By separating additionally between wage and self-employment, we identify nine different states of employment characterised in Figure 1 below. To operationalise these work status categories, the variables were carefully harmonized across the surveys (see Appendix A).

We understand informal employment as a job-based concept comprising “all remunerative work (i.e. both self-employment and wage employment), that is not registered, regulated or protected by existing legal or regulatory frameworks, as well as non-remunerative work undertaken in an income-producing enterprise” (ILO, 2019). In absence of comparable information regarding the nature (duration/security) of employment contracts, we follow Henley et al. (2009) in using information concerning social security contributions as the main indicator defining formality status among the wage employed. Accordingly, wage workers with any social security withholdings from their salaries (for medical care, retirement provisions, and/or unemployment insurance) for are considered formal. All remaining are classified as informal, under the assumption that those individuals, who are formal would have answered to be identified as such through any of the preceding questions. Among the self-employed, the formality status of the job is determined by the nature of the enterprise. Thus, self-employed workers (with or without hired workers) operating an enterprise that is officially registered to relevant national institution(s) are classified as in formal employment.² By further subdividing informal employment into upper-tier and lower-tier activities, we aim to capture key structural components of the labour market that are consequential for workers' economic welfare and, according to the existing literature, are considered relevant particularly in the sub-Saharan African context. Therefore, we follow Fields' (1990, 2005) conceptual approach emphasising entry barriers.

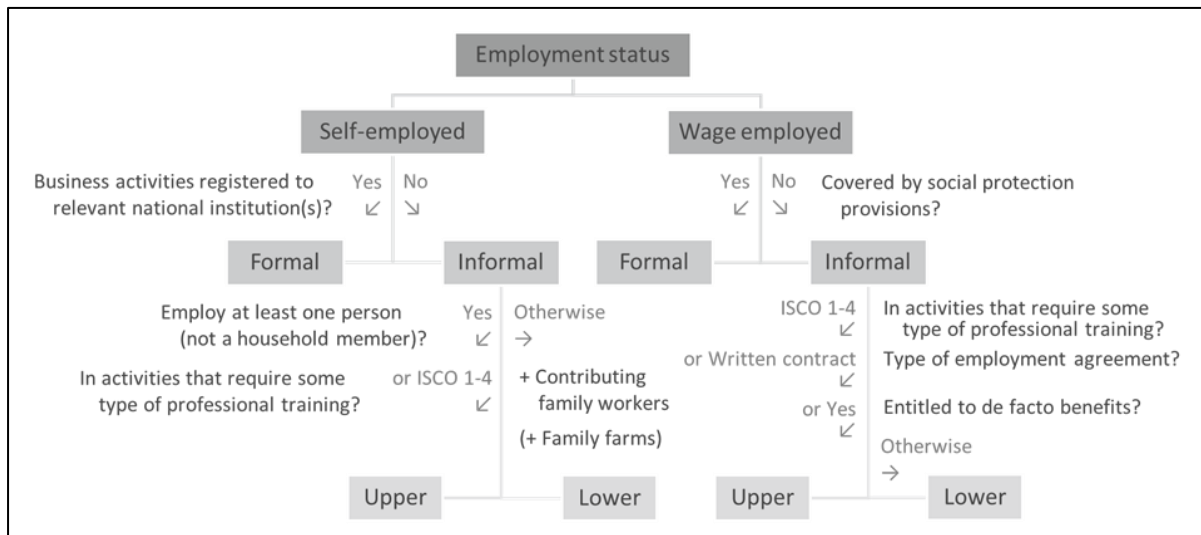
¹ For reasons of cross-country comparability, only information on the current/most recent employment status (past week/7 days) is considered. We do not account for workers holding multiple jobs.

² Due to differences in survey design, the registration to national institution(s) relevant to identify the formality status of enterprises is reported non-uniformly across countries. In South Africa and Uganda, we define businesses that are registered for income tax and/or VAT are classified as formal. In Ghana, enterprise registered with any government agency are defined as formal. In Tanzania, an approximation has been used (see Appendix A).

Accordingly, self-employed workers with unregistered business activities who either employ at least one person (who is not a household member), or are in activities that require some type of professional training (defined as ISCO groups 1-4, covering managers, professionals, technicians, and clerks) are classified as upper informal, while other non-professional own-account workers with unregistered business activities are classified as lower informal. All contributing family workers are classified as lower informal, irrespective of the nature of the enterprise. Workers in smallholder agriculture (family farms) would be classified as lower informal but will be excluded from most of the analysis.

Among the wage workers not covered by social protection provisions, those in professions that require some type of professional training (ISCO 1-4) are classified as upper informal. In addition, we check whether workers report having a written employment agreement and/or are entitled to de facto benefits such as paid sick or maternity leave. This subdivision between formal and upper-tier informal wage employment speaks to the assumption that workers may be willing to trade off social security for higher pay but still demand other benefits that other workers in the firm are receiving (Levy, 2018). The remainder are classified as lower informal.

Figure 1: Work status classification



Source: Authors' own construct.

While many individuals may engage in more than one economic activity, this study only considers the income from the main job. We convert labour earnings reported for different time periods to monthly earnings based on reported working times. All income data is deflated to 2010 prices and, for reasons of cross-country comparability, converted to international dollars using the World Bank's purchasing power parity (PPP) conversion factor for private consumption. Income levels above the 99th percentile of the distribution are considered outliers and replaced by the cut-off value. Only individuals working and reporting strictly positive cash income are included, whereas in-kind income is not taken into consideration. Agricultural income generated by family farms is excluded from the analysis as data on agricultural revenues and costs is relatively noisy.

3.2 Empirical methodology

This study relies on four tools to characterise the nature of employment by work status, assess patterns of mobility, and investigate the relationship between specific transition pathways and individual characteristics and earnings. First, we highlight some key characteristics of each work status group providing a static descriptive assessment. Second, we use basic transition matrices to

examine the conditional transitions probabilities between different employment states. Third, we use a dynamic discrete choice model to analyse how these mobility patterns vary across workers with different characteristics. Fourth, we study the association between changes in work status and changes in labour earnings.

Employment transitions

We use multinomial logistic regression to analyse the dynamics of employment status. Specifically, we are interested in estimating the extent to which the probability of being in a specific employment status at time $t = 1$ depends on the initial status in employment at time $t = 0$ and on observed individual and household attributes, including workers' age, gender, education, and geographic location (rural/ urban).³

Following standard practices in the literature (Gong et al., 2004; Liu, 2015; Maloney, 1999), the statistical model can be formulated as follows. Suppose that each individual i is observed at two points in time $t = \{0, 1\}$, where S_{i0} denotes the initial work status and S_{i1} denotes the final work status. If there are K possible response states indicating the individuals employment status, then $Pr(S_{it} = k | X_{it})$, with $k = \{1, \dots, K\}$, is the probability that individual i has response k at time t given X_{it} , where X_{it} is a column vector of covariates for that observation.

The multinomial model can then be expressed as

$$\pi_{i1k} = \Pr(S_{i1} = k | S_{i0}, X_{i0}) = \left[1 + \sum_{j=1}^K e^{S'_{i0}\theta_j + X'_{i0}\beta_j} \right]^{-1} e^{S'_{i0}\theta_k + X'_{i0}\beta_k} \quad (1)$$

where π_{i1k} represents the transition probability from employment status S_{i0} at baseline to destination state $S_{i1} = k$ at the end of the period, θ_k is the vector of regression parameters for the employment status at baseline, and β_k is the vector of regression parameters for the other covariates on outcome state k . As in any regular multinomial logit regression, the probability $\pi_{i1(K+1)}$ is specified as the residual probability (see Liu, 2015).

In our analysis, the response has six states: i) formal self-employment, ii) upper-tier informal self-employment, iii) lower-tier informal self-employment, iv) formal wage employment, v) upper-tier informal wage employment, vi) and lower-tier informal wage employment. For identifiability, lower-tier informal wage employment is set as the reference category.

By functional transformation, the multinomial logit model can be expressed as a generalized linear model, given by

$$\log\left(\frac{\pi_{itk}}{\pi_{it(K+1)}}\right) = S'_{i0}\theta_k + X'_{i0}\beta_k \quad \text{where } k = 1, \dots, K. \quad (2)$$

We compare our results to an alternative specification using an ordered logit specification. Here, we group together workers in wage and self-employment and define the individual's formality status, F_i , on an ordinal scale, taking on three possible values: lower-tier informal employment,

³ We experimented with adding additional controls on the worker's marital status and composition of the household to the analysis. While these are characteristics previously shown to be associated with the probability to work and the sector-choice of workers (McKay et al., 2018; Sarkar et al., 2019), adding these to our analysis did not significantly affect our findings on the key variables of interest. Results are available from the authors upon request.

$F_i = 1$; upper-tier informal employment $F_i = 2$; and formal employment, $F_i = 3$. We split the sample by the formality work status at time $t = 0$ and express the formality status at time $t = 1$ as a function of the same set of initial worker attributes X_{i0} .

Labour income dynamics

We further analyse the link between employment mobility and changes in earnings. To do so, we regress the change in the logarithm of individual labour earnings between time $t = 0$ and time $t = 1$, Δy_{i1} , on the individual's initial log earnings, y_{i0} , initial work status, S_{i0} , and the set of initial worker characteristics, X_{i0} . Our main interest is in estimating the relationship between transitions in work status and changes in worker's earnings. For this purpose, we add an interaction term between the initial and the final work status, $(S_{i0} \times S_{i1})$. Our dynamic income model is given by

$$\Delta y_{i1} = \alpha + \delta y_{i0} + \vartheta_0 S_{i0} + \vartheta_1 (S_{i0} \times S_{i1}) + \varphi X_{i0} + \varepsilon_i \quad (3)$$

where ε_i denotes the estimation error. In combination, the coefficient estimates for ϑ_0 and ϑ_1 can be interpreted as the earnings penalty/premium associated with staying in or moving to a particular work status.

Initial employment and attrition

We focus our analysis on the employment dynamics observed among individuals who are active in the labour market. Accordingly, our study does not cover entry or exit dynamics. We concentrate specifically on the subset of workers who were employed in non-farm activities in the initial panel wave, as opposed to working on a family farm or being unemployed.⁴ These are likely a non-random sample of the active population (Heckman, 1981). Furthermore, we may have a sample selection problem if attrition is endogenous for estimating employment transitions, as some of the existing literature suggests (Cappellari and Jenkins, 2004, Sarkar et al., 2019).

In order to correct for the double selection problem due to initial employment and panel attrition, we adopt a similar empirical strategy as Sarkar et al. (2019) drawing on previous studies. This method essentially builds on Heckman's two-stage estimation procedure using the inverse Mills ratio to correct for the selection bias. We specify the initial employment decision and sample retention in the following set of equations:

$$E_{i0}^* = \gamma' Z_{i0} + v_{i0} \quad \text{with} \quad E_{i0} = I(E_{i0}^* > 0) \quad (4)$$

$$R_{i1}^* = \psi' W_{i0} + u_{i1} \quad \text{with} \quad R_{i1} = I(R_{i1}^* > 0) \quad (5)$$

where E_{i0}^* captures the latent propensity of non-farm employment in $t = 0$, R_{i1}^* is the latent propensity of panel retention from $t = 0$ to $t = 1$, Z_{i0} and W_{i0} are vectors of baseline characteristics, and $I(\cdot)$ are binary indicator functions equal to one if the underlying latent propensity exceeds some unobserved value (which can be set to zero without loss of generality) and equal to zero otherwise.

Since initial employment and sample attrition are likely to be correlated, we adopt the framework by Sarkar et al. (2019) following the approach suggested by Vella (1998) and estimate questions (4) and (5) simultaneously using a bivariate probit model. We then calculate the selection correction terms for initial employment and sample retention (for the derivation, see Sarkar et al., 2019), and

⁴ We focus our analysis on transitions between the six specified employment states. An expansion of this analysis including family farm activities (under lower-tier informal self-employment) and unemployment as additional destination states is provided in Appendix B.

control for these when estimating equations (2) and (3).⁵ For identification, Z_{i0} and W_{i0} should include some explanatory variables that affect initial employment and panel retention respectively, but are excludable from the main employment transitions equation.

4 Descriptive analysis

A key strength of our analysis is in its comparative nature. The countries included in our study cover different regions – Western Africa (Ghana), Eastern Africa (Tanzania and Uganda) and Southern Africa (South Africa) – as well as the different levels of development – including middle-income (South Africa), lower middle-income (Ghana), and low-income (Tanzania and Uganda) countries. In this section, we assess the differences and commonalities in the composition of employment across these four countries, both from a static and dynamics perspective.

4.1 Composition of employment

The summary statistics presented in **Error! Not a valid bookmark self-reference.** show the aggregated distribution of workers in employment by work status. As widely established in the literature, the composition of the workforce in South Africa differs remarkably from the employment structure observed in poorer sub-Saharan African countries. In South Africa, we observe that 60.6 per cent of those in non-farm employment are formally employed (see **Error! Not a valid bookmark self-reference.**). By contrast, in Ghana and Tanzania only about 20 per cent and in Uganda 16.6 per cent of those employed in non-farm activities are in formal employment, which means that about 80 per cent of the non-farm employment in these three countries is informal. Specifically, in South Africa the vast majority of all non-farm workers are in formal wage employment (56.7 per cent), while in the three other countries, workers in lower-tier informal self-employment present the largest group, accounting for more than 40 per cent of all non-farm employment. The latter share would be yet substantially larger if family farms were included in the analysis (forming part of the lower-tier segment of informality), raising the informality rate in Tanzania and Uganda to just above 90 per cent (see Table B.1 in Appendix B).

The main difference thus consists in the relative absence of lower-tier informal self-employment in South Africa compared to the three other countries. This difference can be attributed to two factors: First, South Africa's economy provides relatively more employment opportunities in the formal economy. Second, in South Africa a larger share of workers can afford to be openly unemployed (23.3 per cent) compared to the three poorer countries (below 2 per cent), where workers revert to survivalist self-employment strategies in absence of other job opportunities and sufficiently developed social protection systems (see Table B.1 in Appendix B).

We observe no large changes in these country-level employment structures between survey waves (see Table B.2 in Appendix B). In Ghana, South Africa, and Uganda, the share of individuals in self-employment in the balanced panel (workers employed non-farm activities in both panel waves) moderately increased, while it slightly decreased in Tanzania. Ghana, South Africa, and Tanzania show an increase in the aggregate rate of formal employment among the balanced panel by 1.5 to 2.5 percentage points (ppts). In Ghana and Tanzania, this was mainly driven by a rise in the share of formal wage employment, mirrored by a decline in the upper-tier segment of informal wage work. By contrast, in South Africa and Uganda, we see an expansion in formal self-employment, accompanied by a moderate decline in formal wage employment. These relatively small changes in

⁵ The inclusion of the sample selection terms marginally affects the size of the coefficient estimates (mostly at the two- or three-digit level) but does not change any of the conclusions presented in this paper.

aggregate shares tend to mask substantial mobility of workers across employment categories, which will be discussed in the next section.

Table 1: Distribution of workers by work status (per cent)

a) Proportion of employment by work status

			Ghana	South Africa	Tanzania	Uganda
Self-employed	Formal		8.9	4.0	9.1	3.1
	Informal	Upper	11.9	5.4	3.9	5.8
		Lower	41.8	3.7	43.4	42.6
Wage employed	Formal		13.2	56.7	11.7	11.7
	Informal	Upper	5.3	8.9	3.3	10.8
		Lower	18.9	21.4	28.6	26.0
TOTAL			100	100	100	100

b) Proportion of formal vs. informal employment

		Ghana	South Africa	Tanzania	Uganda
Formal		22.1	60.6	20.7	14.9
Informal	Upper	17.2	14.3	7.2	16.6
	Lower	60.7	25.1	72.0	68.6
TOTAL		100	100	100	100

c) Proportion of upper-tier informality in informal employment

			Ghana	South Africa	Tanzania	Uganda
Upper informal in total informal employment			22.1	36.4	9.1	19.4
Upper informal in informal self-employment			22.2	59.7	8.2	11.9
Upper informal in informal wage employment			22.0	29.4	10.5	29.4

Note: For each country, summary statistics are compiled for the initial wave of panel study under study. Workers employed on family farms have been excluded from the analysis.

Source: Authors' own calculations based on survey data from GSPS 2009/10, NIDS 2014/15, TZNPS 2010/11, and UNPS 2010/11.

Table 2 below presents three key characteristics of workers by work status. Across countries, informal employment is more common among younger workers. This particularly applies to lower informal wage employment, including casual and time-limited jobs without social protection provisions. Moreover, in all four countries, women (relative to their share in total employment) tend to be underrepresented in formal wage and self-employment and to be importantly overrepresented in lower informal employment. This observation matches the finding that “women in the informal economy are more often found in the most vulnerable situations, for instance as domestic workers, home-based workers or contributing family workers, than their male counterparts” (ILO, 2018: 21).

The level of education is another important factor associated with the level of informality (see also ILO, 2018). As Panel c) in Table 2 illustrates, workers with secondary or tertiary education are overrepresented in formal employment, while workers who have either no education or only completed primary schooling are dominantly found in informal employment. However, there are important differences between workers in self- and wage employment and between the upper-tier and lower-tier segments of informality. Across countries, the share of workers with secondary and tertiary education tends to be highest among those in formal wage employment, followed by the upper-tier informal wage employed. Interestingly, we find that the educational gap between those upper- and lower-tier informality is more pronounced among those in wage employment compared to the self-employed. This can be interpreted to reflect tangible barriers of entry regarding requisite qualification to attain these jobs. This may be attributed to other types of entry barriers in this sector, such as access to credit and capital.

Table 2: Average worker characteristics by work status

a) Average age (years)

a) Average age (years)

			Ghana	South Africa	Tanzania	Uganda
Self-employed	Formal		41.5 (0.64)	41.3 (0.72)	37.5 (0.54)	34.8 (1.29)
	Informal	Upper	40.3 (0.53)	40.2 (0.54)	34.7 (0.83)	38.8 (0.95)
		Lower	40.8 (0.29)	38.6 (0.69)	31.0 (0.33)	34.8 (0.47)
Wage employed	Formal		43.5 (0.53)	37.5 (0.16)	40.5 (0.52)	36.1 (0.71)
	Informal	Upper	36.4 (0.82)	35.6 (0.37)	33.4 (1.00)	34.7 (0.84)
		Lower	38.4 (0.46)	35.4 (0.25)	30.3 (0.35)	30.5 (0.53)
TOTAL			40.5 (0.19)	37.2 (0.12)	32.7 (0.20)	34.1 (0.29)

b) Share of female workers (per cent)

			Ghana	South Africa	Tanzania	Uganda
Self-employed	Formal		47.9 (3.24)	45.8 (3.00)	36.0 (2.83)	39.7 (6.60)
	Informal	Upper	50.0 (2.67)	40.3 (2.25)	25.4 (3.73)	29.4 (4.25)
		Lower	81.7 (1.08)	58.4 (2.96)	63.9 (1.31)	47.9 (1.85)
	Wage employed	Formal		33.7 (2.41)	41.5 (0.75)	32.5 (2.15)
Informal		Upper	30.0 (3.55)	57.0 (1.65)	35.5 (4.33)	34.3 (3.80)
		Lower	33.8 (1.91)	44.3 (1.09)	34.0 (1.55)	26.8 (2.12)
TOTAL			56.8 (0.90)	44.2 (0.55)	46.7 (0.87)	38.9 (1.19)

c) Share of workers with secondary or tertiary education (per cent)

			Ghana	South Africa	Tanzania	Uganda
Self-employed	Formal		18.6 (2.80)	69.3 (2.78)	2.4 (0.91)	42.9 (6.74)
	Informal	Upper	10.0 (1.82)	28.6 (2.07)	3.3 (1.56)	37.3 (4.59)
		Lower	8.3 (0.95)	33.0 (2.82)	1.7 (0.38)	12.2 (1.29)
	Wage employed	Formal		52.2 (3.02)	63.7 (0.74)	36.4 (2.29)
Informal		Upper	37.8 (4.04)	55.5 (1.66)	39.9 (4.76)	59.7 (4.09)
		Lower	19.1 (1.76)	28.0 (0.98)	1.2 (0.36)	11.8 (1.66)
TOTAL			19.0 (0.83)	52.5 (0.55)	7.1 (0.47)	25.8 (1.13)

Note: For each country, summary statistics are compiled for the initial wave of panel study under study. Standard errors of mean values in parentheses.

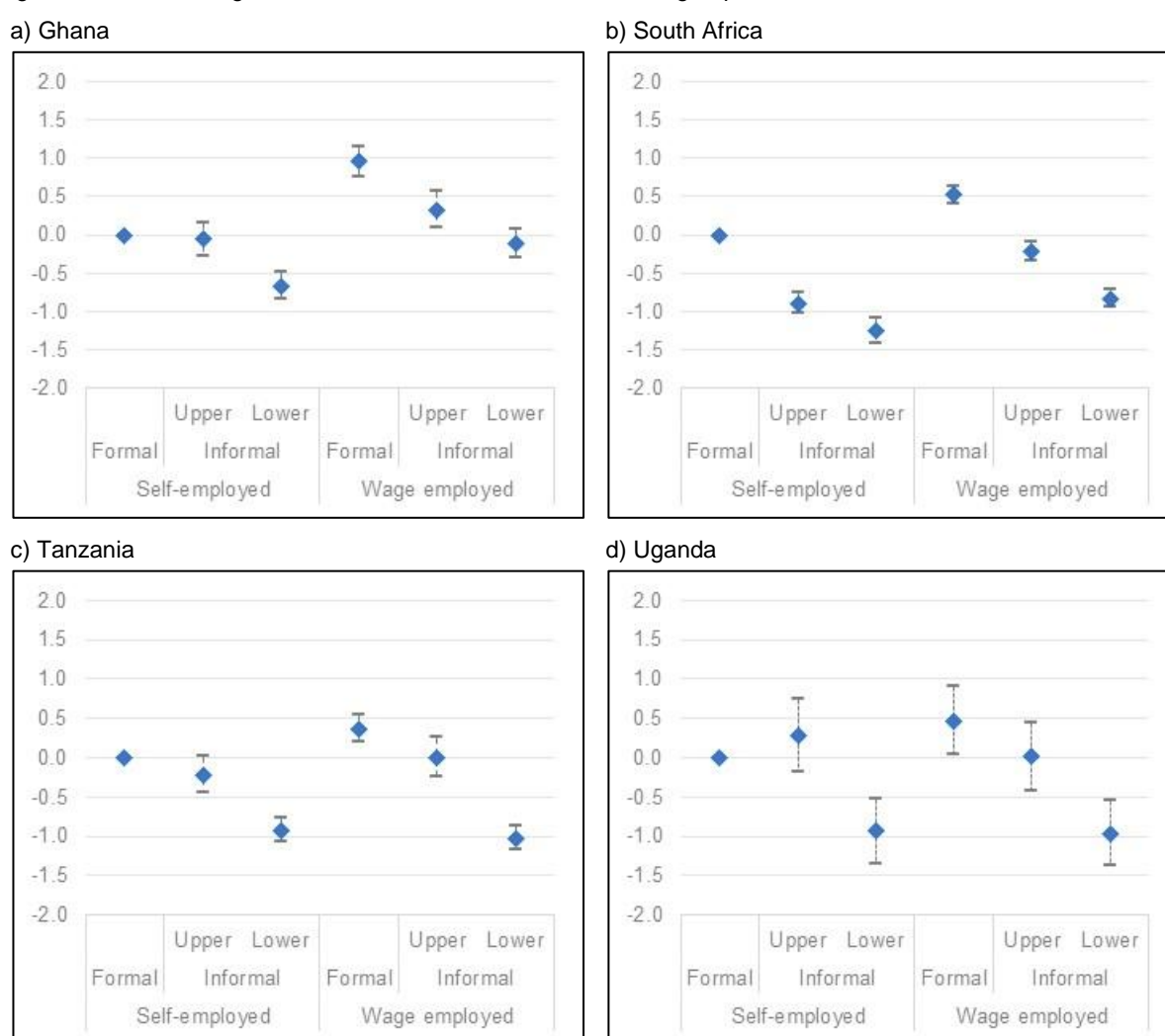
Source: Authors' own calculations based on survey data from GSPS 2009/10, NIDS 2014/15, TZNPS 2010/11, and UNPS 2010/11.

Next, we analyse the within-group and between-group variance (ANOVA) in labour incomes to test for statistically significant differences in mean earnings across the defined work status groups. Figure 2 reports the coefficient estimates of the fitted one-way ANOVA model, regressing log-earnings on the work-status dummies with formal self-employment being the base category. As

expected, labour incomes tend to be higher in formal employment than in informal employment. Moreover, considerable heterogeneity exists within the two informal sub-segments, with earnings being significantly higher in upper-tier informal activities compared to the lower-tier. This pattern is consistent across countries.

Comparing the average labor incomes between survey waves for the balanced sample (see Table B.3 in Appendix B), we find positive annual growth rates in all four countries, which are highest in Ghana (2.8 per cent) and lowest in South Africa (0.6 per cent). The relatively high growth rates (between 2.6 and 6.1 per cent) in the lower-tier segment of informality are good news for the workers in this most disadvantaged segment of the labor market, but must be interpreted in relation to the low average earning levels in this segment (see Figure 2).

Figure 2: ANOVA of log mean labour income across work status groups



Note: Coefficient estimates and 95% confidence interval of fitted one-way ANOVA model of log mean labour income using formal self-employment as the base category.

Source: Authors' own calculations based on survey data from GSPS 2009/10, NIDS 2014/15, TZNPS 2010/11, and UNPS 2010/11.

4.2 Employment transitions

We now move from the preceding static assessment to a dynamic perspective of analysis. The transition matrices in Table 3 report the probabilities with which workers are observed in a certain employment state at the end of the period, conditional on their initial employment state.

Accordingly, the elements in the main diagonal of the transition matrices give the probabilities of staying in the same employment state, while the elements outside the main diagonal give the probabilities of moving to a different employment state. The share of stayers, defined as the proportion of workers who remain in their work status, is calculated as the product of the highlighted diagonals and the initial share of workers in the respective category. We observe the highest employment mobility in Ghana, where just about 50 per cent of all initially employed individuals were observed in the same work status at the end of the period. In the other three countries, the same was true for about 60 per cent of all workers.⁶

The transition patterns vary considerably across countries. However, a commonality observed is that employment stability tends to be highest among the formally wage employed. This may partly be attributed to these jobs being regulated and protected by existing legal standards. In South Africa and Tanzania, around 80 per cent of all workers in formal wage employment remain in this employment state from one survey wave to the next. This share is somewhat lower in Ghana, at 65.1 per cent, and lowest in Uganda, at 48.3 per cent.

Labour turnover tends to be higher in formal self-employment, with important differences observed across countries. In South Africa, among the formally self-employed, 50.8 per cent stay in this state, 13 per cent move into formal wage employment, 23 per cent move into upper-tier informality, and only 13.2 per cent move into lower-tier informality (being either self- or wage employed). On the contrary, only around 30 per cent of the formally self-employed in Ghana, Tanzania and Uganda remain in formal self-employment from one survey wave to the next, while up to 40 per cent move into lower-tier informal self-employment. While these movements may partly be explained by reporting errors, business instability is assumed to also play a major role.

Furthermore, in Ghana, Tanzania, and Uganda, we observe high stability within lower-tier informal self-employment, with around two thirds of the respective workers staying in this segment. The “stickiness” in this segment reflects the limited alternative job opportunities available to workers in this group. Notably, when including family farming under lower-tier informal self-employment activities in the destination state, we observe an even higher level of persistence in this segment of the labour market in these three countries. Hence, a non-negligible share of workers may draw on a combination of self-employment in agriculture and in (short-term) lower-tier non-agricultural informality (see Table B.4 in Appendix B).

In Tanzania and Uganda, we observe a similar level of stagnation within lower-tier informal wage employment, with about 80 per cent of the respective workers either remaining in this category or moving into lower-tier informal self-employment. In Ghana and South Africa, higher mobility out of lower-tier informal wage employment into formal wage employment is observed, suggesting that for about 20 per cent of all workers in this group lower-tier informal wage employment can present a steppingstone into formal employment relationships. This may imply that workers move into such sectors to gain work experience (either voluntarily or due to the limited supply of formal (or regular) job opportunities) before moving into better-paying activities. It may also be a reflection of information asymmetries, where employers first employ workers informally to test their abilities before providing formal contracts.

Furthermore, in all four countries, those in upper-tier informality are more likely to move into formality compared to those in lower-tier informality. While this applies to both those in self- and wage employment, the difference tends to be more pronounced among the wage employed.

⁶ The highest share of stayers is observed in South Africa at 64 per cent. This drop to just below 60 per cent if unemployment is added as an additional destination state (see Table B.4 in the Appendix).

Table 3: Transition matrices across work status groups

a) Ghana

				WAVE t=1						Share of stayers
				Self-employed			Wage-employed			
				Formal	Informal		Formal	Informal		
					Upper	Lower		Upper	Lower	
WAVE t=0	Self-employed	Formal		29.5	15.8	37.1	2.1	3.4	12.1	2.6
		Informal	Upper	14.7	44.7	18.9	4.0	2.5	15.2	5.3
			Lower	6.8	6.3	67.2	2.1	2.0	15.8	28.1
	Wage-employed	Formal		4.3	3.6	10.6	65.1	6.9	9.6	8.6
		Informal	Upper	2.5	3.2	25.8	33.0	11.8	23.7	0.6
			Lower	7.5	6.4	26.9	21.7	5.3	32.2	6.1
			TOTAL			9.1	11.2	35.8	21.0	4.5

b) South Africa

				WAVE t=1						Share of Stayers
				Self-employed			Wage-employed			
				Formal	Informal		Formal	Informal		
					Upper	Lower		Upper	Lower	
WAVE t=0	Self-employed	Formal		50.8	13.3	8.1	13.0	9.7	5.1	2.0
		Informal	Upper	16.1	23.5	17.6	12.2	11.1	19.5	1.4
			Lower	8.8	31.1	33.0	4.9	7.6	14.6	1.1
	Wage-employed	Formal		2.2	1.1	0.4	83.3	7.1	5.9	37.2
		Informal	Upper	6.0	1.0	3.1	50.1	25.1	14.7	2.2
			Lower	4.6	3.8	3.4	27.1	12.8	48.4	10.2
	TOTAL			5.2	3.4	2.8	63.4	9.9	15.3	64.0

c) Tanzania

				WAVE t=1						Share of stayers
				Self-employed			Wage-employed			
				Formal	Informal		Formal	Informal		
					Upper	Lower		Upper	Lower	
WAVE t=0	Self-employed	Formal		31.7	18.2	43.1	2.0	0.0	5.0	2.9
		Informal	Upper	15.2	23.1	34.8	4.5	0.0	22.5	0.9
			Lower	11.0	4.2	66.5	3.8	1.9	12.7	28.2
	Wage-employed	Formal		1.5	1.0	3.1	79.7	4.4	10.2	9.3
		Informal	Upper	6.5	2.9	6.0	45.8	5.8	33.1	0.2
			Lower	2.4	3.6	20.3	14.0	0.8	58.8	17.9
TOTAL				8.6	6.0	31.8	22.2	1.8	29.7	59.3

d) Uganda

				WAVE t=1						Share of stayers
				Self-employed			Wage-employed			
				Formal	Informal		Formal	Informal		
					Upper	Lower		Upper	Lower	
WAVE t=0	Self-employed	Formal		20.7	15.0	51.1	2.7	0.0	10.5	0.6
		Informal	Upper	13.9	39.8	36.8	0.0	4.6	5.0	2.3
			Lower	4.4	7.2	73.2	2.2	1.0	11.9	31.2
	Wage-employed	Formal		5.2	0.0	1.7	58.0	27.0	8.1	6.8
		Informal	Upper	5.7	6.1	5.9	20.0	48.3	14.1	5.2
			Lower	2.7	4.3	13.9	4.7	6.2	68.3	17.7
			TOTAL			5.7	7.9	34.1	14.8	13.3

Note: Each row indicates work statuses in the base period, and each column in transition matrixes indicates work status in the next period; transition matrix rows sum to 100. The likelihood of staying in the same employment status conditional on the base year employment status is highlighted in grey. The share of stayers (proportion of workers who remain in their work status) is calculated as the product of highlighted diagonals and initial size.

Source: Authors' own calculations based on survey data from GSPS 2009/10-2013/14, NIDS 2014/15-2017, TZNPS 2010/11-2012/13, and UNPS 2010/11-2011/12.

5 Regression analysis

We begin this section by a short discussion of the estimation results of the two selection equations capturing initial employment and panel retention. Subsequently, we focus on the dynamics in employment status and labour income.

5.1 Initial employment and attrition

Apart from South Africa, where we observe a substantial share of the active workforce being openly unemployed, the share of workers employed in smallholder agriculture as a percentage of total employment remains very high for Tanzania and Uganda followed by Ghana (see Table B.1 in Appendix B). In consequence, the subset of workers who were employed in non-farm activities in the initial panel wave, as opposed to working on a family farm or being unemployed, ranges between 78.1 per cent in South Africa, 51.7 per cent in Ghana, 34 per cent in Uganda, and 33.2 per cent in Tanzania (see Table B.1 and B.5 in Appendix B). Panel retention rates among the active workforce are highest in South Africa (75.9 per cent) and lowest in Tanzania (69.1 per cent).

For the propensity of initial employment in non-farm activities, we use a binary variable identifying household heads (as opposed to other household members) as an instrument. Across countries, heads of household are significantly more likely to be initially employed in non-farm activities (see Table B.6 in Appendix B), while the variable is validly excludable from the main employment transition equation (see Table B.7 in Appendix B).

Finding a valid instrument for the propensity of panel retention that is available across countries proved difficult. Following a similar approach as Schotte et al. (2018), for South Africa, Tanzania, and Uganda – where at least one previous wave of panel data is available (see Section 3.1 and Appendix A for details) – we use a binary variable indicating whether the respondent was a sample member in the previous survey wave. These original members are more stable survey members compared to those who joined the survey only in $t = 0$ (see Table B.6 in Appendix B). In Ghana, identification relies on the non-linear form of the inverse Mills ratio.

Our regression results show that in Ghana, Tanzania, and Uganda, initial non-farm employment is positively associated with higher levels of education. In South Africa, where the unemployment rate is highest among those with secondary schooling, we observe a U-shaped relationship between the initial employment status and education. We find no clear pattern between education and panel retention. Both initial non-farm employment and panel retention are higher in urban than in rural areas and display an inverted U-shaped relationship with respect to age (Table B.6 in Appendix B).

5.2 Employment transitions

Table 4 presents the conditional transition probabilities estimated from the multinomial logit regression. The average marginal effects in each column are calculated by destination work status in $t = 1$. The reference status that is used as both transitions' starting point and destination is lower-tier informal wage employment. We pool the data for all countries so that the displayed results present cross-country average marginal effects. To ensure that the somewhat different employment structure and transition patterns observed in South Africa compared to the three other countries are not driving the results (see Section 4), we re-estimate the regression excluding South Africa, finding largely similar results (see Table B.8 in Appendix B). We also test how our results change when including family farms (under lower-tier informal self-employment) and unemployment as additional destination states (see Table B.9 in Appendix B), and further explore the coefficient estimates of our control variables using an ordered logit specification (see Table B.10 in Appendix B).

Table 4: Employment transitions

Multinomial logistic regression			Number of obs	=	7,816
Average marginal effects on work status in $t = 1$			Log likelihood	=	-7851.1291
Base outcome: Lower-tier informal wage employed			Pseudo R-squared	=	0.3353
VARIABLES	(1) Formal self- employed	(2) Upper-tier informal self- employed	(3) Lower-tier informal self- employed	(4) Formal wage employed	(5) Upper-tier informal wage employed
Work status in $t = 0$ (base: lower-tier informal wage employed)					
(1) Formal self-employed	0.332*** (0.029)	0.105*** (0.020)	0.172*** (0.022)	-0.198*** (0.022)	-0.068*** (0.009)
(2) Upper-tier informal self-employed	0.121*** (0.012)	0.263*** (0.037)	0.103 (0.068)	-0.165*** (0.018)	-0.034*** (0.012)
(3) Lower-tier informal self-employed	0.104*** (0.016)	0.062** (0.025)	0.312*** (0.028)	-0.153*** (0.008)	-0.062** (0.024)
(4) Formal wage employed	-0.023*** (0.006)	-0.034*** (0.004)	-0.063*** (0.001)	0.475*** (0.025)	-0.024 (0.018)
(5) Upper-tier informal wage employed	0.005 (0.013)	-0.016*** (0.005)	-0.039*** (0.012)	0.140*** (0.022)	0.130*** (0.016)
Level of education (base: no schooling)					
Primary	0.023*** (0.006)	0.001 (0.004)	0.011 (0.026)	0.053*** (0.018)	0.001 (0.009)
Post-primary	0.050*** (0.005)	0.004 (0.005)	-0.025 (0.028)	0.121*** (0.023)	0.011 (0.026)
Secondary	0.052*** (0.003)	-0.002 (0.012)	-0.032 (0.031)	0.196*** (0.025)	0.008 (0.022)
Post-secondary	0.096*** (0.020)	-0.000 (0.012)	-0.084*** (0.028)	0.246*** (0.028)	0.026 (0.022)
Tertiary	0.107*** (0.038)	-0.016 (0.020)	-0.094*** (0.013)	0.342*** (0.024)	-0.001 (0.019)
Age	0.006 (0.004)	0.000 (0.002)	0.003 (0.002)	0.003 (0.003)	-0.002 (0.002)
Age squared (x0.01)	-0.006 (0.005)	-0.001 (0.003)	-0.003 (0.002)	-0.002 (0.004)	0.002 (0.003)
Female	-0.021*** (0.005)	-0.016*** (0.001)	0.052*** (0.011)	-0.037*** (0.009)	0.029*** (0.003)
Urban	0.007 (0.009)	-0.013*** (0.004)	-0.010 (0.012)	0.029*** (0.006)	-0.027*** (0.007)
Country (base: Ghana)					
South Africa	-0.010 (0.006)	-0.020*** (0.003)	-0.135*** (0.013)	0.122*** (0.006)	0.037*** (0.007)
Tanzania	0.017*** (0.002)	-0.036*** (0.005)	-0.040*** (0.015)	0.112*** (0.018)	-0.033*** (0.004)
Uganda	-0.018*** (0.002)	0.007*** (0.002)	0.008 (0.010)	-0.106*** (0.005)	0.099*** (0.010)
Sample selection					
Panel retention from $t = 0$ to $t = 1$	-0.007 (0.011)	-0.009 (0.013)	0.104*** (0.023)	-0.044** (0.020)	-0.019** (0.009)
Employed (non-farm) in $t = 0$	0.005 (0.014)	-0.025** (0.010)	0.001 (0.009)	0.011 (0.012)	-0.002 (0.021)

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors clustered at the country-level in parentheses.

Source: Authors' own calculations based on survey data from GSPS 2009/10-2013/14, NIDS 2014/15-2017, TZNPS 2010/11-2012/13, and UNPS 2010/11-2011/12.

The average marginal effects on initial work status reported in the upper panel of Table 4 can be read similarly to the conditional transition probabilities of a transition matrix. In this sense, the coefficient estimates on being in employment state $k = \{1, \dots, 5\}$ at time $t = 1$, conditional on being observed in the same state k at time $t = 0$ (main diagonal) give an indication of the degree of persistence or state dependence in employment status that is not explained by differences in education, age, gender, and geographic location (Gong et al., 2004; Liu, 2015).⁷

Interestingly, we find a relatively strong segmentation between wage and self-employment. Workers in formal self-employment are most likely to remain in this state or move into informal self-employment. By contrast, transitions from formal self- to formal wage employment are rare and even less likely to occur than a move from lower informal to formal wage employment. Similarly, we find a high degree of persistence in formal wage employment, which is expected given the prevalence of permanent contracts in this groups.

More surprising is the observation that workers originating from formal wage jobs do not display an elevated likelihood of moving into formal or upper-tier informal self-employment, as some of the literature on Latin America would suggest (Maloney 1999, Bosch and Maloney 2010). Our evidence thus does not lend support to the hypothesis that workers use the human capital acquired in formal wage jobs to set up own businesses and benefit from greater flexibility and independence. While workers in lower-tier informal wage jobs are more likely than other wage workers to move into self-employment, we still observe an important extent of segregation between wage and self-employment even in the lower-tier of informality.

Furthermore, as expected, workers in upper-tier informal jobs have significantly higher chances of moving into formal wage jobs than those in the lower tier. It is worthwhile noting that the same does not apply within self-employment, where lower-tier and upper-tier informal workers display a similar (not statistically different) conditional likelihood to formalise. However, this does not imply that both groups face the same obstacles to formalisation, which in the upper-tier may partly be explained by choice.

In line with our descriptive results discussed in Section 4, we find that even after controlling for differences in education, location, age and initial employment status, women have a lower likelihood than men of being in formal wage or self-employment, and are more likely to engage in lower-tier informal self-employment. Interestingly, they are also more likely than men to work in upper-tier informal wage jobs, which excludes them from the social protection benefits associated with formal wage employment. This may be explained by a higher preference for more flexible job arrangements but may also be attributable to the difficulty of females finding jobs in the formal economy. Our results obtained using the ordered logit specification (see Section 3) furthermore indicate that women, on average, face a higher likelihood of dropping out of formal jobs and more often than men slip into upper-tier and particularly lower-tier informality (see Table B.10 in Appendix B).

⁷ Given that we only have two observations per individual, we cannot account for unobserved individual heterogeneity affecting both the initial employment status and transition probabilities. The presented coefficient estimates on the initial employment status must be interpreted in this light. In this regard, “genuine state dependence” would imply that the past employment status itself has an effect on the future state – for example, through signalling or scarring effects to future employers, network effects, or the accumulation or depreciation of human capital. However, as first shown by Heckman (1981, 1991), the detected patterns may also be attributed to sorting effects: First, unobserved individual characteristics and preferences may explain part of the persistence in workers’ employment status choice (see also Gong et al., 2004). Controlling for this “unobserved heterogeneity,” however, would require longer-running panel data that allows for the inclusion of worker-level fixed effects.

Moreover, we find that higher levels of education are associated with a higher likelihood of working formally (see Table 4). While the results for the pooled sample do not show any significant difference in the likelihood of being in upper-tier versus lower-tier informal employment by level of education, this finding changes when excluding South Africa from the sample (see Table B.8 in Appendix B). Here we observe that having primary or post-primary education (compared to no schooling) is associated with a higher likelihood of being in upper-tier rather than lower-tier informal self-employment. Moreover, having post-primary or higher levels of education is associated with a higher likelihood of being in upper-tier rather than lower-tier informal wage employment. Interestingly, as can be seen from Table B.10 in Appendix B, we find that the correlation between educational attainment and formality status is stronger in wage employment than in self-employment.⁸ From this we conclude that the lack of schooling presents an important barrier to attaining formal or upper-tier informal wage jobs, while on the business side other barriers such as access to credit may play an additional role. In line with these findings, the results from the ordered logit specification show that higher levels of education are positively associated with upward movements and negatively associated with downwards movements in formality status (see Table B.10 in Appendix B).

5.3 Labour income dynamics

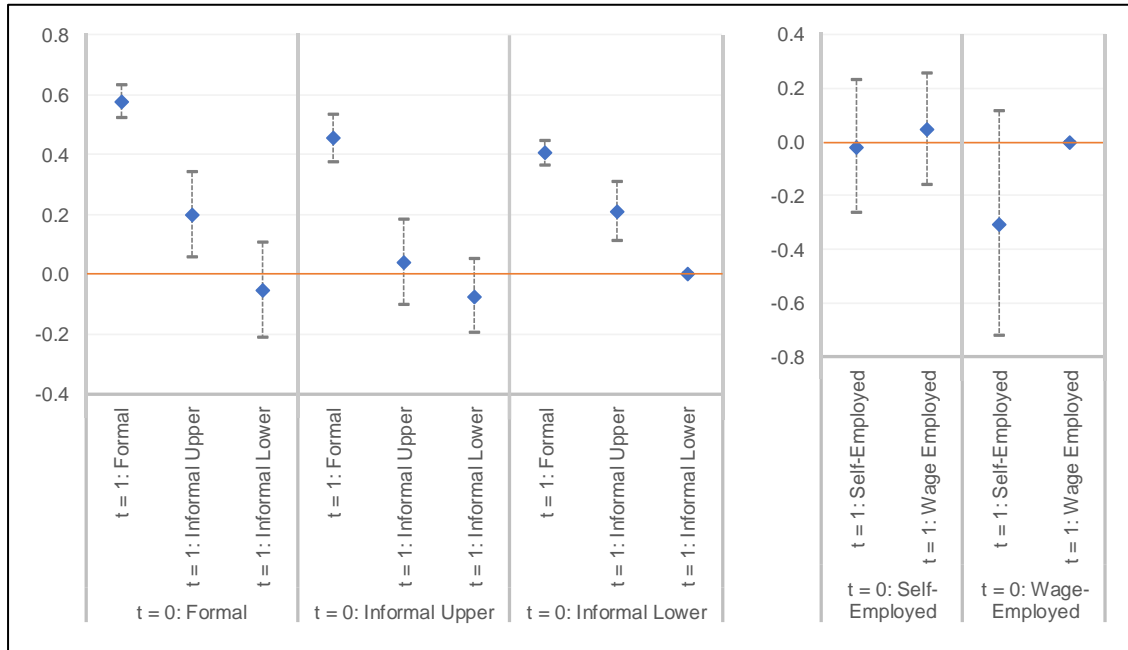
Last, we investigate changes in labour earnings by initial and destination employment state. The main estimated effects are displayed in Figure 3 (see Table B.11 in Appendix B for the full regression results). Again, to keep the number of transition categories manageable for illustrative purposes, we separately control for formality status (formal vs. upper-tier informal vs. lower-tier informal employment) and employment status (wage vs. self-employment). The results for the detailed transition paths (grouping workers by the possible combinations of formality and employment status) are displayed in Figure C.2 in Appendix C.

We observe that workers who transition from lower-tier informal employment to upper-tier informal employment on average experience a rise in earnings relative to those who stay in the same category. As expected, this positive earnings effect is larger for those who move into formal employment. Interestingly, our results show a yet somewhat larger positive change in earnings for those who move into formality out of upper-tier informality. *Ceteris paribus*, the largest inter-temporal change in earnings is experienced by those who were initially in formal employment and sustained this status over time. This may partly be attributable to unobserved individual characteristics of workers in this group but may also reflect premium on experience in this labour market segment. Interestingly, while workers moving from formal to upper-tier informal employment experience a less favourable change in earnings compared to those who remain formal, they still tend to be better off than those who were already initially in upper-tier informal employment and maintained this status. These patterns are observed for workers in both wage and self-employment (see Figure C.2 in Appendix C).

⁸ To further investigate this relationship, we create a binary variable to identify individuals who at least completed secondary schooling. To keep the number of transition categories manageable for illustrative purposes, we re-estimate the main transitions equation controlling separately for formality status (formal vs. upper-tier informal vs. lower-tier informal employment) and employment status (wage vs. self-employment) and interact both categorical variables with the binary indicator for secondary schooling. The average marginal effect of completed secondary education on the destination formality state are presented in Figure C.1 in Appendix C. We find indicative evidence that secondary education has a larger marginal effect on the likelihood to formalise among the initially wage employed compared to the self-employed and is particularly relevant for those in initial upper-tier informality, though both these differences are not statistically significant. Matching the findings from the ordered logit model, secondary education reduces the risk of transitioning from formal employment to upper-tier informality and increases the chances of transitioning from the lower-tier segment to the upper-tier segment of informality

Furthermore, we observe that transitions from self- to wage employment are not significantly associated with an earnings premium, while transition from wage to self-employment often come with an earnings penalty. While a preference for autonomy and independent work may partly compensate this negative income effect, we infer that few workers voluntarily trade a relatively more stable wage job for lower paying and more insecure self-employment.

Figure 3. Labour income dynamics



Note: Each point shows the estimated marginal effect on changes in log earnings by initial and destination employment state, with 'Informal Lower' and 'Wage Employed' being the base categories. The dashed lines show the 95 per cent confidence intervals.

Source: Authors' own calculations based on survey data from GSPS 2009/10-2013/14, NIDS 2014/15-2017, TZNPS 2010/11-2012/13, and UNPS 2010/11-2011/12.

6 Conclusion

This paper examines patterns of labour market transitions between formal and informal employment and across different forms of informality in four sub-Saharan African countries. Using panel data from Ghana, South Africa, Tanzania, and Uganda, the analysis offers a comparative perspective on the composition of employment and documents the transition patterns across work statuses, separating between wage and self-employment.

Our analysis reveals an important extent of heterogeneity in the transition patterns observed for workers in upper-tier versus lower-tier informality. In Ghana, Tanzania, and Uganda, we observe high persistence within lower-tier informal self-employment, with around two thirds of the respective workers staying in this segment. The “stickiness” in this segment reflects the limited alternative job opportunities available to workers in this most disadvantaged group, who tend to remain locked in a situation of inferior pay and conditions. In Tanzania and Uganda, we observe a similar level of stagnation within lower-tier informal wage employment, while in Ghana and South Africa, higher mobility out of lower-tier informal wage employment into formal wage employment is observed. Here we find that for about one out of five workers in this group, lower-tier informal wage employment can present an entry point to formal employment relationships. Nonetheless, for the majority of workers, informal work especially in the lower-tier rather presents a “dead end” rather than a “stepping stone.”

Across all countries, upper-tier informality presents a more dynamic state with a higher proportion of workers formalising than in the lower-tier. However, when controlling for differences in educational attainment and other worker characteristics, the gap in the likelihood of moving into formal self-employment from either upper-tier or lower-tier informal self-employment shrinks and turns insignificant. On the contrary, workers in upper-tier informal wage jobs have significantly higher chances of moving into formal wage jobs than those in the lower tier. This result may partly be explained by formal employers using informal employment relationship as a screening device to overcome information asymmetries and test workers' abilities before providing formal contracts.

As expected, employment stability tends to be highest among the formally wage employed. This can be attributed to these jobs being regulated and protected by existing legal standards. By contrast, formal self-employment is a much more dynamic state, with particularly high mobility into lower-tier informal self-employment being observed in Ghana, Tanzania and Uganda. While these movements may partly be explained by reporting errors, business instability probably plays a major role. Last, we find a relatively strong segmentation between wage and self-employment. That is, transitions between self- and wage employment are comparatively rare and mainly occur in the lower-tier of informality. Exiting formal or upper-informal wage employment for self-employment is not common, particularly among better educated workers.

Our analysis shows that the more nuanced distinction between formal, upper-tier informal and lower-tier informal employment is also consequential in terms of earning dynamics. As expected, labour incomes are on average higher in formal employment than in informal employment. Moreover, across countries considerable heterogeneity exists within the two informal sub-segments, with earnings being significantly higher in upper-tier informal activities compared to the lower-tier. In line with this earnings differential, we find that workers who transition from lower- to upper-tier informal employment on average experience a rise in earnings relative to those who stay in the same category. This positive earnings effect is yet larger for those who move into formal employment. We observe that transitions from self- to wage employment are not significantly associated with an earnings premium, while transition from wage to self-employment often come with an earnings penalty. While a preference for autonomy and independent work may partly compensate this negative income effect, we assume that few workers voluntarily trade a relatively more stable wage job for lower paying and more insecure self-employment.

Summarising, two main findings emerge from this paper that have important implications for policy. Firstly, we find evidence of significant heterogeneity within the informal economy in Sub-Saharan Africa, with clear differences in characteristics between lower-tier and upper-tier informal workers, whether in wage work or self-employment, as well as significant earnings gains for workers who make the transition from lower-tier employment to upper-tier employment. In addition, we find strong evidence of segmentation between wage employment and self-employment in the informal economy. Policymakers need to recognise this heterogeneity in informal work, and devise policies that are not necessarily a "one size fits all" approach to the informal economy. Secondly, given the limited alternative job opportunities available particularly to those in lower-tier informal self-employment, our findings suggest that specific policy measures that seek to enhance the livelihoods of workers in this most disadvantaged segment would be more relevant in the Sub-Saharan context, as compared to policies that aim to reduce the regulatory barriers to formalisation.

References

- Aryeetey, E., Osei-Akoto, I., Darko-Osei, R., & Udry, C. (2011). Ghana Socioeconomic Panel Survey, Report of the Baseline Survey. ISSER/YALE Technical Report.
- Basu, A. K., Chau, N. H., Fields, G. S., & Kanbur, R. (2018). Job creation in a multi-sector labour market model for developing economies. *Oxford Economic Papers*, 71(1), 119-144.
- Banerjee, A. V., & Duflo, E. (2007). The economic lives of the poor. *Journal of economic perspectives*, 21(1), 141-168.
- Bosch, M., & Maloney, W. (2010). Comparative Analysis of Labor Market Dynamics Using Markov Processes: An Application to Informality. *Labour Economics*, 17(4), 621-631.
- Cappellari, L., & Jenkins, S. P. (2004). Modelling low income transitions. *Journal of applied econometrics*, 19(5), 593-610.
- Chen, M. A. (2012). The informal economy: Definitions, theories and policies. WIEGO Working Paper, Vol. 1, No. 26.
- Commander, S., Isachenkova, N., & Rodionova, Y. (2013). Informal employment dynamics in Ukraine: An analytical model of informality in transition economies. *International Labour Review*, 152(3-4), 445-467.
- Cunningham, W. V., & Maloney, W. F. (2001). Heterogeneity among Mexico's microenterprises: An application of factor and cluster analysis. *Economic Development and Cultural Change*, 50(1), 131-156.
- De Mel, S., McKenzie, D., & Woodruff, C. (2010). *Who are the microenterprise owners? Evidence from Sri Lanka*. In: Lerner, J., & Schoar, A. (eds), *International Differences in Entrepreneurship*. Chicago: University of Chicago Press.
- De Vreyer, P., & Roubaud, F. (Eds.) (2013). *Urban labor markets in sub-Saharan Africa*. Washington, DC: World Bank.
- Fields, G. S. (1990). Labour market modelling and the urban informal sector: Theory and evidence [Electronic version]. In: Turnham, D., Salomé, B., & Schwarz, A. (eds.), *The informal sector revisited*. Paris: Organisation for Economic Co-operation and Development.
- Fields, G. S. (2005). A guide to multisector labor market models. Social Protection Discussion Paper Series, No. 0505. Washington, DC: World Bank.
- Fields, G. S. (2019). *Employment and development: How work can lead from and into poverty*. Oxford: Oxford University Press.
- Gong, X., Van Soest, A., & Villagomez, E. (2004). Mobility in the urban labor market: A panel data analysis for Mexico. *Economic Development and Cultural Change*, 53 (1), 1-36.
- Grimm, M., Knorringa, P., & Lay, J. (2012). Constrained gazelles: High potentials in West Africa's informal economy. *World Development*, 40(7), 1352-1368.
- Günther, I., & Launov, A. (2012). Informal employment in developing countries: Opportunity or last resort?. *Journal of Development Economics*, 97(1), 88-98.

- Hart, K. (1973). Informal Income Opportunities and Urban Employment in Ghana. *Journal of Modern African Studies*, 11(1), 61-89.
- Heckman, J. (1981). Statistical Models for Discrete Panel Data. In: Manski, C., & McFadden, D. (eds.), *Structural Analysis of Discrete Data with Econometric Applications*, 114-179. London: MIT Press.
- Heckman, J. (1991). Identifying the Hand of the Past: Distinguishing State Dependence from Heterogeneity. *American Economic Review*, 81(2), 75-79.
- ILO (2018). *Women and men in the informal economy: A statistical picture* (third edition). Geneva: International Labour Office.
- ILO (2019). ILO Thesaurus: informal employment. Available at: <http://ilo.multitools.net>.
- Kanbur, R. (2017). Informality: Causes, consequences and policy responses. *Review of Development Economics*, 21(4), 939-961.
- La Porta, R., & Shleifer, A. (2014). Informality and development. *Journal of Economic Perspectives*, 28(3), 109-26.
- Levy, S. (2018). Under-rewarded Efforts: The Elusive Quest for Prosperity in Mexico, Washington D.C.: IDB.
- Liu, X. (2015). *Methods and applications of longitudinal data analysis*. Elsevier.
- Maloney, W. F. (1999). Does informality imply segmentation in urban labor markets? Evidence from sectoral transitions in Mexico. *World Bank Economic Review*, 13, 275-302.
- Maloney, W. F. (2003). Informal self-employment: Poverty trap or decent alternative. In: Fields, G. S., & Pfeiffermann, G. (eds.), *Pathways out of poverty: Private firms and economic mobility in developing countries*. Boston, MA: Kluwer Academic Publishers.
- Maloney, W. F. (2004). Informality revisited. *World Development*, 32(7), 1159-1178.
- McKay, A., Pirttilä, J., & Schimanski, C. (2018). The elasticity of formal work in African countries. Tampere Economic Working Papers, No. 120.
- Nguimkeu, P. (2014). A structural econometric analysis of the informal sector heterogeneity. *Journal of Development Economics*, 107, 175-191.
- Paulson, A.L., & Townsend, R.M. (2005). Financial constraints and entrepreneurship: Evidence from the Thai financial crisis. *Economic Perspectives*, 29, 34-48.
- Perry, G., Maloney, W., Arias, O., Fajnzylber, P., Mason, A., & Saavedra-Chanduvi, J. (2007) *Informality: Exit and Exclusion*. Washington, DC: World Bank.
- Rodrik, D. (2016). Premature deindustrialization. *Journal of Economic Growth*, 21(1), 1-33.
- Rogerson, C. M. (1996). Urban poverty and the informal economy in South Africa's economic heartland. *Environment and Urbanization*, 8(1), 167-179.

- Sarkar, S., Sahoo, S., & Klasen, S. (2019). Employment transitions of women in India: A panel analysis. *World Development*, 115, 291-309.
- Schotte, S., Zizzamia, R., & Leibbrandt, M. (2018). A poverty dynamics approach to social stratification: The South African case. *World Development*, 110, 88-103.
- Slonimczyk, F., & Gimpelson, V. (2015). Informality and mobility: Evidence from Russian panel data. *Economics of Transition*, 23(2), 299-341.
- Tonin, M. (2013). Informality. In: Cazes, S., & Verick, S. (eds.), *Perspectives on Labour Economics for Development*. Geneva: International Labour Office.
- Ulyssea, G. (2010). Regulation of entry, labor market institutions and the informal sector. *Journal of Development Economics*, 91(1), 87-99.
- Vella, F. (1998). Estimating models with sample selection bias: a survey. *Journal of Human Resources*, 127-169.

Appendix

A. Data sources and operational criteria to define work status

This study is based on household survey data collected in four Sub-Saharan African countries: Ghana, South Africa, Tanzania, and Uganda. For reasons of data availability and cross-country comparability, we focus the analysis on two waves of panel data in each of the four countries under study (listed in Table A.1). Detailed study descriptions are provided below.

Table A.1: Data sources

Country	Survey	t = 0	t = 1
Ghana	Ghana Socio-Economic Panel Survey (GSPS)	Wave I (2009/10)	Wave II (2014)
South Africa	National Income Dynamics Study (NIDS)	Wave IV (2014/15)	Wave V (2017)
Tanzania	Tanzania National Panel Study (TZNPS)	Wave II (2010/11)	Wave III (2012/13)
Uganda	Uganda National Panel Study (UNPS)	Wave II (2010/11)	Wave III (2011/12)

Source: Authors' own construct.

Ghana – Ghana Socio-Economic Panel Survey (GSPS)

The study uses the first and second waves of the Ghana Socio-economic Panel Survey (GSPS) conducted in 2009/10 (GSPS I) and 2013/14 (GSPS II) respectively. The data from the first wave of the GSPS consists of a nationally representative sample of 5,010 households in 334 enumeration areas (EAs) containing 18,889 household members. The second wave covered a sample of 4,774 households containing 16,356 household members. The second wave also tracked movement of households as well as individual within a household. The GSPS collected data on the demographic characteristics of households, education, health, employment, migration, land information, agricultural production input, livestock and household tools, non-farm enterprise, housing characteristics of household, financial assets, psychological measures, and risk preference, social status and responsibilities.

The GSPS data are regionally representative for the 10 regions of Ghana. A two-stage stratified sample design was used for the survey. The first stage involved selecting geographical clusters from an updated master sampling frame constructed from the 2000 Ghana Population and Housing Census. In all, 334 clusters were selected from the list of EAs in each region, based on a simple random sampling technique. A complete household listing was carried out in 2009 in all the selected clusters to produce a sampling frame for the second stage selection of households. The second stage of selection involved a simple random sampling of 15 of the listed households from each selected cluster (see Aryeetey et al., 2011).

South Africa – National Income Dynamics Study (NIDS)

The South African National Income Dynamics Study (NIDS) is funded by the Department of Planning, Monitoring and Evaluation (DPME) of the Government of South Africa, and implemented by the Southern Africa Labour and Development Research Unit (SALDRU) at the University of Cape Town (UCT). NIDS started in 2008 with a nationally representative sample of over 28,000 individuals in 7,300 households. At present, there are five waves of data available, each of which is spaced approximately two years apart.

As the first national panel study of individuals in South Africa, NIDS provides a rich source of individual-level information on demographic characteristics, labour market participation, the nature of employment, working conditions, and earnings, along with information on other socioeconomic characteristics both at the individual and household level. It attempts to track and re-interview respondents as they move out of their original households.

Importantly, the question to those in self-employment asking how many employees there are at the respondent's place of work, excluding him- or herself has only been included in the fourth survey wave. For this reason, we focus our analysis on the two last waves, collected in 2014/15 (NIDS IV) and 2017 (NIDS V) respectively.

Tanzania – Tanzania National Panel Study (TZNPS)

The Tanzania National Panel Study (TZNPS) is a nationally representative household panel survey that collects information on a wide range of topics including household composition and characteristics, individual demographic characteristics, agricultural production, non-farm income generating activities, and other socioeconomic activities. The data is collected by the Tanzania National Bureau of Statistics (NBS) following a stratified two-stage design (51 design strata corresponding to a rural/urban designation for each of the 26 regions; except for Dar es Salaam, which is purely urban and therefore constitutes only one stratum).

The TZNPS sample for the first round conducted in 2008/09 (TZNPS I) covered 3,265 households. The second and third round of the survey were conducted in 2010/11 (TZNPS II) and 2012/13 (TZNPS III) respectively. All original households were targeted for revisit, including split-off tracking of adults who had relocated to a new location. In the fourth round, collected in 2014/15 (TZNPS IV), the sample was refreshed. Since several questions relevant for this study were only introduced in TZNPS II and most households were rotated out between TZNPS III and TZNPS IV, we focus our analysis on TZNPS II + III.

In Tanzania, the two questions that strictly identify informality (see Figure 1 in Section 3) have only been added in TZNPS III. Therefore, we use multiple correspondence analysis (MCA) to create a formality index (separately for wage and self-employment), which draws on variables commonly associated with formality (details on the methodology are available from the authors upon request). Workers are ranked by their index score and the cut-off value (used to differentiate formal/informal work) is chosen to replicate the formality share observed in TZNPS III using the reference definition of formality. For consistency, we use the proxy definition in both TZNPS II and III.

Uganda – Uganda National Panel Study (UNPS)

The Uganda National Panel Study (UNPS), implemented by the Uganda Bureau of Statistics (UBOS), is representative at the national and regional level. The baseline sample was collected in 2009/10 (UNPS I), covering 3,123 households that were distributed over 322 EAs. Within each stratum (rural/urban), the UNPS EAs were selected out of the 783 EAs that had been visited by the Uganda National Household Survey (UNHS) in 2005/06. This initial sample was re-visited in 2010/11 (UNPS II) and 2011/12 (UNPS III), where households or individuals that had permanently left the original households to known locations were tracked and interviewed. After this, in 2013/14 (UNPS IV), one third of the original EAs were permanently rotated out of sample, and 100 new EAs were introduced (extracted from the updated sample frames developed by UBOS from the 2012 Census), resulting in a final sample of 3,119 households that were re-interviewed in 2015/16 (UNPS V). Since the panel was partly refreshed between UNPS III-IV and data for UNPS V has not yet been made publicly available, we focus on UNPS II + III.

B. Complementary tables

Table B.1: Distribution of workers by work status (per cent), extended definition

a) Proportion of employment by work status (incl. family farms and unemployment)

			Ghana	South Africa	Tanzania	Uganda
Self-employed	Formal		4.6	3.1	3.0	1.1
	Informal	Upper	6.2	4.2	1.3	2.0
		Lower	21.7	2.9	14.4	14.5
Family farms			43.4	0.9	65.5	65.0
Wage employed	Formal		6.8	44.3	3.9	4.0
	Informal	Upper	2.8	7.0	1.1	3.7
		Lower	9.8	16.7	9.5	8.8
Unemployment			4.8	21.0	1.3	1.0
TOTAL			100	100	100	100

b) Proportion of formal vs. informal employment (incl. family farms as informal lower)

			Ghana	South Africa	Tanzania	Uganda
Formal			12.0	59.9	7.0	5.1
Informal	Upper		9.4	14.2	2.4	5.7
	Lower		78.6	25.9	90.6	89.2
TOTAL			100	100	100	100

Note: For each country, summary statistics are compiled for the initial wave of panel study under study.

Source: Authors' own calculations based on survey data from GSPS 2009/10, NIDS 2014/15, TZNPS 2010/11, and UNPS 2010/11.

Table B.2: Change in distribution of workers by work status, balanced panel

a) Change (ppts) in proportion of employment by work status

			Ghana	South Africa	Tanzania	Uganda
Self-employed	Formal		0.8	2.0	-2.8	1.3
	Informal	Upper	-1.5	-0.5	0.7	0.5
		Lower	5.3	0.6	-0.1	0.3
Wage employed	Formal		1.7	-0.4	5.0	-2.7
	Informal	Upper	-2.5	1.6	-2.4	0.0
		Lower	-3.8	-3.3	-0.4	0.7

b) Change (ppts) in proportion of formal and informal employment

			Ghana	South Africa	Tanzania	Uganda
Formal			2.5	1.5	2.3	-1.5
Informal	Upper		-4.0	1.1	-1.7	0.5
	Lower		1.5	-2.6	-0.5	1.0

c) Change (ppts) in proportion of self-employment

			Ghana	South Africa	Tanzania	Uganda
Self-employed			4.6	2.1	-2.2	2.1

Source: Authors' own calculations based on survey data from GSPS 2009/10, NIDS 2014/15, TZNPS 2010/11, and UNPS 2010/11.

Table B.3: Annualized growth (per cent) in average log earnings by work status, balanced panel

			Ghana	South Africa	Tanzania	Uganda
Self-employed	Formal		4.0	0.7	4.2	-0.2
	Informal	Upper	1.5	1.6	0.3	3.2
		Lower	2.8	6.0	2.6	6.1
Wage employed	Formal		3.2	0.4	-0.1	1.4
	Informal	Upper	3.0	0.2	0.8	-2.0
		Lower	3.5	-0.5	-0.6	5.4
TOTAL			2.8	0.6	1.0	2.5

Source: Authors' own calculations based on survey data from GSPS 2009/10, NIDS 2014/15, TZNPS 2010/11, and UNPS 2010/11.

Table B.4: Transition matrices across work status groups, including additional destination states (family farms under lower-tier informal self-employment and unemployment)

a) Ghana

				WAVE t=1							Share of stayers	
				Self-employed			Wage-employed					Unempl.
				Formal	Informal		Formal	Informal				
					Upper	Lower		Upper	Lower			
WAVE t=0	Self-employed	Formal		14.9	8.0	64.2	1.1	1.7	6.2	4.0	1.3	
		Informal	Upper	8.0	24.2	52.4	2.2	1.3	8.2	3.8	2.9	
			Lower	2.8	2.6	83.5	0.8	0.8	6.5	3.1	34.9	
	Wage-employed	Formal		3.9	3.2	16.7	59.0	6.2	8.7	2.2	7.8	
		Informal	Upper	2.0	2.5	40.5	25.7	9.2	18.5	1.6	0.5	
			Lower	5.4	4.5	41.2	15.5	3.8	22.9	6.8	4.3	
	TOTAL			5.2	6.4	59.7	11.9	2.6	10.5	3.7	51.7	

b) South Africa

				WAVE t=1							Share of stayers	
				Self-employed			Wage-employed					Unempl.
				Formal	Informal		Formal	Informal				
					Upper	Lower		Upper	Lower			
WAVE t=0	Self-employed	Formal		42.0	11.0	8.5	10.8	8.0	4.2	15.5	1.7	
		Informal	Upper	13.1	19.1	14.3	9.9	9.0	15.8	18.8	1.0	
			Lower	4.7	16.7	21.2	10.0	4.2	11.4	31.9	0.8	
	Wage-employed	Formal		2.1	1.0	0.7	78.9	6.7	5.5	5.1	44.7	
		Informal	Upper	5.0	0.9	2.7	42.0	21.0	12.4	16.1	1.9	
			Lower	3.9	3.2	3.9	22.0	10.9	39.8	16.2	8.5	
TOTAL				4.7	3.1	3.0	56.9	8.9	13.4	10.1	58.6	

c) Tanzania

				WAVE t=1							Share of stayers	
				Self-employed			Wage-employed					Unempl.
				Formal	Informal		Formal	Informal				
					Upper	Lower		Upper	Lower			
WAVE t=0	Self-employed	Formal		25.8	14.8	53.6	1.6	0.0	4.1	0.0	2.3	
		Informal	Upper	12.9	19.6	45.2	3.8	0.0	18.5	0.0	0.8	
			Lower	5.4	2.7	74.8	3.8	1.0	9.9	2.5	32.5	
	Wage-employed	Formal		1.4	0.9	12.7	71.9	4.0	8.8	0.4	8.4	
		Informal	Upper	5.7	2.5	15.9	40.5	5.1	29.2	1.1	0.2	
			Lower	2.2	2.6	35.9	11.1	0.7	46.7	0.8	13.4	
TOTAL				6.4	4.5	49.6	16.5	1.3	20.4	1.3	57.5	

d) Uganda

				WAVE t=1							Share of stayers	
				Self-employed			Wage-employed					Unempl.
				Formal	Informal		Formal	Informal				
					Upper	Lower		Upper	Lower			
WAVE t=0	Self-employed	Formal		19.8	14.3	53.3	2.6	0.0	10.1	0.0	0.6	
		Informal	Upper	11.2	32.1	49.1	0.0	3.7	4.0	0.0	1.8	
			Lower	2.3	3.8	85.4	1.2	0.5	6.3	0.5	36.4	
	Wage-employed	Formal		5.2	0.0	2.2	57.7	26.8	8.1	0.0	6.8	
		Informal	Upper	5.1	5.5	16.0	17.8	43.1	12.6	0.0	4.7	
			Lower	1.8	2.9	41.4	3.2	4.2	46.1	0.5	12.0	
TOTAL				3.9	5.4	54.3	10.2	9.1	16.7	0.3	62.3	

Note: Each row indicates work statuses in the base period, and each column in transition matrixes indicates work status in the next period; transition matrix rows sum to 100. The likelihood of staying in the same employment status conditional on the base year employment status is highlighted in grey. The share of stayers (proportion of workers who remain in their work status) is calculated as the product of highlighted diagonals and initial size.

Source: Authors' own calculations based on survey data from GSPS 2009/10-2013/14, NIDS 2014/15-2017, TZNPIS 2010/11-2012/13, and UNPS 2010/11-2011/12.

Table B.5: (Non-farm) employment and panel retention rates (per cent)

	Ghana		South Africa		Tanzania		Uganda	
	Employed (non-farm)	Panel retention	Employed (non-farm)	Panel retention	Employed (non-farm)	Panel retention	Employed (non-farm)	Panel retention
Percentage share	51.7	70.8	78.1	75.9	33.2	69.1	34.0	71.9

Note: The dummy variable capturing (non-farm) employment equals one if the individual is working in formal or informal wage or self-employment, and zero if working on a family farm or being unemployed. For individuals who responded to the employment module in $t = 0$, the dummy variable capturing panel retention equals one if the individual was successfully re-interviewed in the employment module in $t = 1$, and zero otherwise.

Source: Authors' own calculations based on survey data from GSPS 2009/10-2013/14, NIDS 2014/15-2017, TZNPS 2010/11-2012/13, and UNPS 2010/11-2011/12.

Table B.6: Probability of employment (non-farm) in the baseline and panel retention

	Ghana		South Africa		Tanzania		Uganda	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
VARIABLES	Employed (non-farm)	Panel retention	Employed (non-farm)	Panel retention	Employed (non-farm)	Panel retention	Employed (non-farm)	Panel retention
Level of education (base: no schooling)								
Primary	0.164* (0.086)	-0.044 (0.080)	-0.040 (0.025)	0.013 (0.024)	0.091*** (0.013)	0.014 (0.013)	-0.052** (0.022)	0.004 (0.020)
Post-primary	0.216** (0.085)	-0.034 (0.079)	-0.047* (0.024)	0.013 (0.023)	0.332*** (0.019)	0.014 (0.018)	0.108*** (0.029)	-0.033 (0.026)
Secondary	0.328*** (0.088)	-0.044 (0.082)	0.001 (0.024)	0.010 (0.024)	0.602*** (0.038)	0.127*** (0.032)	0.089** (0.035)	-0.030 (0.031)
Post-secondary	0.506*** (0.093)	-0.018 (0.090)	0.069*** (0.025)	-0.013 (0.025)	0.556*** (0.045)	0.024 (0.045)	0.356*** (0.040)	-0.041 (0.037)
Tertiary	0.521*** (0.091)	-0.081 (0.089)	0.160*** (0.028)	-0.022 (0.032)	0.593*** (0.063)	0.145*** (0.040)	0.124** (0.063)	-0.111* (0.059)
Age	0.033*** (0.004)	0.052*** (0.003)	0.026*** (0.002)	0.010*** (0.002)	0.004* (0.002)	0.031*** (0.002)	0.010*** (0.003)	0.027*** (0.003)
Age squared (x0.01)	-0.042*** (0.005)	-0.064*** (0.004)	-0.024*** (0.003)	-0.013*** (0.003)	-0.010*** (0.003)	-0.039*** (0.003)	-0.016*** (0.004)	-0.030*** (0.004)
Female	0.189*** (0.017)	-0.015 (0.014)	-0.089*** (0.008)	0.061*** (0.008)	-0.001 (0.011)	-0.074*** (0.010)	-0.115*** (0.015)	0.019 (0.013)
Urban	0.264*** (0.012)	-0.032** (0.015)	0.113*** (0.008)	-0.011 (0.008)	0.375*** (0.007)	-0.092*** (0.011)	0.340*** (0.015)	-0.050*** (0.016)
Head of household	0.268*** (0.018)		0.106*** (0.008)		0.108*** (0.012)		0.140*** (0.017)	
Sample member in previous wave		n.a. n.a.		0.234*** (0.008)		0.044*** (0.012)		0.073*** (0.016)
Wald test of $\rho=0$	15.216		10.086		17.313		9.811	
F-statistic (p-value)	0.000		0.002		0.000		0.002	
Observations	6,718		21,498		7,870		4,315	

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors clustered at the country-level in parentheses.

Source: Authors' own calculations based on survey data from GSPS 2009/10-2013/14, NIDS 2014/15-2017, TZNPS 2010/11-2012/13, and UNPS 2010/11-2011/12.

Table B.7: Excludability of instruments from main employment dynamics equation

	Head of household	Sample member in previous wave
F-statistic (p-value)	0.1800	0.9647

Source: Authors' own calculations based on survey data from GSPS 2009/10-2013/14, NIDS 2014/15-2017, TZNPS 2010/11-2012/13, and UNPS 2010/11-2011/12.

Table B.8: Employment dynamics, excluding South Africa

Multinomial logistic regression			Number of obs	=	3,290
Average marginal effects on work status in $t = 1$			Log likelihood	=	-3703.5239
Base outcome: Lower-tier informal wage employed			Pseudo R-squared	=	0.3039
VARIABLES	(1) Formal self- employed	(2) Upper-tier informal self- employed	(3) Lower-tier informal self- employed	(4) Formal wage employed	(5) Upper-tier informal wage employed
Work status in $t = 0$ (base: lower-tier informal wage employed)					
(1) Formal self-employed	0.251*** (0.015)	0.094*** (0.009)	0.241*** (0.035)	-0.173*** (0.017)	-0.035*** (0.009)
(2) Upper-tier informal self-employed	0.127*** (0.009)	0.278*** (0.052)	0.076 (0.096)	-0.155*** (0.004)	-0.015*** (0.002)
(3) Lower-tier informal self-employed	0.068*** (0.023)	0.025* (0.014)	0.410*** (0.033)	-0.150*** (0.015)	-0.024* (0.015)
(4) Formal wage employed	-0.030*** (0.011)	-0.038*** (0.010)	-0.103*** (0.010)	0.412*** (0.051)	0.061*** (0.007)
(5) Upper-tier informal wage employed	-0.001 (0.023)	-0.008 (0.010)	-0.058 (0.044)	0.105*** (0.034)	0.135*** (0.024)
Level of education (base: no schooling)					
Primary	0.040*** (0.015)	0.008 (0.005)	0.000 (0.050)	-0.014* (0.008)	0.021* (0.011)
Post-primary	0.073*** (0.018)	0.017* (0.009)	-0.095*** (0.031)	0.042 (0.043)	0.052*** (0.015)
Secondary	0.081*** (0.012)	0.015 (0.032)	-0.117** (0.050)	0.120*** (0.038)	0.050*** (0.016)
Post-secondary	0.081** (0.033)	0.022 (0.044)	-0.183** (0.072)	0.170*** (0.066)	0.071*** (0.013)
Tertiary	0.120*** (0.038)	-0.017 (0.043)	-0.205*** (0.046)	0.280*** (0.025)	0.061** (0.028)
Age	0.007 (0.012)	0.005 (0.003)	0.005 (0.007)	0.003* (0.002)	-0.006*** (0.002)
Age squared (x0.01)	-0.007 (0.014)	-0.006** (0.003)	-0.004 (0.008)	-0.001 (0.002)	0.008*** (0.002)
Female	-0.033** (0.014)	-0.031*** (0.005)	0.107*** (0.025)	0.006 (0.025)	0.010 (0.011)
Urban	-0.023 (0.022)	-0.014 (0.018)	-0.036 (0.044)	-0.002 (0.009)	-0.008 (0.019)
Country (base: Ghana)					
Tanzania	0.033*** (0.004)	-0.043*** (0.009)	-0.058*** (0.007)	0.087*** (0.015)	-0.017*** (0.003)
Uganda	-0.020*** (0.000)	0.010** (0.004)	-0.002 (0.007)	-0.060*** (0.010)	0.067*** (0.008)
Sample selection					
Panel retention from $t = 0$ to $t = 1$	-0.057 (0.100)	0.056 (0.042)	0.239*** (0.033)	-0.153** (0.072)	-0.063*** (0.018)
Employed (non-farm) in $t = 0$	-0.044** (0.020)	-0.018 (0.013)	-0.024 (0.045)	-0.020 (0.015)	0.003 (0.021)

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors clustered at the country-level in parentheses.

Source: Authors' own calculations based on survey data from GSPS 2009/10-2013/14, NIDS 2014/15-2017, TZNPS 2010/11-2012/13, and UNPS 2010/11-2011/12.

Table B.9: Employment dynamics, including additional destination states (family farms under lower-tier informal self-employment and unemployment)

Multinomial logistic regression				Number of obs	=	9,696
Average marginal effects on work status in $t = 1$				Log likelihood	=	-10,633.255
Base outcome: Lower-tier informal wage employed				Pseudo R-squared	=	0.3393
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Formal self-employed	Upper-tier informal self-employed	Lower-tier informal self-employed	Formal wage employed	Upper-tier informal wage employed	Unemployment
Work status in $t = 0$ (base: lower-tier informal wage employed)						
(1) Formal self-employed	0.259*** (0.021)	0.082*** (0.019)	0.161*** (0.003)	-0.151*** (0.015)	-0.053*** (0.007)	-0.023** (0.011)
(2) Upper-tier informal self-employed	0.087*** (0.011)	0.192*** (0.024)	0.106*** (0.036)	-0.132*** (0.013)	-0.030*** (0.010)	0.019*** (0.004)
(3) Lower-tier informal self-employed	0.064*** (0.011)	0.035** (0.016)	0.261*** (0.031)	-0.131*** (0.007)	-0.055*** (0.016)	0.062*** (0.020)
(4) Formal wage employed	-0.016*** (0.005)	-0.025*** (0.003)	-0.095*** (0.018)	0.451*** (0.028)	-0.005 (0.016)	-0.057*** (0.001)
(5) Upper-tier informal wage employed	0.008 (0.012)	-0.010** (0.005)	-0.056*** (0.019)	0.126*** (0.018)	0.120*** (0.013)	-0.029*** (0.007)
Level of education (base: no schooling)						
Primary	0.019*** (0.006)	0.000 (0.005)	-0.004 (0.032)	0.028 (0.022)	-0.000 (0.008)	0.032*** (0.006)
Post-primary	0.041*** (0.006)	0.003 (0.005)	-0.043 (0.029)	0.080*** (0.015)	0.008 (0.021)	0.039*** (0.007)
Secondary	0.045*** (0.004)	-0.001 (0.011)	-0.053* (0.029)	0.144*** (0.016)	0.007 (0.018)	0.030*** (0.003)
Post-secondary	0.080*** (0.023)	-0.001 (0.013)	-0.086** (0.034)	0.194*** (0.018)	0.021 (0.017)	0.019** (0.007)
Tertiary	0.092** (0.036)	-0.012 (0.020)	-0.093*** (0.035)	0.281*** (0.027)	0.000 (0.017)	0.002 (0.012)
Age	0.005 (0.004)	0.001 (0.002)	0.002 (0.005)	0.004 (0.005)	-0.001 (0.001)	-0.003 (0.005)
Age squared (x0.01)	-0.006 (0.005)	-0.001 (0.003)	-0.001 (0.006)	-0.003 (0.006)	0.001 (0.002)	0.001 (0.006)
Female	-0.022*** (0.004)	-0.017*** (0.004)	0.052*** (0.018)	-0.034*** (0.007)	0.022*** (0.003)	0.018* (0.010)
Urban	0.007 (0.013)	-0.010 (0.008)	-0.017 (0.014)	0.021*** (0.006)	-0.022*** (0.004)	0.003 (0.008)
Country (base: Ghana)						
South Africa	0.004 (0.008)	-0.005 (0.003)	-0.291*** (0.025)	0.115*** (0.012)	0.036*** (0.006)	0.105*** (0.008)
Tanzania	0.032*** (0.002)	-0.015*** (0.003)	-0.107*** (0.022)	0.117*** (0.018)	-0.020*** (0.003)	-0.033*** (0.003)
Uganda	-0.006*** (0.002)	0.016*** (0.001)	-0.028* (0.016)	-0.064*** (0.006)	0.090*** (0.008)	-0.042*** (0.003)
Sample selection						
Panel retention from $t = 0$ to $t = 1$	-0.005 (0.014)	-0.011 (0.016)	0.073 (0.048)	-0.045*** (0.015)	-0.017*** (0.006)	0.029 (0.022)
Employed (non-farm) in $t = 0$	-0.005 (0.021)	-0.035*** (0.011)	0.082*** (0.019)	-0.025 (0.028)	-0.012 (0.011)	0.025 (0.045)

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors clustered at the country-level in parentheses.

Source: Authors' own calculations based on survey data from GSPS 2009/10-2013/14, NIDS 2014/15-2017, TZNPS 2010/11-2012/13, and UNPS 2010/11-2011/12.

Table B.10: Employment dynamics, using ordered logistic regression

Ordered logistic regression, average marginal effects, split sample						
	(1)		(2)		(3)	
Formality status in $t = 0$	Lower-tier informal		Upper-tier informal		Formal	
Formality status in $t = 1$	Formal	Upper-tier informal	Formal	Lower-tier informal	Upper-tier informal	Lower-tier informal
VARIABLES						
Employment status (base: Wage employed)						
Self-employed	-0.036 (0.026)	-0.013 (0.012)	-0.122** (0.057)	0.114** (0.055)	0.081*** (0.007)	0.185*** (0.034)
Level of education (base: no schooling)						
Primary	0.033 (0.020)	0.019** (0.008)	0.022 (0.021)	-0.033 (0.035)	-0.034* (0.020)	-0.125* (0.073)
Post-primary	0.121** (0.053)	0.056*** (0.011)	0.117 (0.080)	-0.145 (0.109)	-0.069*** (0.016)	-0.197*** (0.055)
Secondary	0.208*** (0.074)	0.077*** (0.005)	0.238*** (0.068)	-0.245*** (0.093)	-0.092*** (0.020)	-0.232*** (0.055)
Post-secondary	0.281*** (0.043)	0.084*** (0.004)	0.328*** (0.076)	-0.300*** (0.096)	-0.114*** (0.024)	-0.263*** (0.059)
Tertiary	0.575*** (0.064)	0.048** (0.023)	0.554*** (0.101)	-0.395*** (0.084)	-0.133*** (0.024)	-0.286*** (0.054)
Age	0.003 (0.004)	0.001 (0.001)	0.013 (0.010)	-0.011 (0.009)	-0.004** (0.002)	-0.007** (0.003)
Age squared (x0.01)	-0.002 (0.004)	-0.001 (0.001)	-0.014 (0.009)	0.012 (0.009)	0.004* (0.002)	0.006* (0.004)
Female	-0.066** (0.031)	-0.023*** (0.005)	-0.014 (0.048)	0.013 (0.044)	0.019*** (0.005)	0.034*** (0.009)
Urban	0.021* (0.012)	0.007 (0.006)	0.046** (0.019)	-0.042** (0.018)	0.001 (0.007)	0.003 (0.013)
Country (base: Ghana)						
South Africa	0.102*** (0.015)	0.036*** (0.009)	0.068*** (0.010)	-0.061*** (0.010)	-0.058*** (0.012)	-0.122*** (0.004)
Tanzania	0.038 (0.031)	0.016** (0.007)	0.027 (0.035)	-0.026 (0.033)	-0.049*** (0.004)	-0.106*** (0.008)
Uganda	-0.040*** (0.010)	-0.021* (0.012)	-0.080*** (0.014)	0.090*** (0.015)	0.013*** (0.001)	0.045*** (0.005)
Sample selection						
Panel retention from $t = 0$ to $t = 1$	-0.144* (0.082)	-0.049*** (0.014)	-0.024 (0.044)	0.021 (0.040)	0.026* (0.014)	0.045 (0.030)
Employed (non-farm) in $t = 0$	0.013 (0.023)	0.004 (0.009)	0.090*** (0.030)	-0.082*** (0.026)	0.025* (0.014)	0.043* (0.024)
Observations	2,869		1,166		3,781	
Pseudo R-squared	0.0728		0.0614		0.1220	

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors clustered at the country-level in parentheses.

Source: Authors' own calculations based on survey data from GSPS 2009/10-2013/14, NIDS 2014/15-2017, TZNPS 2010/11-2012/13, and UNPS 2010/11-2011/12.

Table B.11: Change in log labour earnings

VARIABLES	All countries		Excl. South Africa	
	(1)	(2)	(3)	(4)
Log labour earnings ($t = 0$)	-0.646*** (0.033)	-0.643*** (0.034)	-0.695*** (0.062)	-0.696*** (0.060)
Formality status in $t = 0$ (base: Lower informal in $t = 0$)				
Formal ($t = 0$)	-0.052 (0.083)	-0.051 (0.081)	0.059 (0.089)	0.060 (0.100)
Upper informal ($t = 0$)	-0.068 (0.068)	-0.074 (0.063)	-0.031 (0.141)	-0.019 (0.134)
Formality status in $t = 0$ and $t = 1$ (base: Lower informal in $t = 0$ and $t = 1$)				
Formal ($t = 0$) x Formal ($t = 1$)	0.653** (0.118)	0.629*** (0.102)	0.474* (0.128)	0.452** (0.066)
Formal ($t = 0$) x Upper Informal ($t = 1$)	0.261 (0.144)	0.252 (0.141)	0.023 (0.135)	0.003 (0.150)
Upper Informal ($t = 0$) x Formal ($t = 1$)	0.532*** (0.079)	0.528*** (0.064)	0.577* (0.185)	0.556* (0.161)
Upper informal ($t = 0$) x Upper Informal ($t = 1$)	0.118 (0.051)	0.113 (0.051)	0.155 (0.118)	0.126 (0.126)
Lower Informal ($t = 0$) x Formal ($t = 1$)	0.404*** (0.029)	0.404*** (0.021)	0.399** (0.068)	0.405** (0.060)
Lower informal ($t = 0$) x Upper Informal ($t = 0$)	0.165* (0.058)	0.209** (0.050)	0.120 (0.120)	0.166 (0.062)
Employment status in $t = 0$ (base: Wage employed in $t = 0$)				
Self-employed ($t = 0$)		0.045 (0.106)		-0.091** (0.014)
Employment status in $t = 0$ and $t = 1$ (base: Wage employed in $t = 0$ and $t = 1$)				
Self-employed ($t = 0$) x Self-employed ($t = 0$)		-0.064 (0.216)		0.095 (0.214)
Wage employed ($t = 0$) x Self-employed ($t = 1$)		-0.307 (0.213)		-0.217 (0.387)
Characteristics in $t = 0$				
Level of education (base: No schooling)				
Primary	0.113 (0.066)	0.124 (0.062)	0.177* (0.057)	0.170 (0.074)
Post-primary	0.203** (0.035)	0.214*** (0.026)	0.263** (0.032)	0.251* (0.063)
Secondary	0.390*** (0.043)	0.399*** (0.037)	0.461** (0.062)	0.439** (0.048)
Post-secondary	0.633*** (0.041)	0.644*** (0.039)	0.549** (0.070)	0.533* (0.163)
Tertiary	1.045*** (0.077)	1.050*** (0.073)	0.956*** (0.086)	0.937*** (0.028)
Age	0.011 (0.010)	0.011 (0.008)	0.018 (0.013)	0.016 (0.013)
Age squared (x0.01)	-0.012 (0.013)	-0.011 (0.012)	-0.020 (0.015)	-0.019 (0.018)
Female	-0.246*** (0.018)	-0.250*** (0.021)	-0.262* (0.066)	-0.275* (0.066)
Urban	0.141* (0.045)	0.143* (0.047)	0.232 (0.163)	0.242 (0.154)
Country (base: Uganda)				
South Africa	-0.180** (0.043)	-0.209* (0.078)		
Tanzania	-0.241*** (0.016)	-0.259*** (0.017)	-0.221* (0.058)	-0.232* (0.074)
Uganda	-0.457*** (0.019)	-0.480*** (0.015)	-0.417*** (0.032)	-0.436*** (0.033)

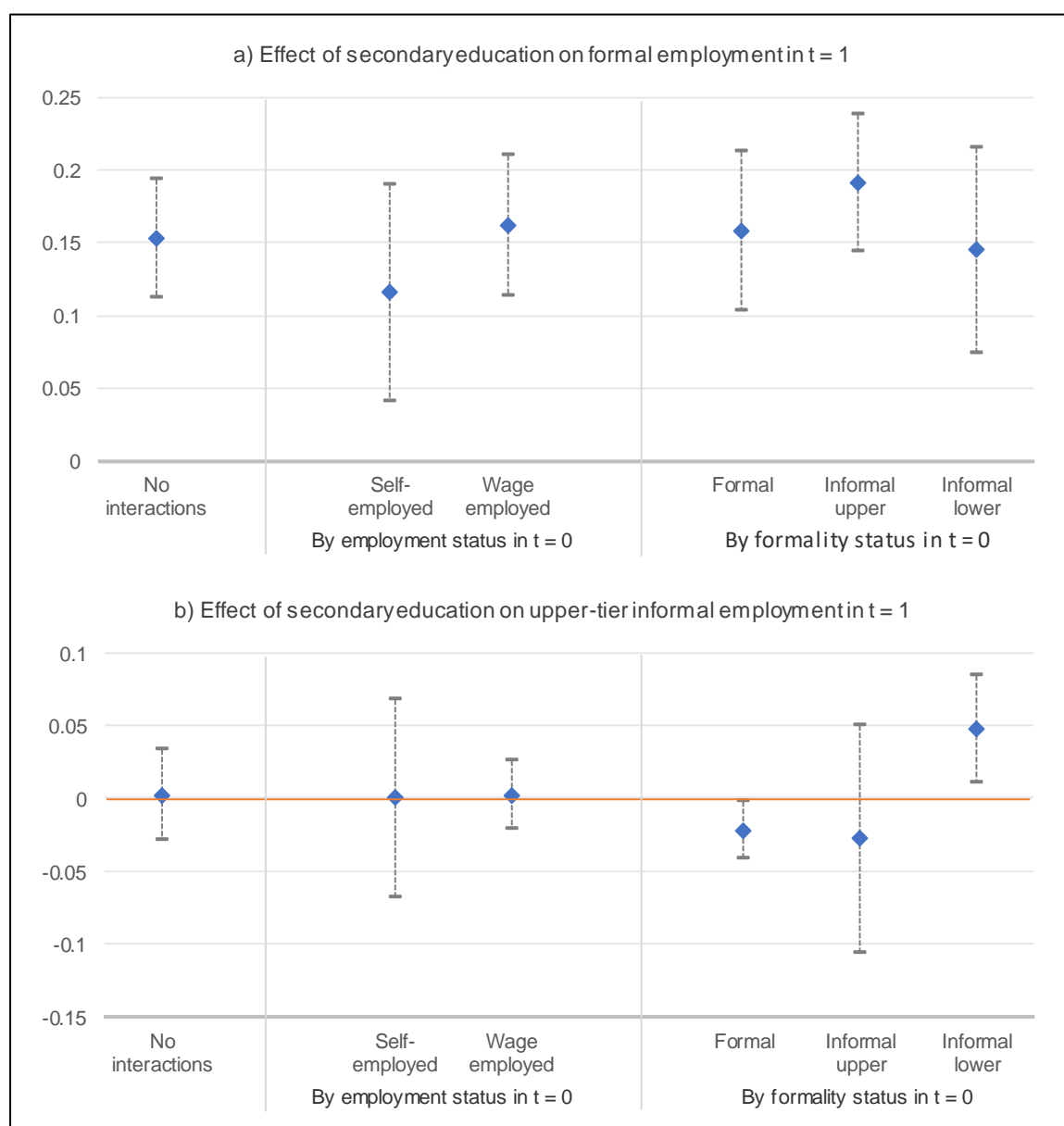
Sample selection				
Panel retention from $t = 0$ to $t = 1$	-0.034 (0.056)	-0.022 (0.044)	-0.210 (0.330)	-0.186 (0.288)
Employed (non-farm) in $t = 0$	-0.014 (0.045)	-0.012 (0.030)	-0.006 (0.152)	0.008 (0.147)
Constant	3.260*** (0.286)	3.281*** (0.254)	3.379** (0.435)	3.409** (0.375)
Observations	7,240	7,240	2,760	2,760
Adj. R-squared	0.371	0.376	0.382	0.384

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors clustered at the country-level in parentheses.

Source: Authors' own calculations based on survey data from GSPS 2009/10-2013/14, NIDS 2014/15-2017, TZNPS 2010/11-2012/13, and UNPS 2010/11-2011/12.

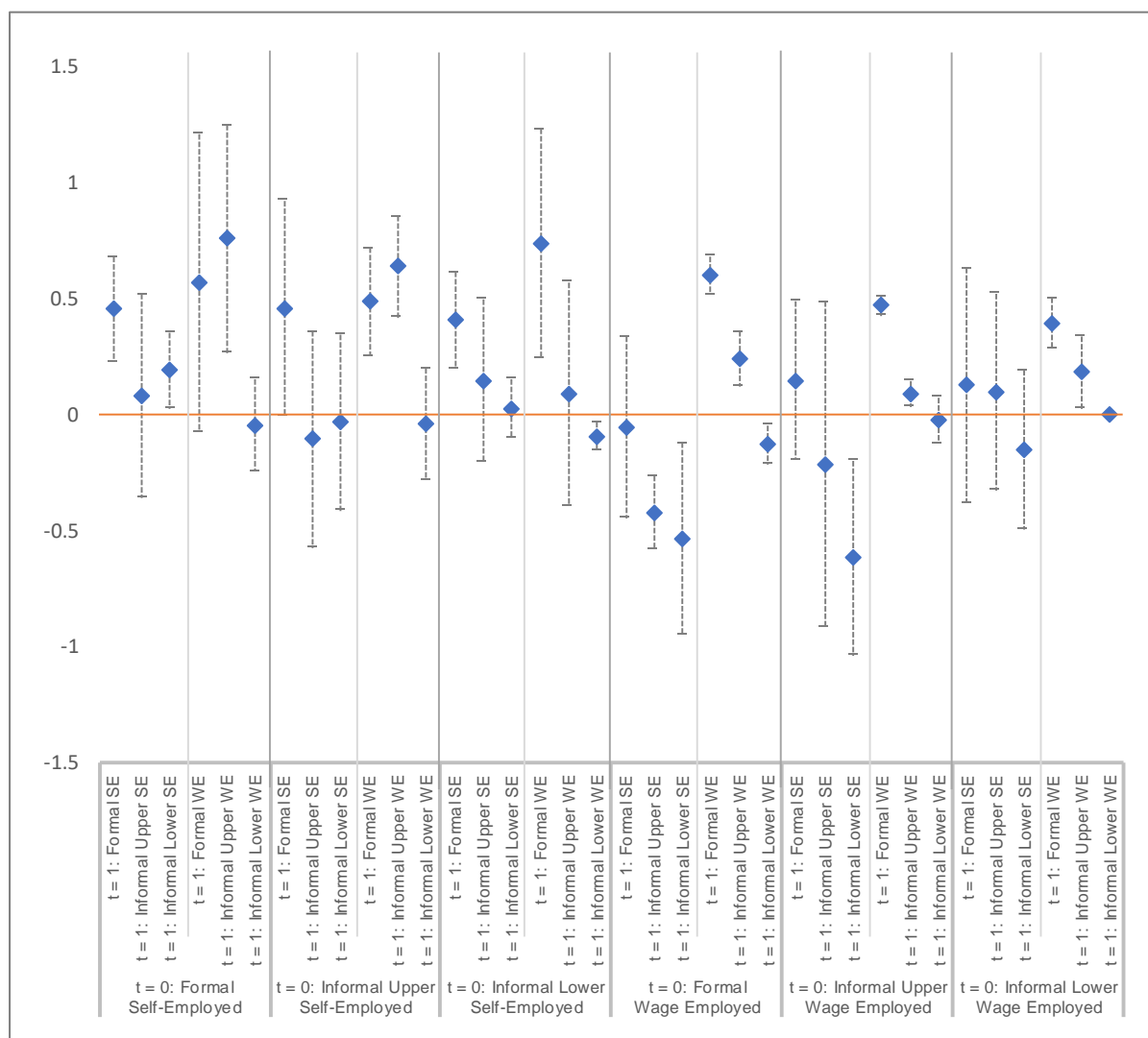
C. Complementary figures

Figure C.1: Average marginal effect of completed secondary education by initial formality status



Source: Authors' own calculations based on survey data from GSPS 2009/10-2013/14, NIDS 2014/15-2017, TZNPS 2010/11-2012/13, and UNPS 2010/11-2011/12.

Figure C.2: Earning dynamics, six employment states



Source: Authors' own calculations based on survey data from GSPS 2009/10-2013/14, NIDS 2014/15-2017, TZNPS 2010/11-2012/13, and UNPS 2010/11-2011/12.