Supplementary material for

"Does the depth of informality influence welfare in urban Sub-Saharan Africa?"

Eva-Maria Egger, Cecilia Poggi, and Héctor Rufrancos

A   Definitions and summary statistics tables

A.1   Definition of formality

Our definition of formality is that of access to social insurance via wage employment, as often applied in the literature to identify formal workers (see for instance Ohnsorge and Yu (2021) and Ulyssea (2020)). The social insurance definition follows the recommendations from the International Conference on Labor Statistics (see ICLS15 (1993); ICLS17 (2003)), and it includes, for example, workers who receive health insurance coverage or pension contributions from their employer (and/or contribute to such schemes). It does not look at social assistance or other

social protection programs, such as cash transfers.[1]

Alternative definitions of formality could include the contract status of an employee, e.g. whether they have a signed contract, or employer characteristics, e.g. employer withholds taxes or has more than 5 employees. However, these definitions are less comparable across surveys, and for our research objective they insufficiently capture the insurance access directly provided to the worker from which household members could benefit as well. This characterisation we propose does not measure a direct benefit from social security instruments, but rather the willingness to forego some income to be covered in case of a life-cycle event. It should be noted, however, that in the countries studied, a wage job with social insurance is in most cases also a job with a signed contract.[2] Our definition of formal employment hence also captures a certain job security.

We acknowledge that most activities in the studied countries are likely to be in self-employment (independent work) which is usually informal. Some self-employment could be considered formal under a different definition of formality, such as tax registration or business size (Dávila, 1994; Maloney, 1999, 2004; Heckman and Pagés, 2004; IDB, 2004; Perry et al., 2007)). However, as our definition is that of social insurance access, formality can only apply to wage employment. Thus, all self-employed activity in our sample is thus considered informal, as for example in Ohnsorge and Yu (2021). Using this definition, we

---

[1]Note that the informal sector we characterise here should not be confused with forms of illegal production, domestic production for own final consumption or underground activities (see OECD et al (2002, p.39).

[2]The share of jobs with social insurance that also have a signed contract with regular payment ranges from 57% in Nigeria to 96% in Tanzania.

choose to focus on accessing the "security" of formal wage employment, without diversifying for any differential form of informal employment (between wage and independent). The reason behind this choice is that self-employment is not easily comparable to formal wage work (Alaniz et al., 2020), due to its livelihood traits (like the returns to capital, the market exposure or the national availability of private or public schemes for protecting the self-employed) and the possibility to be in or switch across activities (i.e. in terms of time availability to assist seasonal agricultural production *versus* those in other forms of entrepreneurship). Self-employment may also be for many an occupation of last resort with very low levels of productivity, risking to be associated with low levels of welfare by construction. As self-employment activities are relatively more common in rural than urban contexts for our sample (also associated to subsistence agriculture in some countries), we choose to focus exclusively on urban areas. This definition of self-employed activities as informal is also in line with most recent comparative studies on the prevalence of informality (Ohnsorge and Yu, 2021).

Lastly, one could wander whether our household-level analysis could also provide further granularity about informality portfolios across different labour market sectors. Although the LSMS data make it possible to identify individual sector of participation, we refrain from proposing this additional analysis in the specific of our estimation strategy, for mainly two reasons. First, choosing to measure discretised sector informality bins would require a discretionary choice of both labour market as well as household-level dynamics imposed to the data (i.e. we would need to choose whose sector of occupation should be attributed first among pri-

mary vs secondary job, or across members).[3] Further, due to the nature of the data at our disposal, we risk due to small country sample sizes to have a heavily spurious specification that we cannot triangulate with other data (due, for instance, to the results being spuriously driven by portfolios' vectors for specific areas that are full zeros).

Table A1 presents an overview of the definitions used in each country. A wage work activity is considered formal with respect to its social insurance status if the employer contributes towards a pension scheme and/or health insurance. All formality definitions can only apply to those of legal working age, which is 16 in all countries. While it is likely that also younger household members could work, they cannot access formal opportunities so that we only consider those aged 16 years and older in our analysis.

A.2 Descriptive statistics

This appendix presents summary statistics of the main variables, especially informality measured at the household level. Table A2 provides the summary statistics of the two new informality measures and of a simple indicator whether the head of the household is informally employed for each country. The average share of

---

[3]As additional example, imagine creating 3-sector informality portfolios in a country sample, for which a total of 9 vectors should be jointly defined as a household-level decision in each period. Defining them contemporaneously as endogenously determined could conflict with classic seasonality patterns visible in urban settings (like construction work). We believe this type of exercise could be the object of further research, which would require the delineating of a specific national labour market context, rather than risking to over-impose arbitrary restrictions to the endogenous employment realisation that we see happening in the data as part of household diversification strategies.

Table A1: Formality definition by country

Country	Social insurance access
Ethiopia	if employer is government, state-owned firm, NGO/charity or political party, employment comes with health insurance and pension
Malawi	if employer provides health insurance and/or pension
Niger	if employer provides health insurance and/or pension
Nigeria	if employer provides health insurance (and/or pension)⋆
Tanzania	if employer provides health insurance and/or offers maternity or paternity leave

NOTE: This table gives the concept combinations which are used to define formal employment. As much as possible we sought to ensure that the concepts were comparable across countries and institutional arrangements.

⋆ Pension cover is only available for cross-sectional analysis. In panel analysis definition relies only on access to health insurance.

informal income varies between 80% in Ethiopia and 88% in Tanzania. Notably, the informal FTE shares show almost the same means and standard deviations as income shares in all countries. Lastly, also the proportion of households with an informally employed household head is comparable to the average informal income and FTE shares. It seems plausible that the relative importance of formal income is distributed similarly across the three different measures.

Table A3 shows for each country the distribution of households across the informal income and FTE share, respectively. In Niger, Nigeria, and Tanzania, 80% or more of all households have a fully informal income. In Ethiopia and Malawi, this proportion is between 70% and 80%. The distribution of informal FTE shares looks almost identical across countries with the exception of Ethiopia

Table A2: Summary statistics of informality measures by country

Country	Variable	Mean	Std. Dev.	N
Ethiopia	Share of informal income	0.8	0.36	1,548
2015/16	Share of informal FTEs	0.78	0.39	1,548
	Informal household head	0.79	0.41	1,548
Malawi	Share of informal income	0.84	0.33	2,154
2016/17	Share of informal FTEs	0.83	0.35	2,154
	Informal household head	0.82	0.38	2,154
Niger	Share of informal income	0.88	0.3	1,263
2014	Share of informal FTEs	0.86	0.32	1,263
	Informal household head	0.86	0.34	1,263
Nigeria	Share of informal income	0.87	0.29	1,156
2014/15	Share of informal FTEs	0.87	0.3	1,156
	Informal household head	0.85	0.35	1,156
Tanzania	Share of informal income	0.88	0.29	1,330
2016/17	Share of informal FTEs	0.88	0.29	1,330
	Informal household head	0.85	0.36	1,330

NOTE: Table reports raw shares of informality measures per country by different definitions of informality.

and Malawi. Here, the informal FTE shares are higher in fully formal activities or when less than half of work effort goes into informal work. The percentage of households with a fully formal income is only around 4% in all countries, a bit higher in Ethiopia. In terms of mixed portfolios, the category of households with between half and 99% of their income earned or FTEs worked in informal work is always around 5%. In contrast, households with a less-than-half informal income or FTE share are relatively more common in our sample. In Ethiopia and Malawi, this share is as high as 12 or 13% if measured as income share.

Table A3: Proportion of urban households in each informality bin, by country

Country	0% informal		1-50% informal		51-99% informal		100% informal	
	Income	FTE	Income	FTE	Income	FTE	Income	FTE
Ethiopia	0.08	0.14	0.13	0.07	0.06	0.05	0.73	0.73
Malawi	0.04	0.10	0.12	0.08	0.05	0.04	0.78	0.78
Niger	0.04	0.08	0.08	0.07	0.05	0.04	0.83	0.82
Nigeria	0.04	0.06	0.09	0.08	0.06	0.05	0.80	0.82
Tanzania	0.04	0.05	0.08	0.07	0.05	0.05	0.83	0.83

NOTE: This table gives in each cell the proportion of urban households from the overall sample contained in each of the bins of informality used in the analysis for both measures of informality as a % of household income, or as a labor FTE share.

Table A4 presents summary statistics by these informality levels of the household and country.

Table A4: Means of household characteristics by country and level of informality share

	Per capita expenditure	Female	Age of head	Household head	Dependency size	Secondary ratio	Household schooling	No. of jobs owns land
0% Informal (= fully formal)								
Ethiopia	3.46	0.23	36.86	3.06	0.48	0.53	0.06	1.35
Malawi	8.77	0.16	39.86	3.74	0.43	0.51	0.01	1.28
Niger	4.40	0.03	44.57	6.00	1.02	0.25	0.00	1.26
Nigeria	9.04	0.14	50.16	4.32	0.74	0.61	0.00	2.05
Tanzania	9.35	0.17	37.47	3.52	0.52	0.41	0.06	1.35
1-50% Informal								
Ethiopia	3.55	0.31	41.12	3.93	0.45	0.46	0.20	2.27
Malawi	7.85	0.15	41.32	4.47	0.62	0.49	0.18	2.34
Niger	4.97	0.15	48.22	6.34	0.77	0.33	0.08	2.37
Nigeria	6.97	0.11	50.16	5.73	0.80	0.53	0.11	3.34
Tanzania	9.84	0.20	40.59	4.04	0.48	0.31	0.24	2.68
51-99% Informal								
Ethiopia	3.24	0.32	46.97	5.43	0.51	0.27	0.18	3.61
Malawi	6.61	0.09	43.12	5.15	0.64	0.42	0.25	3.42
Niger	4.27	0.14	51.62	6.49	0.84	0.23	0.04	2.71
Nigeria	5.49	0.11	51.69	6.21	0.83	0.47	0.08	3.24
Tanzania	7.82	0.24	41.10	5.28	0.77	0.25	0.30	3.49
100% Informal								
Ethiopia	3.03	0.43	41.35	3.89	0.66	0.15	0.17	1.89
Malawi	4.24	0.24	40.11	4.35	0.88	0.18	0.27	2.10
Niger	2.98	0.21	48.70	6.35	1.16	0.09	0.24	2.62
Nigeria	4.97	0.24	52.97	5.18	0.94	0.34	0.17	2.59
Tanzania	5.48	0.34	41.41	4.02	0.77	0.12	0.23	2.50

Notes: This table presents means for each variable by country and depth of informality of the household defined as income share. Female head and age of head refer to the gender and age of the household head, respectively. Dependency ratio is the number of children (0-14 years) and elderly (65 years and older) in the household relative to the number working-age household members. Secondary schooling is measured as the share of household members 18 years and older who completed secondary education. Land ownership is a binary variable equal to 1 if the household reports a land size greater than 0 and that it owns, not rents nor that is communal land. Source: authors' compilation based on LSMS-ISA data.

B Poverty Channel

In this appendix we provide extended results that investigate the impact of the depth of informality on poverty. Poverty is defined with the international poverty line of international US$ 1.90 consumption per person per day (Ferreira et al., 2016). In the dynamic analysis for Nigeria we use the national poverty line for comparability reasons over time. Looking at poverty as an outcome allows us to focus on a relevant threshold within the consumption distribution. Our main results are related to the mean. So we compare those to the relationship between consumption and the depth of informality at the poverty line. The results will offer insights into why a reduction in informality might not always imply poverty reduction (OECD and ILO, 2019).

B.1 Cross Sectional Results

Similar to the analysis we present in Figures 1 and 2 in the main text, we adopt the same approach as in Equation (1) in the main text. For this analysis our dependent variable is a binary indicator for a household being below the poverty line.

Figures B1 and B2 plot the coefficients of the regressions of poverty status of households on their share of income earned from informal sources (Figure B1) or the share of FTEs worked in informal activities (Figure B2). The first coefficient in each figure is from the dummy of the informal status of the household head. In Malawi, Niger, Nigeria and Tanzania households with an informally employed household head are significantly more likely to be poor than households

9

Figure B1: Coefficients of informality measures from regression of probability to be poor. Share of income earned from informal sources. Base category is fully formal income.
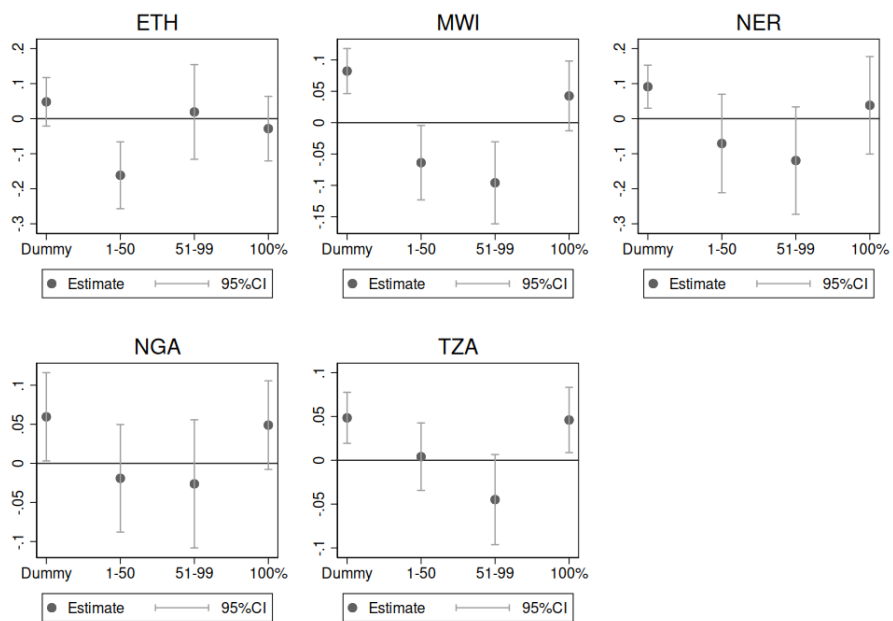
NOTES: The graphs plot coefficients and confidence intervals from two different regressions for each country. The first coefficient is that of the dummy indicating an informal household head from one regression. The other three coefficients are those of the informality bins from the regression as specified in Eq.(1) in the main text. The base category are households with no informal income source.

Source: authors' compilation based on cross-sectional data.

Figure B2: Coefficients of informality measures from regression of probability to be poor. Share of FTEs worked in informal activities. Base category is fully formal activities.
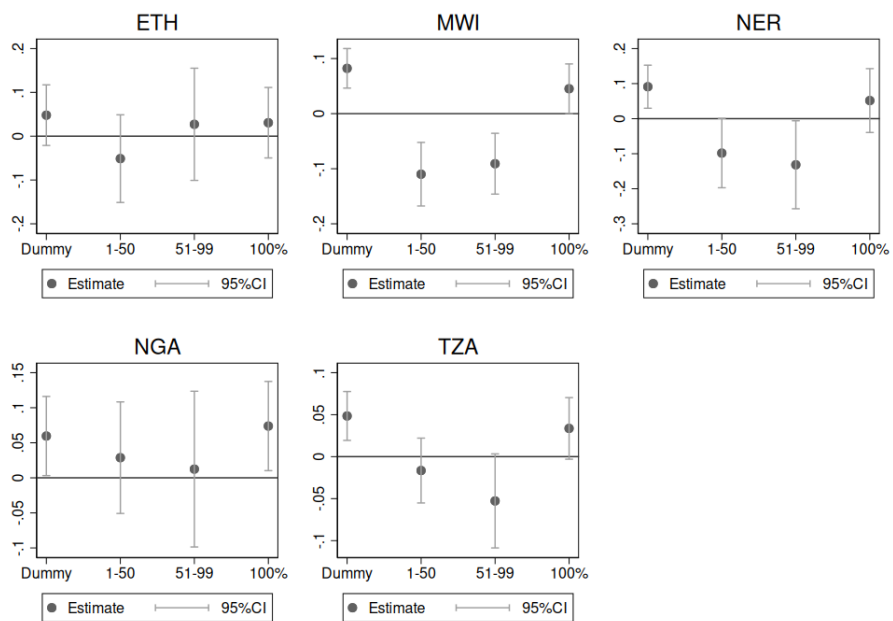
with formally employed heads. In Ethiopia, this relationship is just insignificant.

Looking at the depth of informality, the main finding is that poverty is not in all countries statistically significantly associated with informality and it appears that households with both formal and informal income sources (second and third coefficient) achieve better welfare outcomes than fully informal households (Niger, Nigeria and Tanzania) and sometimes even better than fully formal ones (Ethiopia and Malawi). These correlations might indicate that a greater diversification of income sources is associated with a lower likelihood of being poor. Overall, we observe that the precision of these middle bins is relatively low. This could partially be due to lack of power. Only in Tanzania, households are also significantly more likely to be poor if their income origins to 100 % from informal sources (fourth coefficient).

For FTE shares, this result is also significant for Nigeria. In Figures B1 and B2, only Malawi shows a clear and significant relationship between poverty status and different steps of informality in terms of formal income shares. Households with some share of formal income are all significantly less likely to be poor than households with fully formal or fully informal income. In Ethiopia, this applies to those households with informal income shares between 0 and 50 %, and in Niger to households with more than half of their labour efforts in informal activities. Even though insignificant, similar patterns are observed in Nigeria and Tanzania for income shares, and in Tanzania for labour shares.

B.2 Dynamic Results

In this subsection we replicate our dynamic estimates for our alternate measure of welfare, the probability that a household is below the poverty line.

Table B1 replicates the analysis in Table 1. It is notable that the results are broadly in line with what was previously reported.

Moving on, here we provide estimates for the switching strategy replicating the 'switchers' analysis.

$$Y_{i,t} = \alpha_i + \beta \text{Post}_t + \delta \text{Switcher} \times \text{Post}_{i,t} + X' \gamma_{i,t}$$

where $Y_{i,t}$, is a binary indicator for being poor for the $i^t h$ household in time t, this is predicted by a time invariant indicator whether a household changes its income portfolio, *Switcher*, between t and t-1 and the dummy *Post* indicating the second time period in which a household is observed and the interaction of both indicators. δ is the coefficient of interest showing the average treatment effect on the treated (ATT) of the household's change in status on its welfare. We also control for a vector of household time varying characteristics, X (*viz.* household size and share with secondary schooling). Further controls used in the cross-sectional case such as land ownership and female headship were considered, but ultimately deemed to be potential sources of collider bias in the dynamic set up, as the parameters would be identified only for those switching status. Those 'switchers' amongst the land and household head composition, would therefore likely correlate with the switch in informal household status, expected to bias the

Table B1: Poverty, by status and transition, Nigeria 2010-2015

Portfolio v control	(1) Before	(2) After	(3) Raw Δ	(4) Cond Δ	(5) DiD
IIvFF	0.184***	0.181***	0.145***	0.133***	0.015
	(0.009)	(0.009)	(0.018)	(0.020)	(0.036)
N	1,852	1,852	3,836	3,836	3,836
FIvFF	0.100**	0.075*	0.050	0.041	0.003
	(0.048)	(0.042)	(0.046)	(0.042)	(0.050)
N	40.000	40.000	212	212	212
IFvII	0.110***	0.121***	0.078**	0.079**	0.034
	(0.033)	(0.034)	(0.033)	(0.031)	(0.047)
N	91.000	91.000	314	314	314
FFvII	0.045*	0.030	-0.145***	-0.133***	-0.015
	(0.026)	(0.021)	(0.018)	(0.020)	(0.036)
N	66.000	66.000	3,836	3,836	3,836
MIvMM	0.148***	0.164***	0.065	0.071*	0.025
	(0.046)	(0.048)	(0.046)	(0.042)	(0.065)
N	61.000	61.000	276	276	276
IMvII	0.116***	0.101***	-0.074***	-0.086***	-0.019
	(0.027)	(0.026)	(0.023)	(0.021)	(0.033)
N	138	138	3,980	3,980	3,980
MMvII	0.091***	0.091***	-0.092***	-0.102***	0.002
	(0.033)	(0.033)	(0.030)	(0.028)	(0.034)
N	77.000	77.000	3,858	3,858	3,858
MFvMM	0.040	0.000***	-0.071**	-0.070*	-0.043
	(0.040)		(0.035)	(0.038)	(0.052)
N	25.000	25.000	204	204	204
FMvFF	0.050	0.050	0.012	0.004	0.028
	(0.050)	(0.050)	(0.052)	(0.047)	(0.052)
N	20.000	20.000	172	172	172
JoinersI (MI&FIvFF&MM)	0.129***	0.129***	0.062*	0.057*	0.007
	(0.033)	(0.033)	(0.033)	(0.030)	(0.042)
N	101	101	488	488	488
LeaversI (MF&IFvII&MM)	0.095***	0.095***	-0.084***	-0.078***	0.013
	(0.027)	(0.027)	(0.024)	(0.022)	(0.032)
N	116	116	4,090	4,090	4,090
JoinersM (FM&IMvFF&MM)	0.108***	0.095***	0.035	0.025	-0.015
	(0.025)	(0.023)	(0.026)	(0.025)	(0.036)
N	158	158	602	602	602
LeaversM (MF&MIvFF&MM)	0.116***	0.116***	0.050	0.044	0.015
	(0.035)	(0.035)	(0.032)	(0.030)	(0.048)
N	86.000	86.000	458	458	458

NOTES: This table gives means and estimates of the effect of transitioning as household to/from informality. Groups are defined by their state across the transition gap. So for someone who is always Formal (FF), always Informal (II), always Mix (MM), and permutations, thereof. Cols (1)&(2) provide the raw means for each portfolio group, in the respective time. Cols (3)&(4) provide the gap relative to their respective 'control groups' estimated as a simple intercept shift using OLS. For Cols (4)&(5) the estimates are conditional on household size, Share of school leavers, and 'real-time' fixed effects. Cols (5) are estimated using a household fixed effects model. The data are stacked on a dimensionless 'transition time' that is the gap in time between period 0 & 1, but naturally this duplicates observations in 'real time' in wave 2 in 2012. Errors clustered at household level.

SOURCE: authors' compilation based on Nigeria 2010-2015 data.

estimates of the effect of interest.

Table B2 reports the results of this analysis. We find that the point estimates signs across the different transitions conform to our priors, but notably no estimates are statistically significant. This implies that once selection is accounted for as our strategy does, transitioning does not make any difference to welfare. This suggests that household portfolio changes are likely done to hedge the household to changing circumstances.

Finally, we further replicate the randomisation inference exercise for our dynamic estimates. We report the results of this below in Figure B3. It is remarkable in this case that the randomisation inference suggests that all of our point estimates fall within the null bound-set.

Figure B3: Difference-in-difference estimates of Poverty for switchers, Nigeria
2010-2015 with randomised inference
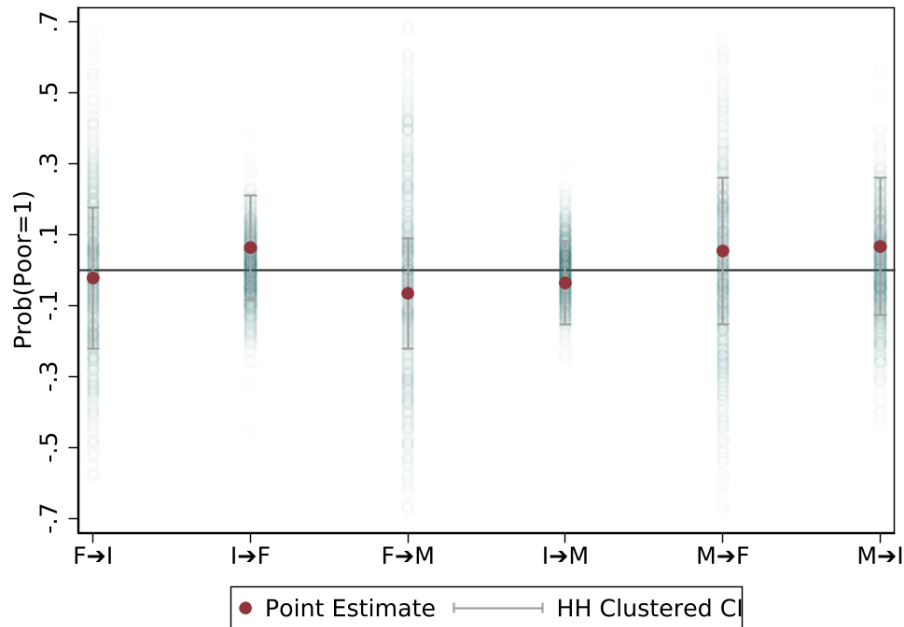
NOTES: The graphs plot coefficients and confidence intervals from Eq.(2) in red. Each point
represents a different hypothesis regression outcome. Where the initials F, I and M represent
Fully Formal, Fully Informal and Mixed portfolios respectively and the → represents the
direction of the switch. So for example, the first estimate I→F is the switch between fully
informal to fully formal. Each of the hollow circles in blue represents a point estimate from the
randomisation inference exercise. The mass of blue points will be clustered along what can be
considered to be the null bounds of the regression, so if a red point estimate is found to lie in this
area, it can be inferred that the point estimate is a true null effect. Conversely, if the red point
estimate lies outside the mass of blue estimates, the point estimate can be said to be different
from zero.

Source: authors' compilation.

Table B2: Difference-in-difference estimates of Poverty for switchers, Nigeria

LHS	Prob(Poor=1)
Fully formal switched to fully informal (H1)	-0.023
	(0.098)
N	76.000
Fully informal switched to fully formal (H2)	0.063
	(0.074)
N	184
Fully formal switched to mix (H3)	-0.066
	(0.074)
N	40.000
Fully informal switched to mix (H4)	-0.036
	(0.059)
N	274
Mix to fully formal (H5)	0.054
	(0.100)
N	48.000
Mix to fully informal (H6)	0.066
	(0.097)
N	128

NOTES: Each cell in this table represents the estimate of the δ parameter from Eq.(2). The functional form presented controls for year fixed effects, household size, dependency ratio, and the share of household members with secondary schooling. Standard Errors are clustered at household level. *, **, *** indicate statistical significance at 1, 5, and 10 % respectively.

SOURCE: authors' compilation based on Nigeria 2010-2015 data.

C   Wealth channel

In this section we investigate if the hedging effect of having a mixed portfolio is mediated through accumulation of asset wealth. We construct an index from communication equipment, transport means and housing characteristics and follow the approach of Smits and Steendijk (2015) for an international wealth index. This index allows comparability across different countries and years as applied in McKenzie (2009). We apply polychoric principal component analysis (Kolenikov and Angeles, 2009) to construct the index and rescale it to range from 0 to 100. As in the main analysis, we estimate:

$$W_i = \alpha + \sigma_{k=1}^{K} \beta_k \text{Inf}_i + X_i' \gamma + Z_l' \theta + \varepsilon_i \tag{1}$$

W_i is the wealth index. Where β_k refers to our informality measure that has been discretised into three bins where in the distribution of our continuous informality measure. The lowest bin represents 0.1 to 50 % of informality, followed by 51 to 99 % of informality, the final bin captures all observations with 100, meaning households whose income is earned fully from or all FTEs worked are in the informal sector. These dummies are all relative to the base category of 0 informal work or income shares, that is the households whom are 100% formal. The logic behind this functional form is to avoid rigidly assuming that there is a linear dose-response function with respect to a household's informality mix. We use only four bins due to the limited sample size. X_i is a vector of household-level controls and includes the sex and age of the household head, the share of

household members with secondary education, the household size, dependency ratio, a dummy whether the household owns any land, and sum of jobs from all working household members. The vector Z_l is a vector of geographic controls that include dummies for the administrative areas of the highest level to capture structural differences between regions.

We find that the wealth results are broadly in-line with the estimates on welfare outcomes considered in the main text. Namely, those households who hedge their portfolio through a mixture of informal, and formal sources, see null or weakly higher wealth than their fully formal or informal counterparts.

Figure C1: Coefficients of informality measures from regression of wealth index. Share of income earned from informal sources. Base category is fully formal income.
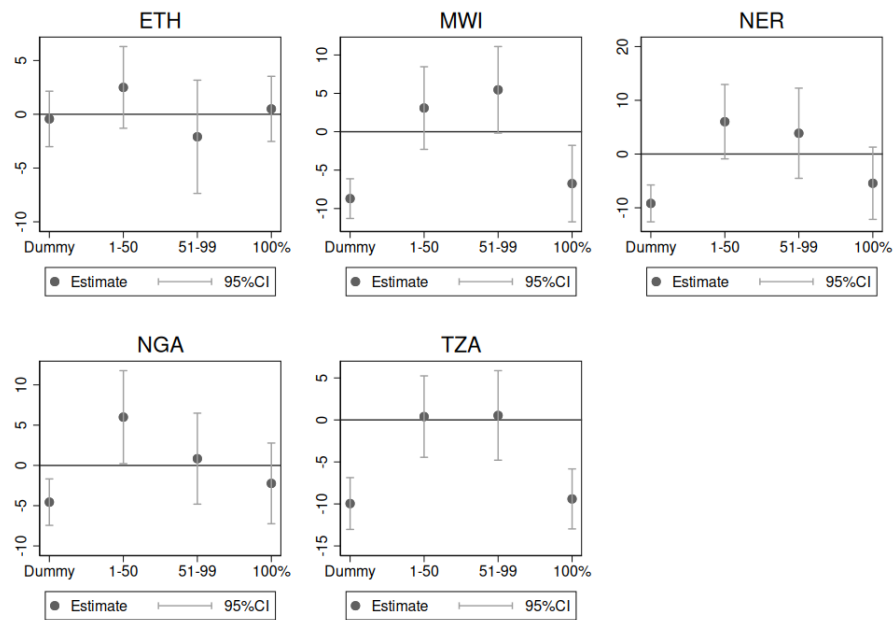
Notes: The graphs plot coefficients and confidence intervals from two different regressions for each country. The first coefficient is that of the dummy indicating an informal household head from one regression. The other three coefficients are those of the informality bins from the regression as specified in Eq.(1) in the main text. The base category are households with no informal income source.

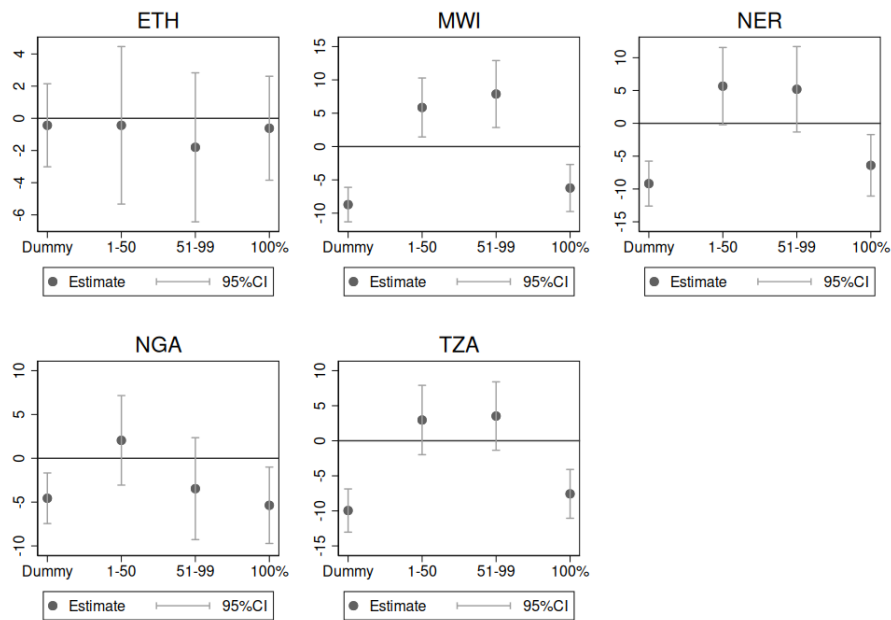Source: authors' compilation based on cross-sectional data.

Figure C2: Coefficients of informality measures from regression of wealth index. Share of FTEs worked in an informal activity. Base category is fully formal work allocation.

Notes: The graphs plot coefficients and confidence intervals from two different regressions for each country. The first coefficient is that of the dummy indicating an informal household head from one regression. The other three coefficients are those of the informality bins from the regression as specified in Eq.(1) in the main text. The base category are households with no informal full-time equivalents (FTEs).

Source: authors' compilation based on cross-sectional data.

D   Dynamic depth of informality: Is it a jobs story?

In this appendix we try to investigate whether the effects of the dynamic analysis are due to a mechanical effect coming from comparing households with more/fewer jobs. We replicate the estimates from Table 3 (main text) for the subsamples, where we hold the number of jobs constant across groups (1-2 jobs, 3-4 jobs and 5+ jobs) and maintain our strategy of using households who will be making the same switch the following period as the controls.

The estimates for number of jobs are remarkable as for the group of 1-2 jobs we find qualitatively similar effects as in the main results presented. The point estimates for log total expenditure still imply a substantial gain for those households moving from informality to full formality. However, we now also find statistically significant estimates which imply that this move is also associated with an increase in the probability of being poor. This point estimate is found to be outside the randomisation inference null interval reported in Figure 3 [-0.155 , 0.162] suggesting that this is not likely due to random chance due to small cell sizes. It should be noted that a similar (in sign) inference can be drawn from the main plot, though it is likely that the lack of significance is coming from the heterogeneity in experience for those households with more than two jobs.

Table D1: Difference-in-difference estimates for switchers by same number of jobs, Nigeria

LHS	1-2 Job in HH P(Poor=1)	3-4 Jobs in HH P(Poor=1)	5+ Jobs in HH P(Poor=1)	1-2 Jobs in HH ln(TotExp)	3-4 Jobs in HH ln(TotExp)	5+ Jo... ln(T...
Fully formal switched to fully informal (H1)	-0.072			-0.151		
	(0.090)			(0.366)		
N	60	8	8	60	8	
Fully informal switched to fully formal (H2)	0.181**			0.333**		
	(0.072)			(0.165)		
N	155	15	4	155	15	
Fully formal switched to mix (H3)	-0.242			0.521		
	(0.222)			(0.339)		
N	35	4	4	35	4	
Fully informal switched to mix (H4)	-0.123	-0.350		-0.264	0.031	
	(0.078)	(0.357)		(0.244)	(0.647)	
N	162	89	16	162	89	
Mix to fully formal (H5)				0.658		
				(0.504)		
N	21	24	3	21	24	
Mix to fully informal (H6)	0.198			-0.050		
	(0.196)			(0.232)		
N	75	41	10	75	41	

NOTES: Each cell in this table represents the estimate of the δ parameter from Eq.(2) where each column is a subsample where we hold the number of jobs within a hou... constrained. Note that due to small cell sizes we are not always able to estimate an effect. The functional form presented controls for year fixed effects, household size, ratio, and the share of household members with secondary schooling. Standard Errors are clustered at household level. *, **, *** indicate statistical significance at 1, respectively.

SOURCE: authors' compilation based on Nigeria 2010-2015 data.

E Full cross section results

This section presents the full estimates tables underlying Figures B1,B2 and Figures 1 and 2 (main text) for each country.

E.1 Ethiopia

Table E1: Welfare and informality status of household head, Ethiopia

	ln(ExpPC)	P(Poor=1)
Informal HH Head	-0.205	0.048
	(0.374)	(0.035)
Household size	0.167**	0.037***
	(0.069)	(0.008)
Female Head	0.988***	0.086***
	(0.325)	(0.029)
Age of household head	0.020**	0.000
	(0.009)	(0.001)
Owns Land	-0.978**	0.091**
	(0.391)	(0.040)
Share of hh members with secondary schooling	-0.284	-0.279***
	(0.671)	(0.044)
Dependency ratio	0.009	0.069***
	(0.198)	(0.025)
Total number of jobs, hh	0.073	-0.004
	(0.092)	(0.010)
R^2	0.04	0.23
N	1,548	1,548
Geog FEs	yes	yes

Notes: *, **, *** indicate statistical significance at 1, 5, and 10 % respectively.
Source: authors' compilation based on data.

Table E2: Welfare and depth of informality of household, Ethiopia

	ln(ExpPC) Sh Inf Inc	ln(ExpPC) Sh FTE Inf	P(Poor=1) Sh Inf Inc	P(Poor=1) Sh FTE Inf
1-50% informal	0.342	-0.636	-0.162***	-0.051
	(0.774)	(1.102)	(0.049)	(0.051)
51-99% informal	0.507	-0.422	0.019	0.027
	(0.559)	(0.575)	(0.069)	(0.065)
100% informal	-0.003	-0.533	-0.029	0.031
	(0.584)	(0.474)	(0.047)	(0.041)
Household size	0.154**	0.167**	0.038***	0.038***
	(0.068)	(0.071)	(0.008)	(0.008)
Female Head	0.971***	0.988***	0.089***	0.087***
	(0.329)	(0.323)	(0.029)	(0.029)
Age of household head	0.019**	0.019**	0.000	0.000
	(0.009)	(0.009)	(0.001)	(0.001)
Owns Land	-0.960**	-0.942**	0.100**	0.094**
	(0.394)	(0.400)	(0.040)	(0.040)
Share of hh members with secondary schooling	-0.319	-0.376	-0.270***	-0.277***
	(0.700)	(0.699)	(0.044)	(0.045)
Dependency ratio	0.027	0.006	0.067***	0.068***
	(0.199)	(0.199)	(0.025)	(0.025)
Total number of jobs, hh	0.058	0.089	-0.005	-0.004
	(0.095)	(0.098)	(0.010)	(0.010)
R^2	0.04	0.04	0.24	0.23
N	1,548	1,548	1,548	1,548
Geog FEs	yes	yes	yes	yes

Notes: *, **, *** indicate statistical significance at 1, 5, and 10 % respectively.

Source: authors' compilation based on data.

E.2 Malawi

Table E3: Welfare and informality status of household head, Malawi

	ln(ExpPC)	P(Poor=1)
Informal HH Head	-0.304***	0.082***
	(0.043)	(0.018)
Household size	-0.130***	0.043***
	(0.010)	(0.006)
Female Head	-0.023	0.037
	(0.039)	(0.023)
Age of household head	0.000	0.001
	(0.001)	(0.001)
Owns Land	-0.177***	0.084***
	(0.045)	(0.029)
Share of hh members with secondary schooling	0.932***	-0.133***
	(0.061)	(0.027)
Dependency ratio	-0.061***	0.060***
	(0.022)	(0.017)
Total number of jobs, hh	0.038***	-0.015**
	(0.012)	(0.007)
R^2	0.49	0.25
N	2,154	2,154
Geog FEs	yes	yes

Notes: *, **, *** indicate statistical significance at 1, 5, and 10 % respectively.
Source: authors' compilation based on data.

Table E4: Welfare and depth of informality of household, Malawi

	ln(ExpPC) Sh Inf Inc	ln(ExpPC) Sh FTE Inf	P(Poor=1) Sh Inf Inc	P(Poor=1) Sh FTE Inf
1-50% informal	0.034	0.146**	-0.064**	-0.110***
	(0.085)	(0.074)	(0.030)	(0.029)
51-99% informal	0.052	0.062	-0.096***	-0.091***
	(0.096)	(0.083)	(0.033)	(0.028)
100% informal	-0.295***	-0.263***	0.043	0.045**
	(0.078)	(0.054)	(0.028)	(0.023)
Household size	-0.134***	-0.135***	0.045***	0.045***
	(0.010)	(0.010)	(0.006)	(0.006)
Female Head	-0.017	-0.019	0.034	0.036
	(0.039)	(0.039)	(0.023)	(0.023)
Age of household head	0.000	0.000	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Owns Land	-0.170***	-0.169***	0.081***	0.080***
	(0.045)	(0.045)	(0.029)	(0.029)
Share of hh members with secondary schooling	0.888***	0.891***	-0.115***	-0.120***
	(0.061)	(0.061)	(0.027)	(0.027)
Dependency ratio	-0.059***	-0.058***	0.060***	0.058***
	(0.022)	(0.022)	(0.017)	(0.017)
Total number of jobs, hh	0.036***	0.032***	-0.013*	-0.011
	(0.012)	(0.012)	(0.008)	(0.008)
R^2	0.49	0.49	0.26	0.26
N	2,154	2,154	2,154	2,154
Geog FEs	yes	yes	yes	yes

Notes: *, **, *** indicate statistical significance at 1, 5, and 10 % respectively.

Source: authors' compilation based on data.

E.3 Niger

Table E5: Welfare and informality status of household head, Niger

	ln(ExpPC)	P(Poor=1)
Informal HH Head	-0.213***	0.091***
	(0.049)	(0.031)
Household size	-0.041***	0.019***
	(0.006)	(0.007)
Female Head	0.014	0.035
	(0.049)	(0.045)
Age of household head	-0.000	-0.000
	(0.002)	(0.001)
Owns Land	-0.018	0.036
	(0.056)	(0.051)
Share of hh members with secondary schooling	0.767***	-0.290***
	(0.086)	(0.064)
Dependency ratio	-0.142***	0.085***
	(0.019)	(0.019)
Total number of jobs, hh	-0.016	0.024**
	(0.011)	(0.011)
R^2	0.39	0.21
N	1,263	1,263
Geog FEs	yes	yes

Notes: *, **, *** indicate statistical significance at 1, 5, and 10 % respectively.

Source: authors' compilation based on data.

Table E6: Welfare and depth of informality of household, Niger

	ln(ExpPC) Sh Inf Inc	ln(ExpPC) Sh FTE Inf	P(Poor=1) Sh Inf Inc	P(Poor=1) Sh FTE Inf
1-50% informal	0.071	0.096	-0.071	-0.098*
	(0.089)	(0.081)	(0.072)	(0.050)
51-99% informal	0.026	0.036	-0.120	-0.132**
	(0.116)	(0.098)	(0.078)	(0.064)
100% informal	-0.163**	-0.163***	0.038	0.052
	(0.083)	(0.063)	(0.071)	(0.046)
Household size	-0.041***	-0.041***	0.019***	0.019***
	(0.006)	(0.006)	(0.007)	(0.007)
Female Head	0.006	0.007	0.038	0.037
	(0.049)	(0.049)	(0.045)	(0.045)
Age of household head	-0.000	-0.000	0.000	0.000
	(0.002)	(0.002)	(0.001)	(0.001)
Owns Land	-0.017	-0.014	0.031	0.029
	(0.056)	(0.056)	(0.051)	(0.051)
Share of hh members with secondary schooling	0.717***	0.723***	-0.259***	-0.264***
	(0.089)	(0.089)	(0.066)	(0.066)
Dependency ratio	-0.142***	-0.141***	0.084***	0.082***
	(0.019)	(0.019)	(0.019)	(0.020)
Total number of jobs, hh	-0.017	-0.017	0.025**	0.027**
	(0.011)	(0.011)	(0.011)	(0.011)
R^2	0.39	0.39	0.21	0.21
N	1,263	1,263	1,263	1,263
Geog FEs	yes	yes	yes	yes

Notes: *, **, *** indicate statistical significance at 1, 5, and 10 % respectively.

Source: authors' compilation based on data.

E.4 Nigeria

Table E7: Welfare and informality status of household head, Nigeria

	ln(ExpPC)	P(Poor=1)
Informal HH Head	-0.268***	0.060**
	(0.058)	(0.029)
Household size	-0.107***	0.032***
	(0.011)	(0.007)
Female Head	-0.031	0.032
	(0.051)	(0.028)
Age of household head	-0.000	0.000
	(0.001)	(0.001)
Owns Land	-0.008	-0.023
	(0.057)	(0.036)
Share of hh members with secondary schooling	0.459***	-0.044
	(0.082)	(0.035)
Dependency ratio	-0.042	0.021
	(0.028)	(0.017)
Total number of jobs, hh	-0.006	-0.002
	(0.012)	(0.007)
R^2	0.41	0.22
N	1,156	1,156
Geog FEs	yes	yes

Notes: *, **, *** indicate statistical significance at 1, 5, and 10 % respectively.

Source: authors' compilation based on data.

Table E8: Welfare and depth of informality of household, Nigeria

	ln(ExpPC) Sh Inf Inc	ln(ExpPC) Sh FTE Inf	P(Poor=1) Sh Inf Inc	P(Poor=1) Sh FTE Inf
1-50% informal	-0.055	0.048	-0.019	0.029
	(0.103)	(0.097)	(0.035)	(0.041)
51-99% informal	-0.087	-0.090	-0.026	0.012
	(0.104)	(0.103)	(0.042)	(0.057)
100% informal	-0.301***	-0.269***	0.049*	0.074**
	(0.091)	(0.075)	(0.029)	(0.032)
Household size	-0.108***	-0.107***	0.033***	0.032***
	(0.011)	(0.011)	(0.007)	(0.007)
Female Head	-0.027	-0.029	0.031	0.031
	(0.051)	(0.051)	(0.028)	(0.028)
Age of household head	-0.001	-0.001	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)
Owns Land	-0.012	-0.005	-0.024	-0.025
	(0.057)	(0.057)	(0.036)	(0.036)
Share of hh members with secondary schooling	0.432***	0.439***	-0.036	-0.040
	(0.084)	(0.083)	(0.036)	(0.035)
Dependency ratio	-0.040	-0.040	0.020	0.020
	(0.028)	(0.027)	(0.017)	(0.017)
Total number of jobs, hh	-0.003	-0.007	-0.003	-0.003
	(0.012)	(0.012)	(0.007)	(0.007)
R^2	0.41	0.42	0.22	0.22
N	1,156	1,156	1,156	1,156
Geog FEs	yes	yes	yes	yes

Notes: *, **, *** indicate statistical significance at 1, 5, and 10 % respectively.

Source: authors' compilation based on data.

E.5 Tanzania

Table E9: Welfare and informality status of household head, Tanzania

	ln(ExpPC)	P(Poor=1)
Informal HH Head	-0.293***	0.048***
	(0.101)	(0.015)
Household size	-0.061***	0.010
	(0.015)	(0.007)
Female Head	-0.011	0.066***
	(0.062)	(0.020)
Age of household head	-0.002	0.002*
	(0.002)	(0.001)
Owns Land	0.103	0.016
	(0.080)	(0.026)
Share of hh members with secondary schooling	0.492***	0.007
	(0.098)	(0.022)
Dependency ratio	-0.241***	0.078***
	(0.034)	(0.022)
Total number of jobs, hh	0.010	-0.002
	(0.013)	(0.008)
R^2	0.23	0.21
N	1,330	1,330
Geog FEs	yes	yes

Notes: *, **, *** indicate statistical significance at 1, 5, and 10 % respectively.

Source: authors' compilation based on data.

Table E10: Welfare and depth of informality of household, Tanzania

	ln(ExpPC) Sh Inf Inc	ln(ExpPC) Sh FTE Inf	P(Poor=1) Sh Inf Inc	P(Poor=1) Sh FTE Inf
1-50% informal	0.321	0.433**	0.004	-0.017
	(0.252)	(0.201)	(0.020)	(0.020)
51-99% informal	0.284	0.232	-0.045*	-0.053*
	(0.247)	(0.199)	(0.026)	(0.029)
100% informal	-0.089	-0.077	0.046**	0.034*
	(0.275)	(0.226)	(0.019)	(0.019)
Household size	-0.064***	-0.063***	0.011	0.011
	(0.015)	(0.015)	(0.007)	(0.007)
Female Head	-0.011	0.003	0.067***	0.067***
	(0.062)	(0.063)	(0.020)	(0.020)
Age of household head	-0.002	-0.002	0.002*	0.002*
	(0.002)	(0.001)	(0.001)	(0.001)
Owns Land	0.098	0.081	0.016	0.016
	(0.079)	(0.079)	(0.026)	(0.027)
Share of hh members with secondary schooling	0.477***	0.490***	0.012	0.011
	(0.111)	(0.111)	(0.022)	(0.022)
Dependency ratio	-0.241***	-0.242***	0.078***	0.078***
	(0.034)	(0.034)	(0.022)	(0.022)
Total number of jobs, hh	0.007	0.006	-0.002	-0.002
	(0.013)	(0.013)	(0.008)	(0.008)
R^2	0.24	0.24	0.21	0.21
N	1,330	1,330	1,330	1,330
Geog FEs	yes	yes	yes	yes

Notes: *, **, *** indicate statistical significance at 1, 5, and 10 % respectively.

Source: authors' compilation based on data.

References

Alaniz, E., T. H. Gindling, C. Mata, and D. Rojas (2020, August). Transforming informal work and livelihoods in Costa Rica and Nicaragua. *UNU-WIDER working paper*.

Dávila, E. (1994). Regulaciones Laborales y Mercado de Trabajo en México. In G. Márquez (Ed.), *Regulación del Mercado de Trabajo en América Latina*. Centro Internacional para el Desarrollo Económico.

Ferreira, F. H. G., S. Chen, A. Dabalen, Y. Dikhanov, N. Hamadeh, D. Jolliffe, A. Narayan, E. B. Prydz, A. Revenga, P. Sangraula, U. Serajuddin, and N. Yoshida (2016). A global count of the extreme poor in 2012: data issues, methodology and initial results. *The Journal of Economic Inequality 14*(2), 141–172.

Heckman, J. and C. Pagés (2004). Law and Employment: Lessons form Latin America and the Caribbean. Technical report.

ICLS15 (1993). Statistics of employment in the informal sector, icls/15/iii. ilo: Geneva. www.ilo.org/public/libdoc/ilo/1992/92b09_385_engl.pdf, Fifteenth International Conference of Labour Statistics.

ICLS17 (2003). General report, report i, icls/17/2003/i. ilo: Geneva. www.ilo.org/wcmsp5/groups/public/—dgreports/—stat/documents/meetingdocument/wcms_087585.pdf, Seventeenth International Conference of Labour Statistics.

IDB (2004). Good Jobs Wanted: Labor Markets in Latin America. Technical report, Wasington, D.C.

Kolenikov, S. and G. Angeles (2009). Socioeconomic Status Measurement with Discrete Proxy Variables: Is Principle Component Analysis a Reliable Answer? *Review of Income and Wealth 55*(1), 128–165.

Maloney, W. (2004). Informality Revisited. *World Development 32*, 1159–1178.

Maloney, W. F. (1999, 05). Does Informality Imply Segmentation in Urban Labor Markets? Evidence from Sectoral Transitions in Mexico. *The World Bank Economic Review 13*(2), 275–302.

McKenzie, D. J. (2009). Measuring inequality with asset indicators. *Journal of Population Economics 18*(2), 229–260.

OECD and ILO (2019). *Tackling Vulnerability in the Informal Economy.* OECD and International Labour Organization Publishing.

OECD et al (2002). Measuring the non-observed economy - a handbook. http://www.oecd.org/sdd/na/measuringthenon-observedeconomy-ahandbook.htm, OECD, the ILO, the IMF, the Interstate Statistical Committee of the Commonwealth of Independent States.

Ohnsorge, F. and S. Yu (2021). *The Long Shadow of Informality: Challenges and Policies. Advance Edition.* Washington, DC: World Bank.

Perry, G. E., W. F. Maloney, O. S. Arias, P. Fajnzylber, A. D. Mason, and J. Saavedra-Chanduvi (2007). Informality: Exit and Exclusion. Technical report, World Bank, Washington, DC.

Smits, J. and R. Steendijk (2015). The International Wealth Index (IWI). *Social Indicators Research 122*(1), 65–85.

Ulyssea, G. (2020). Informality: Causes and Consequences for Development. *Annual Review of Economics 12*(1), 525–546.