ERF Working Papers Series

Introducing the Sudan Labor Market Panel Survey 2022

Caroline Krafft, Ragui Assaad and Ruby Cheung

ECONOMIC RESEARCH FORUM
منتدى البحوث الاقتصادية

Working Paper No. 1647
August 2023

INTRODUCING THE SUDAN LABOR MARKET PANEL SURVEY 2022[1]

Caroline Krafft,[2] Ragui Assaad,[3] and Ruby Cheung[4]

Working Paper No. 1647

August 2023

Send correspondence to:
Caroline Krafft
St. Catherine University
cgkrafft@stkate.edu

[1] The authors appreciate the collaboration of Sudan's Central Bureau of Statistics, the Economic Research Forum, and the participants who responded to the SLMPS. SLMPS 2022 was funded with support from the LSMS+ program at the World Bank and the Growth and Labor Markets | Low Income Countries (GLM | LIC) program at IZA, with underlying support from FCDO. This paper was specifically supported by GLM | LIC funding.
[2] St. Catherine University. Email: cgkrafft@stkate.edu
[3] University of Minnesota, Humphrey School of Public Affairs.
[4] St. Catherine University

First published in 2023 by
The Economic Research Forum (ERF)
21 Al-Sad Al-Aaly Street
Dokki, Giza
Egypt
www.erf.org.eg

Abstract

This paper describes the new Sudan Labor Market Panel Survey (SLMPS) 2022, the first nationally representative survey in Sudan in almost a decade. The paper details the design of the survey, including the topics covered by this multi-purpose household survey and the complexities of the sampling strategy, which over-sampled refugees and the internally displaced. The training, fieldwork, resulting sample, and weights are described. Key demographic and labor market indicators are then compared to other, older nationally representative data sources, both to assess the validity of the SLMPS data and update our understanding of Sudan's labor market. The rich, publicly available data of the SLMPS provide substantial opportunities for researchers to better understand the evolution of Sudan's labor market, economy, and society.

Keywords: Survey, public use data, labor market, sample weights, Sudan

JEL Classification: J00, C81, C83

ملخص

تصف هذه الورقة المسح التتبعي لسوق العمل السوداني الجديد (SLMPS) 2022 ، وهو أول مسح تمثيلي على المستوى الوطني في السودان منذ ما يقرب من عقد من الزمان. وتورد الورقة تفاصيل تصميم الدراسة الاستقصائية، بما في ذلك المواضيع التي تغطيها هذه الدراسة الاستقصائية المتعددة الأغراض للأسر المعيشية وتعقيدات استراتيجية أخذ العينات، التي أفرطت في أخذ عينات من اللاجئين والمشردين داخليا. يتم وصف التدريب والعمل الميداني والعينة الناتجة والأوزان. ثم تتم مقارنة المؤشرات الديموغرافية الرئيسية وسوق العمل بمصادر بيانات أخرى أقدم تمثيلاً على المستوى الوطني، لتقييم صحة بيانات المسح التتبعي لسوق العمل السوداني وتحديث فهمنا لسوق العمل في السودان. توفر البيانات الغنية والمتاحة للجمهور للمسح التتبعي لسوق العمل السوداني فرصًا كبيرة للباحثين لفهم تطور سوق العمل والاقتصاد والمجتمع في السودان بشكل أفضل.

1 Introduction

This paper introduces the 2022 wave of the Sudan Labor Market Panel Survey (SLMPS 2022), the first wave of a planned nationally-representative longitudinal survey of labor market and broader socioeconomic conditions in Sudan. The SLMPS 2022 is part of the Labor Market Panel Survey (LMPS) series conducted by the Economic Research Forum (ERF) since 1998, with four waves completed in Egypt (1998, 2006, 2012, 2018) and a fifth wave planned for 2023, two waves in Jordan (2010 and 2016) and one wave in Tunisia (2014) (Assaad, Ghazouani, Krafft, & Rolando, 2016; Krafft & Assaad, 2021; Krafft, Assaad, & Rahman, 2021; OAMDI, 2016, 2018, 2019).

It should be clear to anyone who has followed developments in Sudan that the successful implementation of SLMPS 2022 faced very long odds. Besides the obvious challenges of political strife, economic instability, a major pandemic, civil conflict, and weather shocks, the team had to overcome major technical challenges, such as the absence of a sampling frame that can be used to draw a nationally-representative sample and how to include the important internally displaced and refugee populations of Sudan. Despite these odds, the Economic Research Forum team and its Sudanese government partners succeeded in implementing a very rich survey of socioeconomic conditions in Sudan. Fieldwork was undertaken by the Central Bureau of Statistics (CBS) in June-September 2022. The resulting SLMPS 2022 is a nationally-representative sample of 4,878 households and 25,442 individuals.

Prior to this, the most recent nationally-representative survey data available for Sudan was from the 2014/15 Sudan Household Budget Survey (SHBS) and the 2014 round of the Multiple Indicator Cluster Survey (MICS5) (Central Bureau of Statistics (CBS) & UNICEF Sudan, 2016). All nationally-representative surveys previous to this, such as the Labor Force Survey or the Harmonized Household Health Survey, date to the period prior to the secession of South Sudan in 2011. The last population census was conducted in 2008. During COVID-19, mobile phone surveys were conducted, but mobile coverage is substantially less than universal in Sudan and mobile phone surveys were limited in length (Central Bureau of Statistics (CBS) & World Bank, 2020; Krafft, Nour, & Ebaidalla, 2022).

Sudan's economic development and challenges have shaped the progress and state of its labor market. The majority of employment remains in agriculture, and even an increasing share from 2009 to 2014 (Ebaidalla & Nour, 2021). Despite decades of agricultural policy efforts, productivity remains low and the potential of the sector underdeveloped (Elbadawi et al., 2022). With the secession of South Sudan, the country lost more than three-quarters of its oil production capacity, but subsequently expanded gold mining, remaining stuck in a "resource curse" (Elbadawi & Suliman, 2018). Recent years have not seen structural transformation, and have even seen declining productivity and real wages outside of agriculture (Etang Ndip & Lange, 2019). Crony capitalism during the Bashir regime may have been one factor limiting economic growth and competitiveness (Ali & Ebaidalla, 2023). Sudan's middle class has been shrinking (Omer & Maglad, 2020). The majority of employment over 2009 and 2014 was in non-wage work, primarily self-employment (Ebaidalla & Nour, 2021).

Sudan has a young population, but one that has struggled with education and insertion into the labor market (Ahmed, Albatal, & Musa, 2022; El-Kogali, Savrimootoo, Feda, & Chugunov, 2021;

El Din & Abd El Rhman, 2022; Etang Ndip & Lange, 2019). While male labor force participation had fallen over time, overall and female participation rose from 2009 to 2014, but women's participation rate remained substantially below that of men's (Ebaidalla & Nour, 2021). Unlike in other MENA countries, there were not large differences in participation by education (Assaad & Marouani, 2021; Ebaidalla & Nour, 2021). How the labor market has further evolved since 2014 is a question the SLMPS 2022 can help answer.

This paper describes the SLMPS 2022 data collection and data. Section 2 covers the household and individual questionnaire design and information on accessing the public use microdata. The complexities of the sample design are covered in Section 3, as well as the sample weights. Section 4 describes the processes of training and data collection. In Section 5, the SLMPS is compared to other, older nationally representative data from Sudan to both validate the results of the survey and update our understanding of Sudan's population and labor market. Section 6 concludes with a discussion of the strengths and limitations of this new data, as well as potential directions for future research.

2 Questionnaires and data availability

2.1 The Questionnaires

The SLMPS questionnaires consist of a household questionnaire and an individual questionnaire, with modules detailed in Table 1. The modules built on and ensured substantial comparability with other LMPSs. ERF received a grant from the Data Production and Methods Unit of the Development Data Group at the World Bank to include specific modules and questions from the Living Standards Measurement Study Plus (LSMS+) surveys that focus on gender disaggregated asset, employment, and entrepreneurship data (World Bank, 2023). The ERF research team worked closely with the LSMS+ team at the World Bank to incorporate these modules.

The household questionnaire includes: (i) identifiers and household location (ii) roster of household members (iii) housing conditions and durable assets (iv) current household member migrants abroad (v) remittances (vi) other income and transfers (vii) shocks and coping mechanisms (viii) non-agricultural enterprises, including information on characteristics, employment of household members and others, assets, expenditures, and revenue (ix) agricultural assets, including land, capital equipment, livestock, crops, and other agricultural income.

The individual questionnaire collects data from all individuals 5 and older (children under five are captured in the household roster). The individual questionnaire elicits information about (i) residential mobility (ii) father's, mother's and sibling characteristics (including siblings abroad) (iv) health and access to healthcare (v) educational attainment and detailed educational history (vi) training experiences (vii) skills (viii) current employment and unemployment (viii) job characteristics for the primary and secondary jobs (ix) labor market history (x) costs and characteristics of marriage, including marriage history (ix) fertility (xii) women's employment (xiii) wages from primary and any secondary jobs (xiv) return migration, refugee, and IDP experiences for Sudanese respondents (xv) modules for immigration and refugees for non-Sudanese respondents (xvi) information technology (xvi) savings and borrowing (xvii) attitudes (xviii) time use (a full 24 hour diary for adults and a shorter module for children) and (xix) a series of questions on rights to parcels of land, livestock, and durable goods.

Table 1. Questionnaire modules

Household	Individual

• Statistical Identification
• Individual Roster
• Housing Information
• Current Migration
• Remittances
• Other Income
• Shocks and Coping
• Nonagricultural Enterprises
  ○ Characteristics
  ○ Outside Employment
  ○ Expenditures
  ○ Assets
  ○ Enterprise Revenue
• Agricultural Assets: Lands and Parcels
• Agricultural Assets: Livestock/Poultry
• Agricultural Assets: Capital Equipment
• Agricultural Crops
• Other Agricultural Income
	• Statistical Identification
	• Residential Mobility
	• Father's Characteristics
	• Mother's Characteristics
	• Siblings
	• Health
	• Education
	• Training
	• Skills
	• Employment
	• Unemployment
	• Job Characteristics
	• Secondary Job
	• Labor Market History
	• Marriage
	• Fertility
	• Female Employment
	• Earnings
	• Earnings in Secondary Job
	• Return Migration (Sudanese)
	• Immigration/Refugees (non-Sudanese)
	• IDPs/Refugees (Sudanese)
	• Information Technology
	• Savings and Borrowing
	• Attitudes
	• Adult Time Use
	• Child Time Use
	• Rights: Parcels
	• Rights: Livestock
	• Rights: Durables

Source: Authors' construction

2.2 Public use microdata access

Public use microdata from the 2022 wave of the SLMPS are available from ERF's Open Access Microdata Initiative (OAMDI) (OAMDI, 2023). Researchers can request the microdata free of charge from www.erfdataportal.com. The other waves of LMPSs are also available from OAMDI, along with a harmonized data set, with a subset of variables from all countries and waves, referred to as the Integrated Labor Market Panel Surveys (ILMPS).

3 Sampling and sample design

A fundamental challenge when designing the SLMPS sample was the lack of a recent, nationally representative sample frame. The last national population census in Sudan was in 2008, before the secession of South Sudan. There had also been limited updating of administrative borders and maps. The first level of administrative geography in Sudan is the state (*wilaya*), and there are 18

states in Sudan.[5] The second level of administrative geography in Sudan is the locality (*mahaliya*), and CBS had updated the borders of localities in 2017 to 189 distinct geographies (each locality nested within a single state). We used the updated borders combined with 2020 population estimates based on remote sensing data to create our sampling frame and draw our sample. We supplemented these sources with additional data to identify refugee and IDP camps and areas for our strata. We describe the details of this sampling frame and design below.

3.1 Summary of sample design: Planned and realized strata and primary sampling units

The planned sample design was a random stratified cluster sample made up of 5,000 households sub-divided into 250 primary sampling units (PSUs). The strata represented in the sample are: (i) refugee camps, (ii) refugee areas (areas with non-camp refugee settlements), (iii) IDP camps, (iv) IDP areas (areas with non-camp IDP settlements), (v) other (non-refugee/non-IDP) rural areas, and (vi) other urban areas. We describe below the details of identifying these different geographies. These area strata were combined with states to create distinct state-area strata for our random stratified cluster sample. As discussed below, we created PSUs within each locality as rectangular geographic areas with similar populations as ascertained by the remote sensing data.

Table 2 shows the planned and realized sample by stratum and state. We successfully collected data from 250 PSUs as planned, and from the various states as planned, with some deviations in the area types sampled, as well as some replacement during fielding, as discussed below. In terms of the area types sampled, two fewer IDP camp PSUs were sampled, one fewer in South Darfur and one fewer in East Darfur than planned (other IDP camps PSUs were sampled in South Darfur but none in East Darfur, however an additional IDP area PSU was sampled in East Darfur as the best replacement[6]). One fewer refugee area PSU was sampled, in South Darfur (other refugee area PSUs were sampled in South Darfur). In both North Darfur and South Kordofan one fewer IDP area PSU than planned was sampled, but in both cases, other IDP area PSUs were sampled in each governorate. A number of governorates added or subtracted one urban other and/or one rural other area, as these were often the backup PSUs. In all but one case, at least some of the planned urban and rural other areas were sampled. In West Darfur, the one planned rural other area was not sampled.[7]

[5] The Abyei area, which under current peace terms is part of both Sudan and South Sudan, was combined with West Kordofan for sampling purposes.
[6] We therefore combine the IDP camp and IDP areas populations in East Darfur when weighting.
[7] We therefore combine the rural other populations of West Darfur and neighboring Central Darfur when weighting.

Table 2. Primary sampling units: Planned and realized, by state-area strata

Strata:	Refugee camp	IDP camp	Refugee areas	IDP areas	Urban other	Rural other	Total
State							
Planned sample							
Khartoum			17		15	2	34
North Darfur		3	1	8	4	1	17
South Darfur		5	3	7	1	1	17
West Darfur		3	2	3	1	1	10
East Darfur	1	1	1	1	6	4	14
Central Darfur		2	1	1		6	10
South Kordofan		1	1	5		3	10
Blue Nile					6	4	10
White Nile	6		1		5	2	14
Red Sea					8	5	13
Kassala	4		1		5	3	13
Gedaref	4		1		4	2	11
North Kordofan				1	10	6	17
Sannar					6	4	10
Aj Jazirah					10	7	17
River Nile					6	4	10
Northern					6	4	10
West Kordofan			1	4		8	13
Total	15	15	30	30	93	67	250
Realized sample							
Khartoum			17		15	2	34
North Darfur		3	1	7	5	1	17
South Darfur		4	2	7	2	2	17
West Darfur		3	2	3	2		10
East Darfur	1		1	2	7	3	14
Central Darfur		2	1	1		6	10
South Kordofan		1	1	4		4	10
Blue Nile					7	3	10
White Nile	6		1		5	2	14
Red Sea					9	4	13
Kassala	4		1		6	2	13
Gedaref	4		1		4	2	11
North Kordofan				1	9	7	17
Sannar					7	3	10
Aj Jazirah					10	7	17
River Nile					5	5	10
Northern					6	4	10
West Kordofan			1	4		8	13
Total	15	13	29	29	99	65	250

Source: Authors' construction based on planned sample and SLMPS 2022 realized sample

3.2 Sample inputs

PSUs were created using publicly available datasets, QGIS, and Stata processing. The 2020 population estimates that underlie our sampling were created by WorldPop (WorldPop, 2020). WorldPop created geoTIFF files with 100 meter x 100 meter (pixel) population estimates.

Estimates are based on the official United Nations population estimates with dasymetric redistribution using random forest modeling and geospatial covariates (Stevens, Gaughan, Linard, & Tatem, 2015; WorldPop, 2020). We used Sudan state and locality boundary files (OCHA Regional Office for Sourthern and Eastern Africa (ROSEA), 2020) to ensure PSUs did not cross administrative boundaries.

We used two additional data sources to identify refugee camps/areas and IDP camps/areas. For refugees, we used geospatial data from UNHCR, updated in January 2021 (OCHA Sudan, 2021). The dataset included points for a variety of location types, such as office, reception center, and crossing/entry point, as well as residential areas. We specifically used the locations for camps, refugee camps, collective self-settlement, open areas, and refugee settlement to identify our refugee camp and refugee area strata. For IDPs, we used the International Organization for Migration (IOM), Displacement Tracking Matrix (DTM), round one from November 2020 (International Organization for Migration (IOM), 2020). The data included points for a variety of location types, all of which were included as they contained households, and which allowed us to distinguish camps from other types of IDP settlements.

3.3 Creating Primary Sampling Units

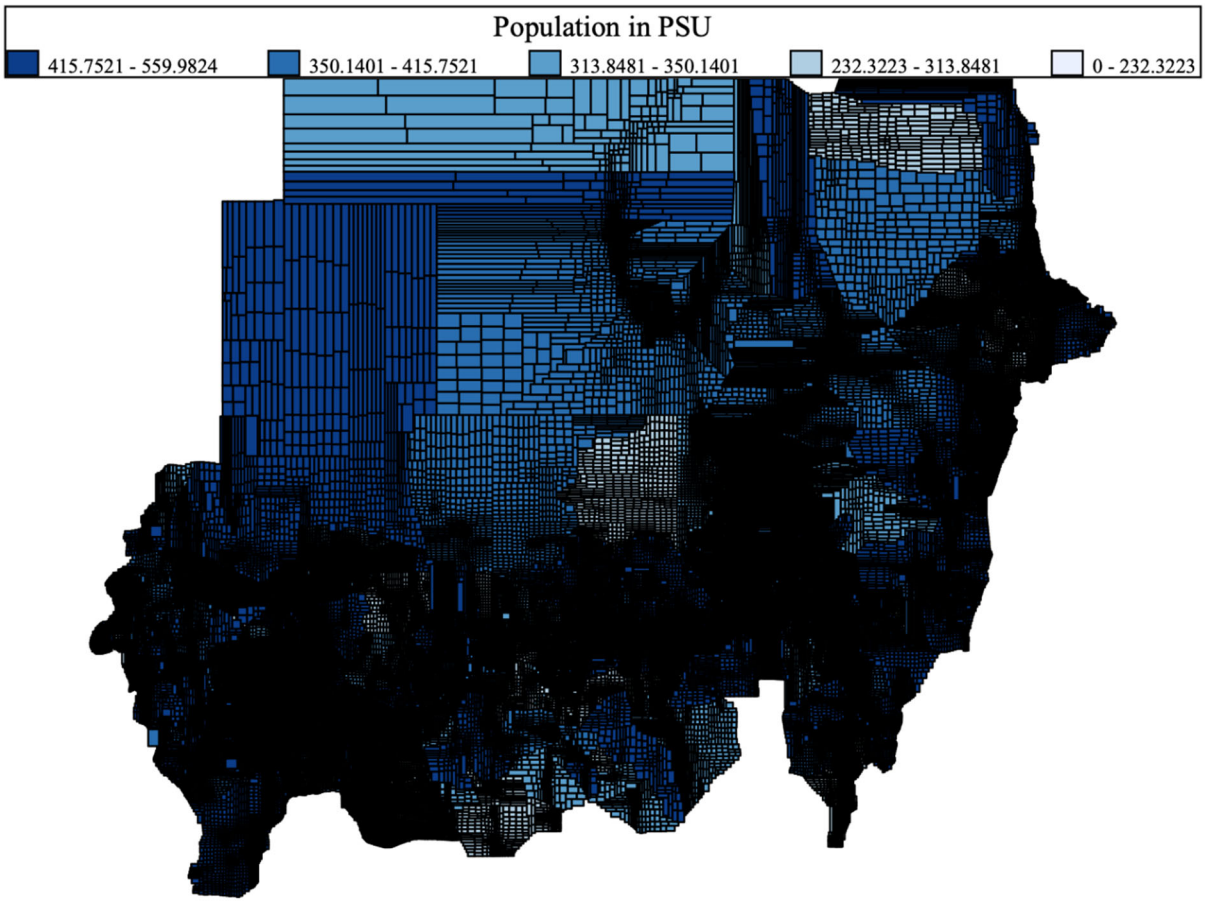
We initially split the population pixels into separate files (using QGIS) for each locality, to make a series of tractable datasets for creating PSUs. Looping over localities, we imported the geospatial data (shapefiles) into Stata.[8] We set a maximum population for the PSU of 560 individuals, as we were aiming for a maximum size of 100 households per PSU for feasibility of listing in the field and the average household size in the 2008 census was 5.6 persons. We then undertook a splitting algorithm in order to generate tractable (rectangular) PSUs for fielding,[9] splitting the locality (and subsequently sub-areas) in half (population-weighted half, not area) on latitude and then longitude, repeatedly splitting so long as the population was above the maximum threshold for each area.

The resulting PSU areas had populations from 232-560 individuals and are illustrated, in full for all of Sudan, in Figure 1. Note that the areas that appear black (PSU borders are black) are those with high population densities.

[8] Using the shp2dta command.

[9] Boundaries were modified in QGIS to respect locality borders and fix any overlaps, making a few PSUs otherwise than rectangular in fielding.

Figure 1. Full sample frame of PSUs, classified by individual population in the PSU



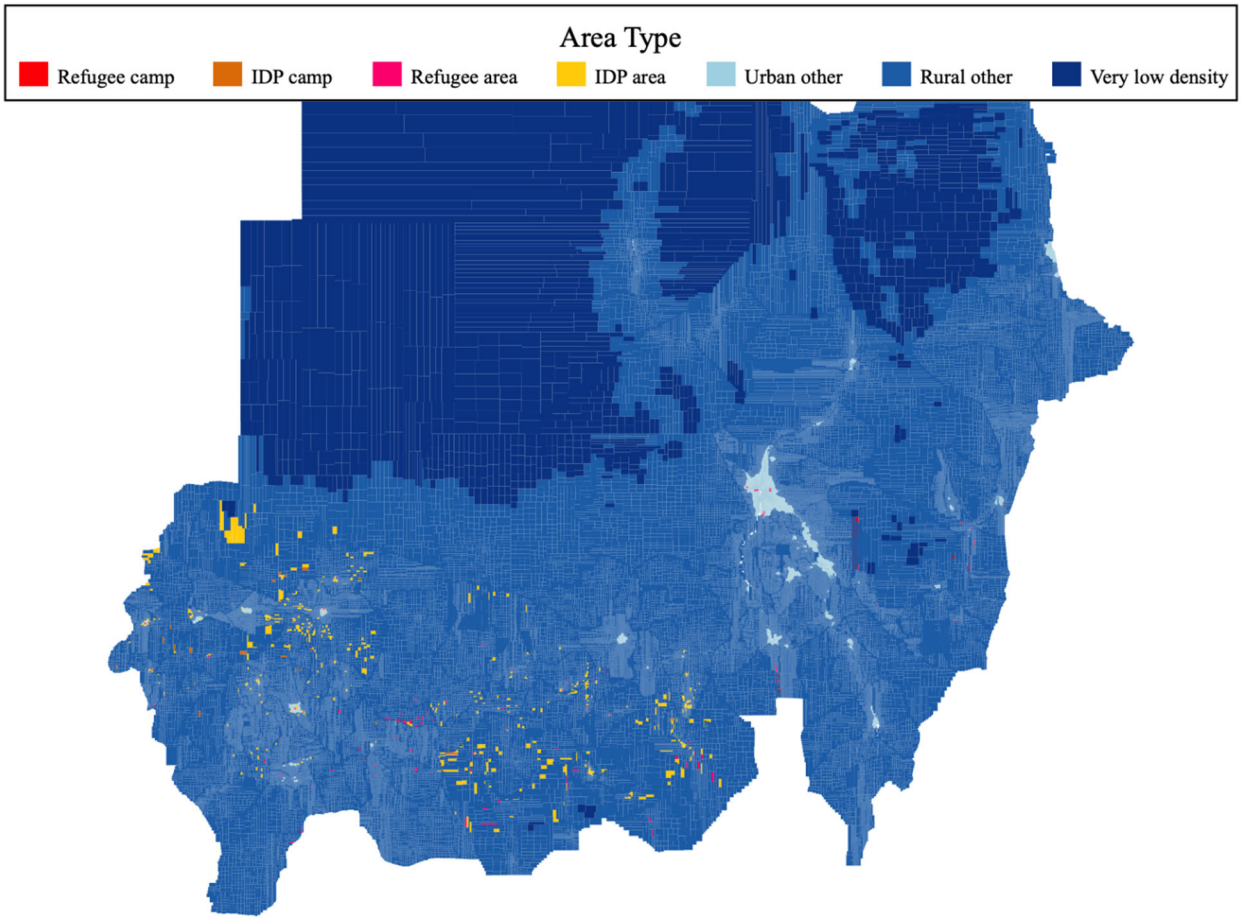
Source: Authors' construction based on WorldPop population estimates (WorldPop, 2020) and locality boundaries (OCHA Regional Office for Sourthern and Eastern Africa (ROSEA), 2020).
Notes: Black lines denote PSU borders, areas with denser population may appear entirely black from the borders.

We then classified the PSUs into strata. We identified the PSUs containing the coordinates (single points) given for IDP and refugee camps,[10] and classified these as the IDP and refugee camp strata. We likewise identified the PSUs containing the coordinates for other (non-camp) types of IDP and refugee settlements, as well as PSUs whose centroid was within two kilometers[11] of a refugee/IDP camp or refugee/IDP settlement (but was not the location of a camp). We classified these as, variously, refugee or IDP areas. We then distinguished the urban or rural nature of other PSUs based on the density. Following UN Statistics guidance (United Nations Statistical Commission, 2020) on measuring the degree of urbanization, we classified as urban any PSU where the population was at least 300 persons per square kilometer. Given the substantial amount of land in

[10] Using the Stata command geoinpoly.
[11] Using the Stata command geonear.

Sudan that is uninhabited, and the practicalities of fieldwork, we also identified unpopulated areas based on a maximum density at the first percentile (less than 0.025 persons per square kilometer). We exclude this land area (which is estimated to have very limited population) from our sampling. The resulting sample frame consisted of 19 refugee camp PSUs, 63 IDP camp PSUs, 868 refugee area PSUs, 1,124 IDP area PSUs, 28,747 urban other PSUs, and 82,077 rural other PSUs, along with 1,149 excluded, very low-density PSUs. The frame is mapped in Figure 2.

Figure 2. Full sample frame of PSUs, classified by area type



Source: Authors' construction based on WorldPop population projections (WorldPop, 2020) and locality boundaries (OCHA Regional Office for Sourthern and Eastern Africa (ROSEA), 2020), data on locations of refugees (OCHA Sudan, 2021) and IDPs (International Organization for Migration (IOM), 2020).

3.4 Sampling PSUs and households

Using this sample frame, we undertook a random stratified sample, probability proportional to size (population) and without replacement in Stata,[12] following the PSU sample sizes presented in Table 2. For fieldwork, we processed the resulting sample in QGIS and prepared it for listing of households on tablets, using QField. With QField, fieldwork supervisors were instructed to canvas the entire PSU (based on the polygon boundaries mapped on their tablet) and place points for all the households in the area, numbering sequentially (estimated per the sample frame to be 50-100 households). The number of households was then entered into the ODK-X (tablet data collection) supervisor program and gave a random sample of 20 households from the number listed, along with a random sample of up to 10 backups (in a random order). If there were 20 or fewer households in the field, all households were sampled. The 20 initial households could be fielded in whatever order was convenient, but if any ultimately were not part of the sample, the backups were to be used in the random order provided.

3.5 GPS coordinates

We received permission from CBS to make available in the public use version of the data GPS coordinates of the centroids of the sampled PSUs. To maintain confidentiality, we follow a displacement practice similar to the Demographic and Health Surveys and Living Standards Measurement Surveys (Burgert, Colston, Roy, & Zachary, 2013; Michler, Josephson, Kilic, & Murray, 2022). From the true centroid of the PSU, we add a random displacement error of up to 5 kilometers in rural areas and up to 2 kilometers in urban areas. In an additional 1% of rural areas, the random displacement error was up to 10 kilometers. We ensured that the displaced centroids remained in the true state and locality. The availability of these GPS coordinates will allow researchers to link the microdata of the SLMPS 2022 with readily available global geospatial data. For instance, the SLMPS 2022 data can be linked to land cover and weather conditions that are available at high geographic resolutions and will thus open up opportunities to study the impact of climate change and desertification on a wide variety of outcomes. It will also be possible to link the SLMPS 2022 data with geolocated data on conflict and civil unrest, allowing for the study of the impact of conflict on human populations.

3.6 Sample weights

From the sample design and data collected in the field, we can generate sample weights that are designed to make the SLMPS nationally representative. This section describes the creation of those weights. We start with the probability of a PSU being sampled. Denote a PSU as p. Denote a stratum as s. Denote a state as t. We sampled probability proportional to size based on the WorldPop 2020 population estimate (the sum of the population across the pixels within the PSU). Denote this PSU-specific population estimate as $E_{p,t,s}$. We sampled a number of PSUs from each stratum-state (see realized cells in Table 2). Denote the number of PSUs completed as $C_{t,s}$ and the total number of PSUs in the stratum-state as $P_{t,s}$. The probability of a PSU being sampled from within a stratum-state is therefore:

[12] Using the Stata command gsample.

$$\pi_{p,t,s} = \frac{E_{p,t,s} * C_{t,s}}{\sum_{p=1}^{P_{t,s}} E_{p,t,s}} \quad (1)$$

That is, the population of the PSU, the number of clusters sampled in the stratum-state, and the total population across all PSUs (not just sampled PSUs) in the stratum-state determine the probability of a PSU being sampled. When the PSU population is higher, or more PSUs in the stratum-state are being sampled, the probability of being sampled is higher. When the total number of PSUs in the stratum-state is higher, the probability of a particular one being sampled is lower.

During fieldwork, the number of households within the PSU was supposed to be listed and recorded into ODK-X. In reality, in PSUs that turned out to have larger populations, supervisors did not fully list the PSU. After the fieldwork was completed, the CBS GIS team used satellite images to estimate the number of households within each PSU.[13] Denote this listed population as $L_{p,t,s}$. Denote the number of successfully completed households (this number accounts for non-response) as $K_{p,t,s}$. This was supposed to be 20 per PSU, but might be smaller if the PSU turned out to be smaller than 20 households, or if there was non-response. In some cases, more than 20 households per PSU were sampled to make up for shortfalls elsewhere. In all cases the weights account for these deviations. The probability of a particular household, within a PSU, being in the sample is thus:

$$\rho_{p,t,s} = \frac{K_{p,t,s}}{L_{p,t,s}} \quad (2)$$

When a larger share of households in the PSU was completed, this probability is higher. It is lower for larger PSUs.

We combine together the probability of the PSU being sampled and the probability of the household being sampled within the PSU to generate the PSU weight as the inverse probability of a particular household being sampled:

$$w_{p,t,s} = \frac{1}{\pi_{p,t,s} * \rho_{p,t,s}} = \frac{1}{\frac{E_{p,t,s} * C_{t,s}}{\sum_{p=1}^{P_{t,s}} E_{p,t,s}} * \frac{K_{p,t,s}}{L_{p,t,s}}} = \frac{(\sum_{p=1}^{P_{t,s}} E_{p,t,s}) * L_{p,t,s}}{E_{p,t,s} * C_{p,t,s} * K_{p,t,s}} \quad (3)$$

Since, as discussed in the fieldwork description below, PSUs that appeared empty or to have fewer than five households were not fielded, these weights, unsurprisingly, sum to more than Sudan's 2020 population. The weights also reflect substantial variability based on the listed populations of a finite number of PSUs. We therefore adjust weights to be equivalent to the sampling frame by stratum-state based on the 2020 population estimates,[14] as follows:

[13] Calculations were based on the ratio of buildings observed on the map to the number listed in the field for the fraction of the PSU the supervisor did list. This ratio and the number of buildings in the listed area was used to generalize to the PSU as a whole based on the fraction of the PSU it covered.

[14] Only national estimates are available for 2022, not state and stratum specific estimates.

$$\widetilde{w}_{p,t,s} = \frac{w_{p,t,s} * \sum_{p=1}^{P_{t,s}} E_{p,t,s}}{\sum_{p=1}^{P_{t,s}} w_{p,t,s}} \quad (4)$$

This essentially multiplies all the weights within a stratum-state by a constant fraction based on the shares within each stratum-state in 2020.

A further complexity arose in terms of nationality and representation of non-Sudanese populations, particularly refugees, as well as Sudanese IDPs. Moreover, the national population presumably grew between 2020, the date of our sample frame, and mid-2022 when we fielded. We therefore adjust $\widetilde{w}_{p,t,s}$ to reflect a 2022 mid-year national population of 46.9 million (United Nations Department of Economic and Social Affairs Population Division, 2022). This population, Q, is divided into four groups, denoted g: refugee households (1.1 million individuals (UNHCR, 2022)), non-refugee non-Sudanese households (250 thousand individuals), IDP households (3.7 million individuals (UNHCR, 2022)), and non-IDP Sudanese households (41.8 million). The adjustment for group g is:

$$\widetilde{w}_{p,t,s,g} = \frac{\widetilde{w}_{p,t,s} * I_g * Q_g}{\sum_{s=1}^{S} \sum_{t=1}^{T} \sum_{p=1}^{P_{t,s}} \widetilde{w}_{p,t,s} * I_g} \quad (5)$$

Where I_g is a dummy variable indicating that an individual is in group g and Q_g is the national population of group g. This modification adjusts the weights to reflect the population of different groups (IDPs, refugees, non-Sudanese, Sudanese) in 2022. These are the final household weights used in analyses.

3.7 Individual non-response and individual weights

Similar to households being visited up to three times in an attempt to collect data, individuals aged five and older were visited up to three times to collect data from the individual him or herself, before a proxy respondent could be taken.[15] Individuals ultimately might have non-response if they refused to respond, or, if on the third visit, the individual was not available and a proxy was not available to respond on their behalf. Of the 21,057 individuals aged five and older in the household roster data, 20,086 (95%) consented to and completed the individual questionnaire. Among the 971 individuals who did not complete the individual questionnaire, 74 were not available on the third visit and no proxy was available. The remaining 897 refused or, if under 18, may have refused themselves or had parents refuse to consent for them to be interviewed. Individuals who did not respond would still be among those listed by the household in the roster, so we have their basic demographic characteristics, but not the detailed individual interview data.

We account for individual non-response and generate individual non-response weights, with covariates (*x)* for sex, age group, marital status, stratum, state, household nationality, household

[15] Parents of children 5 to 17 had the option to respond on their children's behalf at any time.

IDP status, and household refugee status. We interact sex with both age group and marital status as well. We estimate the predicted non-response rate as r_x based on a logit model.

We adjusted the household weight by this non-response to get the individual weight of:

$$w_{p,t,s,g,x} = \frac{\widetilde{w}_{p,t,s,g}}{1 - r_x} \quad (6)$$

Weighting up those individuals who did complete the individual questionnaire and have similar characteristics to those who were particularly likely not to respond.

4 Training and data collection

4.1 Planning and preparations

Planning for the SLMPS 2022 began as early as the fall of 2017 when ERF submitted a proposal to conduct the survey to the Growth and Labor Markets in Low Income Countries (GLM | LIC) research program of the Center for Labor Economics (IZA) in Bonn, Germany. While planning and preparations, such as adaptation work, had begun, in fall 2018, protests against the Bashir regime in Sudan led to the regime's eventual fall and a period of political instability. With the initiation of the transitional period marked by the establishment of the civilian-led government in partnership with the military in 2019, it was possible for ERF to resume efforts to conduct the survey in early 2020. Meetings were held in January 2020 with the Ministry of Finance, the Ministry of Labor and Social Development, and CBS where an agreement in principle to implement the survey was reached. Several consultations with Sudanese researchers were held over the course of 2020 to adapt the LMPS questionnaire last implemented in Egypt in 2018 to Sudanese conditions. A complete version of the questionnaire programmed for implementation on a tablet using ODK-X (Brunette et al., 2017) was available by November 2020.

After a series of delays associated with the onset of the COVID-19 pandemic and the turbulent economic situation in Sudan, including the lingering effects of sanctions on its banking system, a three-way contract involving what had then become the Ministry of Labor and Administrative Reform (MoLAR), CBS, and ERF was signed in April 2021. This led to the formation of a steering and a technical committee involving various agencies of the Sudanese government, which would oversee the implementation of the survey. A training of trainers (TOT) was conducted by ERF researchers and staff in May 2021 to train members of MoLAR and CBS on all aspects of implementing the survey. These trainers then led a two-week training camp organized by CBS for enumerators and supervisors in August 2021. The plan was to start the fieldwork in all 18 of Sudan's states by September 2021.

A series of minor delays pushed the start of the fieldwork into October. Then the military-led coup against the civilian government happened on the 25th of October, 2021, derailing the project timeline. ERF persisted and remained in constant contact with CBS over the ensuing months. As the political situation stabilized somewhat, it was decided to initiate fieldwork in May 2022 after a short online refresher training of the enumerators and supervisors. Fieldwork actually started on June 11, 2022, and most field operations were completed by the end of September 2022. The ERF research team received the anonymized raw data in November 2022 and the first beta version of

the data was ready for analysis by January 2023, just a few months before the collapse of the Sudanese government and the breakout of hostilities on April 15th, 2023.

4.2 Training

As mentioned above, a one-week TOT was organized by ERF for CBS and MoLAR staff in May 2021. The TOT included a review of the purpose of the survey, definition of concepts, fielding procedures, and intensive practical training on administering the questionnaire on the tablet using previously written cases as well as mutual interviewing among trainees. The IT and statistics staff at CBS was also trained on data transmission, data handling, and data processing. A field pre-test of the survey was carried out following the TOT on five PSUs not in the final sample and used to make final modifications to the questionnaire.

In August of 2021, a two-week training camp for enumerators and supervisors was organized in Aj-Jazira state. The trainees included 120 enumerators and 18 supervisors recruited locally in all 18 Sudanese states as well as a team of quality control enumerators recruited centrally. The trainers were made up of CBS and MoLAR staff who were trained in the TOT as well as one representative of the ERF research team. At the end of the training, a day of field practice was organized for the trainees to gain practical field interviewing skills.

Although fieldwork was supposed to start shortly after the completion of the enumerator training, as series of delays pushed it into October when the coup that brought an end to the civilian-led government in Sudan happened on October 25th, 2021. The ensuing political unrest forced a further a postponement of the fieldwork. In April 2022, CBS informed ERF that conditions had stabilized sufficiently to re-launch the work. An online five-day refresher training for the enumerators and supervisors was organized in late May and early June 2022.

4.3 Fieldwork

A total of twenty enumerator teams were formed, one for each state, except for Khartoum and South Darfur states, which had two teams each. Teams varied in size from 3 to 5 enumerators and were led by one supervisor. Fieldwork began June 11, 2022, and was completed by September 26, 2022.[16] Due to the lack of a standard sampling frame, fieldwork proceeded in two steps, listing households and then surveying a random sample of listed households. First, supervisors were supposed to list all households within the borders of the PSU. The borders of the PSU were visualized using QField, a tablet-based geospatial mapping system that allowed the supervisors to mark down the GPS coordinates for each household within the PSU boundaries, give it a number, and take notes on its location.[17] The supervisor then switched to the ODK-X application on the tablet and entered the total number of households. The tablet then provided a random sample of 20 households, which could be fielded in any order, and a backup random sample (to be fielded in the order given, if one of the 20 households refused or otherwise could not be completed).

Data collection for the randomly selected households was undertaken using the ODK-X survey program on Android tablets. Enumerators were instructed to attempt to visit a household up to three times, at least a day a part, before substituting a backup household. Likewise for individuals,

[16] A few interviews were reviewed and revised over the phone thereafter.
[17] A household was considered within the PSU if the front door/entrance of the dwelling was within the boundaries.

up to three visits were to be conducted to attempt to have the individual respond for him or herself. Only if an individual was incapable of responding for him/her self or on the third visit was still not available was a proxy used. Response by a proxy was marked in the questionnaire.

There were substantial electricity and internet connectivity challenges during the fieldwork stage. Power banks and solar chargers were used to recharge tablets. ODK-X uses a SQLite database on the tablet that allowed for local storage of the data even when not connected to the internet. When connectivity was available, data were exported to csv-formatted files and migrated to a Boxcryptor (encrypted) Dropbox folder for each enumerator on the tablet, allowing secure syncing as connectivity permitted with the central CBS office in Khartoum. The central office compiled the data across supervisors and enumerators.

4.4 Deviations from the planned sample in fielding implementation

The realities of the sample frame and the logistics of fielding led to a number of deviations from the planned sample in fielding. While the initial sample was *estimated* to have a reasonable number of households in each PSU based on satellite imaging and population projections, there were cases where a PSU did not, in fact, have any or many households. All PSU locations were reviewed first in the CBS offices to identify locations that were empty or where there appeared to be five or fewer households and these locations were replaced with backup PSUs. There were a variety of reasons why a PSU might have few or no households, including that it consisted of industrial/commercial (not residential) buildings, that it was a mine or grain storage area, or that it had rocks or grain silos that looked like residences.

When office review determined there were at least five or more potential households on the satellite maps, fielding was attempted. However, a number of issues arose in the field as well. Upon visiting, buildings were determined to be non-residential, or were abandoned. Furthermore, a number of locations were determined to be unsafe to field, a status that even changed and fluctuated frequently during the fieldwork. Persistent sandstorms also prevented fielding in specific localities. The rainy season likewise made some locations inaccessible for fielding.

Backup samples were created; initially one urban and one rural backup were provided per state, and further backups were drawn as needed to replace PSUs that could not be fielded. Backups were, if possible, from the same strata and always from the same state. When possible, additional backups were also drawn from the same locality in an attempt to minimize bias. However, there were cases when an entire locality became inaccessible.

Ultimately, 152 PSUs from the original sample of 250 were fielded in the initially planned locations. Nine of the initially planned backups were used. For the remainder, 24 were replaced by the first replacement given, 17 by the second, 17 by the third, 9 by the fourth, 6 by the fifth, 4 by the seventh, and the remaining 12 by various higher order replacements. Repeated replacements tended to occur in localities with a high share of buildings (e.g. mines, grain storage) that the population estimates likely mistook for residences.

The realized sample sizes also varied somewhat per PSU (as discussed above, this is incorporated into the weights). The smallest completed PSU had only five households and the largest 29. The 25^{th} percentile was 19, the median 20, and through the 90^{th} percentile also 20. Smaller PSUs were

generally the result of the PSU having fewer than 20 households. Refusals at the household level were relatively rare; only 81 households refused.

4.5 *Quality control processes*

Quality control took place throughout the fieldwork, along two main axes: quality control call-backs and desk review of data. The CBS central office undertook callbacks targeting one random individual from within the household who had a mobile number for 5% of households. At the household level, half (51%) of households had a mobile, and 42% of individual questionnaire respondents aged 15+ who were asked for their mobile numbers. Households and individuals were selected per enumerator to ensure all enumerators were monitored. The entire household questionnaire was redone and the target individual's questionnaire through the wages section (if applicable). A total of 275 quality control households were successfully completed.

Desk review of the data focused on ensuring data completion. ODK-X stores data in separate tables for different levels of data (household observation; individual observation; crop observation; parcel observation, etc.). These data are entered through linked sub-forms that should be accessed through the parent, household form, to ensure data linkages. The data are also exported into separate tables as csvs. Although ODK-X can sync all tables with a server, given electrical and connectivity limitations, Boxcryptor and Dropbox with exports for each table were used. When households were finalized in the household forms, ODK-X checked all data were present, including all linked forms. However, exporting, copying, and syncing were prone to human error, as well as errors in following the flow of forms. Desk review focused on ensuring all observations at all levels were present, working with enumerators to re-export and sync when there were issues, and undertaking quality control calls and corrections when such issues could not be easily solved. The resulting data were cleaned to address a number of remaining issues (e.g. two enumerators entering the same household number).

5 Key findings and comparisons to other data from Sudan

Ideally, we would validate key statistics from SLMPS 2022 against other, contemporaneous sources from Sudan, as we have done for the ELMPS, JLMPS, and TLMPS (Assaad, Ghazouani, Krafft, & Rolando, 2016; Assaad & Krafft, 2013; Krafft & Assaad, 2021; Krafft, Assaad, & Rahman, 2021). There are no contemporaneous, nationally representative data sources for Sudan; the last representative surveys there were in 2014. We do compare our results to those data sources, as well as the most recent, 2008 census, and most recent, 2011, labor force survey. However, we would not expect statistics to be identical from 2008 or 2014 and 2022; indeed, we present our results more as updates on key labor market and demographic issues than validation.

5.1 Other data sources

In 2008, Sudan completed its last census. At the time, South Sudan had not seceded. A 16.6% sample of the census is available from IPUMS International (Minnesota Population Center, 2020).[18] We kept the states in the 2008 census that ultimately were part of Sudan after the partition.

[18] The authors wish to acknowledge the statistical office that provided the Census data: Central Bureau of Statistics, Sudan.

In 2011, Sudan completed its last labor force survey (SLFS). We do not have access to the microdata, but do have the detailed report (Ministry of Human Resources Development and Labour, 2011) and other research using the SLFS 2011 data and the 2014/15 Sudan Household Budget Survey (SHBS) (Ebaidalla & Nour, 2021). We use statistics from the reports to compare to the SLMPS 2022.

In 2014, Sudan completed its most recent, publicly available, nationally representative survey, the Multiple Indicator Cluster Survey (MICS). The survey collected data on 16,801 households (Central Bureau of Statistics (CBS) & UNICEF Sudan, 2016). We downloaded the MICS microdata and analyze it in comparison to the SLMPS 2022.

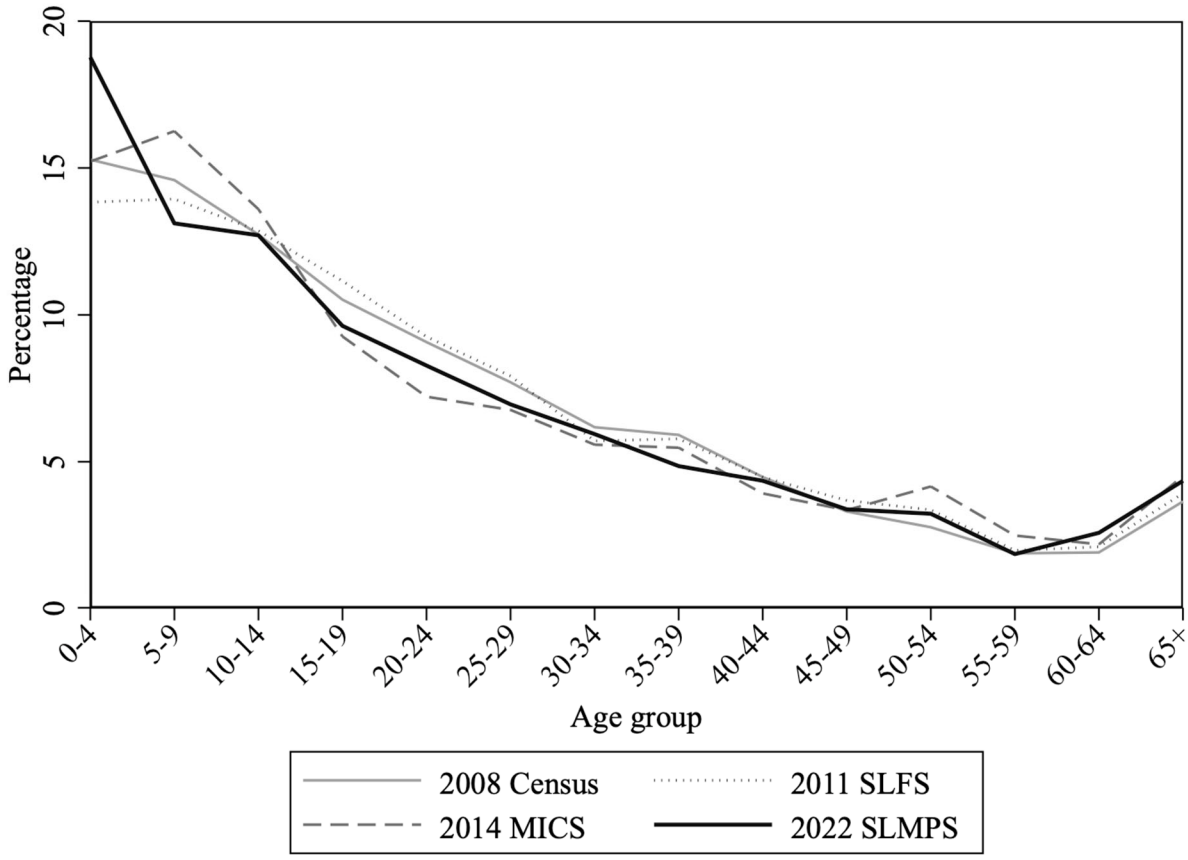
In discussing the SLMPS 2022 results, we also contextualize recent events with research from two COVID-19 era mobile phone surveys (Central Bureau of Statistics (CBS) & World Bank, 2020; Krafft, Nour, & Ebaidalla, 2022; Nour, 2022). Mobile phone ownership is substantially less than universal in Sudan, so these samples are selective, and thus we do not present them in figures and tables.[19]

5.2 Demographic characteristics

With regard to the distribution of the population by age according to the various data sources shown in Figure 3, the SLMPS 2022 results are generally similar to those of the population census and other available surveys for Sudan for the age distribution of the working age population (15+) and for the population 10-14. Where some deviations exists is for the 0-4 population, which appears to be overstated in SLMPS 2022 and the 5-9 population, which appears to be understated. We suspect this is because those aged 5 and above are eligible for inclusion in the individual questionnaire, which implies a lot more data collection time on the part of the enumerators. To avoid the extra work, the enumerators had an incentive to enter five or even six-year-olds as four-year-olds. We attempted to offset the extra effort of filling out individual questionnaires by proposing that enumerators be paid according to the number of data items collected rather than the number of households completed, but CBS deemed that remuneration system too complex. Conversely, in the MICS 2014 survey, the focus is on 0-4 year-olds, and thus enumerators in that survey have an incentive to under-report that age group. This is not uncommon in under-resourced settings (Pullum, 2006; Pullum & Staveteig, 2017). Overall, the data sources agree that Sudan has a disproportionately young population, as fertility has remained around 5 births per woman since the 1990s (Krafft, Assaad, Cortes-Mendosa, & Honzay, 2023).

[19] Only 51% of households in SLMPS 2022 and 42% of adults aged 15+ report owning a mobile phone.

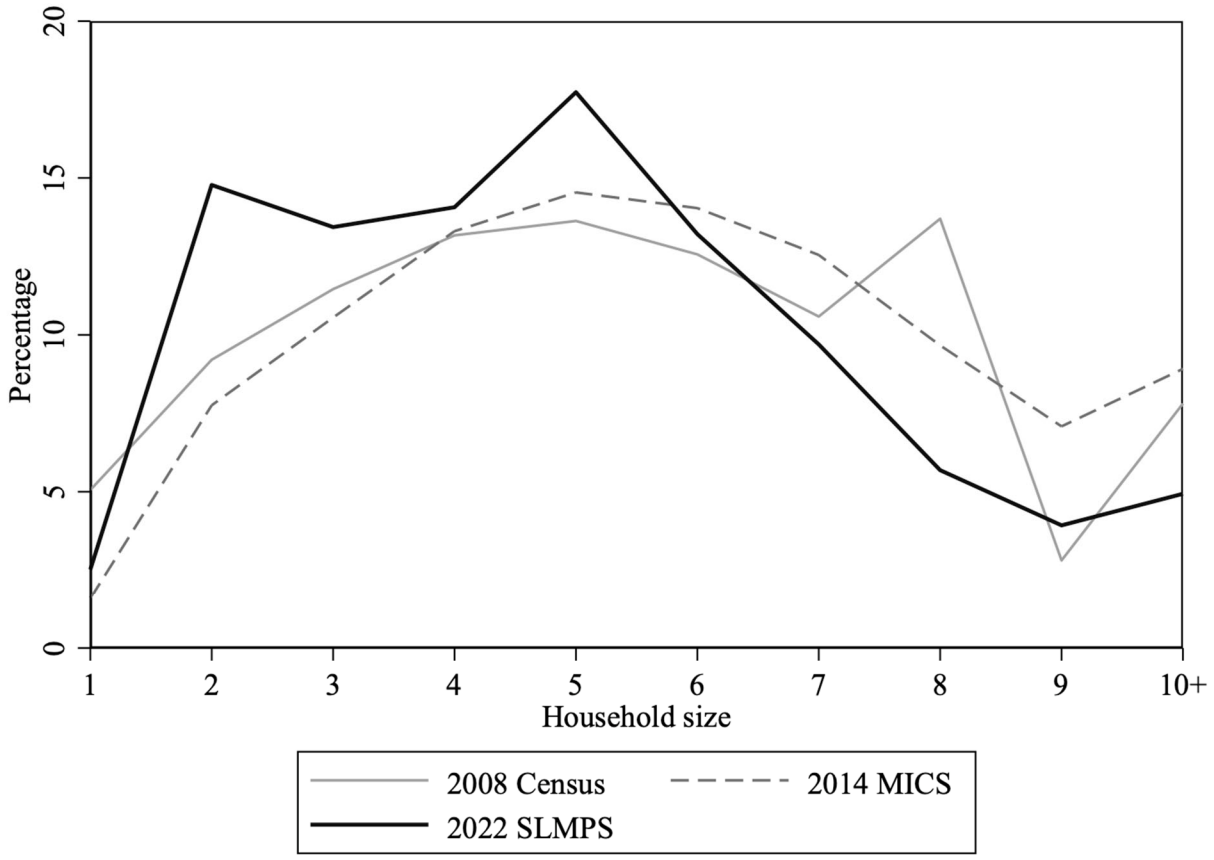
Figure 3. Percentage of the population by age group and data source



Source: Authors' calculations based on SLMPS 2022, MICS 2014, Sudan Census 2008, and SLFS report (Ministry of Human Resources Development and Labour, 2011).

For the distribution of households by household size shown in Figure 4, we can see that the SLMPS 2022 has a tendency to overstate the prevalence of two and five-person households and understate the prevalence of large households of seven plus members relative to the other sources. While the 2008 Census had a mean household size of 5.4, the 2014 MICS had a mean household size of 5.8, and the SLMPS 2023 a mean household size of 4.9. The lower household size result in the SLMPS 2022 could be due to declines in family size over time, or a tendency towards more nuclear households. It could also be due to a more stringent application and training on the definition of a household. However, it is also potentially the result the relatively higher cost in terms of enumerator time and effort of including larger household sizes when a substantial individual questionnaire is included as is the case in SLMPS 2022.

Figure 4. Percentage of households by household size and data source

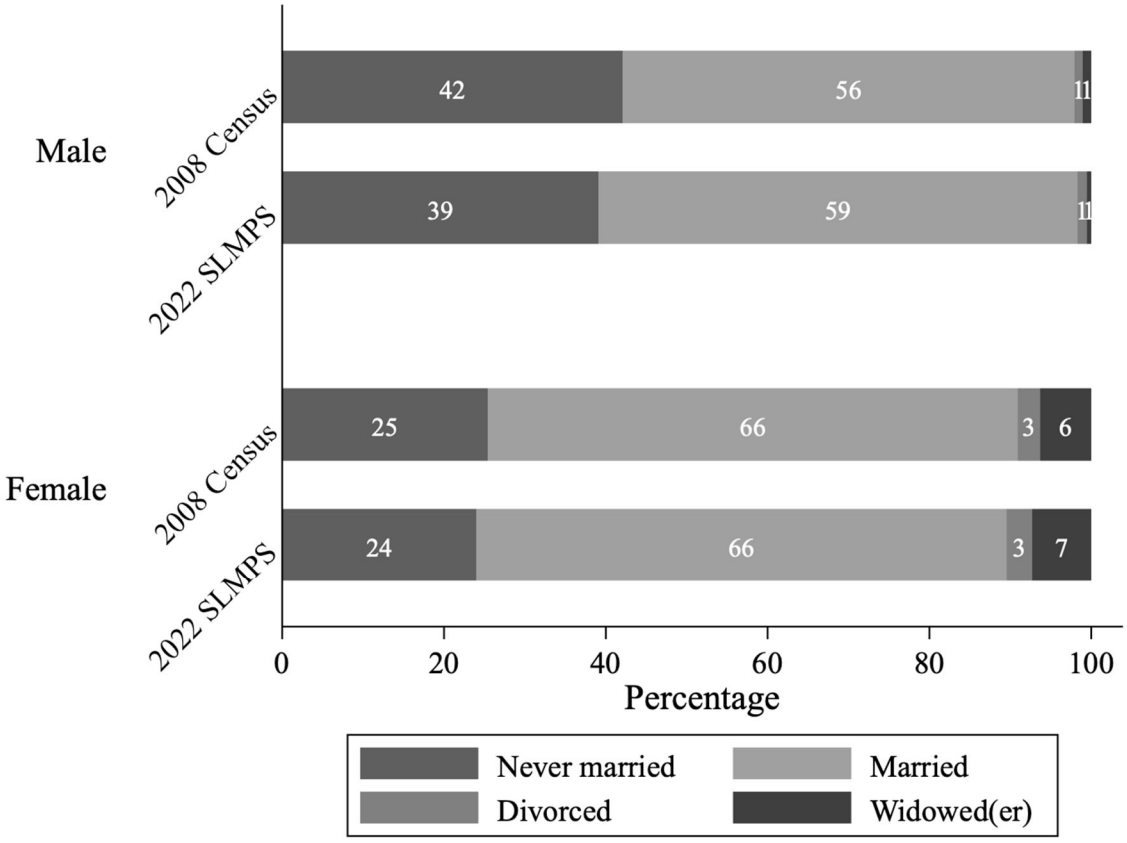


Source: Authors' calculations based on SLMPS 2022, MICS 2014, Sudan Census 2008.
Notes: Household size not available in SLFS report.

The population across the MICS, Census, and SLMPS is similarly divided by sex, with 50% of the population women across all the data sources. Marriage is often early for women in Sudan; the median age of marriage for women was 18 in both the 2008 Census and 2022 SLMPS.[20] Marriage even as early as 15 was common, 25% of girls married by this age in SLMPS 2022 and 23% by this age in the 2008 Census. Fully 40% were married by age 17 in the census, and 41% in the SLMPS. The median age of marriage for men was 25 in the 2008 Census and 26 in the 2022 SLMPS. Figure 5 explores marital status distributions by sex across the two data sources. Distributions are very similar, with 42% of men in the census and 39% in the SLMPS never married, and likewise 25% of women never married in the 2008 Census, and 24% in SLMPS 2022. Very few men were widowed (1%) or divorced (1%) in both data sources. Similar shares of women were divorced (3%), and widowed (6% Census, 7% SLMPS).

[20] Calculations of age of marriage account for right-censoring (those not yet or never married) using the Kaplan-Meier estimator. Age of marriage estimates are based on the population aged 15+ at the time of survey.

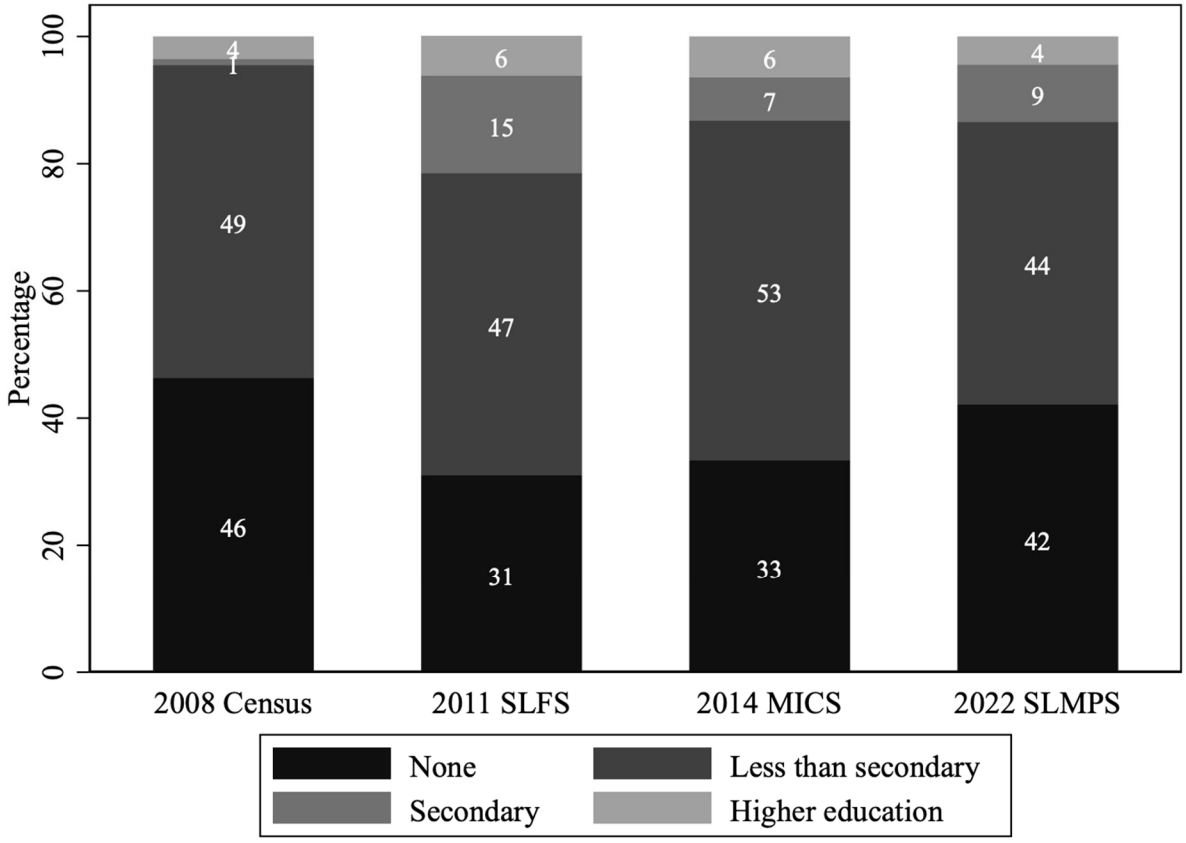
Figure 5. Marital status (percentage of individuals aged 15+), by data source and sex



Source: Authors' calculations based on SLMPS 2022, Sudan Census 2008.
Notes: Marital status not available in SLFS report. Marital status only available for women aged 15-49 in MICS 2014 and therefore not shown.

The distribution of individuals 6+ by educational attainment shown in Figure 6 exhibits a lot more variation across source than other statistics. The proportion of those without schooling varies from 46% in the 2008 Census to 31% in the 2011 LFS, with MICS 2014 closer to the LFS at 33%, and the SLMPS 2022 closer to the census at 42%. There is somewhat less variation in the proportion of those with less than secondary education, which varies from 44% in SLMPS 2022 to 53% in MICS 2014, but substantial variation in the proportion of those with secondary degrees, which varies from 1% in the 2008 Census to 15% in the 2011 LFS, with MICS 2014 and SLMPS 2022 occupying an intermediate position at 7-9%. Finally, the proportion of those with higher education has somewhat less variation across sources, 4% in both the Census and SLMPS and 6% in LFS and MICS.

Figure 6. Education level (percentage of individuals aged 6+), by data source

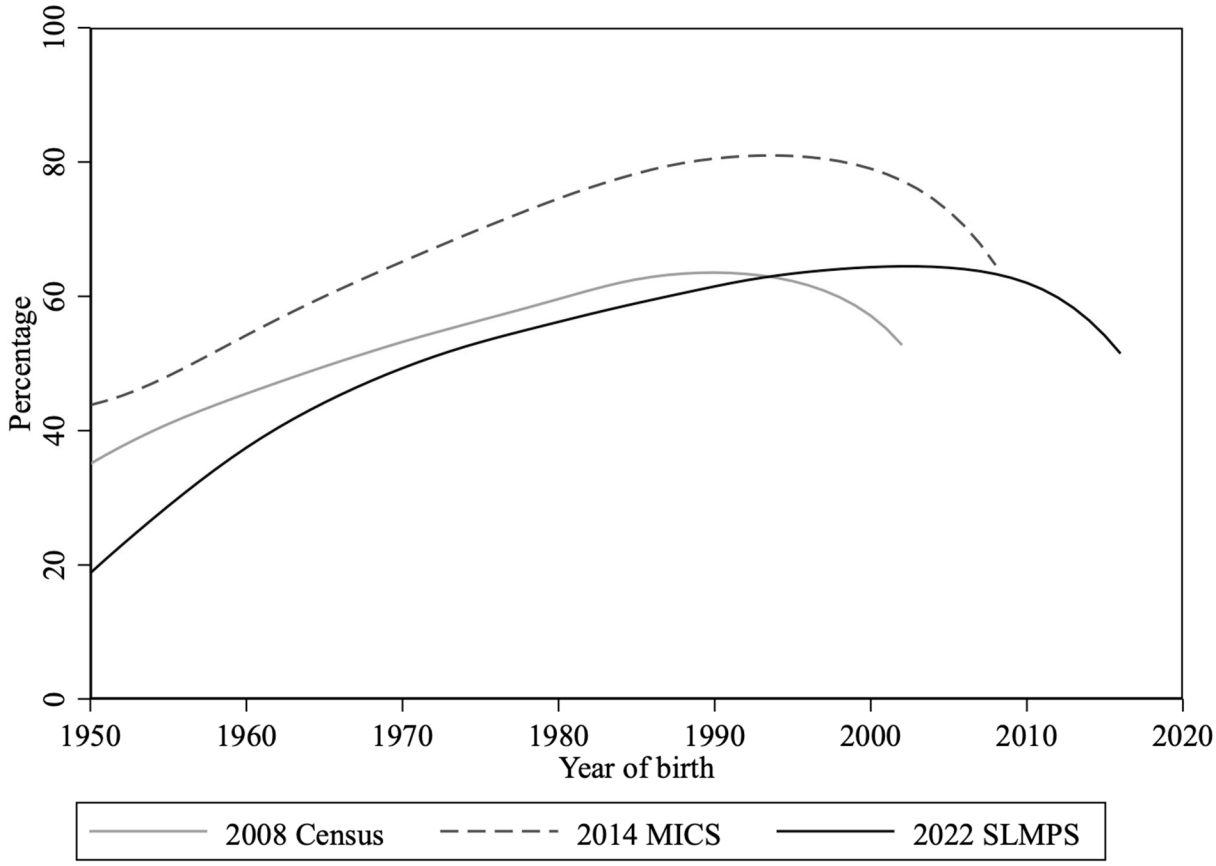


Source: Authors' calculations based on SLMPS 2022, MICS 2014, Sudan Census 2008, and SLFS 2011 report (Ministry of Human Resources Development and Labour 2011.
Note: given very different rates of reporting pre-primary across data sources, we include preschool as the highest level with no school.

The differences between the LFS and MICS, on the one hand, and the SLMPS 2022, on the other could be due to a worsening education situation over time, with more late entry as well as more non-enrollment. There were prolonged disruptions to schooling in Sudan, starting with the 2018 revolution and political instability in 2019, followed by pandemic lockdowns in 2020 and 2021 (UNICEF, 2022). We investigate this by examining the proportion of individuals 6+ who have ever attended school by year of birth in the three sources for which we have microdata, namely the Census, MICS, and SLMPS 2022 (Figure 7). Because we are comparing by year of birth, rates should be the same across data sources, except for the youngest ages in each survey, who may experience delayed entry. As shown in Figure 7, the 2014 MICS shows consistently higher ever attendance rates for all cohorts of birth than either the 2008 Census or the SLMPS 2022, which could indicate something about the MICS sample being more weighted toward advantaged areas. The SLMPS shows a rapid rise in attainment for cohorts born from 1950 to 1970 and then a deceleration for cohorts born from 1970 to 2005. A similar deceleration is observed in the Census

data, but only much later in the MICS data. In the SLMPS and Census data, the maximum share of children ever attending school peaked at slightly above 60%. The reversals observed in each data set for younger cohorts are likely attributable to delays in school entry to beyond age 6.

Figure 7. Ever attended school (percentage of individuals aged 6+), by year of birth and data source, birth years 1950-2016



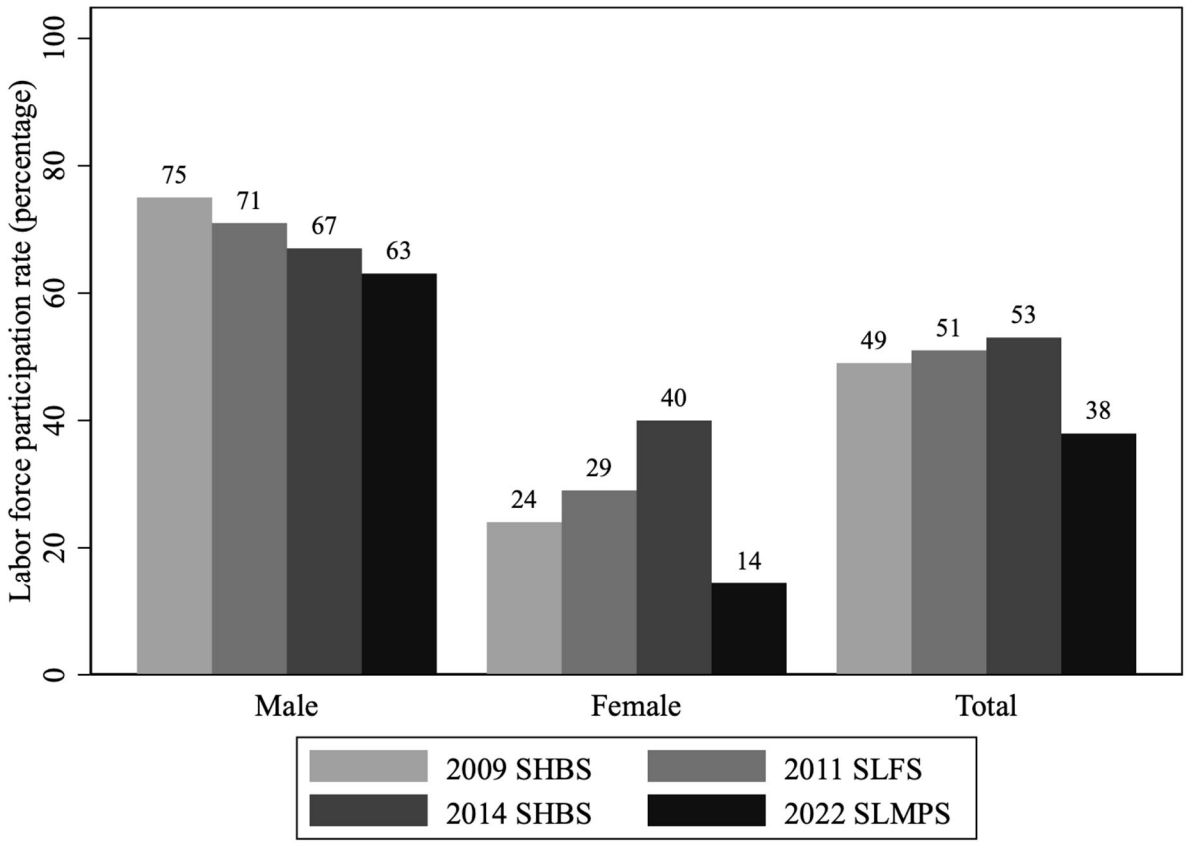
Source: Authors' calculations based on SLMPS 2022, MICS 2014, and Sudan Census 2008.
Notes: Single year of age data not available in SLFS 2011 report. Lowess smoother with bandwidth five.

5.3 Labor market statistics

In Figure 8, we present the labor force participation rate, by sex, estimated across data sources, for ages 15-64. Male labor force participation has been declining steadily over time, from 75% in the 2009 SHBS to 71% in the 2011 SLFS, 67% in the 2014 SHBS, and 63% in the SLMPS 2022. Female participation rose from 2009 (24%) to 2011 (29%) and 2014 (40%), but then declined in the SLMPS to 14%. The net effects for the total population are a slight rise in labor force participation from 49% in 2009 to 53% in 2014 and then 38% in the 2022 SLMPS. Some of the drop in labor force participation may be due to the challenging economic and political situation in Sudan in recent years. There may also be some under-reporting of labor force participation,

particularly for women.[21] The COVID-19 MENA Monitor Surveys, which reached mobile phone owners (a selected sample), found in August 2021 a male labor force participation rate of 68%, and a female labor force participation rate of 45% (Krafft, Nour, & Ebaidalla, 2022).

Figure 8. Labor force participation rate (percentage of individuals aged 15-64), by sex and data source



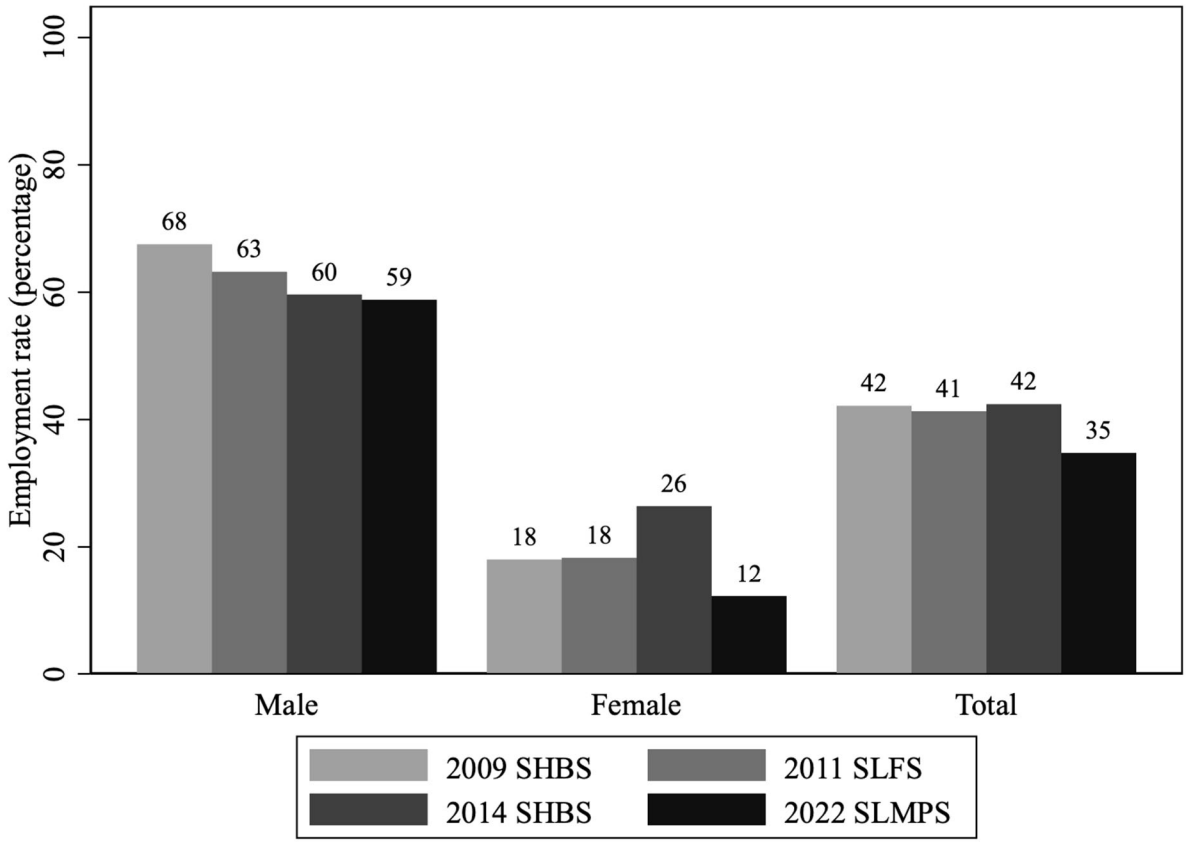
Source: Authors' calculations based on SLMPS 2022, Sudan Census 2008, and report with SHBS 2009, SLFS 2011, and SHBS 2014/15 statistics (Ebaidalla & Nour, 2021).
Notes: Not available in MICS 2014. Market employment, standard, search-required, 7-day reference period definition used for SLMPS 2022.

We turn next to examining the employment-to-population ratio (Figure 9). For men, employment declined from 68% in 2009 to 63% in 2011, 60% in 2014, and 59% in SLMPS 2022. The COVID-19 MENA Monitor captured the same male employment rate of 59% in August 2021 (Krafft, Nour,

[21] Krafft, Assaad, Cortes-Mendosa, and Honzay (2023) demonstrate substantially higher female employment rates using SLMPS 2022 when using measures incorporating participation in household enterprises. This pattern of under detection of women's participation is a common challenge in MENA labor force surveys (Assaad & Krafft, 2023). Assaad, Krafft, and Jamkar (2023) further explore women's work and time use. Assaad, Krafft, and Wahby (2023) further explore labor market dynamics from 2015-2021.

& Ebaidalla, 2022). Women's employment was 18% in both 2009 and 2011, rose to 26% in 2014, and then fell to 12% in SLMPS 2022. This is below the 31% employment rate in the COIVD-19 MENA Monitor (Krafft, Nour, & Ebaidalla, 2022). This disparity may be due to the sample of phone-owning women being much more select than men, or under-reporting of women's employment. Overall employment was thus 41-42% over 2009-2014, but fell to 35% in SLMPS 2022.

Figure 9. Employment-to-population ratio (percentage of individuals aged 15-64), by sex and data source

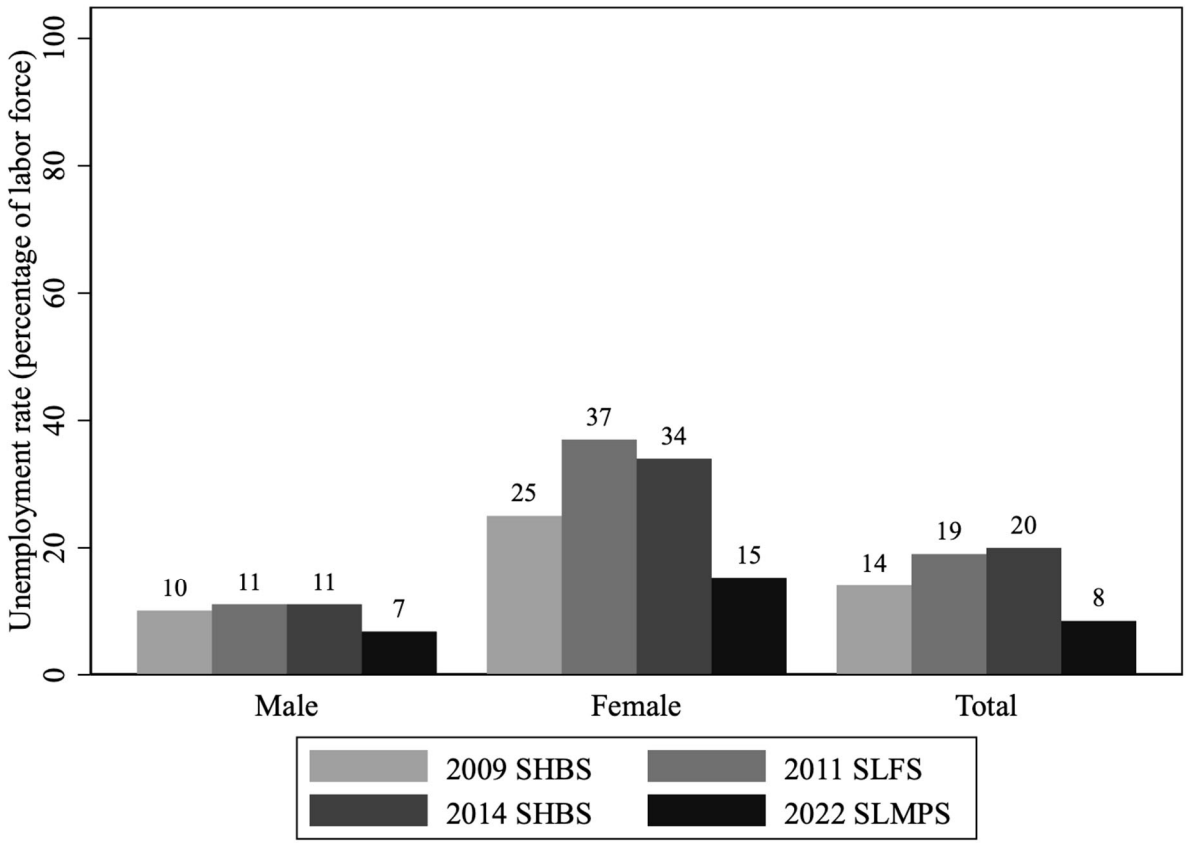


Source: Authors' calculations based on SLMPS 2022, MICS 2014, Sudan Census 2008, and report with SHBS 2009, SLFS 2011, and SHBS 2014/15 statistics (Ebaidalla & Nour, 2021).
Notes: Not available in MICS 2014. Market employment, 7-day reference period definition used for SLMPS 2022.

Figure 10 presents the unemployment rate as a percentage of the labor force. While the unemployment rate was 10-11% in other data sources for men, in SLMPS 2022 it had fallen to 7%. The female unemployment rate was also substantially lower in SLMPS 2022, at 15%, compared to 25-37% in other data sources. Overall, the unemployment rate in the SLMPS 2022 is 8%, substantially lower than the other data sources such as the 2014 SHBS (20%). The differences

in unemployment could be due to decreases in unemployment and particularly search behavior, if discouragement has become an issue in 2022.[22] They could also be due to asking the individual him or herself and thus obtaining a better measure of search.

Figure 10. Unemployment rate (percentage of labor force aged 15-64), by sex and data source



Source: Authors' calculations based on SLMPS 2022, MICS 2014, Sudan Census 2008, and report with SHBS 2009, SLFS 2011, and SHBS 2014/15 statistics (Ebaidalla & Nour, 2021).
Notes: Not available in MICS 2014.

6 Conclusions

6.1 Summary of SLMPS data and comparison

The SLMPS 2022 data provide an updated, nationally representative sample for Sudan, where the most recent preceding survey was the SHBS 2014/15 and the MICS 2014. The data allow for a substantial update on Sudan's economy and society, albeit before conflict broke out in April 2023.

[22] Other work explores the unemployment rate by different definitions of unemployment, including a broader definition without the search requirement (Krafft, Assaad, Cortes-Mendosa, & Honzay, 2023).

ERF and CBS overcame a number of challenges to generate this publicly available data covering a wide variety of topics, with detailed household and individual components.

Demographic indicators from SLMPS 2022 are fairly comparable to those from most recent surveys and census in Sudan. Sudan is a very young country, with 45% of the population under age 15 and 63% under age 25 according to the SLMPS 2022. The SLMPS 2022 age distribution reveals a potential exaggeration of the population 0-4 at the expense of the population 5-9, which is probably a result of how the eligibility criteria for the individual questionnaires (age 5+) interacted with the way enumerators were being remunerated. This may also account for the possible over-reporting of small-sized households and the under-reporting of large-sized households. The main lesson to draw from this is to insist that in future surveys enumerators are paid by data item collected and not by household completed. The distribution by marital status appears to be comparable to the distribution obtained in the 2008 population Census, the only other source from which we could ascertain this distribution for the population as a whole. Early marriage remains a persistent issue for girls in Sudan, with 25% of girls married by age 15 and 40% by 18 in the SLMPS 2022, and little change since the 2008 Census.

With regard to educational attainment, the distribution obtained in the SLMPS is closer to that obtained in the 2008 Population Census than those obtained from the 2011 LFS and the 2014 MICS, at least for the lower levels of educational attainment. Comparing the percentage of the population 6+ who ever attended school by year of birth shows that SLMPS 2022 is indeed much closer to the Census than to MICS 2014, which gives higher rates of ever attendance throughout the distribution of year of birth. This suggests that the MICS 2014 sample may have been skewed toward more advantaged areas in Sudan. All three sources, however, show a decline in ever attendance rates, starting with the 1990 birth cohort in the 2008 Census, the 1995 cohort in the 2014 MICS, and the 2005 cohort in SLMPS 2022. Even with delayed entry into school, these results reveal a deterioration in school attendance rates in Sudan among younger cohorts.

Labor market indicators reveal continued declines in male employment and labor force participation rates, but the figures and trends are comparable across SLMPS 2022 and previous surveys. Male unemployment rates are lower in SLMPS 2022, but this could reveal an increase in discouragement among youth as the economy enters into protracted crisis. Both employment and unemployment rates among women are much lower in the SLMPS 2022 than they are in previous surveys. This could the result of a dramatic loss of employment opportunities for women, as well as a substantial increase in discouragement. The household questionnaire of the SLMPS survey shows that women who declare themselves to be non-employed do engage in a number of household activities that qualify as employment under the current international definitions of employment, but the perceived decline in women's employment ascertained by SLMPS 2022 may still indicate substantial declines in employment opportunities outside the home for women in these times of crisis.[23]

[23] See Krafft, Assaad, Cortes-Mendosa, and Honzay (2023) and Assaad, Krafft, and Jamkar (2023) for further analyses of gender and employment. See also Krafft, Assaad, Kilic, and Moylan (2023) for an investigation of gender norms around women's work and the prioritization of men's employment when jobs are scarce.

6.2 Limitations

The SLMPS 2022 was carried out under very difficult economic and political conditions, which have undoubtedly affected the quality of the data that were collected. While some callbacks were carried out by enumerators at the headquarters of CBS to check on data quality, it was not possible to undertake field quality control of the state teams. The vast distances and poor travel conditions prevented the management of the survey from carrying out many field visits or to check on the progression of the work. In some cases, supervisors and enumerators did not follow the instructions given to them. For instance, supervisors were instructed to enumerate all households located in a PSU irrespective of their number, but many stopped counting once they reached a count of 100, necessitating the use of other less reliable methods to ascertain the number of households in the PSU. In some instances, where PSUs contained both refugee camps and other non-camp populations, supervisors decided to enumerate only the non-camp populations. There were also many instances of PSU deemed to be too unsafe or infeasible to reach, necessitating substitutes be used. We attempted to address these deficiencies by adjusting the weights so that various population groups, such as refugees and IDPs, produce the known totals for these populations, but this does not mean that some biases have not crept into the data.

6.3 Future directions for research

Despite all the challenges encountered in completing the SLMPS 2022, the data will serve as an invaluable resource as the only nationally-representative socio-economic data on Sudan in the turbulent period since 2014. Sudan's economy and labor market have been substantially altered by the fall of the Bashir regime in 2019 and, following the SLMPS 2022, the conflagration of hostilities in April 2023. In fact, the Sudanese economy has been in a state of perpetual crisis since the secession of South Sudan in 2011, which was only exacerbated by the political unrest that broke out in 2018.

Given the continuation of hostilities and the absence of an effective central government, data collection may not be feasible for 2023 and potentially longer. The SLMPS 2022 therefore provides an essential baseline of socio-economic conditions in the country at a crucial time in its history. Given the longitudinal design of the survey, we will have an opportunity to revisit the same households when the political situation stabilizes, to obtain an accurate update on their situation. This kind of data will not only be invaluable for relief and reconstruction efforts, but also creates a unique data set for the study of the effect of conflict on human populations.

The SLMPS 2022 data provide unique insights into a number of issues, given not only the large number of topics covered, but also features such as the over-sampling of the displaced, and GPS coordinates for each PSU allowing for the linking of the SLMPS with geospatial data. Past LMPSs in Egypt, Jordan, and Tunisia have served as the foundations for research on topics such as school-to-work transitions (Alazzawi & Hlasny, 2022; Amer, 2019; Amer & Atallah, 2022; Assaad & Krafft, 2021; Assaad, Krafft, & Salemi, 2023); labor market dynamics (Assaad & Krafft, 2016; Yassine, 2015); gender, family formation and the labor market (Assaad, Krafft, & Selwaness, 2022; Krafft & Assaad, 2020; Selwaness & Krafft, 2021); marriage markets (Assaad & Krafft, 2015a; Assaad, Krafft, & Rolando, 2021); household enterprises (Krafft, 2016; Krafft, Assaad, Rahman, & Cumanzala, 2020; Rizk & Rashed, 2022); the wellbeing of the displaced and the

impacts of displacement on host communities (Al-Hawarin, Assaad, & Elsayed, 2021; Assaad, Ginn, & Saleh, 2023; Elmallakh & Wahba, 2022; Fallah, Krafft, & Wahba, 2019; Krafft, Sieverding, Salemi, & Keo, 2019); inequality (Assaad, Krafft, Roemer, & Salehi-Isfahani, 2018; El Enbaby & Galal, 2020; Said, Galal, & Sami, 2022); labor and social policy (Selwaness & Ehab, 2022; Wahba & Assaad, 2017); education (Assaad & Krafft, 2015b; Krafft & Alawode, 2018); and more. Now that the SLMPS 2022 data are publicly available, all these topics and more can be explored for Sudan.

References

Ahmed, M. O. M., Albatal, D., & Musa, O. A. H. (2022). *Sudan's Youth Bulge: Challenges, Opportunities, and Aspirations*. *Economic Research Forum Working Paper Series No. 1595*. Cairo, Egypt.

Al-Hawarin, I., Assaad, R., & Elsayed, A. (2021). Migration Shocks and Housing: Short-Run Impact of the Syrian Refugee Crisis in Jordan. *Journal of Housing Economics*, *53*(101761).

Alazzawi, S., & Hlasny, V. (2022). Youth Labor Market Vulnerabilities: Evidence from Egypt, Jordan and Tunisia. *International Journal of Manpower*.

Ali, M. E. M., & Ebaidalla, E. M. (2023). *Political Islam and Crony Capitalism in Sudan: A Case Study of "Munazzamat Al-Dawa Al-Islamiyya." Economic Research Forum Working Paper Series No. 1633*. Cairo, Egypt.

Amer, M. (2019). School-to-Work Transition in Jordan, 2010-2016. In C. Krafft & R. Assaad (Eds.), *The Jordanian Labor Market Between Fragility and Resilience*. Oxford, UK: Oxford University Press.

Amer, M., & Atallah, M. (2022). The School to Work Transition and Youth Economic Vulnerability in Egypt. In C. Krafft & R. Assaad (Eds.), *The Egyptian Labor Market: A Focus on Gender and Vulnerability* (pp. 115–155). Oxford, UK: Oxford University Press.

Assaad, R., Ghazouani, S., Krafft, C., & Rolando, D. J. (2016). Introducing the Tunisia Labor Market Panel Survey 2014. *IZA Journal of Labor & Development*, *5*(15), 1–21.

Assaad, R., Ginn, T., & Saleh, M. (2023). Refugees and the Education of Host Populations: Evidence from the Syrian Inflow to Jordan. *Journal of Development Economics*, *164*(103131).

Assaad, R., & Krafft, C. (2013). The Egypt Labor Market Panel Survey: Introducing the 2012 Round. *IZA Journal of Labor & Development*, *2*(8), 1–30.

Assaad, R., & Krafft, C. (2015a). An Empirical Analysis of the Economics of Marriage in Egypt, Morocco, and Tunisia. In C. Monga & J. Y. Lin (Eds.), *The Oxford Handbook of Africa and Economics: Policies and Practices* (Vol. 2, pp. 533–556). Oxford, UK: Oxford University Press.

Assaad, R., & Krafft, C. (2015b). Is Free Basic Education in Egypt a Reality or a Myth? *International Journal of Educational Development*, *45*, 16–30.

Assaad, R., & Krafft, C. (2016). *Labor Market Dynamics and Youth Unemployment in the Middle East and North Africa: Evidence from Egypt, Jordan and Tunisia*. *Economic Research Forum Working Paper Series No. 993*. Cairo, Egypt.

Assaad, R., & Krafft, C. (2021). Excluded Generation: The Growing Challenges of Labor Market Insertion for Egyptian Youth. *Journal of Youth Studies*, *24*(2), 186–212.

Assaad, R., & Krafft, C. (2023). *Connecting People to Projects: A New Approach to Measuring Women's Employment in the Middle East and North Africa*. *World Bank Policy Research Working Paper Series (Forthcoming)*.

Assaad, R., Krafft, C., & Jamkar, V. (2023). *Gender, Work, and Time Use in Sudan*. *Economic Research Forum Working Paper Series (Forthcoming)*. Cairo, Egypt.

Assaad, R., Krafft, C., Roemer, J., & Salehi-Isfahani, D. (2018). Inequality of Opportunity in Wages and Consumption in Egypt. *Review of Income and Wealth*, *64*(S1), S26–S54.

Assaad, R., Krafft, C., & Rolando, D. J. (2021). Evaluating the Impact of Housing Market Liberalization on the Timing of Marriage: Evidence from Egypt. *Population Studies*, *75*(3), 343–361.

Assaad, R., Krafft, C., & Salemi, C. (2023). Socioeconomic Status and the Changing Nature of School-to-Work Transitions in Egypt, Jordan, and Tunisia. *ILR Review*, *76*(4), 697–723.

Assaad, R., Krafft, C., & Selwaness, I. (2022). The Impact of Marriage on Women's Employment in The Middle East and North Africa. *Feminist Economics*, *28*(2), 247–279.

Assaad, R., Krafft, C., & Wahby, S. (2023). *Labor Market Dynamics in Sudan through Political Upheaval and Pandemic*. *Economic Research Forum Working Paper Series (Forthcoming)*. Cairo, Egypt.

Assaad, R., & Marouani, M. A. (2021). Economic Growth and Labour Market Outcomes in North Africa: An Overview of Developments in Algeria, Egypt, Sudan and Tunisia since 2000. In *Regional Report on Jobs and Growth in North Africa 2020* (pp. 23–53). Cairo: ERF and ILO/ADWA'.

Brunette, W., Sudar, S., Sundt, M., Larson, C., Beorse, J., & Anderson, R. (2017). Open Data Kit 2.0: A Services-Based Application Framework for Disconnected Data Management. *MobiSys 2017 - Proceedings of the 15th Annual International Conference on Mobile Systems, Applications, and Services*.

Burgert, C. R., Colston, J., Roy, T., & Zachary, B. (2013). *Geographic Displacement Procedure and Georeferenced Data Release Policy for the Demographicc and Health Surveys*. *DHS Spatial Analysis Reports No. 7*. Calverton, Maryland.

Central Bureau of Statistics (CBS), & UNICEF Sudan. (2016). *Multiple Indicator Cluster Survey 2014 of Sudan, Final Report*.

Central Bureau of Statistics (CBS), & World Bank. (2020). *Socioeconomic Impact of COVID-19 on Sudanese Households*.

Ebaidalla, E. M., & Nour, S. M. (2021). Economic Growth and Labour Market Outcomes in an Agrarian Economy: The Case of Sudan. In *Regional Report on Jobs and Growth in North Africa 2020* (pp. 152–182). International Labour Organization and Economic Research Forum.

El-Kogali, S. E. T., Savrimootoo, T. J., Feda, K., & Chugunov, D. (2021). *The Impact of Twin Shocks on Learning Outcomes and Education Financing in Sudan*. *Economic Research Forum Working Paper Series No. 1507*. Cairo, Egypt.

El Din, S., & Abd El Rhman, D. (2022). *The Pitfalls of the Education System in Sudan: The Challenges of Transition, Transformation, Inclusivity, and Future of Work*. *Economic Research Forum Policy Research Report No. 42*. Cairo, Egypt.

El Enbaby, H., & Galal, R. (2020). Inequality of Opportunity in Earnings and Assets in Egypt. *Applied Economics*, *52*(38), 4146–4161.

Elbadawi, I., Elbashir, A., Osman, A., Elobaid, A. H., Eltahir, E., Alhelo, A., & Suliman, K. M. (2022). *Sudan's Challenges and Opportunities: A Renaissance Project for Sudan: From Poor Agriculture to Agro-Industrial Growth and Sustainable Development*. *Economic Research Forum Policy Research Report No. 40*. Cairo, Egypt.

Elbadawi, I., & Suliman, K. M. (2018). *Macroeconomics of Gold Economy in Sudan*. *Economic Research Forum Working Paper Series No. 1203*. Cairo, Egypt.

Elmallakh, N., & Wahba, J. (2022). Syrian Refugees and the Migration Dynamics of Jordanians: Moving in or Moving Out? *Economic Development and Cultural Change*.

Etang Ndip, A., & Lange, S. (2019). *The Labor Market and Poverty in Sudan*. Washington, DC: World Bank.

Fallah, B., Krafft, C., & Wahba, J. (2019). The Impact of Refugees on Employment and Wages in Jordan. *Journal of Development Economics*, *139*, 203–216.

International Organization for Migration (IOM). (2020). Sudan Displacement Data - IDPs [IOM DTM]. *HDX*. Retrieved December 28, 2022 from https://data.humdata.org/dataset/sudan-displacement-data-idps-iom-dtm

Krafft, C. (2016). *Understanding the Dynamics of Household Enterprises in Egypt: Birth, Death, Growth, and Transformation*. *Economic Research Forum Working Paper Series No. 983*. Cairo, Egypt.

Krafft, C., & Alawode, H. (2018). Inequality of Opportunity in Higher Education in the Middle East and North Africa. *International Journal of Educational Development*, *62*, 234–244.

Krafft, C., & Assaad, R. (2020). Employment's Role in Enabling and Constraining Marriage in the Middle East and North Africa. *Demography*, *57*, 2297–2325.

Krafft, C., & Assaad, R. (2021). Introducing the Jordan Labor Market Panel Survey 2016. *IZA Journal of Development and Migration*, *12*(08), 1–42.

Krafft, C., Assaad, R., Cortes-Mendosa, A., & Honzay, I. (2023). *The Structure of the Labor Force and Employment in Sudan*. *Economic Research Forum Working Paper Series (Forthcoming)*. Cairo, Egypt.

Krafft, C., Assaad, R., Kilic, T., & Moylan, H. (2023). *Women's Economic Empowerment in Sudan: Assets and Agency*. *Economic Research Forum Working Paper Series (Forthcoming)*. Cairo, Egypt.

Krafft, C., Assaad, R., & Rahman, K. W. (2021). Introducing the Egypt Labor Market Panel Survey 2018. *IZA Journal of Development and Migration*, *12*(12), 1–40.

Krafft, C., Assaad, R., Rahman, K. W., & Cumanzala, M. (2020). *How Do Small Formal and Informal Firms in the Arab Republic of Egypt Compare? World Bank Policy Research Working Paper Series No. 9423*. Washington, DC.

Krafft, C., Nour, S. S. O. M., & Ebaidalla, E. M. (2022). Jobs and Growth in North Africa in the

COVID-19 Era The Case of Sudan (2018-21). In *Second Regional Report on Jobs and Growth in North Africa (2018-21): Developments through the COVID-19 Era*. Cairo, Egypt: ERF and ILO.

Krafft, C., Sieverding, M., Salemi, C., & Keo, C. (2019). Syrian Refugees in Jordan: Demographics, Livelihoods, Education, and Health. In C. Krafft & R. Assaad (Eds.), *The Jordanian Labor Market Between Fragility and Resilience* (pp. 141–172). Oxford, UK: Oxford University Press.

Michler, J. D., Josephson, A., Kilic, T., & Murray, S. (2022). Privacy Protection, Measurement Error, and the Integration of Remote Sensing and Socioeconomic Survey Data. *Journal of Development Economics*, *158*, 102927.

Ministry of Human Resources Development and Labour. (2011). *Sudan Labour Force Survey 2011*.

Minnesota Population Center. (2020). *Integrated Public Use Microdata Series, International: Version 7.3 [Dataset]*. Minneapolis, MN.

Nour, S. S. O. M. (2022). *The Impact of COVID-19 on Households and Firms in the MENA Region: The Case of Sudan*. *AFD Research Papers No. 250*.

OAMDI. (2016). Labor Market Panel Surveys (LMPS). Version 2.0 of Licensed Data Files; TLMPS 2014. Egypt: Economic Research Forum (ERF).

OAMDI. (2018). Labor Market Panel Surveys (LMPS). Version 1.1 of Licensed Data Files; JLMPS 2016. Cairo, Egypt: Economic Research Forum.

OAMDI. (2019). Labor Market Panel Surveys (LMPS). Version 2.0 of Licensed Data Files; ELMPS 2018.

OAMDI. (2023). Labor Market Panel Surveys (LMPS). Version 2.0 of Licensed Data Files; SLMPS 2022.

OCHA Regional Office for Sourthern and Eastern Africa (ROSEA). (2020). Sudan - Subnational Administrative Boundaries - Humanitarian Data Exchange. *HDX*.

OCHA Sudan. (2021). UNHCR Refugee Camps - Humanitarian Data Exchange. *HDX*.

Omer, O. E., & Maglad, N. E. (2020). Income Inequality and Middle Class in Sudan: Some Statistical Facts 1996-2011. *Journal of Development and Economic Policies*, *23*(1), 29–49.

Pullum, T. W. (2006). *An Assessment of Age and Date Reporting in the DHS Surveys, 1985-2003*. *DHS Methodological Reports No. 5*. Calverton, Maryland.

Pullum, T. W., & Staveteig, S. (2017). *An Assessment of the Quality and Consistency of Age and Date Reporting in DHS Surveys, 2000-2015*. *DHS Methodological Report No. 19*.

Rizk, R., & Rashed, A. (2022). Trends and Determinants of Women's Entrepreneurship in Egypt. In C. Krafft & R. Assaad (Eds.), *The Egyptian Labor Market: A Focus on Gender and Vulnerability* (pp. 174–196). Oxford, UK: Oxford University Press.

Said, M., Galal, R., & Sami, M. (2022). Evolution of Wages, Inequality, and Social Mobility in

Egypt. In C. Krafft & R. Assaad (Eds.), *The Egyptian Labor Market: A Focus on Gender and Vulnerability* (pp. 89–114). Oxford, UK.

Selwaness, I., & Ehab, M. (2022). Social Protection and Vulnerability in Egypt: A Gendered Analysis. In C. Krafft & R. Assaad (Eds.), *The Egyptian Labor Market: A Focus on Gender and Vulnerability* (pp. 257–294). Oxford, UK: Oxford University Press.

Selwaness, I., & Krafft, C. (2021). The Dynamics of Family Formation and Women's Work: What Facilitates and Hinders Female Employment in the Middle East and North Africa? *Population Research and Policy Review*, *40*(3), 533–587.

Stevens, F. R., Gaughan, A. E., Linard, C., & Tatem, A. J. (2015). Disaggregating Census Data for Population Mapping Using Random Forests with Remotely-Sensed and Ancillary Data. *PLoS ONE*, *10*(2), 1–22.

UNHCR. (2022). Country - Sudan. Retrieved December 30, 2022 from https://data.unhcr.org/en/country/sdn

UNICEF. (2022, May 17). A Safe Return to School for Students in Sudan.

United Nations Department of Economic and Social Affairs Population Division. (2022). World Population Prospects: The 2022 Revision. *Custom Data Acquired via Website*. Retrieved February 14, 2023 from https://population.un.org/wpp/

United Nations Statistical Commission. (2020). *A Recommendation on the Method to Delineat Cities, Urban and Rural Areas for International Statistical Comparisons*.

Wahba, J., & Assaad, R. (2017). Flexible Labor Regulations and Informality in Egypt. *Review of Development Economics*, *21*(4), 962–984.

World Bank. (2023). LSMS+. Retrieved July 21, 2023 from https://www.worldbank.org/en/programs/lsms/initiatives/lsms-plus.

WorldPop. (2020). Sudan - Population Counts - Humanitarian Data Exchange. *HDX*. Retrieved December 28, 2022 from https://data.humdata.org/dataset/worldpop-population-counts-for-sudan

Yassine, C. (2015). Job Accession, Separation and Mobility in the Egyptian Labor Market Over the Past Decade. In R. Assaad & C. Krafft (Eds.), *The Egyptian Labor Market in an Era of Revolution* (pp. 218–240). Oxford, UK: Oxford University Press.