ECONOMIC RESEARCH FORUM

منتدى البحوث الاقتصادية

Working Paper Series

ERF 25TH ANNIVERSARY

INTRODUCING THE EGYPT LABOR MARKET PANEL SURVEY 2018

Caroline Krafft, Ragui Assaad,
and Khandker Wahedur Rahman

Working Paper No. 1360

INTRODUCING THE EGYPT LABOR MARKET PANEL SURVEY 2018

Caroline Krafft,[1] Ragui Assaad,[2] and Khandker Wahedur Rahman[3]

Working Paper No. 1360

October 2019

We acknowledge the general support of the World Bank, the International Labour Organization, Agence Française de Développement, UN Women, and the Arab Fund for Economic and Social Development for the Egypt Labor Market Panel Survey 2018, on which this paper is based.

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

[1] Department of Economics and Political Science, St. Catherine University, 2004 Randolph Ave., St. Paul, MN, 55105. Email: cgkrafft@stkate.edu
[2] Humphrey School of Public Affairs, University of Minnesota, 301 19th Avenue S, Minneapolis, MN, 55455. Email: assaad@umn.edu
[3] Department of Applied Economics, University of Minnesota, 1994 Ruttan Hall, St. Paul, MN, 55108. Email: rahma120@umn.edu

First published in 2019 by
The Economic Research Forum (ERF)
21 Al-Sad Al-Aaly Street
Dokki, Giza
Egypt
www.erf.org.eg

Abstract

This paper introduces the 2018 wave of the Egypt Labor Market Panel Survey (ELMPS), previously fielded in 1998, 2006, and 2012. These publicly-available longitudinal data are nationally representative, tracking both households and individuals over two decades. In this paper, we describe key characteristics of the 2018 wave, including sampling, fielding, and questionnaire design. We examine patterns of attrition and present the construction of weights designed to correct for attrition, as well as to ensure the sample remains nationally representative. We compare the ELMPS data to other Egyptian data sources, namely the 2017 Census and various rounds of the Labor Force Survey. The data provide important new insights into Egypt's labor market, economy, and society.

Keywords: Survey, panel data, public use data, sample weights, labor market, Egypt
JEL Classifications: J00, C81, C83

1. Introduction

The 2018 wave of the Egypt Labor Market Panel Survey (ELMPS) is the fourth wave of a longitudinal survey carried out by the Economic Research Forum (ERF) in cooperation with the Egyptian Central Agency for Public Mobilization and Statistics (CAPMAS).[5] The 2018 wave follows previous waves in 1998, 2006 and 2012.[6] Over its twenty-year history, the ELMPS has become the mainstay of labor market and human development research in Egypt and has served as a model for similar longitudinal surveys in Jordan (2010 and 2016) and Tunisia (2014).

ERF is committed to publicly available microdata. The ELMPS 2018 data, harmonized with previous waves of the ELMPS and the October 1988 Special Round of the Egypt Labor Force Survey, as well as the JLMPS 2010 and 2016 and TLMPS 2014 waves, are publicly available through ERF's Open Access Microdata Initiative (OAMDI) at www.erfdataportal.com, as of October 25, 2019. Data are available to researchers and for academic uses.

As some of the few publicly available, nationally-representative microdata sources on labor markets and human development in the Middle East and North Africa (MENA) region, the ERF labor market panel surveys (LMPSs) have served as the primary data sources for a large number of academic and policy studies. These include studies of education and life course transitions (Assaad, Binzel, & Gadallah, 2010; Assaad & Krafft, 2015a; Biltagy, 2012; Heyne & Gebel, 2016); inequality (Assaad et al. 2018; Belhaj Hassine 2011; Hendy 2013); migration and return migration (Bertoli and Marchetta 2015; Binzel and Assaad 2011; Marchetta and Ferrand 2013; Wahba 2015a); women and work (Assaad & Arntz, 2005; Crandall, Vanderende, Fai, Dodell, & Yount, 2016; El-Hamidi & Said, 2014; Miyata & Yamada, 2017); and work informality and job quality (Assaad & Wahba, 2015; Barsoum, 2015; Radchenko, 2014; Selwaness & Zaki, 2017), to name only a few of the topics studied. The panel design of the survey offers substantial advantages over pooled cross sectional data by allowing for a more accurate assessment of change over time that controls for both observable and unobservable individual and household characteristics. It also allows for a unique perspective on life course transitions by allowing researchers to link life course outcomes like education, marriage, child-bearing, migration, and employment to the individual's household characteristics during childhood and adolescence.

The surveys have supported not only peer-reviewed academic research but also several edited volumes targeted at broad audiences and published by ERF in cooperation with top international publishers (Assaad 2009, 2002; Assaad and Boughzala 2018; Assaad and Krafft 2015b; Krafft and Assaad 2019). They also served as a the primary data source for a number of policy discussions and reports (Krafft and Assaad 2015; World Bank 2013a; b). The LMPSs complement the official quarterly Labor Force Survey (LFS), with a less frequent, but much richer source of longitudinal data covering a far wider variety of topics and covering them in greater depth.

As a longitudinal survey, the ELMPS attempts to track households included in the previous waves and interview all their remaining and new members. The survey also tries to locate any individuals who may have split from these households between waves, and attempts to interview them, as well as any other individuals found in the households they formed or joined. In every wave of the survey, a refresher sample of 2,000-3,000 households is added to maintain the representativeness

[5] ERF received generous support from a number of donors to undertake the 2018 wave of the ELMPS. These include the World Bank, the International Labour Organization, Agence Française de Développement, UN Women, and the Arab Fund for Economic and Social Development.

[6] For additional information on the Tunisia Labor Market Panel Survey (TLMPS) 2014, see Assaad, Krafft, and Ghazouani (2016). For additional information on the Jordan Labor Market Panel Survey (JLMPS) 2010, see Assaad (2014). For more information on JLMPS 2016, see Krafft and Assaad (2018).

of the overall sample and to allow for a more in-depth examination of phenomena of interest. For instance, the refresher sample of the 2012 wave of the ELMPS oversampled high-migration areas to allow for a more detailed examination of the patterns and effects of international migration from Egypt (Assaad and Krafft 2013; Binzel and Assaad 2011; El-Mallakh and Wahba 2017; Wahba 2015b). The 2016 wave of the Jordan Labor Market Panel Survey (JLMPS) added a 3,000 household refresher sample that oversampled neighborhoods with a high proportion of non-Jordanian households to allow for an in-depth examinations of the effects of the Syrian refugee influx on Jordanian society and the situation of migrants and refugees in Jordan (Al-Hawarin, Assaad, and Elsayed 2018; Assaad, Ginn, and Saleh 2018; El-Mallakh and Wahba 2018; Fallah, Krafft, and Wahba 2019; Krafft and Assaad 2018; Malaeb and Wahba 2018). The focus we selected for the 2018 wave of the ELMPS was economic vulnerability among Egypt's poorest communities. Accordingly, we added a refresher sample of 2,000 households that oversampled rural communities that were among the "1,000 poorest villages" of Egypt, as ascertained by the most recent national poverty map available to us.[7]

The final sample included 15,746 households and 61,231 individuals. Of these households, 13,793 households included members from 2012 (10,042 panel and 3,751 split households) and 1,953 were refresher households. Among individuals, 53,040 were in households that included at least one individual interviewed in 2012 (i.e., either panel or split households), while 8,191 were in refresher households. Of the 49,186 individuals included in the 2012 sample, 39,153 (79.6%) were successfully re-interviewed in 2018. Of the 37,140 individuals in the 2006 sample, 22,901 (61.7%) were successfully tracked over three waves. Finally, of the 23,997 individuals included in the 1998 wave, 10,145 (42.3%) were successfully tracked over four waves. We present a detailed discussion of sample attrition patterns in Section 2 and the creation of weights to address such attrition in Section 3. We also discuss the design of the refresher sample and the calculation of the weights for it. In the subsequent section, we compare the (weighted) results of the ELMPS on key demographic and labor market indicators to those of other data sources, namely Egypt's 2017 Census and various rounds of the LFS. First, however, we discuss the design of the questionnaires, sample, and fielding practices.

2. Data collection and sample attrition

2.1. Questionnaires

Each wave of the survey attempts to maintain consistency for the indicators measured in previous waves while adding additional modules and questions to examine new issues or allow more in-depth examination of existing issues. Accordingly, the 2018 wave devoted more attention to the measurement of the instability of employment, focusing in particular on job turnover among casual workers. It also provided more detailed information on health, gender role attitudes, food security, hazardous work, community infrastructure and the cost of housing. It incorporated specific questions on vulnerability, coping strategies and access to social safety net programs.

The 2018 wave has two primary questionnaires, a household questionnaire and an individual questionnaire. The modules in these two questionnaires are listed in Table 1. They are for the most part the same as those in the previous waves of the survey with a few exceptions. The "tracking splits" module in the household questionnaire allows interviewers to ascertain whether the composition of the household has changed since the 2012 wave and inquire about new members present in the household as well as those who may have split to form new households. The "shocks

[7] We used the 2013 poverty map, based on the 2012/2013 HIECS and 2006 Census (with projected population numbers) prepared by CAPMAS in cooperation with the World Bank and UNDP.

and coping module" is also new in the 2018 wave and enquires about both idiosyncratic and community level shocks that the household may have been exposed to, household food security, and coping mechanisms that the household may have used to respond to shocks. The main changes in the individual questionnaire relative to the 2012 wave were a substantial expansion of the health module, a reconfiguration of the labor market history module to better capture past periods of non-employment[8] and the addition of a module on attitudes.

Table 1. Questionnaire modules

Household	Individual
• Statistical Identification	• Statistical Identification
• Tracking Splits	• Residential Mobility
• Individual Roster	• Father's Characteristics
• Housing Information	• Mother's Characteristics
• Current Migrants	• Siblings
• Transfers from Individuals	• Health
• Other Sources of Income	• Education
• Shocks and Coping	• Past Seven Days Subsistence & Domestic Work
• Household Non-Farm Activities	• Employment in the Past Seven Days
• Agriculture Assets: Lands	• Unemployment
• Agriculture Assets: Livestock/Poultry	• Employment in the Past Three Months
• Agriculture Assets: Equipment	• Characteristics of Main Job
• Agricultural Crops	• Secondary Job
• Other Agricultural Income	• Labor Market History
	• Marriage
	• Fertility
	• Female Employment
	• Earnings
	• Earnings in Secondary Job
	• Return Migration
	• Information Technology
	• Savings & Borrowing
	• Attitudes

Source: Authors' construction based on ELMPS 2018 questionnaire

2.2. Data collection

Data collection was tablet-based. The ODK-X (previously named ODK2) tools were used, given their ability to easily handle hierarchical data structures, such as multiple births to individuals within households, and validate across these different structures (Brunette et al. 2017). A training of the trainers was held at CAPMAS in January 2018. Enumerators were trained in April 2018, and data collection began in end-April. The bulk of data collection finished by July 2018, but some teams continued to collect data until November 2018 in order to complete the sample. For fieldwork, teams were governorate-based and composed of a supervisor and from three to five

[8] See Assaad, Krafft, and Yassin (2018) for a discussion of ways to improve the collection of retrospective labor market data.

enumerators (numbers varied depending on the sample size in the governorate). All enumerators were women.

Based on experience from past surveys, we eschewed a distinct enumeration round (a phase of locating and listing individuals interviewed in the previous wave prior to fielding). From our experience in the 2012 wave, we found that the data from the enumeration phase could not simply be used directly in fielding, as individuals may have split between enumeration and fielding, and locating households twice added to cost, time, and attrition. In fielding the ELMPS 2012, 1,680 individuals who were enumerated in 2011 were simply not found in the main fieldwork phase in 2012, and we lacked information on whether they split together, died, moved abroad, or otherwise (Assaad and Krafft 2013). A similar problem arose in Jordan in 2016, where, although we designed the questionnaire to track splits at fielding as well as enumeration, many such splits were not actually fielded, leading to the loss of 616 split households and 647 individuals from the sample (Krafft and Assaad 2018).[9] Instead of implementing a separate enumeration phase, data on the status of all 2012 members was collected as part of the main fieldwork in 2018. We processed the data regularly (multiple times per week) throughout fielding to extract split households that needed to be tracked and added them to the server database for fielding. This dynamic process also allowed us to track repeat splits, i.e., cases where individuals split together, but once located, were found to have further sub-divided into additional new households.

During fieldwork, field quality control took place. We undertook an innovative quality control process where, rather than asking the quality control teams to entirely re-do certain questionnaires, we randomly deleted a certain number of modules from a sample of completed questionnaires and asked the quality control teams to repeat the data collection for those modules. More critical modules had a higher probability of being selected for deletion. For example, we took a 5% deletion rate on section 4.0, which covers assets and housing characteristics, as these questions primarily have yes/no answers and presented fewer data quality challenges. More central modules, such as unemployment detection, were assigned a higher probability of deletion. Because certain variables such as economic activities and occupations were entered in text form and post-coded at the office, this provided a further opportunity for quality control by checking for inconsistencies and possible data entry errors.

2.3. The 2012 sample: Attrition from 2012 to 2018

A key goal of the ELMPS 2018 was to track ELMPS 2012 households and all their members, in order to be able to generate a panel of households and individuals (a panel which now spans 1998, 2006, 2012, and 2018). Two types of attrition could occur during this process. First, entire households who were included in 2012 might be lost in 2018. We refer to the loss of entire households as Type I attrition. If a 2012 household was found, some of its members may have left to form a split household (for example, young persons may marry and leave their natal household to form a new household). Although fieldworkers collected as much information as possible from the original 2012 household (as located in 2018) in order to locate split households, sometimes split households could not be located. We refer to the loss of a split household as Type II attrition. In this section, we describe and then model these two types of attrition. We then use these models to estimate the predicted probability of each type of attrition as a function of 2012 observable characteristics and use these probabilities to create weights to correct for attrition, as detailed in Section 3 below.

[9] The repetition of enumeration in fielding did allow us to determine that 208 individuals had naturally attrited, i.e. died, migrated, or moved to group housing between enumeration and fielding (Krafft and Assaad 2018).

2.3.1. Attrition of entire households (Type I attrition)

Table 2 presents the status of the households from the 2012 wave in 2018. There were 12,060 households fielded in 2012. Of these, 10,042 (83.3%) were located in 2018; they may have been located in a different location or with a different composition, but at least one of their 2012 members was found. Among the 2012 households that were not located in 2018, data was collected from neighbors, if possible, on the disposition of the household. For 188 households, we know that the entire household died out. For another 64 households, we know the entire household left the country.[10] We refer to these two groups of households (252 or 2.1% of the original 2012 households) as having naturally attrited. If we had been drawing a completely new sample, they could not have been included. Thus, these households are excluded from the sample when modelling attrition or calculating the Type I attrition rate.

Table 2. Status of 2012 households in 2018

	Number	Percentage
Initial households	12,060	100.0
Household located	10,042	83.3
Natural attrition	252	2.1
Household died out	188	1.6
Household left the country	64	0.5
Type I attrition	1,766	14.6
Household refused	204	1.7
Household found but not completed	291	2.4
Household not found	1,271	10.5
Type I attrition rate		15.0

Source: Authors' calculations based on ELMPS 2012 and 2018

Type I attrition occurs when a 2012 household was not found or refused to respond in 2018, since (to the best of our knowledge) they should have been included in the sample and were not. Some 204 households refused (1.7% of the initial households) and a further 291 were found but could not be successfully completed (2.4%). A further 1,271 households were not found at all (10.5% of original 2012 households). Together, 1,766 households attrited. After excluding natural attrition, this results in a Type I attrition rate of 15.0%.

This 15.0% Type I attrition rate represents an improvement compared to previous rounds of the ELMPS and the JLMPS. For example, in ELMPS 2006 the Type I attrition rate was 23.5%, in ELMPS 2012 it was 17.3%, and in JLMPS 2016 it was 38.1% (Assaad and Krafft 2013; Barsoum 2009; Krafft and Assaad 2018). Besides the dedication of the fielding team, we attribute this lower rate of attrition to integrating the enumeration step with the main fieldwork, since in both past JLMPS and ELMPSs additional households were lost between enumeration and fielding (Assaad and Krafft 2013; Krafft and Assaad 2018).

We present our model of Type I attrition in Table 3. We present odds ratios from a logit model that includes 2012 characteristics (since only households who did not attrite would have 2018 characteristics) on the 2012 household sample, excluding households that experienced natural attrition. Characteristics include household composition, location (governorate fully interacted with urban/rural), housing type, and 2012 household head characteristics (age, sex, marital status, sex-marital status interaction, education, and labor market status), and household wealth quintile. Importantly, the model indicates the extent to which attrition was random versus related to

[10] Or the sample frame, if they moved to the Frontier governorates.

observable household characteristics (and thus suggests, although does not calculate, the relationship with unobservables as well). Overall, the pseudo R-squared of the model is moderate, 12.2%, indicating that while 2012 household characteristics do predict Type I attrition, they do so only to a modest extent. As a point of comparison, the JLMPS 2016 Type I model pseudo R-squared was 14.7% (Krafft and Assaad 2018).

Table 3. Type I attrition logit model: odds ratios for probability of attrition

Number of household members	
No. of Children 0-5 in HH	0.938
	(0.037)
No. of Children 6-14 in HH	0.859***
	(0.032)
No. of Males 15-64 in HH	0.833***
	(0.040)
No. of Females 15-64 in HH	0.918
	(0.045)
No. of Males 65+ in HH	0.595**
	(0.095)
No. of Females 65+ in HH	0.693**
	(0.090)
Single sex households (mixed sex omit.)	
All male	1.543
	(0.454)
All female	1.407*
	(0.209)
Governorate (Cairo (urban) omit.)	
Alex. # urban	1.811***
	(0.197)
Port-Said # urban	0.310**
	(0.114)
Suez # urban	0.740
	(0.156)
Damietta # urban	0.654
	(0.179)
Damietta # rural	0.517***
	(0.098)
Dakahlia # urban	0.151***
	(0.043)
Dakahlia # rural	0.181***
	(0.039)
Sharkia # urban	0.628*
	(0.114)
Sharkia # rural	0.225***
	(0.047)
Kalyoubia # urban	0.966
	(0.168)

Kalyoubia # rural	0.484***
	(0.086)
Kafr-Elsheikh # urban	0.414***
	(0.101)
Kafr-Elsheikh # rural	0.251***
	(0.054)
Gharbia # urban	0.830
	(0.145)
Gharbia # rural	0.219***
	(0.048)
Menoufia # urban	0.487**
	(0.117)
Menoufia # rural	0.154***
	(0.052)
Behera # urban	0.287***
	(0.074)
Behera # rural	0.312***
	(0.063)
Ismailia # urban	0.658*
	(0.131)
Ismailia # rural	0.223***
	(0.052)
Giza # urban	0.787
	(0.123)
Giza # rural	0.657*
	(0.117)
Beni-Suef # urban	0.592**
	(0.117)
Beni-Suef # rural	0.647*
	(0.118)
Fayoum # urban	0.681
	(0.135)
Fayoum # rural	0.179***
	(0.056)
Menia # urban	0.314***
	(0.079)
Menia # rural	0.264***
	(0.056)
Asyout # urban	0.454***
	(0.089)
Asyout # rural	0.417***
	(0.083)
Suhag # urban	0.556**
	(0.114)
Suhag # rural	0.137***
	(0.036)

Qena # urban	0.561*
	(0.128)
Qena # rural	0.348***
	(0.073)
Aswan # urban	0.404***
	(0.099)
Aswan # rural	0.244***
	(0.072)
Luxur # urban	0.738
	(0.266)
Luxur # rural	0.427
	(0.192)
Housing type (own or benefit omit.)	
Old rent	1.352***
	(0.104)
New rent	2.636***
	(0.272)
Head age (<25 omit.)	
25-34	0.956
	(0.145)
35-44	1.012
	(0.163)
45-54	0.799
	(0.138)
55+	0.642*
	(0.114)
Head sex (male omit.)	
Female	0.907
	(0.160)
Head marital stat. (married omit.)	
Single	0.392**
	(0.141)
Divorced	0.608
	(0.265)
Widow(er)	1.177
	(0.306)
Head marital stat. and sex int.	
Female # Single	2.643
	(1.384)
Female # Divorced	2.210
	(1.140)
Female # Widow(er)	0.778
	(0.246)
Head education (illit. omit.)	
Reads & Writes	1.049
	(0.142)

Less than Intermediate	1.037
	(0.100)
Intermediate	1.075
	(0.100)
Above Intermediate	0.936
	(0.154)
University	1.326*
	(0.146)
Head labor mkt. status (Government employee omit.)	
Out of manpower	1.589**
	(0.257)
Out of labor force	0.974
	(0.120)
Unemployed.	0.848
	(0.191)
Public enterp.	1.133
	(0.169)
Priv. formal wage	1.018
	(0.113)
Priv. inf. reg. wage	1.065
	(0.118)
Priv. irreg. wage	0.872
	(0.104)
Employer	0.811
	(0.096)
Self-emp./UFW ag.	0.605
	(0.181)
Self-emp./UFW non-ag.	0.886
	(0.111)
Wealth quintile (poorest omit.)	
Second	0.896
	(0.090)
Third	0.883
	(0.092)
Fourth	1.045
	(0.111)
Richest	1.176
	(0.136)
Constant	0.564*
	(0.129)
Pseudo R-sq.	0.122
N (households)	11808

Source: Authors' calculations based on ELMPS 2012 and 2018
Notes: *p<0.05; **p<0.01; ***p<0.001

Examining the specific 2012 characteristics which were predictors of attrition, having more children aged (as of 2012) 6-14, having more men aged 15-64 as well as 65+, and having more women aged 65+ significantly decreased the odds of Type I attrition. Single sex, compared to mixed sex households, had significantly higher odds of attrition with similar odds ratios for all men and all women, but only the odds ratio for all women was significant. There were significant geographic differences compared to (all-urban) Cairo; with the exception of a significantly higher odds of attrition in (all-urban) Alexandria, all other governorate-location combinations had lower, often significantly so, odds of attrition. Compared to households that owned their house, or had it as a benefit, in 2012, those who rented had significantly higher odds of attrition, particularly so for new rent (non-rent-controlled) housing contracts.

Turning to household head characteristics, households with heads 55+ were significantly less likely to attrite compared to households with heads less than 25, but there were not significant differences for other age groups. There were not significant differences for female headed households compared to male. Households headed by single individuals were significantly less likely to attrite than those for married individuals, although this was true primarily for the reference category of men; the female and single as well as female and divorced interactions resulted in an increase in attrition, although not significant. Compared to households with illiterate heads, those with most other levels of education had a similar probability of attrition, with the exception of households with university-educated heads, who were more likely to attrite. Most head labor market statuses were not significantly different from the reference, government employee category. The exception was a significantly higher probability of attrition for heads who were out of the manpower basis, meaning heads who were either disabled or 65 years or older.

2.3.2. Attrition of split households (Type II attrition)

Among the 2012 households that were found, some individuals who were present in 2012 may not have been present in 2018. They may have been absent for a variety of reasons, including splitting from their original household to form a new household. Table 4 shows the status of individuals who were present in 2012 households, whose 2012 household was found. Of the 49,186 individuals present in 2012, 42,340 were in households that were found in 2018. However, only 34,325 of those individuals were still in their 2012 households in 2018. Of the 8,015 individuals not in their 2012 households, 2,171 were absent due to natural attrition; 1,497 died, 552 emigrated either outside the governorates covered or outside the country, and 122 moved to group housing and were thus outside the sampling frame. The remaining 5,844 individuals split to form households within Egypt.

When individuals split to form new households, sometimes they did so together, for example, if two brothers left their natal household in Aswan to come work in Cairo and shared a flat. Since our sampling unit was households, we identified the "split households" composed of individuals who split together. The 5,844 individuals who split made up 4,598 split households, meaning that, on average each split involved 1.27 individuals leaving together. Of these split households, 3,751 were found and 847 were not, resulting in a Type II attrition rate of 18.4%. This is a substantial improvement over the 30.3% rate of ELMPS 2012 or the 50.5% rate in JLMPS 2016 (Assaad and Krafft 2013; Krafft and Assaad 2018). Even more so than in the case of Type I attrition, we attribute this reduction in the Type II attrition rate to eliminating enumeration as a separate step in fieldwork, as we previously lost many splits between enumeration and fielding.

Table 4. Status of individuals and split households in 2018, conditional on 2012 household being found

	Number	Percentage
Individuals present in 2012 in original households found in 2018	42,340	100.0
Individuals still in original households in 2018	34,325	81.1
Individuals no longer in original households in 2018	8,015	18.9
Natural attrition through death and migration or leaving sample frame	2,171	5.1
Individuals known to have died	1,497	3.5
Individuals who emigrated or left for a gov. outside scope of survey	552	1.3
Individuals who moved to group housing	122	0.3
Individual splits to form households within Egypt	5,844	13.8
Potential split households (individuals who split together)	4,598	
Split households found	3,751	81.6
Split households not found (attrited)	847	18.4
Type II attrition rate		18.4
Individuals from 2012 in split households found	4,828	82.6
Individuals from 2012 in split households not found (attrited)	1,016	17.4
Total individuals from 2012 who were found	39,153	

Source: Authors' calculations based on ELMPS 2012 and 2018

We present in Table 5 odds ratios from a logit model of Type II attrition, estimated for the sample of 4,598 split households. The pseudo R-squared for the Type II model, at 22.3%, is higher than for the Type I attrition model, indicating the loss of split households is less random than households overall. The determinants we include are characteristics of the 2012 households as well as those of the "head" (or most senior member) of the split household. While most "split households" were made up of one individual, the presence of additional household members of all ages and both sexes predicted a significantly lower probability of attrition. Compared to those that were in (urban) Cairo, split households were significantly less likely to attrite everywhere else. As in the case of Type I attrition, splits from households living in "new rent" housing, compared to owned housing, were significantly more likely to attrite. Compared to heads who were under 15 in 2012 (so would have been under 21 in 2018), those 15-24 in 2012 were significantly less likely to attrite. Other age groups were not significantly more likely to attrite than the reference category. Female-headed households (mostly single women) were significantly less likely to attrite than male-headed ones. Those split households whose heads were single or divorced/widowed were significantly less likely to attrite than those who were married in 2012. There were no significant interactions between the sex of the split household head and their marital status. There were also no significant differences in attrition by split household head education or labor market status in 2012. There were significant wealth interactions, splits from households that were in the fourth and especially richest quintile were more likely to attrite. Overall, Type II attrition appears to be driven more by demographic and geographic factors than by education or labor market characteristics.

Table 5. Type II attrition logit model: odds ratios for probability of attrition

Number of household members	
No. of Children 0-5 in HH	0.088***
	(0.047)
No. of Children 6-14 in HH	0.002***
	(0.002)
No. of Males 15-64 in HH	0.009***
	(0.007)
No. of Females 15-64 in HH	0.032***
	(0.021)
No. of Males 65+ in HH	0.038**
	(0.047)
No. of Females 65+ in HH	0.007***
	(0.009)
Governorate (Cairo (urban) omit.)	
Alex. # urban	0.210***
	(0.053)
Port-Said # urban	0.012***
	(0.012)
Suez # urban	0.028***
	(0.015)
Damietta # urban	0.050***
	(0.032)
Damietta # rural	0.083***
	(0.034)
Dakahlia # urban	0.069***
	(0.027)
Dakahlia # rural	0.053***
	(0.016)
Sharkia # urban	0.122***
	(0.045)
Sharkia # rural	0.100***
	(0.028)
Kalyoubia # urban	0.300**
	(0.111)
Kalyoubia # rural	0.094***
	(0.032)
Kafr-Elsheikh # urban	0.049***
	(0.026)
Kafr-Elsheikh # rural	0.050***
	(0.017)
Gharbia # urban	0.172***
	(0.058)
Gharbia # rural	0.098***

	(0.028)
Menoufia # urban	0.072***
	(0.032)
Menoufia # rural	0.082***
	(0.033)
Behera # urban	0.086***
	(0.036)
Behera # rural	0.083***
	(0.026)
Ismailia # urban	0.123***
	(0.043)
Ismailia # rural	0.111***
	(0.035)
Giza # urban	0.245***
	(0.071)
Giza # rural	0.348***
	(0.102)
Beni-Suef # urban	0.086***
	(0.034)
Beni-Suef # rural	0.161***
	(0.044)
Fayoum # urban	0.015***
	(0.011)
Fayoum # rural	0.112***
	(0.034)
Menia # urban	0.072***
	(0.028)
Menia # rural	0.061***
	(0.017)
Asyout # urban	0.095***
	(0.030)
Asyout # rural	0.107***
	(0.029)
Suhag # urban	0.050***
	(0.021)
Suhag # rural	0.035***
	(0.011)
Qena # urban	0.106***
	(0.040)
Qena # rural	0.086***
	(0.023)
Aswan # urban	0.071***
	(0.027)
Aswan # rural	0.074***
	(0.026)
Luxor # urban	0.199*

	(0.146)
Luxor # rural	0.125***
	(0.069)
Housing type (own or benefit omit.)	
Old rent	1.036
	(0.159)
New rent	3.021**
	(1.190)
Head age (<15 omit.)	
15-24	0.075*
	(0.099)
25-34	0.079
	(0.104)
35-44	0.115
	(0.151)
45+	0.127
	(0.171)
Head sex (male omit.)	
Female	0.269*
	(0.147)
Head marital stat. (married omit.)	
Single	0.122***
	(0.033)
Divorced/Widow(er)	0.293*
	(0.175)
Head marital stat. and sex int.	
Female # Single	1.640
	(0.526)
Female # Divorced/Widow(er)	0.837
	(0.587)
Head education (illit. omit.)	
Reads & Writes	1.355
	(0.350)
Less than Intermediate	1.131
	(0.193)
Intermediate	1.142
	(0.201)
Above Intermediate	1.811
	(0.590)
University	1.498
	(0.334)
Head labor mkt. status (Government employee omit.)	
Out of manpower	1.042
	(0.832)
Out of labor force	1.546

	(0.377)
Unemployed.	1.668
	(0.444)
Public enterp.	1.028
	(0.498)
Priv. formal wage	1.755
	(0.526)
Priv. inf. reg. wage	1.618
	(0.432)
Priv. irreg. wage	1.524
	(0.432)
Employer	1.612
	(0.645)
Self-emp./UFW ag.	0.677
	(0.274)
Self-emp./UFW non-ag.	1.203
	(0.408)
Wealth quintile (poorest omit.)	
Second	1.278
	(0.174)
Third	1.122
	(0.163)
Fourth	1.436*
	(0.223)
Richest	2.125***
	(0.363)
Constant	7218.760***
	(8790.485)
Pseudo R-sq.	0.223
N (households)	4598

Source: Authors' calculations based on ELMPS 2012 and 2018

Notes: *p<0.05; **p<0.01; ***p<0.001. For 46 split households, failures were completely determined.

2.4. Panel sample

The ELMPS waves track individuals over two decades, including a large number of individuals who were in multiple waves. Table 6 presents the panel, with an observation being a unique individual, and shows the number and percentage of individuals in different combinations of waves, e.g. in 1998 & 2006 & 2012 & 2018 or in 2018 only. There were 10,145 individuals tracked from 1998 all the way through 2019 (out of 23,997 in the 1998 wave). Since refresher samples were added in each wave, the number of individuals in multiple waves increases over time. For instance, there were 12,756 individuals in 2006 & 2012 & 2018 (but not in 1998), who can be analyzed as a 2006-2018 panel along with the 10,145 individuals in the 1998-2018 panel.

Table 6. Panel sample, 1998-2018

Attrition from 1998 to 2018	Number	Percentage
In 1998 & 2006 & 2012	3,073	3.6
In 1998 & 2006	4,143	4.8
In 1998 only	6,636	7.7
In 2006 & 2012	2,796	3.2
In 2006 only	4,227	4.9
In 2012 only	4,164	4.8
In 1998 & 2006 & 2012 & 2018	10,145	11.8
In 2006 & 2012 & 2018	12,756	14.8
In 2012 & 2018	16,252	18.8
In 2018 only	22,078	25.6
Total	86,270	100.0

Source: Author's calculations based on ELMPS 1998-2018

2.5. The 2018 refresher sample

In addition to tracking the households that were present in the 2012 round into 2018, the ELMPS 2018 added a refresher sample. The refresher sample was focused on over-sampling the 1,000 poorest villages in Egypt, as a key theme of the 2018 wave was economic vulnerability. A poverty map created by CAPMAS in 2013 with the assistance of the World Bank and UNDP was used to identify the 1,000 poorest villages.[11] A stratified random sample was used, with strata defined by governorate (the first level of administrative geography in Egypt), urban/rural location, and within rural areas, poor (i.e. among the 1,000 poorest villages) versus non-poor. Table 7 presents the distribution of the refresher sample clusters across these various strata. It is important to note that the 1,000 poorest villages were all in rural areas and only in some governorates. We sampled the poorest villages from all governorates that had them. Likewise, there are no rural areas in Cairo, Alexandria, Port Said, or Suez. Since the ELMPS sample frame excludes the frontier governorates, the refresher sample likewise excludes them. Overall, we planned for a refresher sample of 2,000 households distributed over 200 geographic clusters (primary sampling units or PSUs). Of those, we allocated 60 clusters to urban areas, 100 to poor rural areas, and 40 to non-poor rural areas. Within strata, sampling was carried out according to the principle of probability proportional to size.

[11] The 2013 poverty map was never formally published, but was made available to the ELMPS team through CAPMAS.

Table 7. Refresher sample clusters by governorate, urban/rural location, and poor vs. non-poor

	Urban			Rural			Total		
Governorate	**Poor**	**Non-poor**	**Total**	**Poor**	**Non-poor**	**Total**	**Poor**	**Non-poor**	**Total**
Cairo		10	10					10	10
Alexandria		7	7					7	7
Port Said		1	1					1	1
Suez		1	1					1	1
Damietta		1	1		1	1		2	2
Dakhalia		5	5	4	3	7	4	8	12
Sharkia		4	4		3	3		7	7
Kalyoubia		4	4	4	1	5	4	5	9
Kafr-Elsheikh		1	1		1	1		2	2
Gharbia		2	2		2	2		4	4
Menoufia		1	1	4	2	6	4	3	7
Behera		2	2	14	2	16	14	4	18
Ismailia		1	1		1	1		2	2
Giza		7	7	8	2	10	8	9	17
Beni-Suef		1	1	6	1	7	6	2	8
Fayoum		1	1	1	2	3	1	3	4
Menia		1	1	2	3	5	2	4	6
Asyout		3	3	19	4	23	19	7	26
Suhag		2	2	19	4	23	19	6	25
Qena		2	2	15	5	20	15	7	22
Aswan		2	2	3	2	5	3	4	7
Luxor		1	1	1	1	2	1	2	3
Total		60	60	100	40	140	100	100	200

Source: Authors' calculations based on ELMPS 2018

Each cluster was designed to sample ten households. A list of 12 households per cluster was generated in case some households could not be located and ten were to be collected. There was some non-response. Overall, of a planned 2,000 refresher households, 1,953 were sampled. The non-response rate ranged from a low of 20% (only two of ten households collected, which happened in two clusters), followed by three clusters with a rate of 60%, five at 70%, 14 at 80%, 26 at 90%, 117 at 100%, and 16 clusters at 110% along with 17 clusters at 120%, where the "additional" households were fielded beyond the intended ten. The non-response rate was slightly higher (response rate of 94.2%) in urban than rural areas, but no other non-response patterns were discernable.

3. Calculation of Attrition and Sample Weights

In this section, we discuss how the attrition modeling and sampling were incorporated into creating weights for the sample, to ensure that the data remained nationally representative. The starting point for weights for the 2018 sample was their household weights in 2012. These weights were adjusted to account for Type I attrition. The weights of split households were derived from their 2012 households, but account for Type II attrition as well as whether any other households may have merged (i.e., share adjustment for component households). These weights are the panel weights, which were brought together with the weights for the refresher sample (which were based on refresher sample design) and then expanded to the national population. This section details all the specifics of these calculations.[12]

3.1. Weights for Panel and Split Households

The main idea behind the weights for the panel component of the sample is to generate weights that weight "up" remaining households whose observable characteristics were similar to those households that attrited. From the Type I attrition model, we estimate $\Pr(A_h)$, the probability of Type I attrition (attrition of the entire household) for 2012 household h. There may also be splits from that household; refer to a split household as s. For such households we calculate:

$$\begin{aligned} \Pr(A_{hs}) &= 1\text{-}\Pr(h \text{ found } \& \ s \text{ found}) \\ &= 1\text{-}\Pr(h \text{ found})*\Pr(s \text{ found} \mid h \text{ found}) \end{aligned} \tag{1}$$

To adjust for attrition among the 2012 households that were found in 2018, we compute a response adjustment factor, r_h, for original households:

$$r_h = \frac{1}{1 - \Pr(A_h)} \tag{2}$$

and r_{hs} for split households:

$$r_{hs} = \frac{1}{[1 - \Pr(A_{hs})] * c_s} \tag{3}$$

where c_s is the number of component households. Component households are the number of originating households in the population (not the sample) that contribute individuals to the new, split household. For example, when two individuals leave their natal households to get married, they come from two households that existed in 2012; this would be a case of two component households. Essentially, this household has "double" the probability of selection and this must be accounted for in weighting by using a share correction (dividing by component households) (Himelein 2014). If a split contains only members from a 2012 household or born since 2012, there is only one component household.[13]

The calculation of the 2018 weights for 2012 households accounts for both the household weight in 2012 and attrition. Denote the 2012 weight (whether cross section or panel) as w_{2012}. We therefore calculate 2018 panel weights as $w_{2018} = w_{2012} * r_{h(s)}$. We normalize the weights (dividing by the mean weight) to have a mean of one. When we subsequently normalize the refresher weights, this allows 2012 and refresher households to contribute equally to the sample, on average.

3.2. Weights for refresher sample

In addition to the 2012 sample, the 2018 wave included a refresher sample over-sampling the 1,000 poorest villages. This sub-section describes the creation of the refresher sample weights. These weights can stand alone, and thus serve as further validation of the full sample, as well as feed into the 2018 combined sample weights. Since our sampling strategy was a two-stage stratified cluster sample, sampling households, weights were initially calculated on the household level.

[12] Notation is as in Krafft & Assaad (2018).
[13] When calculating panel weights rather than wave-specific cross-sectional weights the share correction was omitted.

Each stratum in governorate *g*, urban/rural location *l,* and poor vs non-poor area *s* is made up of a certain number of clusters $P_{g,l,s}$ as indicated in Table 7. Given that each cluster p in the refresher sample is designed to have 10 households, the total number of households per stratum is given by:

$$h_{g,l,s} = \sum_{p=1}^{P_{g,l,s}} 10$$

To account for deviations from the planned 10 households per cluster, we accounted for the cluster level non-response rate. Households start with a weight, w_p, such that they represent the planned number of households per cluster, based on the observed number of households in the cluster, m_p:

$$w_p = \frac{10}{m_p} \tag{4}$$

We have population counts from the April 2017 Egyptian Population Census for each of our strata, which we use to weight our sample. Denote the census count of population in a given stratum as $c_{g,l,s}$. We therefore calculate household weights as:

$$w_{p,g,l,s} = w_p \ \frac{c_{g,l,s}}{h_{g,l,s}} \tag{5}$$

Using the Population Census to ex-post weight households necessarily generates expansion weights that yield the same number of households as in the population. However, they do not necessarily yield the same number of individuals. We also allow for individual non-response (refusal) as part of the individual consent process (71 individuals in the refresher and 490 in the panel refused or could not be completed). Individuals would still be listed in the roster, collected at the household level, but not have the individual interview data. To account for potentially age-group (e) and sex-specific (x) non-response we calculated an age and sex-specific non-response rate of $r_{x,e}$. We then adjusted the household weight by this non-response to get an individual weight, as:

$$w_{p,g,l,s,x,e} = \frac{w_{p,g,l,s}}{1 - r_{x,e}} \tag{6}$$

3.3. Combined sample weights

To combine the refresher and panel samples into a single panel with a unified set of weights, first we divided the weights in each group (refresher and from 2012) by their means to have a mean of one. The process for the combined weights follows a very similar structure to the latter part of the refresher weight construction. If $\widetilde{w}_{g,l}$ is the normalized weight for a household (either refresher or panel) in a particular governorate-urban/rural location combination, we then calculate the household combined sample weight as:

$$w_{g,l} = \widetilde{w}_{g,l} \frac{c_{g,l}}{\sum \widetilde{w}_{g,l}} \tag{7}$$

Where $c_{g,l}$ is the number of households in the 2017 Census for that same governorate-urban/rural location combination. Again, we allow for sex and age group specific non-response ($r_{x,e}$) and adjust to create individual weights as:

$$w_{g,l,x,e} = \frac{w_{g,l}}{1 - r_{x,e}} \tag{8}$$

4. Comparisons with other data sources for Egypt

In this section we compare several demographic and labor market indicators from the ELMPS waves to other data sources and evaluate the representativeness of the ELMPS data. We use the LFS rounds from 2001-2018 and Egypt Population Census 2017[14] as comparators. In examining age distributions, we also use Egypt's 2006 Population Census to examine changes over time (Minnesota Population Center 2018). We obtained the key labor market statistics of LFS 2001-2007 from the International Labor Organization's ILOSTAT (ILO 2019) and analyzed the microdata for 2008-2017.[15,16] We used ILOSTAT for years 2001-2007 for two reasons, first because calculations from the microdata were inconsistent with the published figures, and second because this time period covered ages 15-64, the focus of our microdata analysis, but subsequently were 15+. Key labor market indicators were missing for years 2003 and 2004 and unemployment and thus employment were missing in 2005. In addition, we discuss the published LFS numbers from the four quarters of 2018 and first two quarters of 2019 (CAPMAS 2019a; b), since the microdata were not yet publicly available.[17]

The different surveys have slightly different universes. While the ELMPS covered all of Egypt except the Frontier governorates, the Census and LFS include the Frontier governorates as well. The Frontier governorates include Red Sea, El-Wadi El-Gedid, Matrouh and North and South Sinai. Since these governorates had 1.7% of Egypt's population in the 2017 Census, we do not expect this to substantially alter comparability (CAPMAS 2019c).

We first undertake comparisons in terms of demographic characteristics, such as age, household size, marital status, and education. When we focus on demographic outcomes, we specifically compare ELMPS 2018 with LFS 2014, LFS 2017, and Census 2017.[18] We include two rounds of the LFS because, as we will show below, we observed a substantial difference in demographic outcomes starting in LFS 2015 that continued through LFS 2017.

We make use of the LFS to assess and compare labor market outcomes, specifically employment, unemployment, and labor force participation rates.[19] We analyze these outcomes by sex for ages 15-64. We follow the "work for pay or profit" definition of the 19th International Conference of Labor Statisticians (ILO 2013) to calculate these statistics. This is also referred as the market definition of employment. Under this definition, if someone worked for at least one hour in the past week as either a wage worker, employer, self-employed worker, or unpaid family worker,

---

[14] Based on the published Census report (CAPMAS 2017).
[15] Harmonized LFSs are available from ERF's OAMDI (OAMDI 2019).
[16] Employment rates were missing from the ILOSTAT. We, however, were able to calculate the employment rate (e) from the labor force participation rate (l) and unemployment rate (u) using the following formula: $e=l(1-u)$.
[17] The labor force participation rate and employment rate for year 2018 are for 15+ population. While the numerator of unemployment rate is total size of unemployed population between ages 15 to 64, the denominator is size of labor force for 15+ population. We therefore do not include these statistics in our graphs.
[18] For additional comparisons of previous waves of the ELMPS, see Assaad and Krafft (2013).
[19] We do not compare labor market outcomes with the Census, since previous research has demonstrated that labor market outcomes are poorly measured in the Census (Assaad 1997).

producing goods for market, then that individual was considered to be employed. On the other hand, if someone was willing to work, had searched for work actively in the preceding three months, was available to start working in the following two weeks but did not even work for an hour over the past one week, then that person was considered to be unemployed. A person was considered to be in the labor force if she or he was employed or unemployed.

We also present the 95 percent confidence intervals (CIs) of each labor market indicator calculated from the ELMPS. The CIs take into account the contemporaneous cluster and strata of the ELMPS. Comparing the CIs to other statistics allows us to assess the degree to which any differences in results are likely to be purely due to sampling variability. In addition, we also separately add the same statistics for the ELMPS 2018 refresher sample. Since the ELMPS 2018 refresher is an independent sample, covering the same time frame and using the same questionnaire, fielding practices, and enumerators, this is a particularly valuable test of whether any differences in results are due to the panel sample (and potentially attrition).

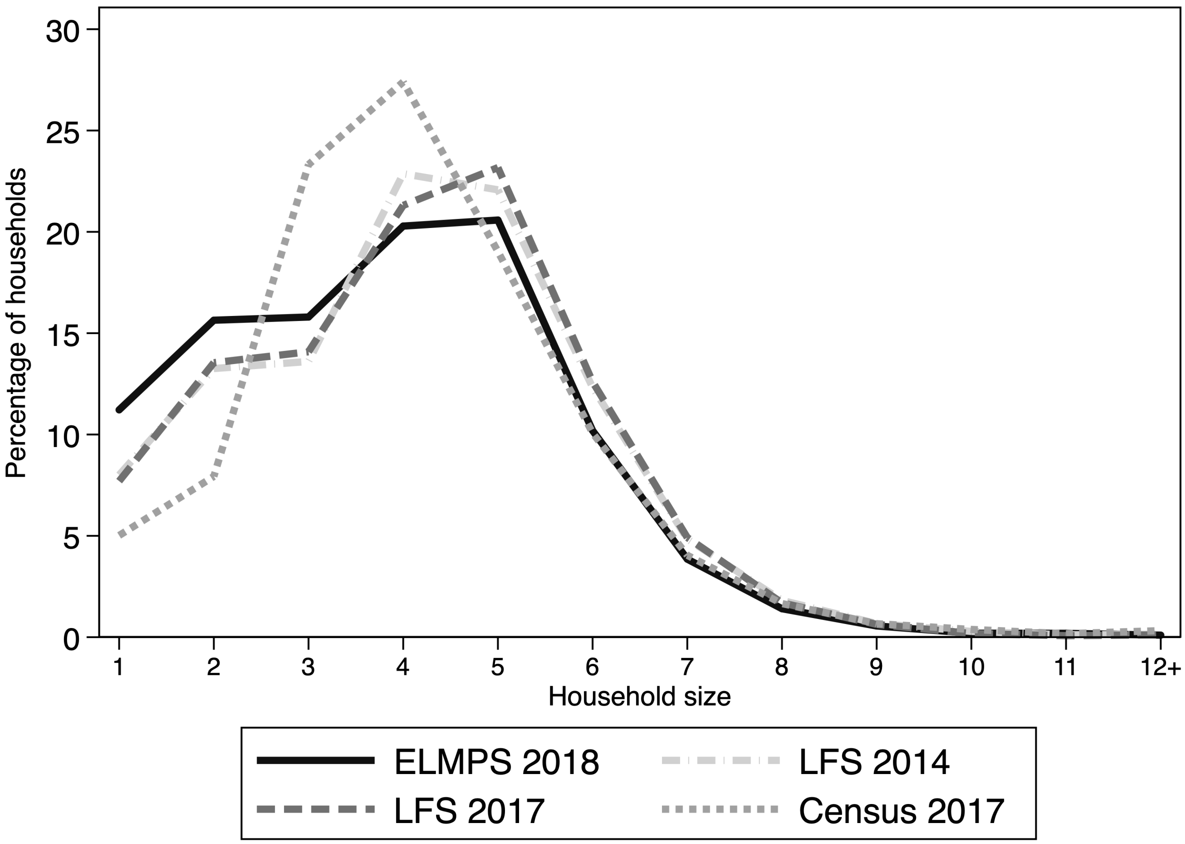
4.1. Demographic comparisons

In this section, we compare the demographics of ELMPS 2018 to other sources. We start with the size of the population, and then discuss its household size, age structure, marital status, and educational attainment.

The 2017 Population Census enumerated 23.1 million households (within the governorates covered by ELMPS), which contained 93.2 million individuals (48.1 million men and 45.1 million women). Given the design of the weights, both the overall ELMPS and refresher sample generate the exact same number of households. From the whole ELMPS sample, when we calculated the expansion-weighted population numbers, after adjusting for individual non-response, we found 88.6 million individuals (44.0 million men and 44.6 million women). When we calculated the refresher sample expansion-weighted population numbers, after adjusting for individual non-response, we found 91.4 million individuals (45.8 million women and 45.7 million men). Both the refresher sample and the full sample found fewer individuals, and especially fewer men, than the Census.

One of the reasons for different numbers of individuals across data sources may be different definitions of households implemented in the different data sources, an issue observed in ELMPS 2012 as well (Assaad and Krafft 2013). As Figure 1 shows, ELMPS 2018 sampled more households of one person (11%) than the Census (5%), with the LFSs between the two (8%). Likewise, while the ELMPS found 16% of households had one individual, the Census found 8% of households did so, and the LFSs 13-14%. Correspondingly, the Census found 23% of households had three people, while the ELMPS found 16% and LFSs 13-14%. Likewise, the Census found more households (27%) were four persons than the ELMPS (20%) or LFSs (21-23%). The sampling of households with 5 members or above in ELMPS 2018 was almost identical to that of the Census.

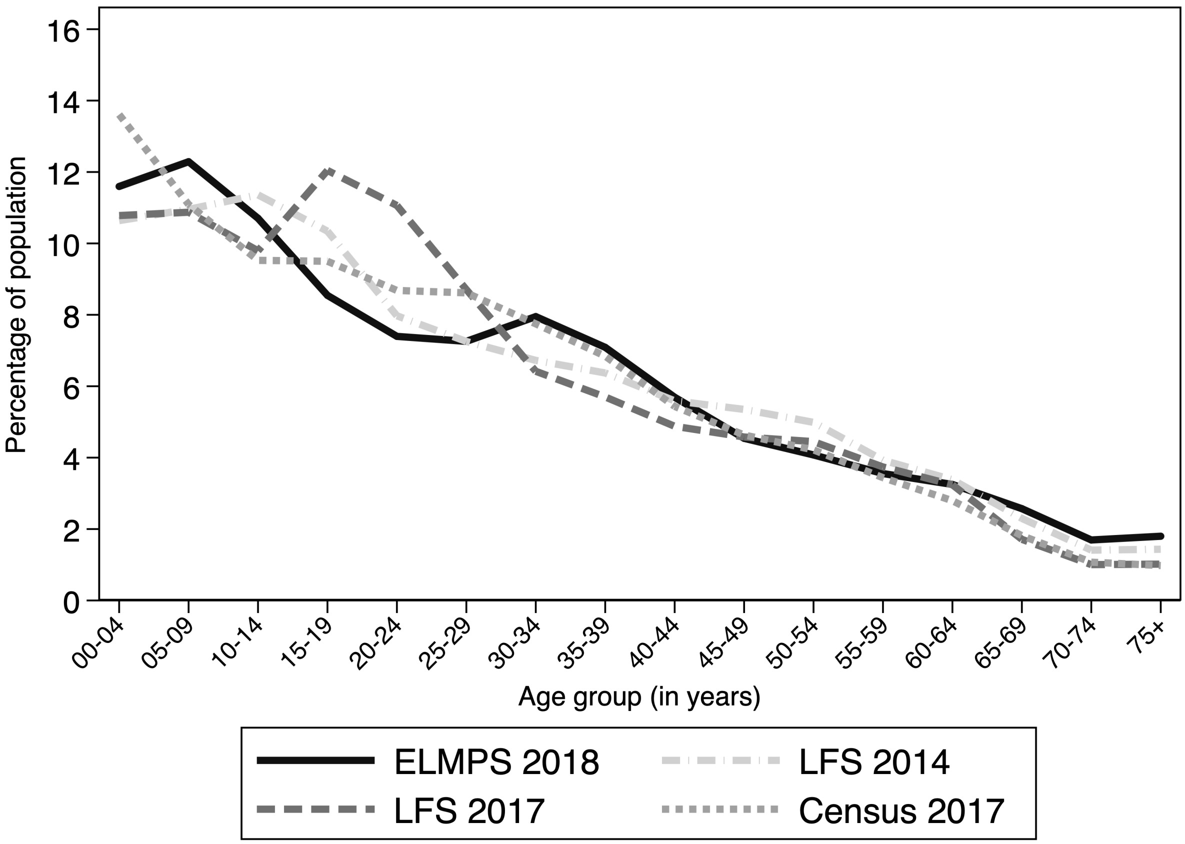
Figure 1. Comparison of household size (percentage of households), by 2017/18 data sources



Source: Authors' calculations from LFS 2014 and 2017, ELMPS 2018, and Egypt Population Census 2017 Table 1-9 (CAPMAS 2017).

Egypt's population is young and has a well-known "youth bulge," whereby a decline in mortality followed by a delayed decline in fertility led to a particularly large youth population (Assaad and Roudi-Fahimi 2007; Krafft and Assaad 2014; Robinson and El-Zanaty 2006). The youth bulge were young adults by 2018 (Krafft, Assaad, and Keo 2019). Figure 2 presents the distribution of population by five-year age groups for the ELMPS 2018, Census 2017, and LFS 2014 and 2017 rounds. While for ages starting around 40, the data had similar age distributions, at younger ages they diverged substantially. The ELMPS 2018 and Census 2017 had a similar share of 30-39 year-olds, but the LFS 2014 and 2017 had slightly fewer. The LFS 2017 was strikingly different than other sources at ages 15-24, showing a large youth bulge at those ages, which was not reflected in the LFS 2014. The LFS 2014 was somewhat more similar to the ELMPS 2018, with some more 20-24 year-olds and somewhat fewer children under 10. The 2017 Census showed less of a "trough" at ages 20-24 than the ELMPS, slightly fewer children 5-9, and more 0-4. These different age compositions are important to keep in mind when considering labor statistics that are strongly related to age, for instance, unemployment rates.

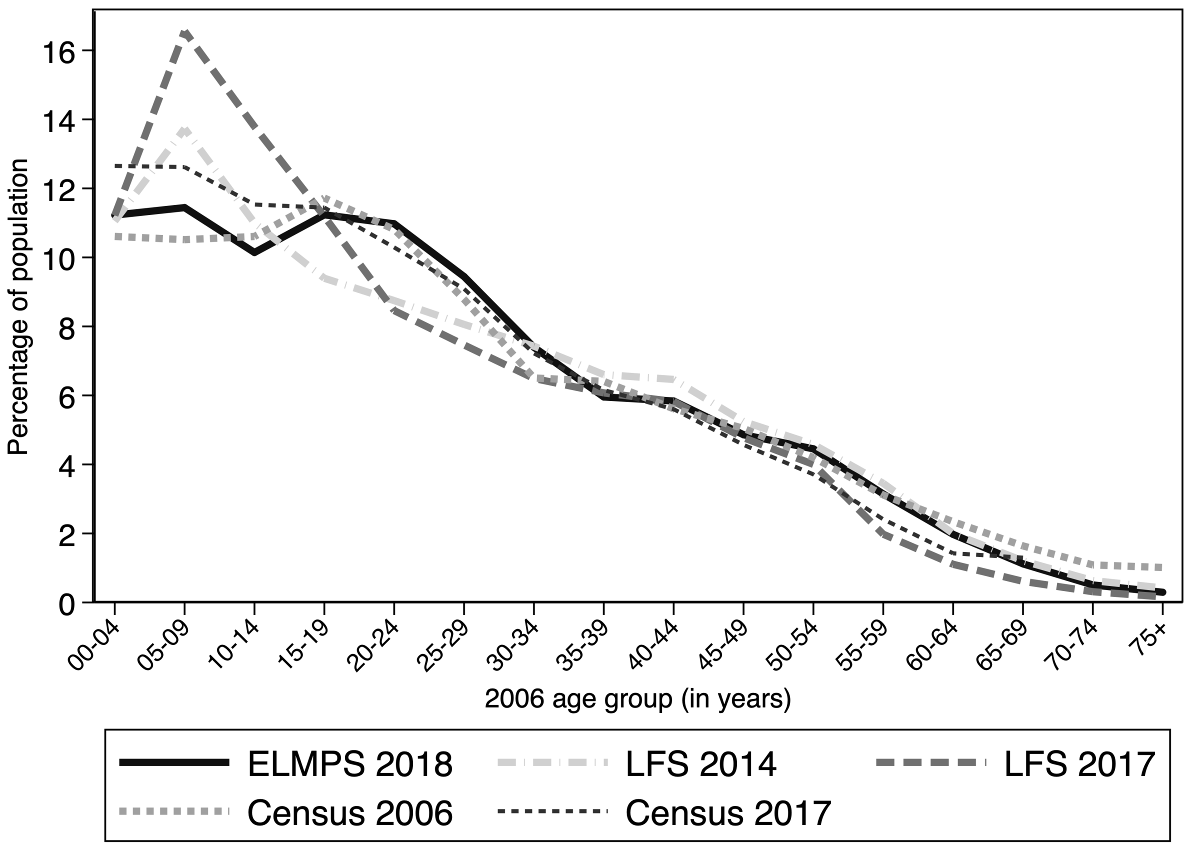
Figure 2. Comparison of age distribution (percentage in 5-year age group), by data source



Source: Authors' calculations from LFS 2014, 2017, ELMPS 2018, and Egypt Census 2017 Table 1-2 (CAPMAS 2017).

Figure 3 compares the age distribution of the sources in Figure 2 with the 2006 Population Census. To compare the LFS and ELMPS, we calculate what individuals' ages would have been in 2006 and exclude those who were born after 2006. Because we only have reports, not microdata, for the Census 2017, we subtract two age groups (ten years, so off by one year). The spike in ages 15-24 observed above in the LFS is prominent as a spike in ages 5-9 in 2006, and while larger for the 2017 wave, was also visible in the 2014 wave. The 2006 and 2017 censuses also diverged in terms of the share of the population at younger ages, with the 2017 Census having a higher share from ages 0-14 in 2006, but not thereafter. Both the Censuses and the ELMPS show a very similar youth population, starting at 15-19 in 2006. The Census 2006 shows it as a youth bulge, with fewer youth thereafter, whereas the higher share 0-14 means the bulge is less pronounced in Census 2017. Overall, the ELMPS 2018 is very consistent with the 2006 Census, and the 2017 Census for ages 15+ (in 2006), but diverged at younger ages.

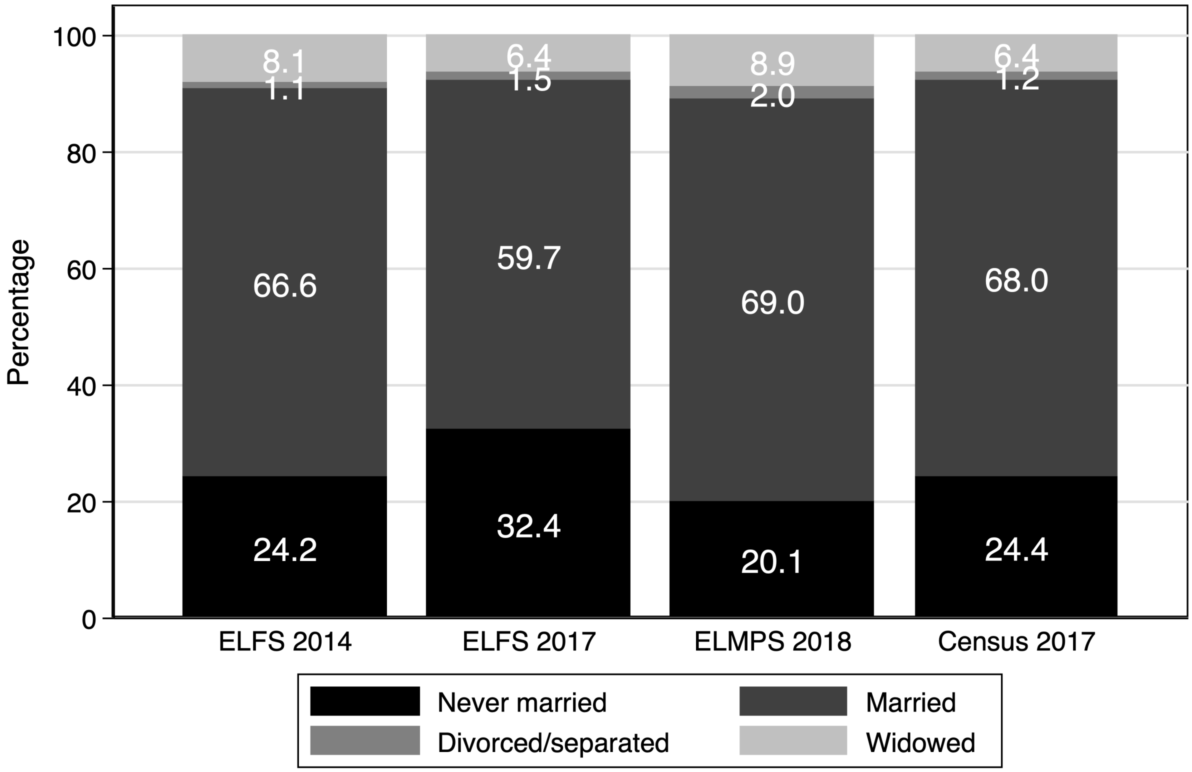
Figure 3. Comparison of age distribution, age in 2006 (percentage in 5-year age group), by data source



Source: Authors' calculations from LFS 2014, 2017, ELMPS 2018, and Egypt Census 2017 Table 1-2 (CAPMAS 2017).

Figure 4 presents the distribution of population aged 18 years or older by marital status. ELMPS 2018 and Census 2017 align in terms of the share of the population married (68-69%). ELMPS found more divorced (2.0%) individuals than the Census (1.2%) and more widowed (8.9% ELMPS versus 6.4% Census). These variations are similar to those across LFSs. As a result, ELMPS 2018 found fewer never married individuals (20.1%) than the Census 2017 (24.4%), LFS 2014 (24.2%) or LFS 2017 (32.4%). These differences may, however, be due to differences in the age composition of each data source, as, for instance, LFS had far more 15-24 year-olds.

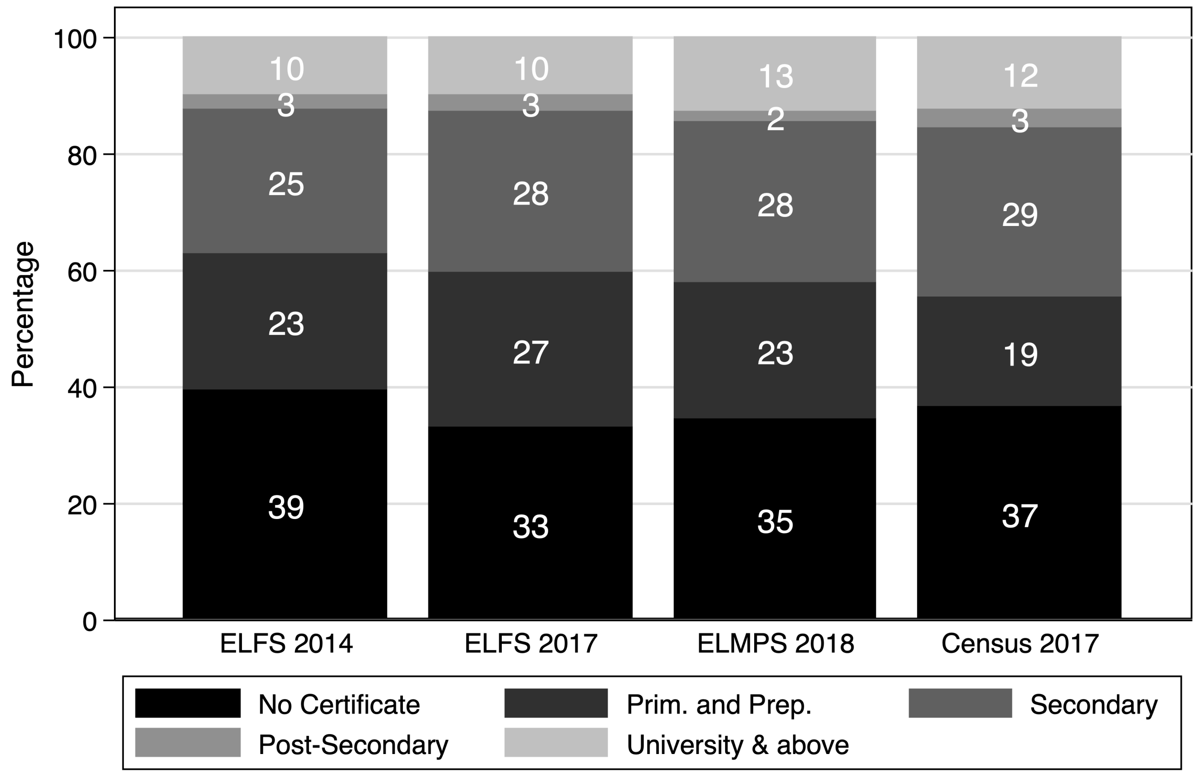
Figure 4. Marital status (percentage of individuals aged 18 years or above), by data source



Source: Authors' calculations from LFS 2014 and 2017, ELMPS 2018, and Egypt Census 2017 Table 1-10 (CAPMAS 2017).

Figure 5 presents the distribution of the population aged 10 years or older by educational attainment and data source. Patterns are generally similar. For instance, ELMPS 2018 sampled individuals with no certificate (35%) slightly less than the Census (37%). ELMPS found a similar share of university graduates (13%) as the Census (12%) and more than the LFSs (10%) although this may be driven by age composition differences. The share in secondary was similar across ELMPS 2018, Census 2017, and LFS 2017 (28-29%) although LFS 2014 was lower (25%). Differences were largest in terms of primary and preparatory education, 19% in the Census, 23% in ELMPS 2018 and LFS 2014, and 27% in LFS 2017.

Figure 5. Educational attainment (percentage of individuals aged 10 years or above), by data source



Source: Authors' calculations from LFS 2014 and 2017, ELMPS 2018, and Egypt Census 2017 Table 1-3 (CAPMAS 2017).

4.2. Labor market outcome comparisons

In this section we turn to comparisons of labor market outcomes between the ELMPS and LFS. We first look at trends over time and comparisons across data sources. We then, given some differences in labor market aggregates as well as demographics, show labor market outcomes by age and sex, in order to disentangle composition differences from differences in outcomes after accounting for composition.[20]

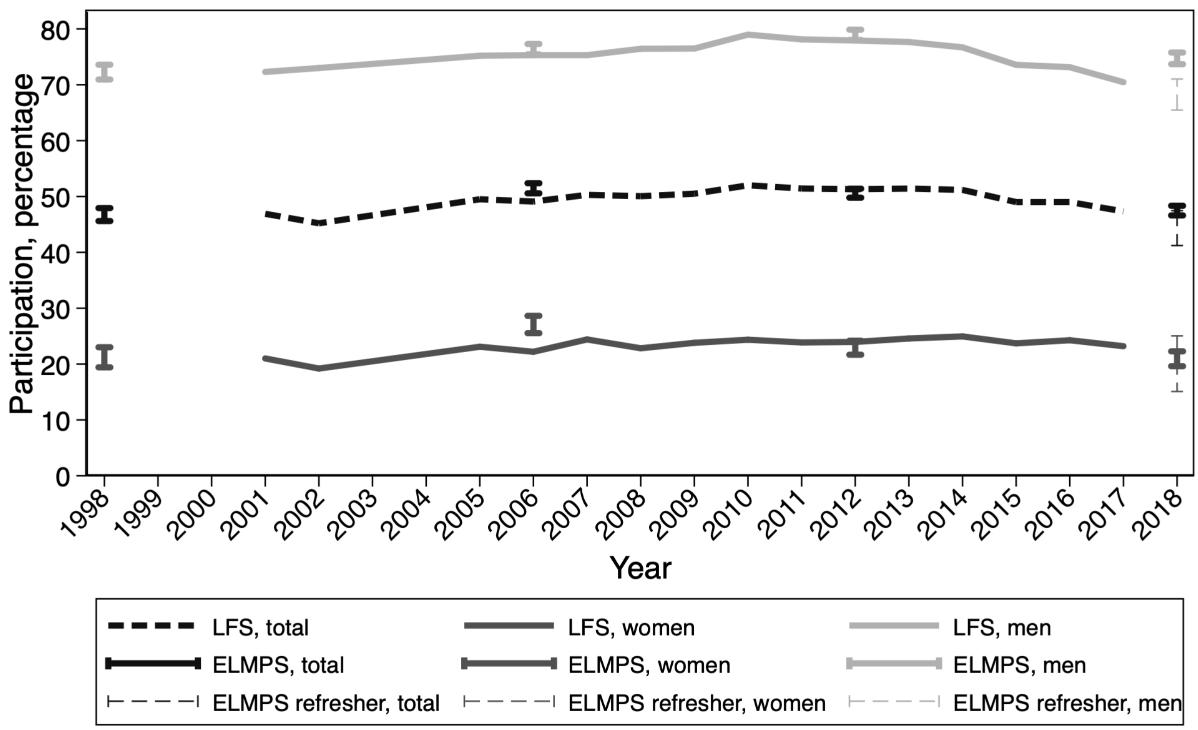
4.2.1. Labor force participation

Figure 6 presents the labor force participation (LFP) rate of the population aged 15-64 years by sex and in total. The figure includes LFS 2001-2017 rates and the rates and CIs from ELMPS 1998, 2006, 2012, and 2018. While total LFP rose from 1998 to 2006, it plateaued in the late 2000s and began falling thereafter. The drop in LFP has been particularly pronounced in the LFS in recent years, dropping from 49% in 2016 to 47% in 2017. The ELMPS 2018 found a similar LFP (47%) to the 2017 LFS. Since the LFS published statistics are for 15+ they are not directly comparable, however, comparing the 2018 and first two quarter of 2019 numbers to those of the 2017 LFS shows a continuing decline in labor force participation rates among ages 15+ (CAPMAS 2019a; b). Examining results by sex, the ELMPS female LFP rate is a bit below and the male rate a bit above the trend from the LFS. The ELMPS has historically found higher participation rates than

[20] Krafft, Keo, and Assaad (2019) explore patterns of labor supply in detail using the ELMPS.

the LFS, which may be in part due to the focus on an individual responding him or herself, rather than using a proxy respondent (Assaad and Krafft 2013). The confidence intervals for women but not for men and not quite for overall overlap with the full ELMPS point estimates, suggesting that there may be some small differences in sample.

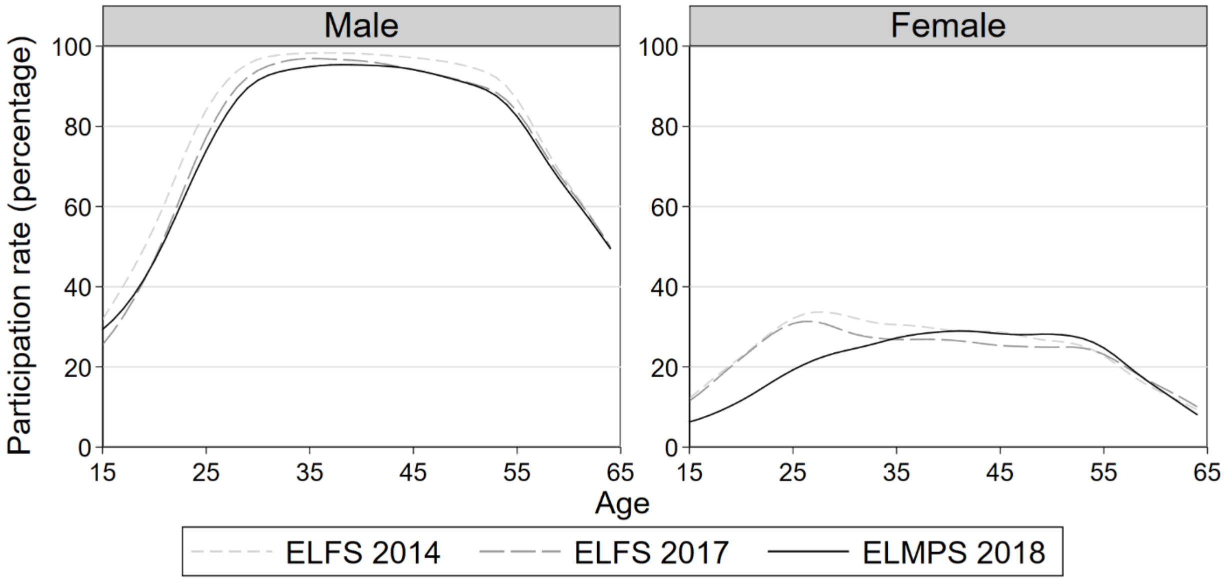
Figure 6. Labor force participation rate (percentage), by sex and data source, ages 15-64, 1998-2018



Source: Authors' calculations from LFS 2008-2017, ELMPS 1998-2018. LFS 2001-2007 from ILOSTAT (ILO 2019).

Figure 7 presents the LFP rate by age, sex and data source. Generally, the pattern for men is similar across sources. For men, LFS 2017 had a very slightly lower participation rate for ages 15-20 and a slightly higher participation rate for ages 25 to 45 than ELMPS 2018. In the LFS 2014, participation was higher for all ages. The pattern across time, from 2014 to 2017 and 2018 shows a clear decline in participation across ages. For men, differences in overall rates were likely driven by differences in composition, given how similar patterns are by age. For women, participation from LFS 2017 was higher than ELMPS 2018 for ages 15-35, while the ELMPS participation rate was higher than the LFS, but only slightly so, for ages 35 to 60. We return to these differences for women below, in examining their types of employment.

Figure 7. Labor force participation rates (percentage) by sex, age and data source

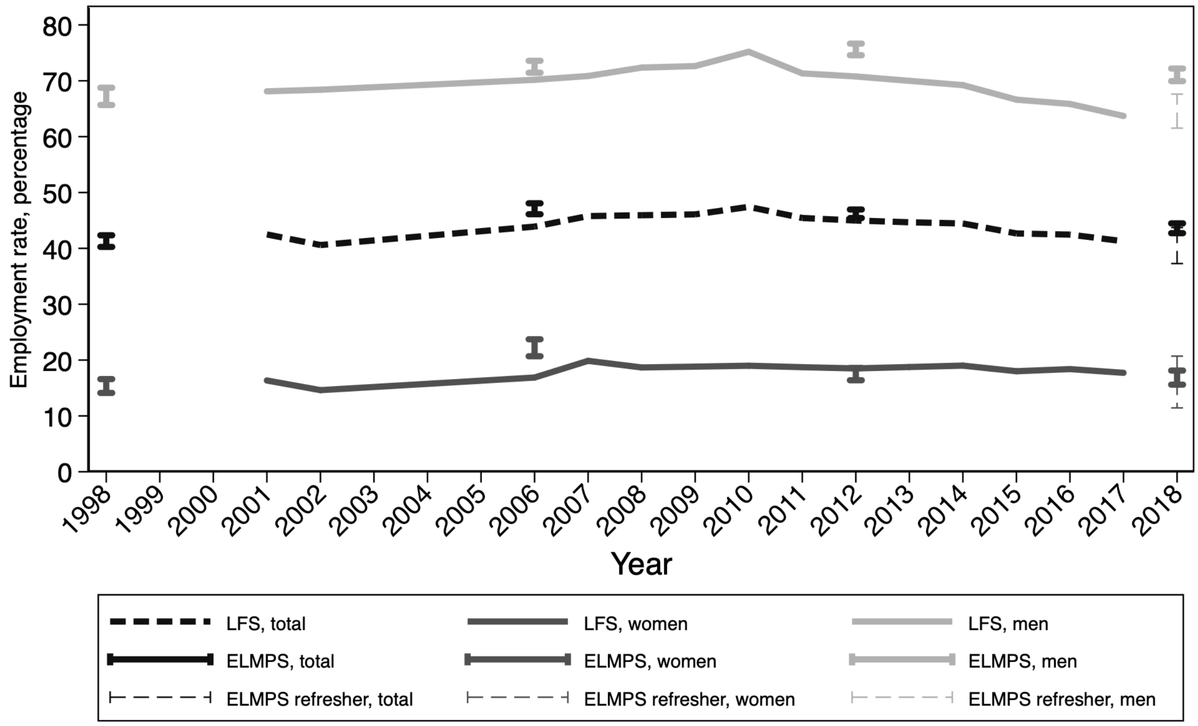


Source: Authors' calculations from LFS 2014, 2017, and ELMPS 2018.
Notes: Running mean smoother, bandwidth 0.4.

4.2.2. Employment rates

Figure 8 presents employment rates by data source. Employment rates follow a similar pattern to LFP rates in rising through approximately 2010 and then declining thereafter (the trend continued into 2018/19 (CAPMAS 2019a; b)). The decline, particularly among men, is a notable new trend, although must be interpreted with some caution given the composition patterns by age in the LFS. The results with the ELMPS 2018 compared to LFS again show similar women's employment, slightly higher total employment, and higher men's employment. The refresher confidence intervals overlap the full sample point estimates in 2018 for total and women but not for men. For men's employment and total employment ELMPS 2006 and 2012 were also higher than LFS.

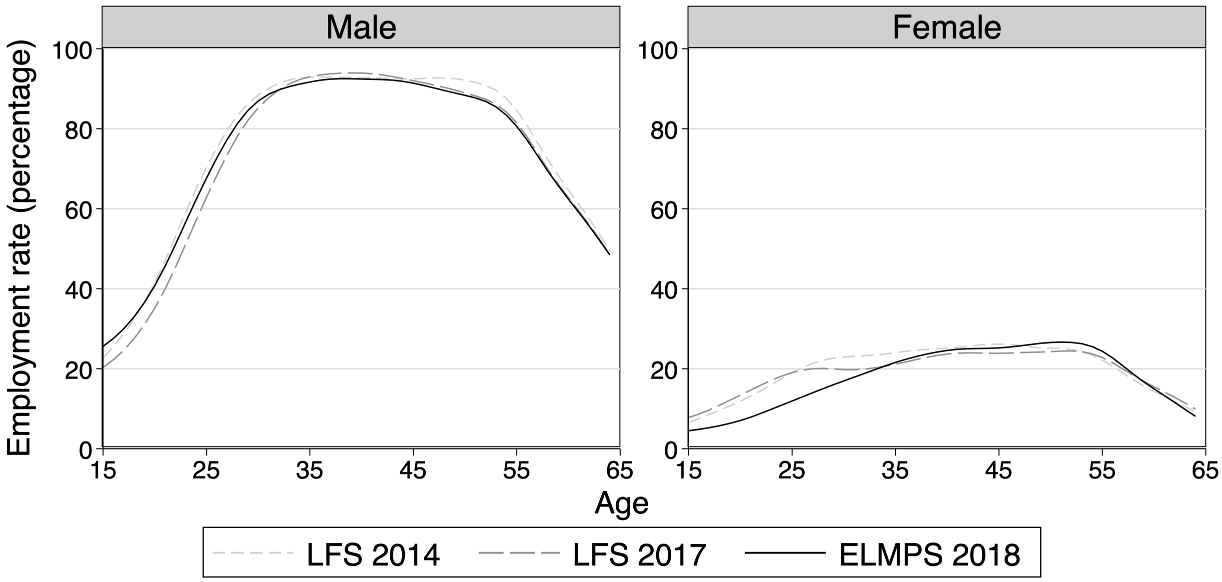
Figure 8. Employment rates (percentage) by sex and data source, ages 15-64, 2001-2018



Source: Authors' calculations from LFS 2008-2017, ELMPS 1998-2018. LFS 2001-2007 from ILOSTAT (ILO 2019).

In Figure 9, we see that employment rates for men, by age, were very similar in the ELMPS 2018 and LFS 2014 and 2017. The drop in prime-age male employment observed in ELMPS 2018 (Krafft, Assaad, and Keo 2019) is thus confirmed by other sources. There were only slight differences with the LFS 2017 having lower youth employment rates for men, and the LFS 2014 higher employment rates for older men. For women, while employment rates at ages 35 and older were generally similar, with ELMPS 2018 finding slightly more employment at these ages than LFS 2017, there was again a difference at younger ages in employment rates, with ELMPS 2018 estimating lower employment rates for young women, a point we revisit in examining the nature of employment, below.

Figure 9. Employment rates (percentage) by sex, age and data source

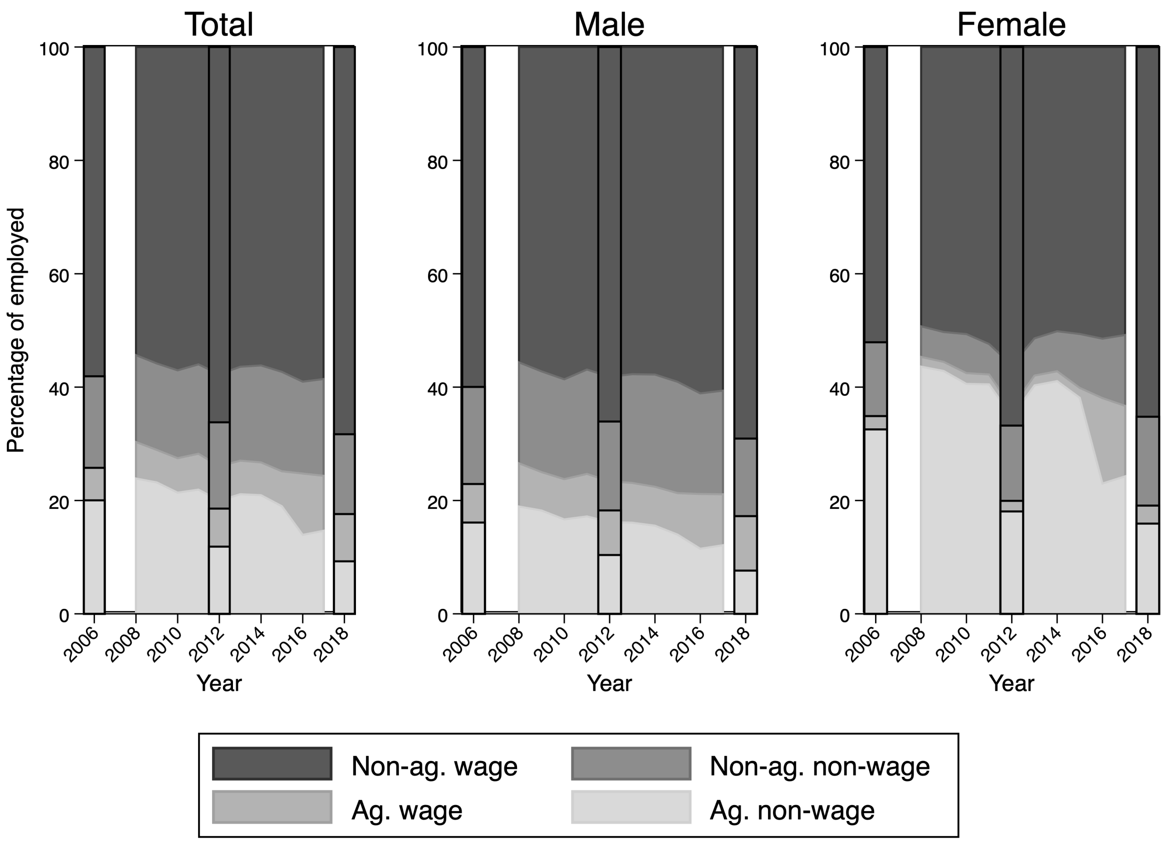


Source: Authors' calculations from LFS 2014, 2017, and ELMPS 2018.
Notes: Running mean smoother, bandwidth 0.4.

4.2.3. Types of employment

Examining the structure of employment by sex, industry and data source illustrates important patterns that may explain differences in overall labor market indicators. Figure 10 examines the structure of employment by industry, specifically comparing agriculture vs. non-agriculture and employment status (wage vs. non-wage). All data sources agree that the share of workers in agricultural non-wage work has declined over time. However, when the ELMPS data are overlapped with the time trends for the LFS, they show different employment statuses and industries, especially for women in recent years. In particular, recent LFSs have classified far more women as agricultural wage workers and fewer as agricultural non-wage workers. In general, fewer individuals were classified as agricultural non-wage workers, especially among women. It may be that the LFS employment detection questions are picking up more subsistence work as market work. This phenomenon may explain why more young women were classified as employed in the LFS than ELMPS.

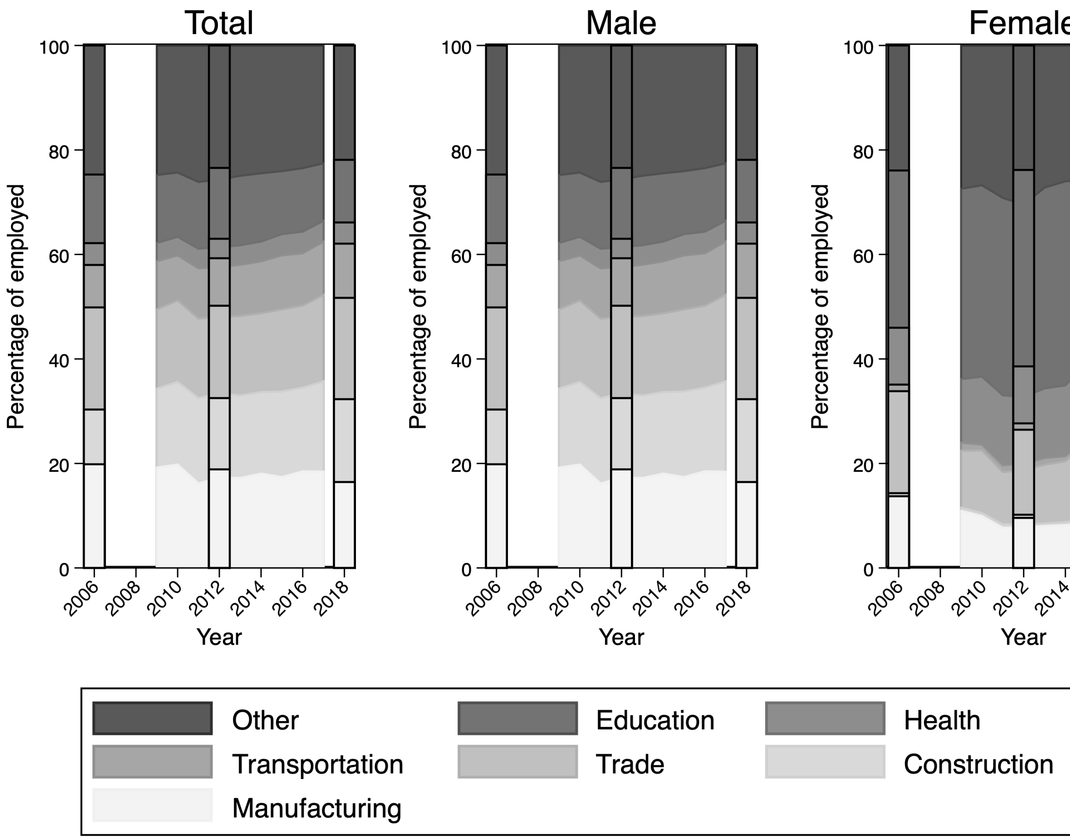
Figure 10. Structure of employment in terms of industry (agriculture vs. non agriculture) and employment status (wage vs. non-wage), by sex and data source, employed, ages 15 to 64, 2006-2018



Source: Authors' calculations from LFS 2008-2017 and ELMPS 2006, 2012, and 2018.
Notes: Bars at 2006, 2012, and 2018 from ELMPS.

Figure 11 shows the structure of employment for those employed outside of agriculture, by sex, comparing data sources. Slightly fewer years of the LFS are included (2008 is excluded) due to different industry coding systems being used over time. Within non-agricultural activities, the LFS and ELMPS show similar distributions of employment by industry. Results were most similar overall and for men, but differences among women may be due to the relatively low employment rates of women and thus more sampling variability. Women were most likely to be employed in education, followed by the "other" sector, which includes a number of services. The vast majority of women were employed within these two sectors. Men were distributed across a wider variety of sectors, with a rising share in construction particularly.

Figure 11. Structure of employment by industry, sex, and data source, employed, not in agriculture, ages 15 to 64, 2006-2018

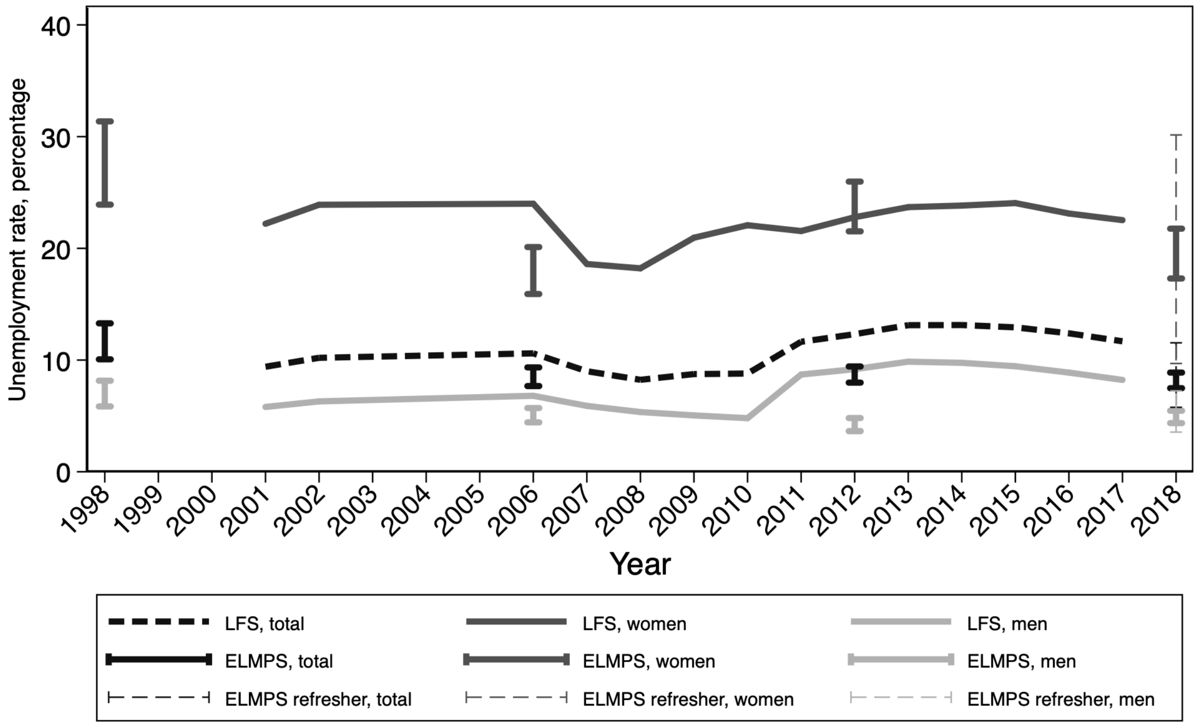


Source: Authors' calculations from LFS 2009-2017 and ELMPS 2006, 2012, and 2018.
Note: Trade is "Wholesale and retail trade." Bars at 2006, 2012, and 2018 from ELMPS.

4.2.4. Unemployment rates

Figure 12 presents unemployment rates by sex and data source. It is important to keep in mind here that differences in composition as well as employment and unemployment can contribute to differences in unemployment rates. Unemployment fell from 1998 to the 2000s, and further in the 2006-2008 period, before rising in 2009 and thereafter with the start of the global financial crisis and Egypt's 2011 uprising. Since peaking in 2015, unemployment rates have been declining in the LFS. ELMPS 2018 shows consistently lower unemployment rates than LFS 2017. These rates were, however, consistent with declines in the 2018 LFS and first two quarters of 2019, which found overall unemployment dropped from 10.6% in Q1 of 2018 to 9.9% in Q2, 10.0% in Q3, 8.9% in Q4, 8.1% in Q1 of 2019, and 7.5% in Q2 of 2019. The ELMPS 2018 estimate of 8.2% unemployment overall (8.7% in the refresher sample) is in line with this trend. It is, however, important to keep in mind that discouraged unemployment – where individuals were ready and want to work but were not searching – rose between 2012 and 2018 (Krafft, Assaad, and Keo 2019). Women have consistently had higher unemployment rates than men and have experienced smaller declines in unemployment over time; they also had more growth in discouraged unemployment (Krafft, Assaad, and Keo 2019).

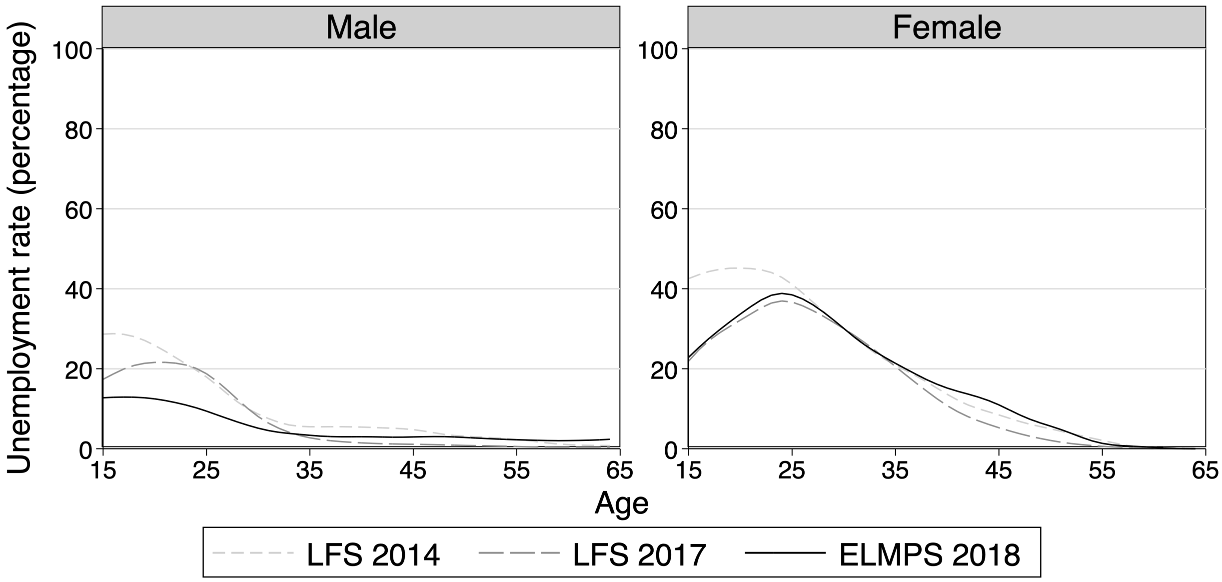
Figure 12. Unemployment rates (percentage) by sex and data source, ages 15-64, 2001-2018



Source: Authors' calculations from LFS 2008-2017, ELMPS 1998-2018. LFS 2001-2007 from ILOSTAT (ILO 2019).

Figure 13 examines the unemployment rate by age. Historically, unemployment has been an educated new entrant phenomenon (Krafft and Assaad 2015; Krafft, Assaad, and Keo 2019), and that generally continues to be the case. Here, ELMPS 2018 and LFS 2017 show similar patterns of women's unemployment, with only small differences at older ages, while LFS 2014 shows a higher rate of young women's unemployment from 15-24. The ELMPS and LFSs both show a rise in younger (aged 15-19) men's unemployment, suggesting some shift among men of unemployment towards the younger and less educated (Krafft, Assaad, and Keo 2019). LFS 2014 clearly shows higher unemployment among the youngest for both men and women.

Figure 13. Unemployment rate (percentage) by sex, age and data source



Source: Authors' calculations from LFS 2014, 2017, and ELMPS 2018.
Notes: Running mean smoother, bandwidth 0.5.

5. Conclusions

Research and policy depend on high-quality, publicly available data. Particularly in the MENA region, including Egypt, there are substantial economic and social challenges that can be informed by high-quality data. Past LMPS and especially ELMPS waves have played a key role in research and policy; we hope ELMPS 2018 will likewise contribute. Having two decades worth of panel data is particularly valuable, unique in the region, and rare even globally. The breadth of inter-linked topics, spanning from employment to fertility and gender role attitudes, has proved a rich source for cutting-edge research. As well as maintaining comparability with other waves and surveys, this wave had an enriched focus on economic vulnerability, a particularly important topic in Egypt in light of recent economic reforms as well as rising poverty rates (The Economist 2019; World Bank 2019).

This paper has detailed the key features of the new wave, including the questionnaires, fielding practices, sampling, attrition, and weighting. We have compared the ELMPS 2018 to other contemporaneous sources in Egypt, including the LFS rounds and 2017 as well as 2006 Censuses. On most measures the ELMPS was consistent with other sources, and when differences occurred, further investigations identified key issues potentially driving differences, such as the high share of youth in recent LFS waves.

The longitudinal, retrospective, and rich data of the ELMPS have already begun to illustrate important developments in Egypt's economy and society, such as fertility, after rising in 2012-2014, once more decreasing in 2018 (Krafft, Assaad, and Keo 2019). The rise in non-participation in the economy (Krafft, Assaad, and Keo 2019), and especially the high and increasing share of youth not in education nor employment (Amer and Atallah 2019) are concerning developments which can be better understood through rigorous research using the ELMPS. These are just a few of the ways ELMPS can contribute to our understanding of Egypt's economy and society.

References

Al-Hawarin, Ibrahim, Ragui Assaad, and Ahmed Elsayed. 2018. "Migration Shocks and Housing: Evidence from the Syrian Refugee Crisis in Jordan." *Economic Research Forum Working Paper Series No. 1213*. Cairo, Egypt.

Amer, Mona, and Marian Atallah. 2019. "The School to Work Transition and Youth Economic Vulnerability in Egypt." *Economic Research Forum Working Paper Series (Forthcoming)*. Cairo, Egypt.

Assaad, Ragui. 1997. "The Employment Crisis in Egypt: Current Trends and Future Prospects." *Research in Middle East Economics* 2 (1): 39–66.

———, ed. 2002. *The Egyptian Labor Market in an Era of Reform*. Cairo, Egypt: American University in Cairo Press.

———, ed. 2009. *The Egyptian Labor Market Revisited*. Cairo, Egypt: American University in Cairo Press.

———, ed. 2014. *The Jordanian Labor Market in the New Millenium*. Oxford, UK: Oxford University Press.

Assaad, Ragui, and Melanie Arntz. 2005. "Constrained Geographical Mobility and Gendered Labor Market Outcomes Under Structural Adjustment: Evidence from Egypt." *World Development* 33 (3): 431–454.

Assaad, Ragui, Christine Binzel, and May Gadallah. 2010. "Transitions To Employment and Marriage Among Young Men in Egypt." *Middle East Development Journal* 2 (1): 39–88.

Assaad, Ragui, and Mongi Boughzala, ed. 2018. *The Tunisian Labor Market in an Era of Transition*. Oxford: Oxford University Press.

Assaad, Ragui, Samir Ghazouani, Caroline Krafft, and Dominique J. Rolando. 2016. "Introducing the Tunisia Labor Market Panel Survey 2014." *IZA Journal of Labor & Development* 5 (15): 1–21.

Assaad, Ragui, Thomas Ginn, and Mohamed Saleh. 2018. "Impact of Syrian Refugees in Jordan on Education Outcomes for Jordanian Youth." *Economic Research Forum Working Paper Series No. 1214*. Cairo, Egypt.

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

———. 2015a. "Is Free Basic Education in Egypt a Reality or a Myth?" *International Journal of Educational Development* 45: 16–30.

———, ed. 2015b. *The Egyptian Labor Market in an Era of Revolution*. Oxford, UK: Oxford University Press.

Assaad, Ragui, Caroline Krafft, John Roemer, and Djavad Salehi-Isfahani. 2018. "Inequality of Opportunity in Wages and Consumption in Egypt." *Review of Income and Wealth* 64 (S1): S26–S54.

Assaad, Ragui, Caroline Krafft, and Shaimaa Yassin. 2018. "Comparing Retrospective and Panel Data Collection Methods to Assess Labor Market Dynamics." *IZA Journal of Development* *and Migration* 8 (17): 1–34.

Assaad, Ragui, and Farzaneh Roudi-Fahimi. 2007. “Youth in the Middle East and North Africa: Demographic Opportunity or Challenge?” *Population Reference Bureau Policy Brief*. Washington, DC.

Assaad, Ragui, and Jackline Wahba. 2015. “Flexible Labor Regulations and Informality in Egypt.” *Economic Research Forum Working Paper Series No. 915*. Cairo, Egypt.

Barsoum, Ghada. 2009. “Methodological Appendix 1: The Egypt Labor Market Panel Survey 2006: Documentation of the Data Collection Process.” In *The Egyptian Labor Market Revisited*, edited by Ragui Assaad, 259–284. Cairo, Egypt: American University in Cairo Press.

———. 2015. “Striving for Job Security: The Lived Experience of Employment Informality among Educated Youth in Egypt.” *International Journal of Sociolocy and Social Policy* 35 (5): 340–358.

Belhaj Hassine, Nadia. 2011. “Inequality of Opportunity in Egypt.” *The World Bank Economic Review* 26 (2): 265–295.

Bertoli, Simone, and Francesca Marchetta. 2015. “Bringing It All Back Home – Return Migration and Fertility Choices.” *World Development* 65: 27–40.

Biltagy, Marwa. 2012. “Quality of Education, Earnings and Demand Function for Schooling in Egypt: An Economic Analysis.” *Procedia - Social and Behavioral Sciences* 69: 1741–1750.

Binzel, Christine, and Ragui Assaad. 2011. “Egyptian Men Working Abroad: Labour Supply Responses by the Women Left Behind.” *Labour Economics* 18 (S1): S98–S114.

Brunette, Waylon, Samuel Sudar, Mitchell Sundt, Clarice Larson, Jeffrey Beorse, and Richard Anderson. 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*.

CAPMAS. 2017. “Census of Population, Housing, and Establishments 2017 (Arabic).” Cairo, Egypt.

———. 2019a. “Egypt Statistics Quarterly ‘Unemployment’ Rate.” Retrieved September 6, 2019. https://www.capmas.gov.eg/Pages/IndicatorsPage.aspx?Ind_id=5778.

———. 2019b. “Egypt Statistics Quarterly Economic Activity Rates.” Retrieved September 6, 2019. https://www.capmas.gov.eg/Pages/IndicatorsPage.aspx?Ind_id=5778.

———. 2019c. *Egypt in Figures 2019*. Cairo, Egypt: CAPMAS.

Crandall, AliceAnn, Kristin Vanderende, Yuk Fai, Sylvie Dodell, and Kathryn M. Yount. 2016. “Women’s Age at First Marriage and Postmarital Agency in Egypt.” *Social Science Research* 57: 148–160.

El-Hamidi, Fatma, and Mona Said. 2014. “Gender-Based Wage and Occupational Inequality in the New Millenium in Egypt.” *The Journal of Developing Areas* 48 (1): 21–41.

El-Mallakh, Nelly, and Jackline Wahba. 2017. “Return Migrants and the Wage Premium: Does the Legal Status of Migrants Matter?” *Economic Research Forum Working Paper Series No.*

1133. Cairo, Egypt.

———. 2018. "Syrian Refugees and the Migration Dynamics of Jordanians: Moving in or Moving Out?" *Economic Research Forum Working Paper Series No. 1191*. Cairo, Egypt.

Fallah, Belal, Caroline Krafft, and Jackline Wahba. 2019. "The Impact of Refugees on Employment and Wages in Jordan." *Journal of Development Economics* 139: 203–216.

Hendy, Rana. 2013. "Assessing the Effects of Trade Liberalization on Wage Inequalities in Egypt: A Microsimulation Analysis." *The International Trade Journal* 27: 63–104.

Heyne, Stefanie, and Michael Gebel. 2016. "Education Effects on the School-to-Work Transition in Egypt: A Cohort Comparison of Labor Market Entrants 1970 – 2012." *Research in Social Stratification and Mobility* 46: 37–49.

Himelein, Kristen. 2014. "Weight Calculations for Panel Surveys with Subsampling and Split-off Tracking." *Statistics and Public Policy* 1 (1): 40–45.

ILO. 2013. "Resolution Concerning Statistics of Work, Employment, and Labour Underutilization Adopted by the Nineteenth International Conference of Labour Statisticians (October 2013)."

———. 2019. "ILOSTAT."

Krafft, Caroline, and Ragui Assaad. 2014. "Beware of the Echo: The Impending Return of Demographic Pressures in Egypt." *Economic Research Forum Policy Perspective No. 12*. Cairo, Egypt.

———. 2015. "Promoting Successful Transitions to Employment for Egyptian Youth." *Economic Research Forum Policy Perspective No. 15*. Cairo, Egypt: Economic Research Forum.

———. 2018. "Introducing the Jordan Labor Market Panel Survey 2016." *Economic Research Forum Working Paper Series No. 1186*. Cairo, Egypt.

———, ed. 2019. *The Jordanian Labor Market Between Fragility and Resilience*. Oxford, UK: Oxford University Press.

Krafft, Caroline, Ragui Assaad, and Caitlyn Keo. 2019. "The Evolution of Labor Supply in Egypt from 1988-2018: A Gendered Analysis." *Economic Research Forum Working Paper Series (Forthcoming)*. Cairo, Egypt.

Malaeb, Bilal, and Jackline Wahba. 2018. "Impact of Refugees on Immigrants' Labor Market Outcomes." *Economic Research Forum Working Paper Series No. 1194*. Cairo, Egypt, Egypt.

Marchetta, Francesca, and Clermont Ferrand. 2013. "Return Migration and the Survival of Entrepreneurial Activities in Egypt." *World Development* 40 (10): 1999–2013.

Minnesota Population Center. 2018. "Egypt 2006. Integrated Public Use Microdata Series, International: Version 7.1." Minneapolis, MN: IPUMS.

Miyata, Sachiko, and Hiroyuki Yamada. 2017. "Do Female Gender Role Attitudes Affect Labour Market Participation in Egypt?" *The Journal of Development Studies* 52 (6): 876–894.

OAMDI. 2019. "OAMDI, 2019. Egypt - Labor Force Survey (LFS). Various Rounds." Cairo, Egypt: Economic Research Forum.

Radchenko, Natalia. 2014. “Heterogeneity in Informal Salaried Employment: Evidence from the Egyptian Labor Market Survey.” *World Development* 62: 169–188.

Robinson, Warren, and Fatma El-Zanaty. 2006. *The Demographic Revolution in Modern Egypt*. Lanham, MD: Lexington Books.

Selwaness, Irène, and Chahir Zaki. 2017. “Assessing the Impact of Trade Reforms on Informal Employment in Egypt.” *The Journal of North African Studies* 20 (3): 1–24.

The Economist. 2019. “Egypt Is Reforming Its Economy, but Poverty Is Rising.” *The Economist*.

Wahba, Jackline. 2015a. “Selection, Selection, Selection: The Impact of Return Migration.” *Journal of Population Economics* 28: 535–563.

———. 2015b. “Through the Keyhole: International Migration in Egypt.” In *The Egyptian Labor Market in an Era of Revolution*, edited by Ragui Assaad and Caroline Krafft, 198–217. Oxford, UK: Oxford University Press.

World Bank. 2013a. *Opening Doors: Gender Equality and Development in the Middle East and North Africa*. Washington, DC: World Bank.

———. 2013b. “Jobs for Shared Prosperity: Time for Action in the Middle East and North Africa.” Washington, DC: World Bank.

———. 2019. “Egypt Economic Monitor: From Floating to Thriving: Taking Egypt’s Exports to New Levels.” Washington, DC: World Bank.