R·E·L·I·E·F CENTRE
مركز ريليف

Data.

Data Audit:
Requirements for
Improving Measures and
Outcome Indicators in
Lebanon

Contents

© Annelise Andersen

Introduction

This report provides an analysis of the social, economic and political data sources available in Lebanon and their quality, highlighting the challenges of accessing such sources and the risks associated with using them. We explore and examine the data that have been compiled by international organizations, local NGOs and the Government of Lebanon since the commencement of the Syrian crisis in 2011. It is intended that this report will provide the foundation for a better understanding of the realities and complexities surrounding the "data scene" in Lebanon and how research can be realistically and pragmatically carried out. We hope this will help the Institute for Global Prosperity (IGP) and RELIEF teams not only avoid duplication of effort, but also improve research coordination and data collection for the RELIEF project.

This report addresses two main points contained in the terms of reference provided to us by the RELIEF project, including:

(1) **Data quality.** We identify the sources of data available in Lebanon (including official and unofficial sources), assess how prone to error, manipulation, disruption etc. data sources are, identify current and emerging data supply challenges and gaps, and provide a realistic assessment of the efforts required in terms of time, cost and quality for collecting, collating and using such data.

(2) **Identifying how data quality and outcome measures and indicators can be improved.** We examine the purpose of decisions regarding data supplies and measures, describe data processing with respect to existing and emerging methods, models, simulations and visualisations, review how data and measures either comply with or shape evolving international standards.

Methodology

As part of the research process we interviewed a range of academics and policy-makers concerned with humanitarian and domestic policy data issues in Lebanon. Interviewees were drawn from United Nations agencies, local universities, and government employees. We conducted semi-structured interviews with eleven key informants who are based in Lebanon. These interviews helped us gain a better understanding of data access and quality issues surrounding publicly and privately held data sources.

The data landscape in Lebanon

The data landscape in Lebanon is characterized by weak statistical capacity and limited state resources, resulting in insufficient, low-frequency and unreliable data across a range of social, economic and environmental issues. Lebanon ranks second-to-last in the World Bank's statistical capacity indicators among similar countries. In its *Systematic Country Diagnostic for Lebanon*, the World Bank highlighted the availability and quality of data as a foundational constraint to achieving the twin goals of ending extreme poverty and promoting shared prosperity (Le Borgne and Jacobs, 2016). Weak statistical capacity and limited data availability hinder social and economic analysis and estimates, thus impacting on critical decision-making. Various large academic projects involving Lebanon, such as SAHWA,[1] have resorted to collecting their own data in order to address various shortcomings. Indeed, Thomas Piketty, in his recent analysis of inequality in the Middle East, highlighted the problem of data in the region in terms of lack of access and its poor quality:

> All in all, it is very difficult to have an informed public debate about inequality trends – and also about a large number of substantial policy issues such as taxation and public spending without proper access to such data. While the lack of transparency on income and wealth is an important issue in many, if not most, areas of the world, it appears to be particularly extreme in the Middle East, and arguably raises a problem of democratic accountability in itself, independent from the levels of inequality observed. (Piketty et al., 2018: 137)

Further, in their latest situation report on the fiscal situation in Lebanon the International Monetary Fund (IMF, 2018) explicitly commented on data quality issues:

> Finally, there is a long-standing need to improve data quality. This could improve access to international investment, and enhance evidence-based policy-making. At a minimum, the quality, frequency, and timeliness of national accounts and balance of payment statistics needs to be improved; data on employment, unemployment and wages need to be frequently collected and published; trade in goods and services data needs to be enhanced; quality of indicators to monitor economic activity need to be improved; and inter-agency dialogue and sharing of information should be strengthened. (IMF, 2018)

However, despite the bleak picture, vital data have been collected in Lebanon by various official and non-official entities working in a range of social and economic policy areas. While the aim of this report is not to analyse the reasons behind weak statistical capacity, it is necessary to note that the confessional power-sharing system of Lebanon has resulted in the politicisation of demographic and social data. As a result, the most recent population census in Lebanon dates back to 1932. The lack of a population census has compromised the existence of a common sampling frame that representative household and multi-purpose surveys are generally drawn from. Apart from the difficulty of demographic analysis, the lack of a common sampling frame has resulted in surveys drawn from different sampling frames, thus affecting the comparability of data over time and across surveys. This does not necessarily mean that data quality has been compromised, but rather, in many instances, it has paved the way for more innovative and "Lebanon-specific" sampling techniques. The following section describes the entities responsible for data collection in Lebanon, their main data publications, and the datasets collected by various agencies. Although there is an abundance of datasets available, they are rather fragmented, uncoordinated, and are either not comparable internationally or are very difficult to access, thus hindering the research process. A key question for researchers embarking on large-scale projects, such as RELIEF, is: do they (a) set up and collect their own quantitative survey data; or (b) adopt more ethnographic qualitative methods that avoid generating political sensitivities with local actors?

---

[1] See http://www.sahwa.eu/SAHWA-PROJECT/About-SAHWA. This project will survey 10,000 young people in Algeria, Egypt, Lebanon, Morocco and Tunisia to assess the attitudes of Arab youth in these countries. The questionnaire covers all the research questions of the project and will create a significant comparative dataset for the region.

Report overview: Key challenges of data quality and data access in Lebanon

There is consensus among policy-makers and academics that the data situation in Lebanon needs to be improved in order to enhance the decision-making capacities of the Government and multi-lateral agencies, and to enable academics to conduct research properly.

It is important to be aware that public services such as health, education and energy supplies in Lebanon have come under increased pressure as a result of the presence of Syrian refugees. However, the refugee crisis has served to expose the policy and political neglect of these services over the past two decades and difficulties were present long before 2011.

There is a policy culture in Lebanon of avoiding measuring social, economic and environmental issues in order to avoid addressing them: "What doesn't get measured, doesn't get done." This is a result of long-term social, political and sectarian tensions that the country has experienced, making detailed statistics and information on the population highly politically sensitive. Added to this, many ministries are constrained by a lack of technical capacity for processing large amounts of data. This is a problem globally, but it is particularly acute in settings that face large-scale security and humanitarian challenges. In Lebanon a range of issues persist, including:

- a lack of representative household data published on a periodic basis;
- decision-making and policy-making are often based on outdated or obsolete data;
- prolonged time lags between data collection and publication of results, making any decisions and impact assessments difficult;
- data are incomparable across years due to different sampling techniques or calculation methods;
- access to raw data held by the Central Administration for Statistics (CAS) is almost impossible and should only be pursued with local support and those who work directly with CAS or have pre-existing relationships;
- no government body endorses any statistics as official unless approved by CAS;
- prior permission to collect data and conduct research in general should be sought from relevant ministries – it is advisable that any outside academics notify ministries of their research plans and how they will contribute to addressing social and economic challenges in Lebanon;
- any collaborations with UN agencies should be immediately embarked on as these will take time to develop. RELIEF themes (RTs) should ensure that they have conducted a thorough evidence review of pre-existing data and literature before approaching agencies. From our experience, and those of our key informants, they must also propose why it is in the interest of the agency to collaborate with RELIEF. What services can RTs offer to such an agency that they do not already possess or are able to access through existing networks?

Addressing RELIEF requirements

RTs need to provide useful research proposals that outline how they will build local policy capacity for using data. This will take time and active facilitation from American University of Beirut (AUB) partners.

RTs should go through each of the relevant data sources we have identified in this document in order to familiarise themselves with the methodological challenges and the "politics of the data terrain" in Lebanon. This will also enable them to understand more about the political economy of certain social and economic issues in the country. This is particularly pertinent for surveys such as the Arab Barometer Survey and the World Values Survey, which can be easily accessed and analysed. Background analyses of these data sets by RTs would help contextualize information for RELIEF projects. They would also provide ideal qualitative question guides for fieldwork in each RT field site.

It is advisable that RTs undertake in-depth literature reviews in order to help them decide on the indicators they wish to use and also to check whether research evidence already exists. For instance,

there are a number of large-scale studies on the impacts of the refugee crises on host communities conducted by the World Bank and partners that provide an ideal base.

We suggest caution in becoming overly focused on the use of new and "innovative" statistical techniques when applied to a context where existing data is not fully reliable and social and political on-the-ground dynamics are rapidly changing. Even with advanced statistical techniques, many basic issues with data quality, accessibility and reliability cannot be overcome in the short to medium term. A cursory glance at studies conducted by academics and major research institutions over the past ten years in Lebanon attests to the fact that complex statistical methods are not necessarily best suited or appropriate to analysing social and economic issues in Lebanon.

"Innovative sampling techniques" in this report refers to new representative surveys of Lebanese host communities, such as the rapid poverty assessment (not yet published). When conducting research on refugees, sampling frameworks should resemble those adopted by UNHCR, WFP and UNICEF. This is the registration database of refugees held by UN agencies. Guidance should be sort from these agencies when designing any surveys involving refugees.

In order to address this we suggest focusing on small-area case studies incorporating in-depth qualitative and ethnographic methods, as outlined in a recent World Bank report (see Brixi et al., 2015) This can be complemented with small quantitative pilot surveys in each area, as explained below. Small-scale pilot surveys will also demonstrate the feasibility of such methods and the potential costs of a larger representative survey.

We are aware that RTs have selected Ouzaii, Ras Beirut, Bar Elias and Saida as research sites. Clear and robust methodological justification should be given as to why these sites have been selected. From our knowledge and the information presented in this report, the majority of datasets identified do not contain detailed and current information across policy issues on these sites.

The benefit of implementing a project-specific survey is that it avoids lengthy negotiations and reliance on UN agencies or the Government to provide data that might not be robust, or creating unnecessary political sensitivities with ministries and local policy actors. It also provides a unique baseline from which to demonstrate changes over the course of the project and creates a data legacy.

The multi-sector survey can be implemented at various geographical levels and complimented through the use of new forms of data extraction, such as text mining of local media and government reports.

We understand the RELIEF project wishes to examine inequality in Lebanon. RTs should come to a consensus regarding how to measure inequality and what sorts of inequalities they will examine, whether income, health, or environmental, for example. A conceptual framework for inequality in Lebanon must be devised that involves a thorough examination of the existing studies on inequalities in the country. Care should also be taken when using income and financial data in Lebanon, as the recent World Inequality Report highlighted issues of data validity.
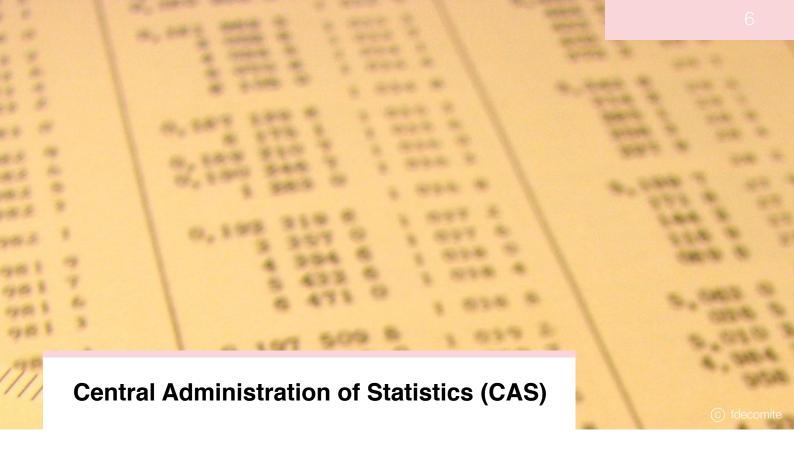
Important update

The Information Management Unit (IMU) of the United Nations in Lebanon has recently published an online document outlining issues with the data available in their registry and with data issues in general.[2] IMU have also added a document which provides all the variables and questions used in assessments and surveys submitted to the UN portal.[3]

The humanitarian data exchange also provides access to 137 datasets relevant to social, economic and environmental factors/issues in Lebanon. GIS maps are also available for buildings and natural resources. It is advisable that the RTs read through the site and the datasets. However, care should be taken when using this site as many datasets have not been validated.[4]

---

[2] Available at: https://docs.google.com/spreadsheets/d/1mdTuM- g4M4mnrG45rKKIp_Z710PJ5gznbHV6T2Vhwdg/edit#gid=0 poor quality'.

[3] Available at: https://docs.google.com/spreadsheets/d/1cWscZpSj_xWO7d1h_1PBrsIXC6CMmgQzUeRUfRudumQ/edit#gid= 599019923.

[4] See: https://data.humdata.org/group/lbn?#dataset-filter-start.

Central Administration of Statistics (CAS)

A "Central Directorate of Statistics" was first established in Lebanon in 1962 under the Ministry of Public Planning and was mandated to produce and coordinate national-level statistics. However, as a result of the civil war starting in 1975, the Ministry of Public Planning was abolished. In 1979, the Central Administration of Statistics was established as a public Administration within the Presidency of the Council of Ministers (PCM), whose mission is to "collect, process, produce and disseminate social and economic statistics at the national level" (CAS, 2013). It was not until after the civil war that CAS reinitiated its work in 1994, having lost all its archives as a result of the war.

The main publications from CAS (outlined in the following sections) include the national accounts, the consumer price index (CPI), household surveys, an annual statistical yearbook, and several reports concerning social, economic, and environmental issues. CAS's publications are plagued by significant publication delays – there is substantial time lag between the time of survey and the publication of relevant data. This is mainly due to the limited human and financial resources of the CAS. Moreover, the post-civil war institutional structure, whereby CAS was established under the PCM as opposed to the abolished Ministry of Public Planning, has affected its ability to coordinate, oversee, and plan data collection across official entities. No government department will endorse any statistics as "official" in the Lebanese context unless they are approved by CAS.

National accounts

CAS is the entity responsible for publishing the national accounts for Lebanon, which represent a country's gross domestic product (GDP), broken down by activity and expenditure component – vital for any macroeconomic analysis of the economy. However, the national accounts of Lebanon are published with a significant time lag due to the limited resources of CAS. For example, the 2014–15 national accounts were only released in 2017. Up until then, the previous figures available were dated 2011. In the absence of timely national accounts, the World Bank (2014) has developed two indicators to better analyse the state of the economy, namely a coincident indicator (WB-CI) and a leading indicator (WB-LI). It is also important to acknowledge that there have been significant improvements in the quality of the national accounts as a result of the introduction of new data and an enhanced framework adhering to international standards. Revised national accounts for the period 2004–2011 were published in 2013, and figures are available on the CAS website[5] for the period 1997–2015.

[5] Available at: http://www.cas.gov.lb/index.php/national-accounts-en.

Consumer price index

With technical assistance from the IMF, the consumer price index (CPI) has been published quarterly by CAS since 1999. A more reliable CPI adhering to international standards was published in 2013, when CAS rebased Lebanon's CPI data to December 2013, providing a much more comprehensive breakdown of prices and applied an updated weighting scheme to the inflation basket (Matta, 2014). CPI publications are available on the CAS website[6] for the period 1998–2017. As of 2008, the CPI is published on a monthly basis, and includes a breakdown of the market baskets of goods and services by region. This data does not cover refugees. WFP and UNHCR data should be used for refugee groups.

Census of buildings and dwellings and establishments (CBDE)

CAS has conducted two CBDE censuses, one in 1996 and another in 2004. Both censuses are crucial in the absence of a population census to derive a sampling frame. The 1996 census served as a sampling frame for the Living Conditions of Households Survey in 1997 (discussed below). The 2004 census resulted in two sampling frames (one of buildings and the other of establishments), which were used to draw samples for several multi-purpose socioeconomic surveys to follow, most importantly the Household Budget Survey of 2004. Aggregated at the governorate and *kaza* (*qadaa or district*) levels, the 2004 census captured the following key variables, which may be relevant to RELIEF research teams:

- the number and distribution of buildings, including the number of floors, the outside building material, year of construction, and main use of the building;
- the number and distribution of dwellings;
- the occupation of dwellings, their use and area;
- the number and distribution of establishments by economic activity and legal status (CAS, 2013).

It is unlikely that the RELIEF project will be able to access CBDE data. One new source of information is the UN Inter-Agency Information Management Map Hub.[7] Key informants commented that GIS data on refugee issues is "either lacking, overlapping or uncoordinated".

Household Budget Survey

The availability of microdata is also limited. Only three representative household surveys have been carried out by CAS (or in coordination with CAS) since 1994.

Population and Housing Survey (1994–1996)

The oldest Population and Housing Survey dates back to the years 1994–1996, as a joint production between the United Nations Fund for Population Activities (UNFPA) and the Ministry of Social Affairs (MoSA), in collaboration with CAS. The total sample size included 65,000 households (of which 61,580 responded), with the sample indicative at the national level, the governorate level, and the kaza level (UNDP and MOSA, 2007).

The 1994–1996 Population and Housing Survey resulted in the Mapping of Living Conditions study in 1998, a joint study by the United Nations Development Programme (UNDP) and MoSA. The study was the first to assess poverty through the basic needs approach, rather than simply measuring income. Eleven indicators were compounded into an index of living conditions to measure poverty. The Mapping of Living Conditions in 1998 was primarily utilized for developing programmes to reduce poverty and tackle disparities (UNDP and MOSA, 2007).

---

[6]Available at: http://www.cas.gov.lb/index.php/economic-statistics-en.
[7]Available at: https://unhcr.carto.com/me.

National Survey of Household Living Conditions (2004)

The second National Survey of Household Living Conditions, a multipurpose survey, was carried out in 2004 by CAS and UNDP. The survey sampled 14,000 households, with results indicative at the national and governorate level, but not the kaza level (CAS, UNDP and MOSA, 2004).

The 2004 multipurpose survey resulted in two main publications. First, the *Comparative Mapping of Living Conditions between 1995 and 2004*, produced in 2006 by UNDP and MoSA, compared the changes in deprivation levels between the 1994–1996 and the 2004 survey at the national and governorate levels (UNDP and MOSA, 2007). Second, the *Poverty, Growth and Inequality in Lebanon in 2008* provided an overview of the characteristics of poor households in Lebanon, and calculated a national poverty line based on household expenditure (UNDP, 2008).

We suggest that RTs examine these two publications in order to develop their evidence background and contextual information. However, care should be taken as the data are from 2004. The infrastructure and built environment has developed greatly since then.

Household Budget Survey (HBS) 2011–2012

The latest household budget survey dates back to 2011–2012. The 2012 HBS was implemented by CAS with technical assistance from the World Bank. Data was collected between September 2011 and November 2012. However, the results of the survey were published later, in 2016, due to technical difficulties related to the very high non-response rate. The sample size covered 4,805 households (with only 2,476 participating households), and is considered only representative at the national level.

More importantly, the poverty estimates derived from the HBS 2012 are incomparable to previous household surveys (namely poverty estimates derived from the 2004 multipurpose survey) due to vast differences in sample design, instruments used, and the methodology of constructing the poverty line (Yaacoub et al., 2015). The technical report derived from the HBS 2011–2012 calculates Gini coefficients by region.[8]

Multiple Indicator Cluster Survey (MICS)

MICSs are part of UNICEF's initiative to assist countries in collecting and analysing data in order to fill data gaps for monitoring the situation of children and women. Lebanon has participated in several rounds of MICSs, namely round 2 of MICS (in 2000) for both Lebanese households and Palestinians (in camps and gatherings) in Lebanon, MICS 3 (2005–2009), also for both Lebanese and Palestinians in Lebanon, and MICS 4 in 2011, only for Palestinians in Lebanon. Moreover, Lebanon is expected to participate in the sixth wave of MICS, with data collection planned for 2018. CAS has been the main partner of UNICEF to implement MICSs of Lebanese citizens, while UNICEF has partnered with the Palestinian Central Bureau for Statistics for MICSs carried out in Palestinian camps in Lebanon. The scope of all MICSs includes specific questionnaires for households, women and children (refer to Annex I for details on the latest MICS for both Lebanese and Palestinians residing in Lebanon).
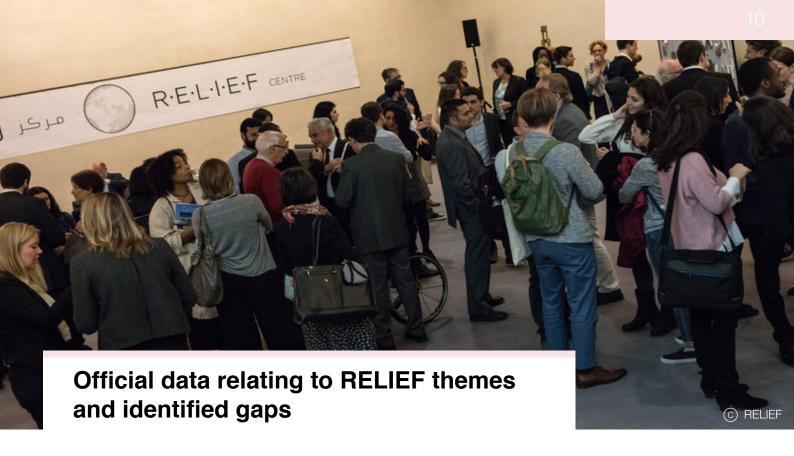
National labour market statistics

Regular labour force surveys are not implemented or available in Lebanon. The official unemployment data in Lebanon dates back to 2009 using round three of the MICS for Lebanese households, in addition to basic labour market data collected in the CAS multi-purpose survey of 2004. To mitigate the absence of this data, the World Bank carried out an employer– employee survey in 2011, comprised

---

[8]The technical report can be accessed at: http://www.cas.gov.lb/images/Excel/Poverty/Measuring%20poverty%20in%20 Lebanon%20using%202011%20HBS_technical%20report.pdf.

of a household questionnaire, and employer (for both formal and informal labour) questionnaire, and an employee questionnaire, on a nationally representative sample of 2,000 households. The survey was used to analyse labour market dynamics in Lebanon. The importance of the dataset is that it allows for an analysis of both demographic and socioeconomic characteristics of labour in Lebanon (both formal and informal), in addition to their employers (World Bank, 2012). Moreover, a Labour Force and Household Living Conditions Survey, implemented by CAS and ILO is in the pipeline. That survey will aim to produce statistically representative data on working conditions at the national, governorate and kaza levels.[9]

---

[9]Available at: http://www.ilo.org/beirut/projects/WCMS_340472/lang--en/index.htm.

© RELIEF

Official data relating to RELIEF themes and identified gaps

In this section we highlight measures and indicators that may be specifically relevant for each RT. However, further work needs to be done by the RTs in terms of research questions, aims and methods in order for the most appropriate outcome measures or indicators to be selected from secondary datasets. They must also explore these datasets to identify relevant indicators.

BLOM Research Indicators and Trends in the Economy (BRITE)

This is a new data portal established and funded by BLOM Bank, Moody's Analytics and Economena Analytics. Interviewees considered this could be a "breakthrough" site in terms of providing reliable and easy to use social and economic statistics. However, this remains to be seen, and the methods used to develop the database are not currently transparent.

The database, though, provides access to 5,000 economic and social indicators in Lebanon from official statistics. This includes financial, real estate, government, education and trade data in Lebanon. The website also allows the use of data visualisation tools and comparisons with international data. BRITE may, potentially, provide a huge time-saving resource of contextual information and data for RELIEF.

Meanwhile, the International Monetary Fund (IMF) and the CAS are currently upgrading their data systems for Lebanon and moving to the Statistical Data and Metadata Exchange (SDMX) over the next 1 to 2 years. The project aims to standardize and modernize mechanisms and processes for the exchange of statistical data and metadata in Lebanon.

Education

Center for Educational Research and Development (CERD)

CERD is the custodian for national education data, and issues annual statistical bulletins. This dataset provides information on the following:

- student population in various sectors (by gender, age, educational cycles, educational areas

and governorates);
* number of teaching staff in various sectors (by gender, academic level, age and governorate);
* number of educational institutes in various sectors and their geographical distribution by educational cycles.

CERD has limited capacity, in terms of human and financial resources, and its major limitation is that data collection is mainly paper-based.

Education outcomes

Lebanon features in the international TIMSS (2007, 2011, and 2015) and PISA (2009, 2015) datasets as well as the World Bank Edstats and the UNESCO Institute for Statistics education dataset. Within these data sources a number of gaps have been identified, including:

* a lack of national standards for measuring learning achievement in the public education system (indicators available are grade-to-grade transition and public examinations);
* outdated national systems (paper-based data collection);
* limited data on assessing learning achievements, and teacher competencies;
* Syrian refugee children are not covered.

Health and health services

In terms of identifying the health status of the Lebanese population and refugees the data scene has historically been a "mixed bag". Since the arrival of Syrian refugees in 2011, the weak nature of data on health has been exposed. Various UN agencies have questioned the validity of sampling methods and surveillance data – particularly vaccination rates – collected by the Government.

However, data quality has vastly improved over the past three years with the EU Instrument for Stability Programme (Strengthening Health) in collaboration with the Ministry of Health (MoH). Dr Walid Ammar (the effective director of the MoH) has been influential in leading this data modernization programme of his department. Key datasets that exist on the health of the population include:

* epidemiological surveillance (EWARN) data covering communicable diseases across the country;[10]
* school-based surveillance, laboratory-based surveillance, influenza surveillance, medical centre and dispensary-based reporting, hospital mortality surveillance, and the National Cancer Registry;
* the Syndicate of Hospitals, which holds all data relating to the private sector in Lebanon and accounts for 85 per cent of health provision. This data is financially sensitive and thus very difficult to access without existing relationships.

Data on inequality and public provision of services in Lebanon

Inequality in Lebanon is particularly pronounced with respect to income and education, but less so on health outcomes (lower by 30.0 per cent, 24.1 per cent, and 6.7 per cent, respectively, than the non-inequality-adjusted Human Development Index).[11]

> Inequality within regions accounts for most of the inequality in Lebanon. Lebanon's Gini coefficient was estimated at 0.361 in 2004–05 using spatially adjusted consumption per capita. It is on par with the average of countries in the Middle East and North Africa and much lower than inequality in countries in Latin America. Inequality was the highest in the North, which also

---

[10]Available at: https://www.moph.gov.lb/en/Pages/2/193/esu.
[11]The latest UN data on inequality in Lebanon is from the 2014 HDI. Available at: http://data.un.org/DocumentData.aspx?id=379.

had the lowest consumption per capita. North and Mount Lebanon had the lowest consumption growth between 1997 and 2004–05. The North region was also distinct because it had the lowest elasticity of poverty to consumption growth, which may be related to highly unequal distribution in this region. (Le Borgne and Jacobs, 2016)

Water provision: The CRI assessment of the Bisri Dam project

This is a survey of 1,200 households in Greater Beirut and Mount Lebanon. It is the only regional survey available and there are no national surveys on accessing safe water in Lebanon. There is substantial variation in access to public water across municipalities. For example, access to the public network is almost universal in Matn, but less than 10 per cent in Aley. Furthermore, the quality of public water is worse for households in the bottom 40 per cent than for wealthier households. For example, service water is saline for 46 per cent of households in the bottom 40 per cent compared with 32 per cent of households in the top 60 per cent. The delivery of public water is also unreliable, with the majority of subscribers experiencing frequent supply interruptions. On average water is supplied by the public network for 6 hours per day, compared with 13 hours per day from private suppliers. Only 20 per cent of households connected to the public network have water available every day in the wet season (December through May) compared to 90 per cent of households using water from other sources. An estimated 82 per cent of subscribers are not satisfied with the public network service (Le Borgne & Jacobs, 2016).

Data on Palestinian refugees in Lebanon

The key data sources concerning the social and economic conditions of Palestinian refugees are provided by the United Nations Relief and Works Agency for Palestine Refugees (UNRWA). Professor Jad Chaaban (AUB) has conducted numerous analyses using such data. In order to access the data it is first necessary to collaborate with UNRWA. Recently the Government conducted the first ever census of Palestinians refugees.[12]

---

[12]Available at: http://www.lpdc.gov.lb/selected-articles/lebanon-conducts-first-ever-census-of-palestinian/401/en.

International datasets that include Lebanon

© NAADVISORY

In this section we provide details of international datasets that include Lebanon and enable comparative indicators to be used.

Global Financial Inclusion (Global Findex)

The Global Findex database was launched in 2011 and contains indicators on financial inclusion and resilience, and on how adults manage their finances, savings, and mitigate risk (Demirguc-Kunt et al., 2014). Lebanon features in the 2011 and 2014 editions. The 2014 Global Findex includes more than a 100 indicators on financial inclusion, namely ownership of financial institution accounts, purposes of account use, remittances, savings behaviour, use of savings methods, sources of borrowing, and how families deal with emergencies. The 2014 financial inclusion survey for Lebanon sampled 1,000 individuals and is nationally representative.[13]

World Enterprise Surveys

The World Bank has been conducting firm-level surveys across the world since 1998 with a centralized database and global methodology. A Lebanese Enterprise Survey was conducted by the World Bank in 2013. The Enterprise Survey is conducted on a representative sample of firms in the non-agricultural formal private economy. The data were collected between April 2013 and September 2014, and included 561 establishments in the manufacturing and services sectors. The data are used to create comparable indicators on the business environment across countries. The main variables include firm characteristics, information about sales and suppliers, competition, infrastructure services, judiciary and law enforcement collaboration, security, government policies, laws and regulations, financing, overall business environment, bribery, capacity utilization, performance and investment activities, and workforce composition.[14]

World Wealth and Income Database

Initially established in 2011, this database is comprised of data series on income in more than 30 countries, with tens of new countries under study. The database combines fiscal data, survey data

---

[13]Data can be accessed via http://databank.worldbank.org/data/reports.aspx?source=1228#.
[14]Data can be accessed via http://www.enterprisesurveys.org/data/exploreeconomies/2013/lebanon.

and national accounts. With respect to Lebanon, surveys, national accounts and personal income tax records were analysed to produce estimates of the national income distribution in Lebanon over the period 2005–2014 (Assouad, 2017).[15]

Arab Barometer

Similar to other regional barometers, the Arab Barometer seeks to produce reliable and comparable data on citizens' attitudes and factors that drive attitudes regarding a number of social, economic and political issues across Arab countries. To date, there have been four waves of the Arab Barometer (2007, 2012, 2014, and 2016). Lebanon has taken part in all four waves, the latest of which included 1,200 Lebanese citizens and an additional 300 Syrian refugees. The survey gauged attitudes on public institutions, elections, media, democracy, social, cultural and religious topics, relationships with other Arab countries, and reactions to the Arab Spring.[16]

World Values Survey (WVS)

The WVS consists of nationally representative surveys conducted in almost 100 countries, containing almost 90 per cent of the world's population, using a common questionnaire to gauge and compare people's beliefs and values.[17] Lebanon took part in wave 6, with data collected in 2013 on a nationally representative sample of 1,200 individuals.[18]

Institutional Profiles Database[19]

This provides an original measure of countries' institutional characteristics through composite indicators built from self-report data. The database was designed to facilitate and stimulate research on the relationship between institutions, long-term economic growth and development.

The 2016 edition of the database follows on from the 2001, 2006, 2009 and 2012 editions. It covers 144 countries and contains 127 indicators, derived from 320 variables describing a broad range of institutional characteristics, structured according to nine functions:

(1) political institutions;
(2) security, law and order, and control of violence;
(3) functioning of public administrations;
(4) free operation of markets;
(5) coordination of stakeholders, strategic vision and innovation;
(6) security of transactions and contracts;
(7) market regulation;
(8) social dialogue and openness; and
(9) social cohesion and social mobility.

---

[15]Data on Lebanon can be accessed via http://wid.world/data/.
[16]Data files for all four waves can be accessed at: http://www.arabbarometer.org/instruments-and-data-files.
[17]Available at: http://www.worldvaluessurvey.org/WVSContents.jsp.
[18]Data files for Lebanon WVS 6 in 2013 can be accessed at: http://www.worldvaluessurvey.org/WVSDocumentationWV6.jsp.
[19]Data for Lebanon (2009, 2012, 2016) can be downloaded from: http://www.cepii.fr/institutions/EN/ipd.asp.

The data landscape of a refugee crisis

The Syrian refugee crisis has changed the data landscape in Lebanon and exposed the fragile nature of public services and pre-existing information systems as insufficient for designing reliable and robust humanitarian policies. In response to this, several international NGOs and research institutes have developed large datasets for programmatic operations and identifying the needs of both Syrian and Lebanese households in areas heavily concentrated with refugees. While this implies a relative abundance of diverse datasets in comparison with pre-crisis levels, a caveat is that this data is far from centralized. Nor is it coordinated, and its quality remains largely unmonitored. In terms of access, very few academics have been able to access the datasets of UN agencies. Regardless of the limitations of CAS's efforts in the collection of data, it was the official entity producing most of the representative data pre-crisis. The current situation is that various UN agencies have ignored CAS sampling frames and applied their own techniques, resulting in varying levels of representativeness and data quality. Furthermore, these datasets are seldom comparable across years or with government datasets.

Apart from the humanitarian needs of Syrian refugees, the majority of research on the impacts of the Syrian crisis have focused on the immediate effects on the Lebanese host community economy. In 2013, The World Bank's Rapid Economic and Social Impact Assessment (ESIA) of the Syrian conflict on Lebanon estimated that the conflict would lead to a 2.85 per cent cut in real GDP growth each year, and a significant increase in government expenditure (World Bank, 2013). On the other hand, a study commissioned by UNDP and UNHCR on the impact of an annual aid package of USD 800 million revealed that every dollar spent on humanitarian assistance had a multiplier effect of 1.6, resulting in a 1.3 per cent growth in GDP (UNDP and UNHCR, 2015).

Verme et al. (2015) noted that seven in ten registered Syrian refugees in Lebanon can be categorised as poor and called for a shift in focus towards economic inclusion and self-reliance for Syrian refugees. A substantial number of reports, surveys and qualitative assessments by NGOs and international agencies have focused on a generalised situation analysis or assessment of a particular region or sector. For example, an assessment of the impact of Syrian refugees on Lebanon and their employment profile was carried about by the International Labour Organization (ILO) in 2013. The assessment surveyed the employment profile of refugees and the impact of their economic participation on host community livelihoods, through purposive sampling. It was, however, no means representative, and focused on four particular regions (Masri and Srour, 2014). A UNHCR study (2017) conducted a cross-sectional survey of Syrian refugees to monitor access and utilization of health services. UNDP also publishes qualitative situation analyses of conflict and inter-community dynamics and relations for various regions (UNDP and Lebanon Support, 2015).

In an attempt to centralize the numerous assessments, surveys and studies that are related to the Syrian crisis, the United Nations Inter-Agency Coordination Unit has created an online assessment registry for completed and ongoing assessments in Lebanon (Figure 1). However, this registry is far from comprehensive as agencies are by no means obliged to register their assessments, but rather do so on a voluntary basis. The information in the registry includes a description of the methodology, sampling, relevant agency, and main findings for each assessment.[20] Raw data can be requested, subject to approval by the relevant agency. It is important to note that there are no data quality standards associated with the database. It is left to users to vet the data for quality and usefulness. It is advisable to seek advice from the principal investigators of each survey or dataset before extracting any information.

---

[20]Available at: http://lebanonfactbook.azurewebsites.net/assessmentregistry/.

Figure 1. Image of Assessment Registry

Lebanon | Assessment Registry

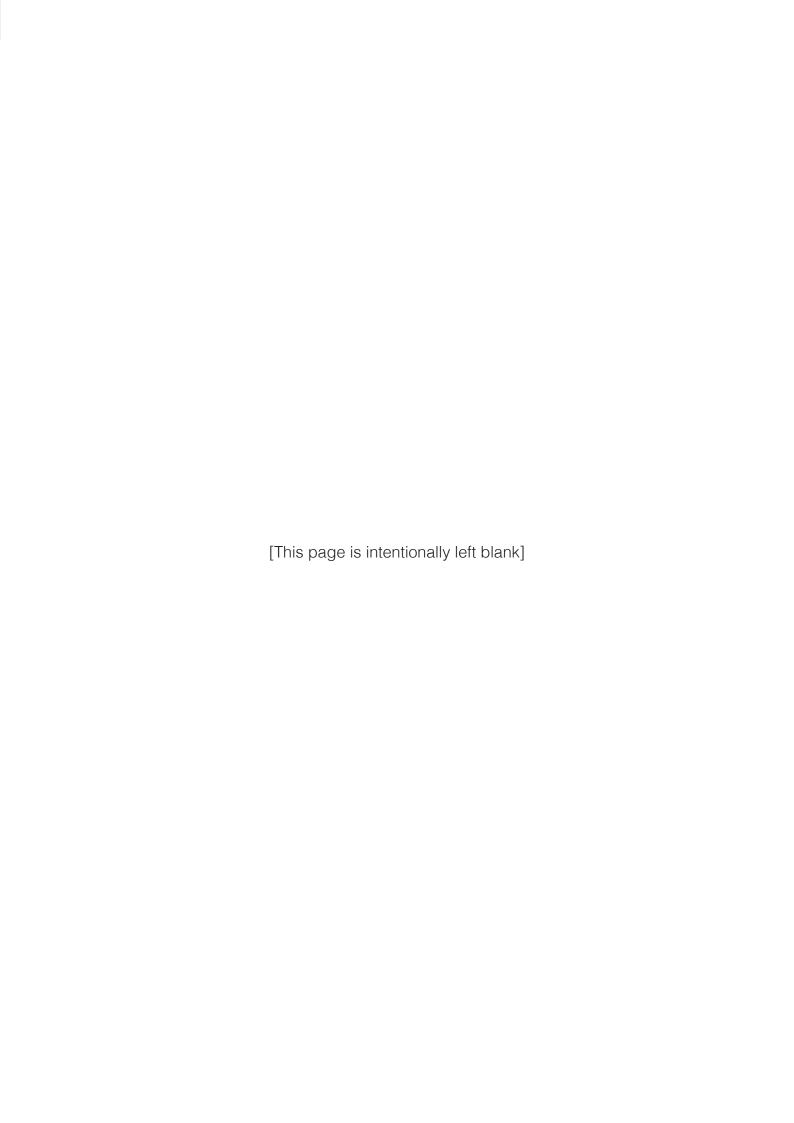Inter-Agency Coordination Lebanon

Syria Refugee Response in Lebanon

The Assessment Registry is an online repository of on-going and past assessments in Lebanon. When researching a topic or planning an assessment, check the existing knowledgebase by searching keywords. The registry includes a brief description of the methodology, key findings and a link to the final report. Click here to register your own assessment.

Contact brownc@unhcr.org for more information.

Check the Factbook for a repository of assessment findings.

registered	name						
July 2017	Baseline Survey 2016	UNICEF	rghoussoub@unicef.org	Report finalized	Baseline Survey 2016 was entirely based on MICS protocols (i.e. questionnaire, tools, data collection tools etc..). The survey was done jointly with MOSA and used 3 questionnaires about children, women and HHs. It had 4 independent population groups: Lebanese residents, Syrian, PRS, PRL.	Household level interviews	Syrian refugees - registered with UNHCR, Lebanese host community, Palestine refugees from Syria, Palestine refugees from Lebanon
July 2017	Health access and utilization survey among Syrian refugees in Lebanon	UNHCR	http://data.unhcr.org/syrianrefugees/download.php?id=12285	Report published	This cross sectional survey was conducted among Syrian refugees living in Lebanon to monitor access to and utilization of key health services. Refugees in Lebanon are predominantly living in urban areas and informal settlements and there are no refugee camps.	Individual level interviews	Syrian refugees - registered with UNHCR
June 2017	Health & Nutrition Knowledge Practices and Coverage Survey	Medair	http://relief.medair.org/assets/uploads/media/2016_KPC_Survey_Report_Final_20Jun2017.pdf	Report published	The Medair 2016 KPC Survey is the primary outcome evaluation data source for the Medair Health Program, and was designed to provide a snapshot in time of the knowledge, practices and coverage of health care services in the catchment area of the program. The objective of the 2016 KPC survey is to measure key health indicators within the Medair catchment area for program evaluation and performance-improvement purposes. By identifying current gaps in knowledge, practice and coverage, this evaluation also highlights areas for further intervention, not only by Medair, but by other partners working in the Bekaa governorate's health sector.	Household level interviews	Syrian refugees - registered with UNHCR, Syrian refugees - not registered with UNHCR, Lebanese host community
June 2017	Impact of EU funded education project	British Council	Roohi.Malik@britishcouncil.org	Report finalized	To assess the impact of the British Council/EU funded project 'Accessing Education'.	Direct observation, Focus group discussion, Key Informant Interviews, Secondary data review	Syrian refugees - registered with UNHCR, Lebanese host community

[This page is intentionally left blank]

Key datasets and databases post-Syrian crisis

Dataset	Description
UNHCR (registry)	Registry of over 1 million registered refugees in Lebanon. Includes contact information, demographics, and social and economic indicators. (Individual records of registered Syrian Refugees in Lebanon.) Access: Subject to permission of UNHCR, and without disclosure of personal contact information.
VASYR (UNICEF)	Multi-sector overview of the situation of Syrian refugees in Lebanon (2013–2017). Representative sample of over 4,000 Syrian refugee households, randomly sampled from the UNHCR registry. Data is comparable across years. Indicators: Demographics, food security, housing conditions, education level of adults and children, assets owned, crowding of dwelling index, dependency ratios, expenditure by category. Access: Subject to UNICEF permission. Link to 2017 report: https://reliefweb.int/report/lebanon/vasyr–2017-vulnerability-assessment-syrian- refugees-lebanon Questionnaire is public and is available at: https://data2.unhcr.org/en/documents/download/61020
National Poverty Targeting Program (NPTP) Database	A database of 500,000 of the poorest Lebanese households in Lebanon (database owned by the Presidency of Council of Ministers). Indicators: Demographics, assets, educational attainment, disabilities, health problems. Access: Permission from PCM rated as difficult.

Dataset	Description
Rapid Poverty Assessment in Lebanon (2016) – UNDP and AUB	Rapid assessment of poverty among Lebanese communities using a nationally representative sample (n = 4,000). Access: Unknown (not yet published).
UNICEF Baseline Survey (2016)	Survey of four distinct communities – host communities, refugees, PRLs, and PRS (n = 20,000+). Entirely based on MICS protocols. Access: Subject to UNICEF approval. Link to report: data.unhcr.org/syrianrefugees/download.php?id=11355020
Three-country Syrian Refugees and Host Communities Survey (SRHCS) in Lebanon, Jordan, and northern Iraq (Kurdistan region) – World Bank, 2015	Assesses the socioeconomic and living conditions of refugees and host communities (representative sample of both communities in Lebanon). Access: Never published, data inaccessible.
Refugee Health Access Survey (2014) – AUB, Johns Hopkins University, UNHCR, MdM, ECHO.	National household sample of Syrian refugees and host communities (n = 2,100). Covers a range of health issues (NCDs, health service usage, mental health, vaccinations). Access: Via the Faculty of Health Sciences, AUB and Johns Hopkins University. Various negotiations will be required with JHU, which possesses the original datasets.
MSF health assessments – Bekaa, Tripoli	Extensive surveys of refugee health were conducted 2013–2015. Access: Not accessible.

Requirements for enhancing data quality, measurement and indicators

It is important to note that in order to fully utilize the data available in the Lebanese context RTs must produce a coherent and realistic research proposal and survey design prior to approaching any ministry or UN agency.

In order to access datasets held by government ministries and UN agencies, RTs must explain why these collaborations are useful for Lebanese policy-makers and UN agencies. What capacity-building elements can the RELIEF team bring to an agency or ministry in terms of data analysis resources? Will they conduct analysis on behalf of the organization while also undertaking their own academic work? How will the research contribute to the lives of refugees and host communities?

Over the past year these questions have been increasingly asked of outside research organizations and institutions wishing to conduct projects in Lebanon. It is clear from the experiences of our key informants and the authors of this report that these are essential prerequisites for conducting research in Lebanon. We advise that RTs decide on a small number of key outcome indicators that are aligned to local policy interests.

While the lack of a national sampling frame has not necessarily compromised the quality of data, it has affected the comparability of data across years given the difference in sampling designs and methodologies. Accessing CAS data will require a considerable investment of time and face-to-face meetings with CAS officials. We would suggest that access to CAS Household Budget Surveys may be difficult, if not impossible, for the RELIEF teams.

Interviews with key informants in UN agencies also suggest that trying to access large UN datasets will also require a considerable time investment that may produce few valid results. Our suggestion is that, in order to fill the data gaps and address the need for specific and locally sensitive indicators and outcome measures, the RELIEF project team should design and build their own representative survey that covers the multiple levels of geographical aggregation.

The creation of a specific UCL–AUB survey would allow RTs to integrate their desired indicators from across policy areas or themes – health, education, socioeconomic, civil society and the built environment. However, in order for indicators to be relevant and applicable at all geographical levels, exhaustive preparatory work must be done in terms of research questions and sample selection, as well as gaining local knowledge through small-scale ethnographic qualitative work. The local partners (AUB) have considerable experience in designing and implementing such surveys. This type of project would also serve to strengthen project cohesion between institutions and RTs, enabling cross-cutting issues and themes to be identified.

This type of multi-sector survey does not exist in the region. The closest project of this nature is the Youth Employment and Civil Society Research Consortium (SAHWA) funded by the European Union and involved the Lebanese American University.[21]

The benefits of implementing such a project-specific survey would include the following:

• It avoids lengthy negotiations and reliance on UN agencies and the Government to provide data which may not be good quality or clean when it is delivered to the RELIEF team. A number of key informants expressed the view that various agencies would be unwilling to transfer data to an outside academic institution as they are themselves concerned with the data quality and methods used for collection.
• It sets a robust baseline from which RELIEF can work and share with other agencies and donors. It becomes a commodity and policy lever for engagement in Lebanon and the UK.
• It would increase collaboration and linkages between RTs.
• It can be easily replicated mid-programme and the end of RELIEF to assess change across the

---

[21]See: http://www.sahwa.eu/.

project cycle.
- It can be designed with potential future linkages to existing government and World Bank datasets when the opportunity arises.
- Ministries, donors and the UN can be invited to include specific measures/indicators to encourage their participation in the RELIEF project.
- Smaller booster surveys can be conducted after the main survey in order to examine specific policy issues and indicators in small local areas such as Ras Beirut, Beqaa and Tripoli.
- Survey data can also be used to run natural experiments in order to examine the effects of particular external events in Lebanon (conflict, refugee movements into certain areas) and modelling to predict socioeconomic trends.

Unconventional data sources and methods

To complement conventional survey methodologies various techniques can be used to extract more unconventional sources of information, such as text mining from newspapers, government speeches and social media such as Twitter. From interviews with key informants, there is currently a major project being undertaken by a consortium from AUB and Qatar University to mine texts in Arabic such as newspapers, political speeches and twitter. We suggest that RELIEF make contact with OMA project to discuss their data needs.

It has been noted that synthetic populations could be created from existing data sources such as the UNHCR registration data to model health trajectories. However, if RELIEF wishes to pursue these kinds of techniques, it will require considerable technical expertise, which should be sought as soon as possible.

References

Assouad, L. (2017). Rethinking the Lebanese economic miracle: The extreme concentration of income and wealth in Lebanon 2005–2014. WID. Working Paper No. 2017/13. Available at: http://wid.world/document/rethinking-lebanese-economic-miracle-extreme-concentration-income-wealth-lebanon–2005–2014-wid-world-working-paper–201713/).

Brixi et al. (2015). Trust, Voice, and Incentives: Learning from Local Success Stories in Service Delivery in the Middle East and North Africa. World Bank

CAS, UNDP and MoSA. (2004). National Survey on the Living Conditions of Households. Central Administration of Statistics (CAS), United Nations Development Program (UNDP) and the Ministry of Social Affairs (MOSA) of the Republic of Lebanon. Available at: http://www.databank.com.lb/docs/Living%20conditions%202004%20-%20CAS.pdf

Central Administration for Statistics (CAS). (2013). About us. Available at: http://www.cas.gov.lb/index.php/about-us-en [accessed 1 March 2018].

Demirguc-Kunt, A., Klapper, L., Singer, D. and Van Oudheusden P. (2014) The Global Findex Database: Measuring Financial Inclusion around the World. Policy Research Working Paper 7255. Washington, D.C.: World Bank.

International Monetary Fund. (2018). Concluding mission statement. February. Lebanon: Staff Concluding Statement of the 2018 Article IV Mission. Available at: https://www.imf.org/en/News/Articles/2018/02/12/ms021218-lebanon-staff-concluding-statement-of-the-2018-article-iv-mission

Le Borgne, E. and Jacobs, J. (2016). Lebanon: Promoting poverty reduction and shared prosperity, systematic country diagnostic. Washington, D.C.: World Bank.

Masri, S. and Srour, I. (2014). Assessment of the impact of Syrian refugees in Lebanon and their employment profile. International Labour Organization.

Matta, S. (2014). New Coincident and Leading Indicators for the Lebanese Economy, World Bank Policy Research Working Paper No. 6950.

UNDP and Lebanon Support. (2015). Between local patronage relationships and securitization: The conflict context in the Bekaa region. Conflict analysis report written by Muzna al- Masr.

UNDP and MOSA. (2007). Development of Mapping of Living Conditions in Lebanon, 1995-2004: A Comparison with the Results of Mapping of Living Conditions in Lebanon, 1998. United Nations Development Program (UNDP) and the Ministry of Social Affairs (MOSA) of the Republic of Lebanon.

UNDP and UNHCR. (2015). Impact of Humanitarian Aid on the Lebanese Economy. Beirut: UNDP. Available at: http://reliefweb.int/report/lebanon/impact-humanitarian-aid-lebanese-economy.

UNDP. (2008). Poverty, Growth and Income Distribution in Lebanon. Available at: http://www.lb.undp.org/content/dam/lebanon/docs/Poverty/Publications/Poverty,%20Growth%20and%20Income%20Distribution%20in%20Lebanon.pdf.

UNHCR. (2017). Health Access and Utilization Survey among Syrian Refugees in Lebanon. Available at: https://reliefweb.int/sites/reliefweb.int/files/resources/LebanonHealthAccessandUtilisationSurvey2017UNHCR.pdf.

Verme, P., Gigliarano, C., Wieser, C., Hedlund, K., Petzoldt, M. and Santacroce, M. (2015). The Welfare of Syrian Refugees: Evidence from Jordan and Lebanon. Available: https://openknowledge.worldbank.org/handle/10986/23228

World Bank. (2012). Lebanon – Good jobs needed: the role of macro, investment, education, labour and social protection policies (MILES) - a multi-year technical cooperation program. Washington, DC: World Bank. Available at: http://documents.worldbank.org/curated/en/230521468089355499/Lebanon-Good-jobs-needed-the-role-of-macro-investment-education-labour-and-social-protection-policies-MILES-a-multi-year-technical-cooperation-program.

World Bank (2013). Lebanon – Economic and Social Impact Assessment of the Syrian Conflict. Washington and Beirut: The World Bank.

World Bank. (2014). Downside Risks Materialize. Lebanon Economic Monitor, Fall 2014. Washington, D.C.: World Bank. Available at: https://openknowledge.worldbank.org/handle/10986/21969.

Yaacoub, N., Daher, M., Jolliffe, D. and Atamanov, A. (2015). Snapshot of Poverty and Labour Market Outcomes in Lebanon Based on Household Budget Survey 2011–2012. Washington, DC: World Bank. Available at: https://openknowledge.worldbank.org/handle/10986/23797.

Annex I: MICS survey details

Name of Dataset	MICS 3 (Lebanese)	MICS 4 (Palestinian refugees in Lebanon)
Publisher	UNICEF and CAS	UNICEF and the Palestinian Central Bureau of Statistics
Date of data collection	2009	May–July 2011
Data of publication	2011	June 2012
Type of survey/data	Household questionnaire Married women aged 15–49 Children under 5 years	Palestinian household questionnaire - Living in Lebanon Palestinian women aged 15–49 years Palestinian children aged 0–14 years
Sample size	15,181 households 7,560 women 3,439 children	5,190 households
Data representation	Nationally representative	Representative of Palestinian Camps and Gatherings
Key indicators	• Household: household listing form, education, water and sanitation, household characteristics, child labour, child discipline, handwashing, salt iodization and water testing, immigration. • Women: background, marriage, child mortality (with birth history), HIV/AIDS, desire for last birth, maternal and new-born health, illness symptoms, contraception, unmet needs, female genital mutilation/cutting, and attitudes towards domestic violence. • Children under 5 years: age, birth registration, early childhood development, breastfeeding, care of illness, immunization, and anthropometry.	• Household: household listing form, education, water and sanitation, household characteristics, child labour, child discipline, handwashing, salt iodization and water testing. • Women: background, marriage, child mortality (with birth history), HIV/AIDS, desire for last birth, maternal and new-born health, illness symptoms, contraception, unmet need, female genital mutilation/cutting, and attitudes towards domestic violence. • Children under 5 years: age, birth registration, early childhood development, breastfeeding, care of illness, immunization, and anthropometry.

Name of Dataset	MICS 3 (Lebanese)	MICS 4 (Palestinian refugees in Lebanon)
Disaggregation variables	Region, age, gender and woman's education level	Gender, region (5 regions), age, education level of mother, wealth proxy indicator
Programmatic use of dataset	Update existing data, production of reports on education, environment, and education in Lebanon by CAS	Update existing data, monitor progress of indicators, and strengthen technical expertise
Accessibility of data	Need permission	Need permission
Name of dataset	http://www.cas.gov.lb/images/Mics3/CAS_MICS3_Household_Questionnaire.pdf. http://www.cas.gov.lb/images/PDFs/SIF/CAS_Population_and_housing_In_Lebanon_SIF2.pdf http://www.cas.gov.lb/images/PDFs/SIF/CAS_Education_In_Lebanon_SIF3.pdf. http://www.cas.gov.lb/images/Mics3/MISC3_new/Environment-Housing%20characteristics%20in%202009.pdf	http://microdata.worldbank.org/index.php/catalog/1994/study-description

البيانات