

Economic Research Forum

Research Initiative for Arab Development (RIAD)

Expert Group Meeting for Data Harmonization in the ERF Region

Summary of the meeting

April 29-30, 2010

The Economic research Forum (ERF) is undertaking a wide research project to advance understanding about the patterns and causes of inequality in the Arab region. A key component of this research is related to collecting and harmonizing household surveys from as many countries in the region as possible and making them comparable to international statistics on household living standards distribution.

ERF has already started the process by securing a number of household income and expenditure surveys and harmonizing these data following current best practice. Nevertheless, the differences in the national household surveys made available to ERF, significantly challenge the compilation of microdata to a common set of concepts and the construction of comparable inequality statistics. The data treatment has shown the difficulty of achieving the harmonization of household surveys in practice.

In this context, ERF organized recently an expert group meeting on data harmonization in the Arab region. The meeting was held in Cairo on April 29-30, 2010 with the participation of academic researchers and representatives from national statistical offices as well as from international organizations.

The aim of this meeting was to exchange ideas and best practices about how to address the conceptual and practical issues related to the production of harmonized household survey microdata and of comparable statistics on income and expenditure distribution.

The meeting was organized in six sessions structured to trigger and frame the discussions on the underlying quality and comparability of household survey data in the hope that it will help to expand the availability of consistent distribution statistics in the Arab region.

Three sessions were devoted to the conceptual and methodological difficulties encountered in producing robust well being distribution statistics; two sessions devoted to the experience of international research projects in assembling microdata from household surveys on a harmonized basis and in addressing the lack of comparability; and one round table session in which representatives from the national statistical offices raised the challenges they faced in improving the quality and reliability of their own household survey data and welcomed the potential collaboration with research institutes to enhance national household statistics in the Arab region.

The content of each of the sessions is summarized as follows:

Session 1:

The meeting opens with welcoming and introductory remarks by Ahmed Galal the managing director of ERF, and a brief description of the meeting objectives and structure by Nadia Belhaj Hassine senior economist at ERF.

The first session involved two speakers, Abu Bakr El-Guindi the President of the Central Agency for Public Mobilization and Statistics (CAPMAS) and Peter Lanjouw lead economist in the World Bank.

Abu Bakr El-Guindi made a Keynote speech in which he highlighted the importance of the topic of the meeting for data users and producers and discussed the challenges of ensuring data comparability in Arab countries due to the use of different measurement methods across countries, and even over years, especially in the case of expenses on durable and semi-durable goods. He also raised the issue of the confidentiality of the surveys data. One of the fundamental principles of official statistics is to refrain from disclosing individual data for the purpose of protecting confidentiality rights of data providers of all kinds. He emphasized nevertheless that access to anonymized survey data may be allowed based upon specific requirements and careful check of the purpose of the data request and intended use.

Peter Lanjouw tackled the issue of the comparability of household consumption statistics and stressed the importance of addressing the lack of data comparability, across countries as well as over time, for producing robust distribution indicators and for tracking welfare changes. He reviewed the main sources of differences in national household surveys and showed how these differences affect the welfare comparisons. While a satisfactory solution for addressing the obstacles to achieve complete harmonization across countries and over time is yet to be found, Lanjouw provided some guidance on how to address the lack of comparability of household consumption survey microdata. He nonetheless emphasized that these solutions need to be applied carefully and clearly documented, and pointed to the importance of checking the robustness of the distribution statistics to the assumptions and choices made.

The discussion that followed focused on two main issues: 1) the data confidentiality problem and 2) the factors which affect data quality and comparability and the pitfalls of which data producers and users must be aware.

Regarding data confidentiality: while there is no law against disseminating anonymized microdata data, rules and regulations governing the data release remain necessary to prevent misuse by some researchers.

Concerning the data quality and comparability issues: the changes in the design of household surveys, coding schemes, questionnaires etc., underlying the differences in data availability and quality, may be explained by the need of extending the range of individuals characteristics and expenditures components collected and the desire of enhancing the response rates. While some solutions, may be adopted to improve the accuracy of comparisons of well being indicators and their distribution, a more promising solution would be to promote the collaboration between National Statistical Offices (NSO) and research institutes to improve the data collection practices and try to solve the practical problems related to the design of the household survey and sample, the questionnaires, the data cleaning and harmonizing etc.

Session 2:

In the second session, two speakers, Jérémie Gignoux researcher at Paris School of Economics and Marwan Khawaja, section chief of social statistics at UN-ESCWA, talked about the issues of measuring some monetary and non monetary dimensions of inequality using imperfect data.

Jeremie Gignoux discussed the measurement of circumstances-related inequalities. A growing literature recognizes that inequality in the distribution of welfare is in part due to differences in basic life chances which are outside the individual control. The measurement of these type of inequities require observations on the individual's predetermined circumstances such as parents education and employment status, race, place of birth etc. These data are missing in the available income and expenditures household surveys in Arab countries. Gignoux presented an approach that combines information from household surveys and Demographic and health Surveys (DHS) to circumvent the problem of data missing. He provided an illustration in the case of Turkey.

Marwan Khawaja focused his presentation on the disparities in child health in 18 Arab countries during the 1990s. He used microdata from different national surveys to investigate the extent of within-country child health disparities and to examine the impacts of gender and socio-economic disadvantages on the health of Arab children. Khawaja highlighted the importance of data availability and quality and their impact on the accuracy and the consistency of his findings. He explained how he dealt with some of these issues.

Session 3:

The third session involved three experts and researchers in household data analysis, Paul Alkemade microdata expert and statistician at Luxembourg Income Study (LIS), Kinnon Scott senior economist at the World Bank and Nicolas Ruiz principal research assistant at OECD. The session was dedicated to review the experience of the LIS, the Living Standards Measurement Study (LSMS) group and the OECD in harmonizing the household surveys.

Paul Alkemade introduced the experience of the LIS in collecting, classifying and standardizing national household income survey data. He highlighted the main challenges related to constructing harmonized micro databases, such as the treatment of missing variables, imputed variables, the disparities in population coverage, and the differences in the level of details in collecting income and expenditure items, individual socio economic characteristics etc.

He explained how LIS dealt with the data confidentiality issue by providing a user friendly and secure remote access to its micro-databases.

Kinnon Scott talked about the methodology of data collection, cleaning, harmonization and dissemination developed by the LSMS group. She exposed the main problems impeding the data accessibility and she gave some recommendations that help to deal with the lack of comparability of survey data. Finally, she presented a brief description of the Comparative Living Standards Project (CLSP) which allows access to harmonized and well-documented micro-datasets from LSMS surveys in a user-friendly way. The core of CLSP is an interactive on-demand query system of

cross-country household-level data sets that responds to users requests with tables and analyses created by the users.

Nicolas Ruiz talked about the effort made by the OECD to improve the harmonization and dissemination of household labor, income and wealth microdata. He provided an overview of the projects currently conducted by the OECD for setting a remote access to harmonized labor force and migration statistics; for the revision of the standards on the conceptual and practical problems related to the production of internationally comparable income distribution statistics; and for creating a multidimensional framework for the joint appraisal of income, consumption and wealth.

Session 4:

The focus of the fourth session was on the harmonization of household income and expenditure surveys in the Arab region. Nadia Belhaj Hassine and Christiane Wisa started the session with a description of the process of data collection and harmonization used by ERF. The presentation highlighted the practical difficulties faced in compiling the micro data from different surveys to some common standards and the challenges of producing across countries and over time comparable income/consumption distribution statistics.

The main issues were about the classification of individuals' socioeconomic characteristics and of consumption items into comparable categories, the treatment of expenses on durables and on housing, and the adjustments of household income and expenses to account for differences in household size and in cost of living.

The discussion that followed was about trying to address some conceptual and practical difficulties related to data harmonization and to suggest some recommendations to deal with the quality and comparability issues underlying the different national household surveys made available to ERF and for improving the data comparability. While it emerged from the discussion that there are no clear cut solutions to all the issues raised and that the complete harmonization is very hard to attain, some guidance based on best practices were provided to address some of the main issues. The participants pointed nevertheless to the importance of clearly setting the hypothesis adopted, of checking their consistency with the structure of the Arab countries economies and of carefully documenting the standardization process. The participants recommended to analyze the sensitivity of the results to the assumptions and choice made.

Session 5:

Session 5 involved two researchers, Djavad Salehi-Isfahani professor at the Virginia Tech University, and Branko Milanovic lead economist at the World Bank. The focus of the session was on the limitations associated with using household and labor survey data in the MENA region.

Djavad Salehi-Isfahani started his presentation by illustrating his experience in obtaining the Iranian household survey. He discussed the relations between statistical offices and researchers and pointed to the difficulty of obtaining microdata for non-institutionalized researchers. Data confidentiality is motivated by the concerns of misuse. Isfahani then switched to talking about the comparability issues in labor and education statistics. He pointed to the importance of knowing the structure and the functioning of the educational and labor institutions in harmonizing

the data. He mentioned that working in a regional perspective is easier because of the cross countries similarities in terms of social and cultural norms, and in terms of the definition of some variables such as family and marriage etc.

Branko Milanovic introduced his presentation by illustrating three different concepts of inequality: The first concept looks at inequality as reflected in differences in average incomes between countries; the second concept uses these same differences in incomes between the countries but weighs each country by its population; and the third concept is global inequality which considers only individuals in the world regardless of their nationality. The analysis of this last concept is based on household survey data and goes back to the individual as the unit of analysis, ignoring country boundaries.

Milanovic showed that these different concepts lead to different conclusions about the evolution of world inequality. He highlighted the advantages and difficulties of each concept focusing on those of concept 3.

Then Milanovic discussed the issue of inequality in the Arab countries. He indicated that the Arab region is underrepresented in the World Income Distribution database since less of one third of the Arab population is included in this dataset.

He compared inequality in eight Arab countries to the rest of the world and showed that inter-personal inequality in these countries is 20% higher than in European Union and in the United States. In a figure depicting the relative position of different Arab income groups compared with the same income groups in Germany and Spain, Milanovic explained that 50% of Jordanians are better off than the poorest in Spain, while 25% of Egyptians and Moroccans are better off than the poorest Spaniards.

Session 6:

The round table session was designed for interaction between the representatives from six National Statistical Offices (NSO) in the Arab region; namely Egypt, Jordan, Lebanon, Libya, Palestine and United Arab Emirates; and the other meeting participants.

The session was moderated by Mustapha Nabli Senior Advisor at the World Bank and started by three presentations around the theme of the collaboration between the NSOs and researchers. The first presentation made by Ola Awad, acting president of PCBS Palestine, detailed the strategy adopted by the Palestinian statistical office to disseminate the household surveys microdata for public use. Where her objective by broaching this subject was to show available and potential means to provide microdata for public use in a manner to ensure confidentiality of data on one hand and make the greatest interest to researchers and all users of statistical data, on the other side (balancing between users needs and data confidentiality). In the second presentation, Baha-Eddine Abdelhamid Mortagui deputy director of CAPMAS Egypt reviewed the process followed by CAPMAS to enhance the quality and the reliability of the national survey data and the challenges faced in this process. Maral Tutelian general director of the CAS Lebanon explained in the third presentation, the importance of harmonizing household surveys within and among NSOs and between NSOs and international agencies. She discussed the efforts of the NSOs in improving the pertinence and robustness of their national statistics and in adopting statistical methods compatible with international standards and national needs. Tutelian emphasized the importance of the collaboration between researchers and NSOs in this area and clarified their respective roles in harmonizing household surveys and in improving the availability and quality of data. She stressed the importance of promoting the collaboration at three levels: 1) the design of the questionnaires with respect to concepts, classifications and methodologies, and to topics pertaining to

national and international needs; 2) the treatment and analysis of data and the drafting of qualitative and quantitative reports; and 3) the dissemination of the results and findings in a timely and user friendly way.

The panel presentations were followed by a discussion in which Salem Abu Aicha director of the GIA Libya, talked about the recent experience of some Arab countries in conducting household surveys and reviewed some of the difficulties encountered. He emphasized the importance of the collaboration with researchers for addressing the challenges facing the NSOs in designing the surveys and in analyzing the data. Ali Tafour Statistics analyst at CSD of United Arab Emirates, talked about the experience of UAE in disseminating the raw household data. Ahmed Mowafi from DOS Jordan emphasized the importance of data dissemination to enhance the use of household survey data and the improvement of their quality.

A rich discussion offered then the opportunity to the meeting participants to ask questions about the strategies employed by the NSOs in designing the household surveys, collecting and treating the data and the areas of potential collaboration with researchers.