Existing data infrastructure and new information technologies to build sustainable data for sustainable development

The case in Israel

International Seminar on World Statistics:
"Sustainable Data for Sustainable Development"
Xi'an, China, October 2015
Existing data infrastructure- administrative data

Israel has rich administrative records:

Population register, income register, education register,
Transportation register, Housing register, Business register,
All registers with identifiers which facilitates linkage,
Geographical mapping at fine (very detailed) resolution.

Next population census planned for 2020 (last in 2008), based on the population register with supplementary samples at small statistical regions to estimate over and under counts.

Next Census of Agriculture planned for 2017 (last in 1981!!). Currently building a framework of all farmers in the country.
Existing data infrastructure - surveys

We run more than 50 surveys:

**Household and Person surveys:** Labor Force, Family Expenditure, Victimization, Social (with added changing topics), Consumer confidence, Longitudinal, PIAAC,…

**Business surveys:** CPI, Business tendency and consumer prices, Industrial production, Job vacancies, Construction, Transportation, Tourism, Commerce, Bio-technology, R&D, Innovation in the business sector, Environmental surveys (waste and emission,…), Direct investment,…

**Surveys in education, Health Surveys,**…
New production: Indices of Quality of Life

Goals

To present the **public** with a broad picture for examining and understanding wellbeing and policy outcomes.

To present the **government** with a broad picture for retrospection and evidence-based policy planning.
Main uses on a national level

**Encourage inter-sectorial partnerships**
- Joint goals for government, the private sector and the general public

**Inspection tool**
- Assess progress relative to targets
- Feedback for examining effectiveness of various processes

**Central component in policy planning**
- Serve as framework for gov’t and other sectors
- Basis for setting measurable targets
Wellbeing Indicator Set

9 Domains

8 indicators per domain

Headline indicator per domain
<table>
<thead>
<tr>
<th>Domain</th>
<th>Leading Ministry</th>
</tr>
</thead>
<tbody>
<tr>
<td>Material Standard of Living</td>
<td>Finance Ministry</td>
</tr>
<tr>
<td>Civic Engagement and Government</td>
<td>Prime Minister’s Office</td>
</tr>
<tr>
<td>Employment and Work-Life Balance</td>
<td>Ministry of Economy</td>
</tr>
<tr>
<td>Personal and Social Well-Being</td>
<td>Ministry of Welfare</td>
</tr>
<tr>
<td>Personal Safety</td>
<td>Ministry of Public Security</td>
</tr>
<tr>
<td>Infrastructure and Housing</td>
<td>Ministry of the Interior</td>
</tr>
<tr>
<td>Health</td>
<td>Ministry of Health</td>
</tr>
<tr>
<td>Environment</td>
<td>Ministry of Environmental Protection</td>
</tr>
<tr>
<td>Education</td>
<td>Ministry of Education</td>
</tr>
</tbody>
</table>
Small Area Estimation (SAE)

- Information required not only at the global (national) level but for small statistical regions.
- Israel is divided into ≈3,000 statistical region of approximately equal size (≈ 3,000 persons).
- Sample sizes in statistical regions too small to allow direct accurate estimates based on data collected for the region. Many regions have no samples.
- Requires the use of SAE models that use administrative data and borrow information from neighbouring regions and past surveys.

Examples: poverty mapping, disease incidence, education achievements, employment rates, environment quality and protection,…
New information technologies- Big data

We are in the process of obtaining sale prices from all big supermarket chains in Israel (will be used for CPI).

We are looking into ways of obtaining data from cellular phone companies. (Potentially used to replace survey of traveling custom.)

Caution: **Big Data → Big Problems** (Pfeffermann, *JSSAM*, 2015)

- Coverage/selection bias (we are talking of official statistics)
- Data accessibility, new legislation?
- Privacy (data protection), disclosure control
- Data storage, new sampling algorithms
- Computation and Analysis, new measures of error
- Linkage (integration) of different files.
**Big data for official statistics - summary remarks**

New expensive computing facilities, new data processing techniques, new linkage methods, new visualization methods, new sampling methods, new analytic methods, new measures of error, new disclosure control procedures, new legislation,…

Big potential advantages: timeliness, much broader coverage (possible coverage bias), no sampling frames, no questionnaire, no interviewers,…

- Constant decline in response rates in traditional surveys,
  \[\Rightarrow\text{use of big data inevitable.}\]

**Good news:** Big data will just grow bigger and bigger.
National Statistical System (NSS)

➢ The last slide is an example of cooperation between the Israel Central Bureau of Statistics (ICBS) and other Government offices.

➢ We are working hard in setting up more general cooperation in the production of official statistics by all members of the NSS.

GOALS:

• Set up a long term cooperation between the ICBS and other offices producing official statistics,

• Establish Statistical code of practice for the production of official statistics following international standards, including training,

• Develop methods for quality control at the ICBS and NSS

• Establish assistance and consultation of the ICBS to other offices.

➢ By law, the ICBS supervises all offices producing official statistics.