



***From physical area to virtual lists:  
Toward an administrative census in Israel***

*Group of Experts on Population and Housing Censuses  
Geneva, Switzerland, 26-28 September 2018*

*Olivia Blum ([blum@cbs.gov.il](mailto:blum@cbs.gov.il))*

# Background

- The **1995** Census of Population and Housing in Israel was planned to be the last **traditional**, door-to-door census, and to be the starting point of a gradual process toward an administrative census.
- In **2008** it was an **Integrated Census**, where a geo-demographic file was generated by improved Central Population Register (CPR), followed by evaluation surveys that corrected the population estimates and provided the socio-economic information.
- In **2012-2014** a **rolling census** did not generate reliable population estimates and was stopped.
- In **2017** a 1<sup>st</sup> phase experiment of a rolling census, with different methodology, was not culminated in satisfactory results, and
- At the beginning of **2018**, a decision has been made to move back to the **administrative census** path.

# Total Population Frame

A conceptual and practical leap toward an administrative census can be done, due to two enablers that are crucial to the total population frame:

1. **All people** in the country have administrative records;
  - Citizens are registered in the CPR and
  - Foreigners are reported in functional records like work permits and visas.
2. **Entries and departures** to and from the country are well recorded.

The administrative sources provide the potential population residing in Israel, along with the ability to identify Israelis who emigrated abroad, and illegal foreigners who overstayed their visa validity.

# Challenges/Questions

*"The 'usually resident population' of a country is composed of those persons who have their place of usual residence in the country at the census reference time and have lived, or **intend** to live, there for a **continuous** period of time of at least 12 months.*

*A 'continuous period of time' means that absences (from the country of usual residence) whose durations are shorter than 12 months do not affect the country of usual residence". (UNECE, 2015)*

- Globally, travelers are prone to double counting.
- There is no symmetry in definitions of in-migration and out-migration.
- Is there a need for standardization of population definitions in censuses (stock at census day) and in the ongoing demographic estimates (periodical flows)?

# *Spatial Address: Who is where?*

Given the population frame , each person is to be ascribed to the smallest spatial entity, up to the dwelling unit level (x,y,z), in order to form households and to enable flexible aggregates.

Address problems in the admin files are of two types: People move from one address to another without reporting, people have more than one address and only one should be defined as a census address.

Correction methodology:

1. “Choosing” address from different administrative files;
2. Small Area Estimation, based on the Geo-Demographic Administrative Census (GDAC) and on a designated survey.

# Administrative Address

- When different admin sources provide different addresses to the same person, choosing the address can be either **single selection** or assigning **probability scores** to each one (summing up to 1) (Zhang, 2018).
- The **selection process** will be based on machine learning, using data from different sources, current and historical, known rules and relations between all relevant entities, and on true data to validate the outgoing algorithms.
- An ongoing **improvement process** will engage additional (qualitative) admin files and additional features learnt from previous process, and will ignore counter-productive files and features.

# Challenges/Questions

- The Israeli CPR carries about 20% erroneous addresses on the individual level, but only 8% on a local authority level (people tend to move within the same locality).
- Similar people move in and out of an address and do not influence the estimates,
- Internal migration can be generated on an aggregated level, yet, out-migration is defined on the individual level.

Where should we define the investment threshold regarding the individual address?

How long should we maintain the **register-like** perception, before turning to the **statistical** estimates, in the census process?

# *The Sample Survey*

The working assumption that the population frame is known, allows for disengagement from the area!

- The population frame can serve as the sampling frame for the survey. The sampling unit will be the individual and the research unit will be his whole household.
- Data collection can be done from afar (internet and telephone).
- The questionnaire flow is tailor-made, by avoiding selectively, questions whose answers are already known from the admin files (like education of young adults and country of birth of most foreign-born residents).
- SAE will be mainly based on the GDAC and the survey results, and correction weights will be attached to the GDAC records where needed.
- PES can be truly independent of the census process and rely on area sampling.

# Challenges/Questions

The SNA2008 definition of the boundaries of official statistics is:

*“Any geographic area or jurisdiction for which statistics are required,” and “the area under the effective economic control of a single government”.*

Administrative censuses (and the planned census in Israel) refer to the population first. They do not refer simultaneously to **population and area**, but rather indirectly to area for which population information is provided in the administrative registers.

- Do the admin files cover the required area, although does not necessarily correspond to the nation-state area?
- Should governments take into account wide margins of uncertainty regarding the population to be served?

# *Concluding Remarks*

- The late shift in the census methodology in Israel brings about many risks to be carefully managed.
- Referring to identified population in a known population frame implies substantial change in the concept of a census.
- Direct data collection can be significantly reduced since administrative information will be used to navigate the identified respondent to provide the missing data only.
- The availability of mass information and private data is to be harnessed to statistical processes and processing, carried out by data scientists.
- The theoretical and socio-economic implications of population and area definitions, and their influence on policy making, should be further discussed.

# Thank You!

