

UN TASK FORCE ON DATA INTEGRATION FOR DISAGGREGATED STATISTICS ON INTERNATIONAL MIGRATION

Jason Schachter, Co-Chair
Chief, Net International Migration Branch
U.S. Census Bureau

Nino Ghvinadze, Co-Chair
Data Analyst, Public Service Development Agency
Georgia

United Nations Expert Group Meeting
New York
July 1-3, 2019

This presentation is released to inform interested parties of ongoing research and to encourage discussion of work in progress. Any views expressed are those of the authors and not necessarily those of the U.S. Census Bureau.



U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov



PUBLIC SERVICE
DEVELOPMENT AGENCY

Task Force Background and Objectives

- UN Country Consultation identified a need for guidance on data integration, particularly for non-register based countries
- Objective: advance methodologies to produce data that are sufficiently disaggregated for the measurement of international migration
- Proposing a UN Task Force on Data Integration with two sub-task forces
 - (1) Macro-data integration
 - Analysis which incorporates results based on data which are aggregates (statistics) of individual-level records
 - (2) Micro-data integration
 - Integration of data based on linkage/matching of individual records
- Complement, not duplicate, efforts of the 2019 UNECE Task Force on Data Integration

Macro-Data Integration

- Practical applications: methods to produce disaggregated estimates of international migration via the integration of migration statistics derived from multiple data sources
- Examples:
 - Use different data sources to produce different sub-components of international migration estimates (sometimes referred to as data “compilation”)
 - Adjust migration estimates from one data source using estimates from one or more other data sources (sometimes referred to as data “triangulation”)
 - Combine different data sources to produce migration estimates at different levels of geography (e.g. survey data for national totals/administrative data for sub-national) or by various characteristics of the population (e.g. Bayesian methods to provide estimates for areas/characteristics with scarce observations)
- Task force will:
 - (a) produce technical materials that provide general guidance on methods to integrate data at the macro level to improve estimates of international migration
 - (b) produce technical materials on macro-data integration for migration statistics

Specific U.S. Example

- Combined different data sources (“big data” and survey data) to produce a final estimate of net migration from Puerto Rico in wake of Hurricane Maria (September 2017)

Measuring the Impact of Hurricane Maria on Migration between Puerto Rico and the Mainland United States



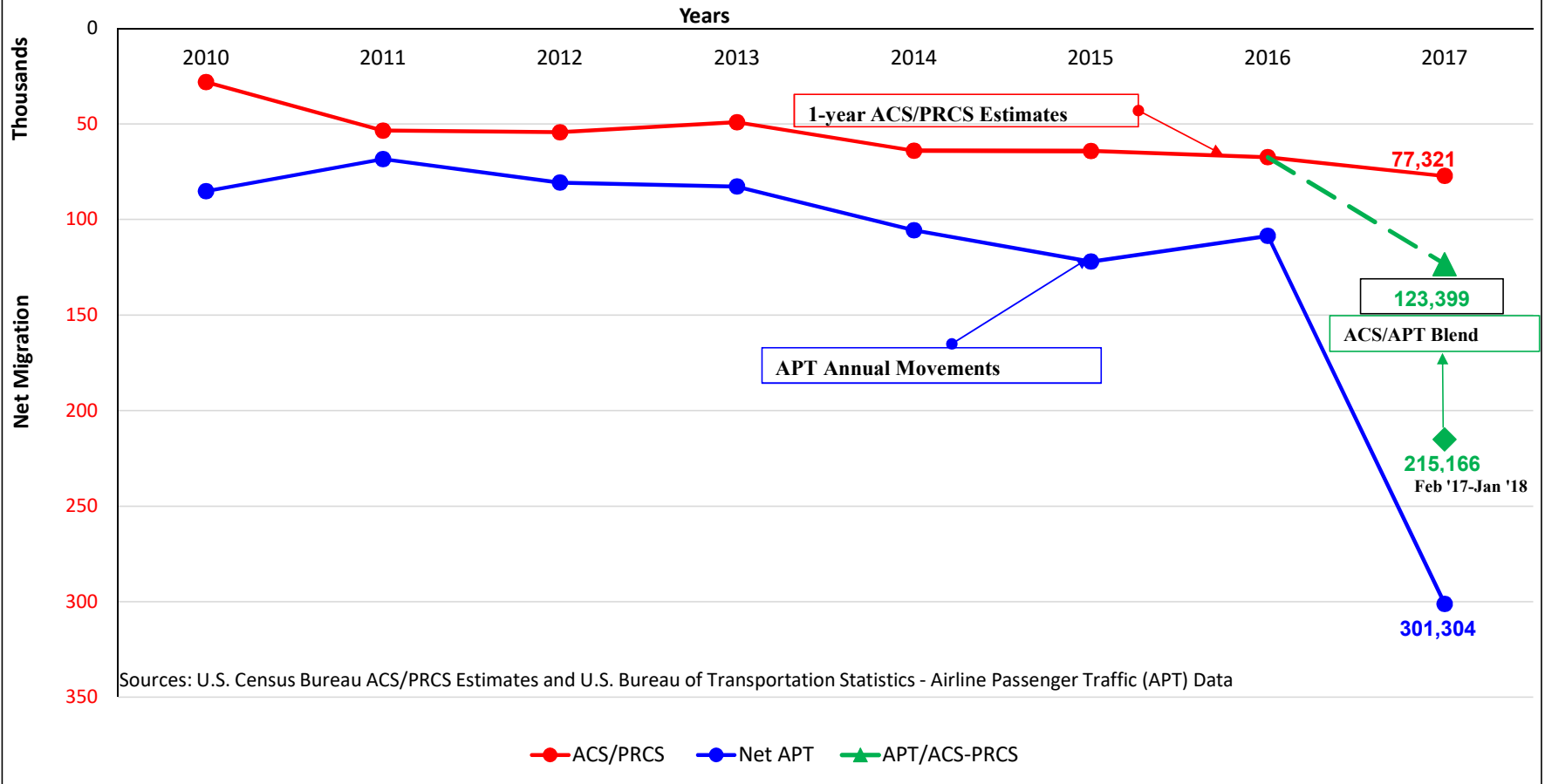
Source: <https://islandsofpuertorico.com/puerto-rico-map/>



U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
[census.gov](https://www.census.gov)

Fig. 1

Net Puerto Rico to U.S. Migration: 2010-2017



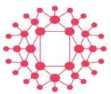
Micro-Data Integration

- The second sub-task force would focus on processes needed for successful micro-data integration, particularly for countries without access to population registers
- Possible topics could include:
 - How to access data from different agencies
 - How to merge data without unique identifiers
 - Methods to validate integrated data and ensure quality control
 - How integrated data could be compiled to follow UN standards and definitions for international migrants
- TF would not duplicate existing guidelines on data integration, but would complement the prior UNECE Task Force, producing information about procedures and general guidance to set up an integrated microdata system in a country



Georgia Example

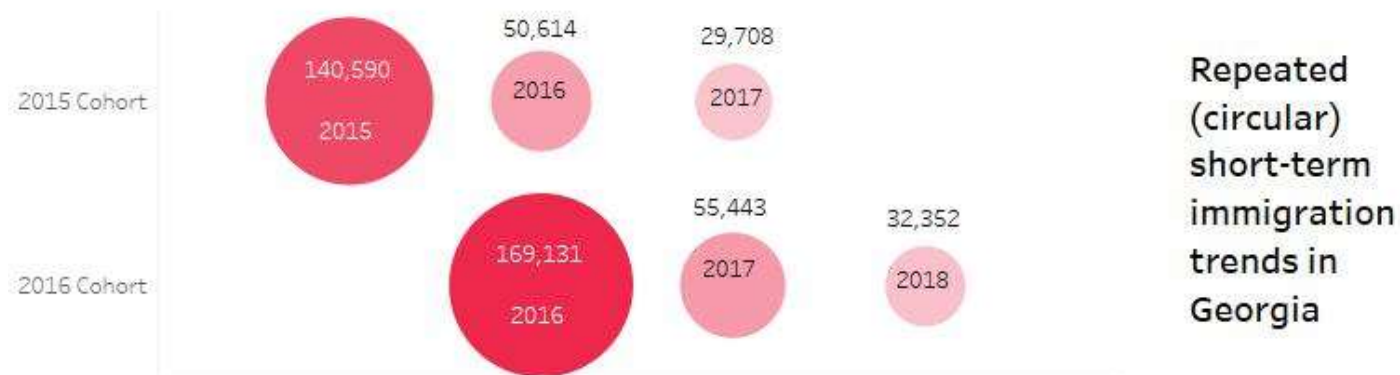
- An analytical system to integrate administrative data on micro level to improve immigration data in a broad sense.
 - Administrative data from 9 agencies is collected;
 - Data update and integration happens on regular basis (weekly/monthly);
 - New data sources can be add (integrated to the system) as it develops;
 - Used for immigration trend analysis as BI tool;
 - 10+ test reports produced.



Example: short-term (circular) movements in Georgia

- 2015 & 2016 cohorts observed for 3 years;
- Cumulative visits from 1 to 6 months in 1-year period;
- Administrative data sources used: border control, residence permits, nationalization.





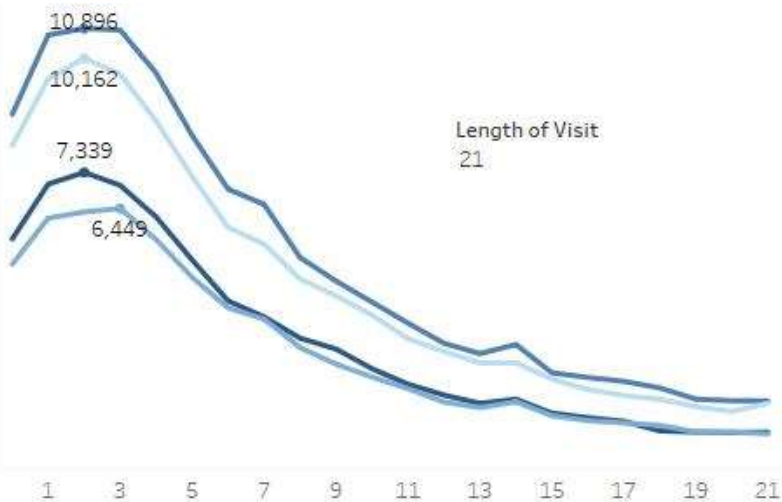
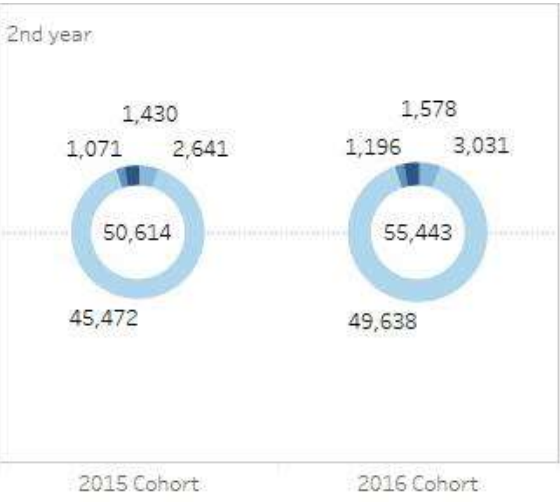
2nd year (2015 & 2016 cohorts)

	2016 Coho..	2015 Coho..
Russian Federati..	23,698	22,345
Azerbaijan	8,796	7,693
Turkey	6,472	6,253
Armenia	4,819	4,581
Ukraine	2,254	1,948
USA	1,101	925
Greece	1,097	1,061
Iran Islamic repu..	958	513
Germany	768	721
Israel	565	463
Kazakhstan	524	391
The Great Britain	420	376
France	362	349
Stateless person	274	276
Other countries	3,335	2,719

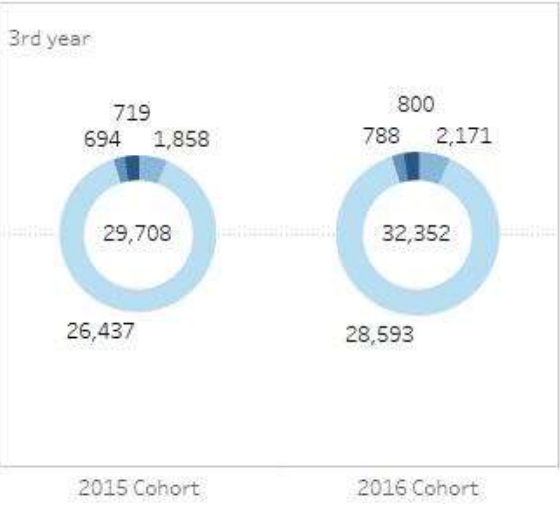
3rd year (2015 & 2016 cohorts)

	2016 Co..	2015 Co..
Russian Federation	14,792	13,579
Azerbaijan	5,215	4,639
Turkey	3,742	3,769
Armenia	2,709	2,652
Ukraine	1,050	907
Greece	666	620
USA	568	478
Germany	458	398
Iran Islamic republic	470	312
Israel	320	265
The Great Britain	247	229
Kazakhstan	225	223
France	211	212
Stateless person	159	145
Other countries	1,520	1,280





- Individual and cumulative lengths of visit;
- Legal status during the visit.



	1-2 months	2-3 months	3-4 months	4-5 months	5-6 months
2015(2nd year)	23,235 45.91%	13,485 26.64%	6,656 13.15%	3,946 7.80%	3,292 6.50%
2015(3rd year)	12,643 42.56%	8,679 29.21%	4,303 14.48%	2,363 7.95%	1,720 5.79%
2016(2nd year)	25,163 45.39%	14,748 26.60%	7,369 13.29%	4,450 8.03%	3,713 6.70%
2016(3rd year)	14,880 45.99%	9,321 28.81%	4,391 13.57%	2,164 6.69%	1,596 4.93%

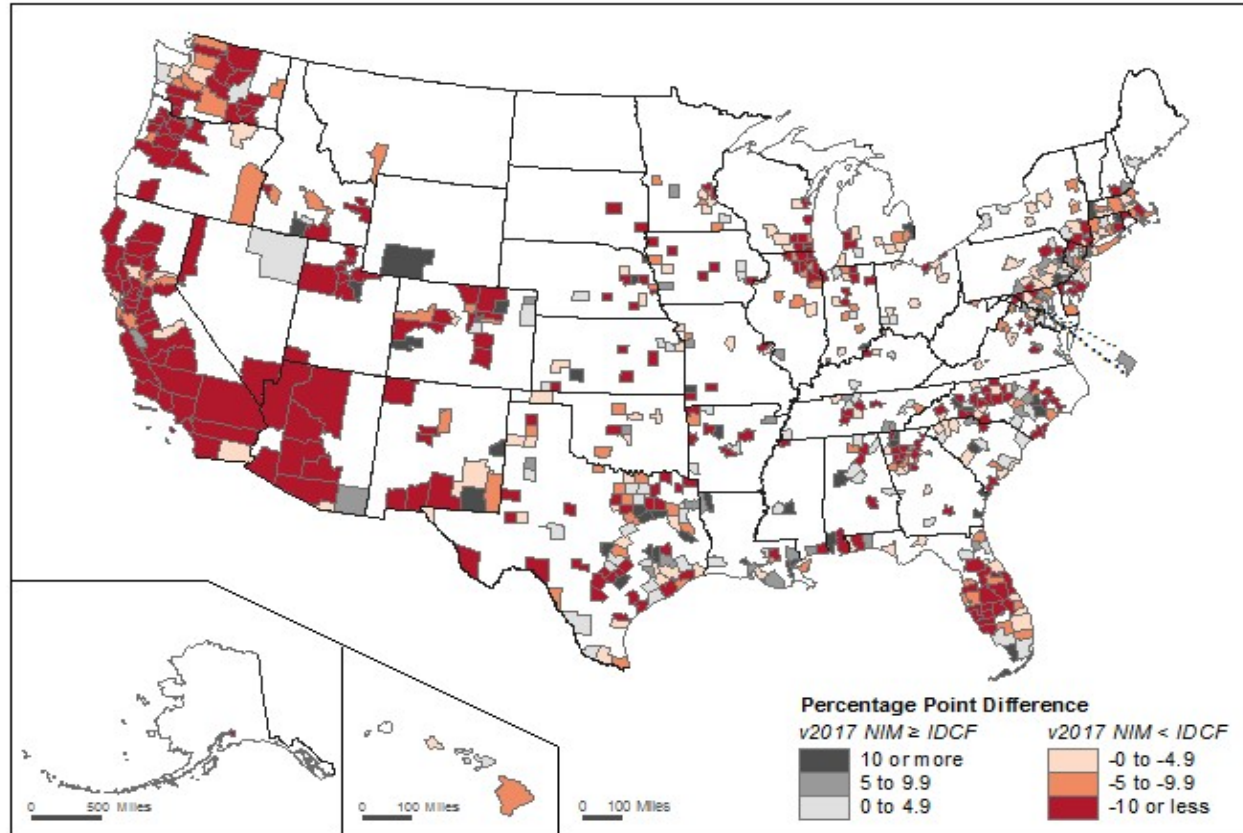
U.S. Example

- Linking multiple data sources to improve measurement of the subnational distribution of international migrants
 - IDCF (Immigrant Demographic Characteristics File)

IDCF

NUMIDENT	IRS	Census 2010
<ul style="list-style-type: none">• Age• Sex• Country of Birth	<ul style="list-style-type: none">• Geography	<ul style="list-style-type: none">• Race• Hispanic origin <p>*Race and Hispanic origin are imputed for the foreign born who entered after 2010.</p>

Percentage Point Difference between Net International Migration (NIM) Estimates and IDCF Data in Hispanic Immigrant Population by County: 2017



Source(s): U.S. Census Bureau 2017 Net International Migration Estimates and IDCF.

Task Force Membership

- Call for member participation
- The Task Force would consist of two co-chairs to guide work on each respective sub-task force (members can serve on both sub-task forces, or just one, as desired)
- Task Force members are asked to contribute actively to the work by providing materials on case studies and drafting some parts of the technical materials
- Work schedule TBD