Economic Commission for Europe

Conference of European Statisticians

Work Session on Migration Statistics
Geneva, Switzerland
18-20 May 2016
Item 9 of the provisional agenda
Asylum seekers and refugees

An overview of refugee and asylum statistics in the United States

Note by the U.S. Census Bureau *

Abstract

As global events continue to increase the total number of persons seeking asylum and refuge throughout the world, the need to accurately measure this population for demographic and policymaking purposes increases. Though the United States is relatively far removed in terms of geographic distance from the epicenter of the current refugee crisis, these events still impact U. S. immigration policy. For example, the Executive Office has significantly increased the total number of refugees allowed to be resettled in the United States in the coming years, from a level of 70,000 in 2015, to 85,000 in 2016, and up to 100,000 by 2017, primarily increasing the number of relocated refugees from Syria[1].

The United States Census Bureau does not collect information on the number of refugees and asylum seekers admitted to the United States, as these rather come from administrative data collected by the U.S. State Department (refugees) and Department of Justice (asylum seekers). Given refugee and asylum status result from an administrative procedure, the annual number admitted or granted these statuses is well documented, though the characteristics and long-term integration of these groups is less clear.

This paper discusses recent historical levels of refugee and asylum seeker admittance to the United States, by country of origin, as well as the data sources used to collect these numbers. It also looks at what information are collected on refugees/asylum seekers by the U.S. Census Bureau surveys and how these groups could be simulated using information on country of origin and year of entry from the American Community Survey (ACS). Focus is not only on the number of refugees by country of origin, but also on their characteristics and possibilities for analysis of integration over time. Finally, we address possible methods to integrate administrative data on refugees with U.S. Census data.

---

[1] United States Department of State, "Proposed Refugee Admissions for Fiscal Year 2016";
http://www.state.gov/j/prm/releases/docsforcongress/247770.htm

* Prepared by Jason Schachter, Eric Jensen, Anthony Knapp, and Chris Galvan

I.   U.S. Definitions of Refugee and Asylum seekers

1.   The United States provides refuge to persons who are unable or unwilling to return to their country of nationality because of persecution or a well-founded fear of persecution on account of race, religion, nationality, membership in a particular social group, or political opinion[2]. Humanitarian protection is primarily granted through two channels: refugee resettlement and asylum status.

2.   Resettlement is the selection and transfer of refugees from a country in which they have sought protection to a third country (e.g. the United States). The U.S. Department of Homeland Security (DHS) determines whether individuals are accepted to the United States as resettled refugees. Before entering the United States their refugee status has already been determined, often by an organization like the UN High Commissioner for Refugees (UNHCR), and they thus become eligible for resettlement. Resettlement can take 18-24 months or longer from referral to arrival in the United States.

3.   This differs from the asylum process, whereby individuals who are already living in the United States, or are seeking admission at a port of entry, apply for and can be granted protection. Asylum cases include those physically present in the United States who apply for asylum, as well as those who request asylum as a defense against removal from the United States. The time needed to be granted asylum status depends on individual cases, but can take as little as several weeks, while also possible to be held up in court for many years.

II.   Recent trends in refugee resettlement and asylum grants

4.   Statistics on the number and origin of resettled refugees to the United States for the fiscal year (FY) are available on a quarterly basis[3]. These numbers come from administrative data collected by the State Department's Bureau of Population, Refugees, and Migration and are publically available on their Refugee Processing Center (RPC) website[4]. The maximum number of resettled refugees is determined by executive order on an annual basis, and is allocated from different geographic origins. For example, in FY 2016 the United States can admit up to 85,000 refugees, of which 34,000 will come from the Near East and South Asia (at least 10,000 from Syria), 25,000 from Africa, 13,000 from East Asia, 4,000 from Europe, and 3,000 from Latin American and the Caribbean[5]. The remaining 6,000 are reserved to respond to emergent situations.

---

[2] See section 101(a)(42) of the Immigration and Nationality Act.
[3] The fiscal year runs from 1 October of the previous calendar year to 30 September of the current calendar year.
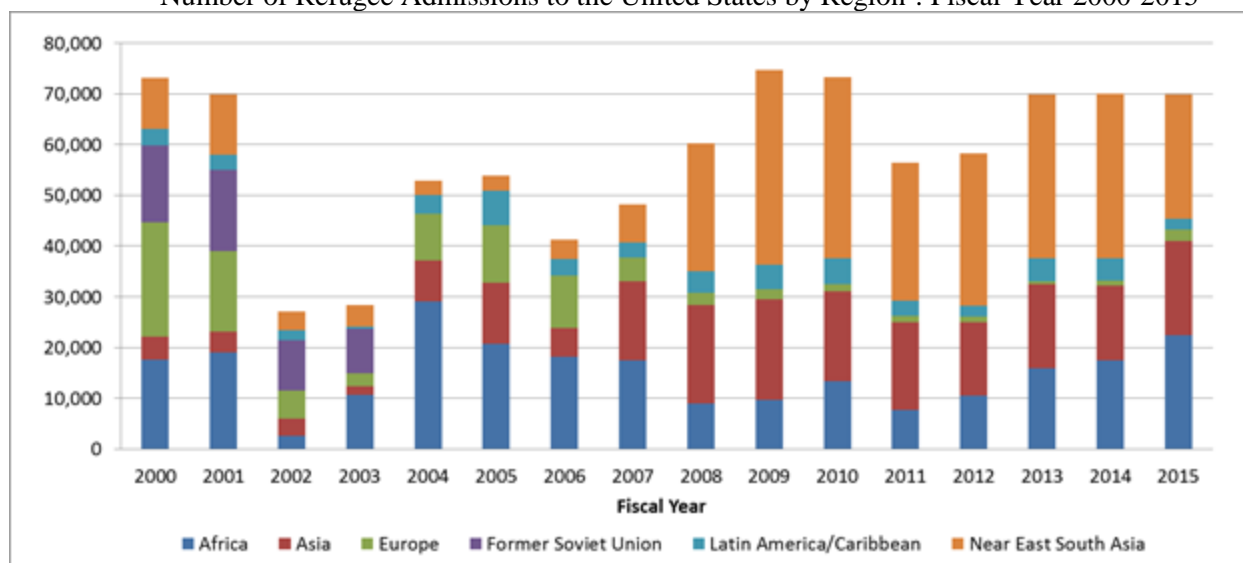[4] https://www.wrapsnet.org/
[5] United States Department of State, "Proposed Refugee Admissions for Fiscal Year 2016"
http://www.state.gov/j/prm/releases/docsforcongress/247770.htm

5.   Asylum seekers are not included in refugee resettlement statistics produced by the State Department, as these cases are processed through different administrative systems, specifically the Department of Justice (DOJ) who adjudicates over individual asylum cases. DHS releases statistics on asylum grants in their annual Yearbook on Immigration Statistics, for which 2013 is the most recent year available.   There are no limits on the number of asylum grants in a given year.

6.   Per RPC data, from 1975 (the earliest data available) through February 2016, the United States admitted 3.3 million refugees for resettlement, 1.5 million of whom were from East Asia[6]. Historically the number of refugees admitted to the United States in a given year (i.e. the flow) has fluctuated, ranging from a high of 207,000 in 1980 (predominately from Southeast Asia), to over 100,000 from 1989-1993 (due to a surge from the former Soviet Union), and as low as 27,000 in 2002 (post-9/11). From 2008 onwards (see Figure 1) refugee numbers have ranged from around 60-70,000/year, though more predominately coming from the Near East/South Asia region. In FY 2015, the United States resettled 69,933 refugees and in FY 2013 (the most recent published data) granted asylum status to 25,199 people[7].

Figure 1
Number of Refugee Admissions to the United States by Region[8]: Fiscal Year 2000-2015

Data Source: Department of State, Office of Admissions - Refugee Processing Center

7.   Looking at the "country of chargeability" of recent resettled refugees, Table 1 shows that refugees from Burma, Iraq and Somalia made up over half of all

[6] Refugees Processing Center, "Admissions & Arrivals Reports: Historical Arrivals Broken Down by Region (1975 – Present)"
[7] United States Department of Homeland Security, "Annual Flow Report on Refugees and Asylees: 2013"
[8] See http://www.state.gov/countries/ for a list of U.S. State Department defined regions

resettlements in FY 2015[9]. Other countries with a large share of refugees came from the Democratic Republic of Congo (DRC), Bhutan, Iran, Syria, Eritrea, Sudan, and Cuba. Countries like Burma, Iran, and the DRC have seen their share of refugees increase in recent years, while Bhutan, Iraq and Cuba have declined. While country of chargeability is usually determined by country of birth, when an applicant is a child, accompanied by or joining a parent, the child may be charged to the foreign state of either parent. As will be seen, this has potential data comparability issues with U.S. census data, particularly if refugees have been living in a second country for a prolonged period of time (e.g. Bhutanese in Nepal or Somalis in Kenya) and have had children born in those countries.

Table 1

Top Countries of Chargeability for Resettlement of Refugees in the United States: FY 2015

Country	Admissions
Burma	18,386
Iraq	12,676
Somalia	8,858
Dem. Rep. Congo	7,876
Bhutan	5,775
Iran	3,109
Syria	1,682
Eritrea	1,596
Sudan	1,578
Cuba	1,527
Ukraine	1,451
Burundi	1,186
Afghanistan	910
Ethiopia	626
Colombia	521
Moldova	333
Russia	281
Central African Republic	270
Rwanda	173
Pakistan	159
Grand Total	**69,933**

Data Source: Department of State, Bureau of Population, Refugees, and Migration Office of Admissions - Refugee Processing Center

8.   Data are also available by country of processing, though only the current fiscal year is available on the RPC website. Obviously, given residence outside

---

[9] Country of chargeability is usually equal to country of birth.

one's country of origin is a prerequisite for resettlement, countries of processing often differ from countries of chargeability. The top countries of processing include Malaysia, Kenya, Turkey, Ethiopia, Nepal, Iraq, and Thailand[10].

9. Finally, RPC provides monthly information about to which states refugees are resettled in the United States, by nationality (see Figure 2)[11]. At this time, data are only publically available for the first quarter of the current fiscal year. Settlement numbers tend to mirror the size of states, with large states like California, Texas, New York, and Ohio receiving the most refugees. Information on nationality illustrates that resettlement patterns tend to follow network migration patterns, attracting migrants from similar countries or regions. For example, half the refugees resettled in Massachusetts and Arizona are from the DRC and Somalia, over half of resettled refugees to California are from Iran and Iraq, 80 percent of all refugees to Indiana are from Burma, and about half of refugees to Maine and Minnesota are from Somalia. While the State Department data show in which states refugees are initially resettled, information on the secondary migration of refugees is available from the Department of Health and Human Services (HHS), which can tell us how the refugee population is redistributed after arrival. According to most recent figures, Minnesota has by far the greatest net domestic migration of refugees (2,500), presumably of Somali origin[12].
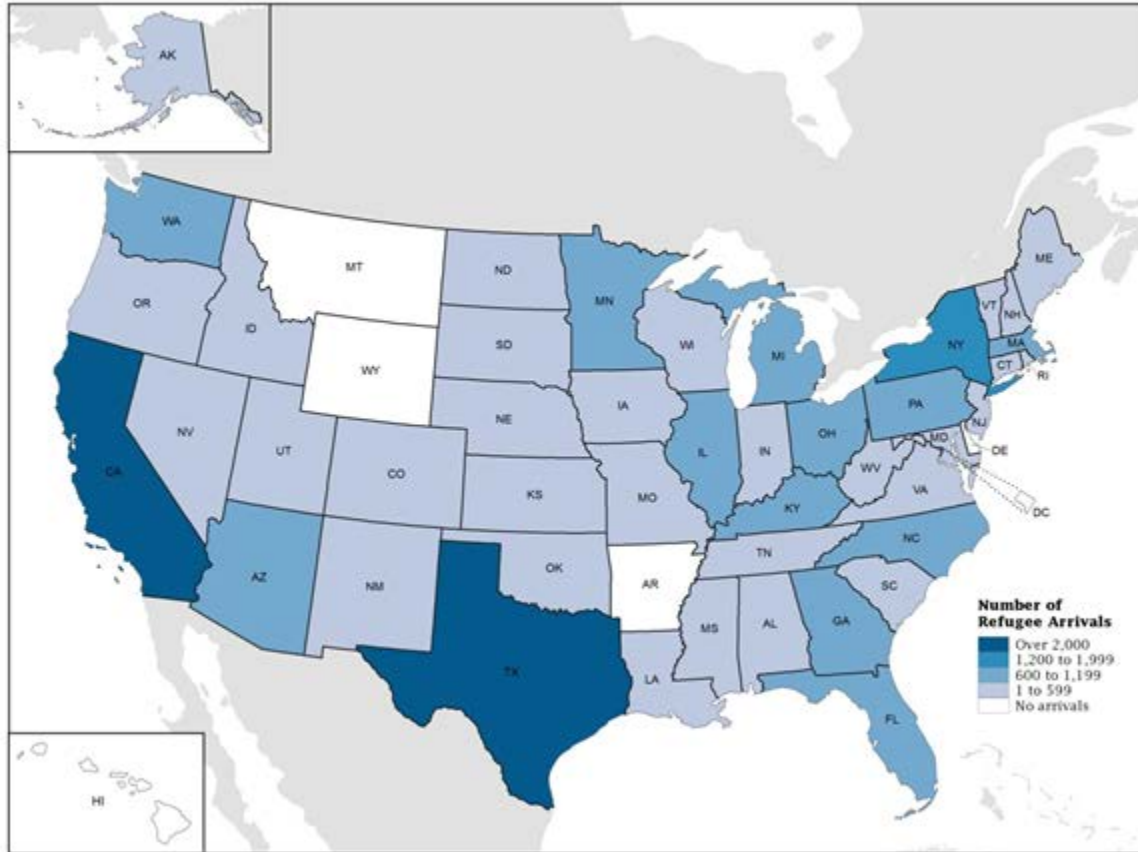
---

[10] Refugees Processing Center, "Admissions & Arrivals Reports: Map Arrivals by Processing Country and Nationality as of March 31, 2016"

[11] Refugees Processing Center, "Admissions & Arrivals Reports: Arrivals by State and Nationality as of March 31, 2016"

[12] Office of Refugee Resettlement. 2014. Statistical Abstract for Refugee Resettlement Stakeholders. Department of Health and Human Services

Figure 2

Geographic Distribution of Relocated Refugees in the United States: Quarter 1 FY 2016

10.   Compared to refugees there is less publically available information on asylum seekers. As mentioned earlier, different branches of the DOJ collect data on asylum seekers and they are publicly released as tables published in DHS' Year Book of Immigration Statistics. Since 2000, the number of asylum grants has remained somewhat steady, ranging from between 20-40,000 per FY. The peak in granted asylum cases was in 2001, when it reached 39,148, though it has been as low as 21,106 in 2010[13]. Statistics on asylees are released for those granted asylum on "affirmative" or "defensive" bases. "Affirmative" cases refer to those physically present in the United States, who apply for asylum via the U. S. Citizenship and Immigration Services (USCIS) (unlike refugees, they are prohibited from working during the application process), while "defensive" cases are those who request asylum as a defense against removal from the United States

---

[13] United States Department of Homeland Security, "Yearbook of Immigration Statistics: 2013 Refugees and Asylees, Table 16

(handled by the Executive Office for Immigration Review, (EOIR)). In 2013, of the 25,200 persons granted asylum status, about 15,300 were affirmative and 9,900 defensive[14]. The largest numbers of affirmative grants came from China (4,000) and Egypt (3,000). Close to half the total number of defensive grants came from China (4,500), followed by Ethiopia (400). No additional characteristics of asylum seekers are released in these table packages.

III. About the data

11.   As mentioned earlier, these administrative data come from a variety of sources depending on who processes the information. The State Department centralizes this information using a proprietary database system called the Worldwide Refugee Admissions Processing System (WRAPS). Figures in WRAPS differ slightly from those published in DHS's Yearbook of Immigration Statistics due to their different treatment of some groups. Refugee admission data include eligible family members granted follow-to-join refugee status, while asylee data are collected separately. Statistical reports including both refugees and asylum seekers would provide a more complete picture of these related populations of interest.

12.   Additional information on the characteristics of refugees is collected during the resettlement application process from a variety of sources. The Migration Policy Institute's (MPI 2015) study on refugees illustrates the potential data that are available. Through utilization of non-publically released data available from WRAPS, MPI was able to collect information on characteristics such as age, sex, native language, English language proficiency, and education. The report notes that this information was self-reported by applicants and the quality of some data was questionable, particularly for the education variable. However, if available, these characteristics, when combined with country of origin information, could increase the likelihood of linking this information with U.S. Census Bureau data to further analyze the refugee population.

IV. US Census data on refugees and asylum seekers

13.   The U.S. Census Bureau measures the migrant population via country of birth and other migration-related questions on many of their data sources, including major household surveys like the American Community Survey (ACS), which serves as its primary source of sub-national information on international migration. In general, the U.S. Census Bureau does not include specific questions on refugees and asylum seekers in its surveys, though there is one exception, the Survey of Income and Program Participation (SIPP), which includes a question on refugee or asylum status. This section outlines what information about refugees can be gleaned from Census survey data, as well as the potential for alternative methods to measure this population group.

---

[14] United States Department of Homeland Security, "Yearbook of Immigration Statistics: 2013 Refugees and Asylees, Table 17

V. The American Community Survey (ACS)

14. The ACS is an annual on-going survey of the U.S. population that collects detailed information previously collected on the decennial census long form. Fully implemented in 2005, it currently surveys about 3.5 million households per year. National level estimates for small populations (e.g. the foreign born) are calculated using annual data, while smaller geographic units (states and counties) are often dependent on 5-year data files. Of particular interest to measurement of refugees are questions on country of birth, year of entry, and to a more limited extent residence one-year ago, citizenship, and receipt of public assistance. While the combination of country of birth and year of entry could be used to simulate refugee flows over time, the relative small size of the refugee population, as well as the sample-based nature of this survey, could make it difficult to adequately measure this group using 1-year ACS data.

15. The number of refugees and asylum seekers on the ACS is unknown because there are no indicators of refugee/asylum status on the survey instrument. However, an estimate of the potential refugee population entering the country in a given year can be derived using the place of birth and year of entry variables from the ACS. These estimates can be compared to refugee cohorts in administrative data released by the U.S. State Department. Because not all immigrants from a specific country during a specific year are refugees, this method will produce the maximum possible number, or ceiling, of potential refugees from that country. Because asylum seekers often already reside within the United States when applying for asylum, this method does not work for their measurement.

16. To illustrate this method, we compare data on refugees resettled in the United States from the Refugee Processing Center (RPC) by country of chargeability for 2013 to estimates of potential refugees from this same cohort in the 1-year 2014 ACS file. Table 2 compares RPC totals for the top ten countries of origin to estimates of the foreign-born population who entered the United States in 2013. The RPC reports that for fiscal year 2013, there were 19,488 refugees from Iraq. The foreign-born population whose place of birth was Iraq who entered the United States in 2013 was 18,870 in the 1-year ACS file, which is not statistically different from the RPC data. Similarly, consistent data were found for the Democratic Republic of Congo, Sudan, and Eritrea, with higher RPC numbers, but still within acceptable ACS estimate variability.

Table 2

Refugees that entered the United States in 2013 from selected Countries of Chargeability/Birth

Country	Refugee Processing Center	1-Year ACS		Percent Difference	Margin of Error
		Estimate	Margin of Error		
Iraq	19,488	18,870	5,203	-3.2	26.7
Burma	16,299	7,710	1,984	-52.7	12.2
Bhutan	9,134	4,830	1,739	-47.1	19.0
Somalia	7,608	5,185	2,278	-31.8	29.9
Cuba	4,205	44,950	5,961	969.0	141.8
Iran	2,578	12,265	2,525	375.7	97.9

Country	Refugee Processing Center	1-Year ACS		Percent Difference	Margin of Error
		Estimate	Margin of Error		
Dem. Rep. Congo	2,563	1,615	1,281	-37.1	50.0
Sudan	2,160	1,665	782	-23.0	36.2
Eritrea	1,824	1,815	1,140	-0.4	62.5
Ethiopia	765	11,340	2,784	1382.2	363.9

Source: Refugee Processing Center and U.S. Census Bureau, Population Division, 2014 1-Year ACS file.
Notes: ACS data uses country of birth. Sudan includes both Sudan and South Sudan. 90% confidence interval shown. Burma is listed as Myanmar in the ACS. Country of chargeability is most often the country of birth. The ACS place of birth variable was used to approximate the country of chargeability. The ACS estimates presented in this paper are subject to sampling and non-sampling error. For more information, see: http://www.census.gov/programs-surveys/acs/technical-documentation/code-lists.html.

17.   ACS figures for Iran, Cuba, and Ethiopia are far higher than their RPC equivalents, which shows the difficulty of using this method for countries with historical migration ties to the United States (Cuba and Iran) or recently high rates of refugee resettlement and subsequent chain migration (Ethiopia). Of greater concern are ACS estimates which are far lower than RPC figures, such as found for Burma (53 percent lower), Bhutan (-47 percent), and Somalia (-32 percent). These findings suggest a number of possibilities, including undercoverage or higher non-response for recent migrants from these countries on the ACS (see Jensen et al. 2015). There also could be a data comparability issue, in that country of birth and country of chargeability are measuring slightly different things, which could be exacerbated in these particular countries. Many refugees from these countries have been settled in second countries for a long period of time prior to resettlement in the United States, such as Malaysia and Vietnam (for Burmese), Nepal (for Bhutanese), and Kenya and Ethiopia (for Somalis). This increases the chance that they have had children born in these countries, not the country of birth of their parents, from which "chargeability" is derived. ACS asks a question on ancestry which allows us to examine this issue. Looking at those with Burmese, Somali, and Bhutanese ancestry born in Thailand, Malaysia, Kenya, Ethiopia, and Nepal, the numbers increase by only a small amount (450 Somalis, 600 Bhutanese, 2,000 Burmese). While this does bridge the gap between ACS and RPC figures for Somali born, it still does not do so for Bhutanese and Burmese born.

18.   These inconsistencies between data sources could also be partly attributable to the inherent nature of a sample survey and its difficulties measuring small sub-groups of population. It is also possible that RPC values are inflated as refugees have a high incentive to report as their country of origin a country which is eligible for being granted refugee status. However, while it is not possible to test this hypothesis about RPC data, it is possible to analyze ACS data with larger sample sizes.

19.   Since the ACS is based on a sample of data, it is often necessary to combine multiple years of ACS files to reach a sample size which adequately measure relatively small populations, like refugees. However, one must exercise caution when using time-bounded variables like year of entry (YOE) in multi-year ACS files, either when using pooled data or ones created by the U.S. Census Bureau (Greico 2015). Due to the annual rolling sample design used in the ACS, persons

are surveyed continuously throughout the year. Thus, when YOE is for the current survey year, the risk of having moved that year is reduced based on what point of the year you are surveyed (e.g. those surveyed in January are far less likely to have moved during that year than those surveyed in December). In effect, this yields lower results (about halved) when the year of entry is the same as the survey year. When data from multiple years are combined, the cumulative effect of this bias is that the earliest survey year in the file contains a disproportionate share of year of entries (more recent years are underrepresented). Similarly, a person surveyed in 2014 for YOE can provide responses for the years 2014, 2013, 2012, 2011 (and earlier), while a person surveyed in 2011 can only provide responses for 2011 and earlier, further biasing results towards earlier years.

20.   Fortunately, there is a relatively simple solution for this issue, which is to only use YOE dates prior to the earliest year of data collection in your pooled/combined data. For example, if pooling three years of ACS data (2012-2014), only look at YOE for 2011 or earlier. In our case, we have utilized the 2010-2014 ACS 5-year file and thus limited analysis for those who entered in 2009[15].

21.   Again, in Table 2a we compare ACS 5-year estimates for 2009 with RPC totals for the top ten countries of origin for that fiscal year. Not surprisingly, the estimates are less variable than the 1-year ACS estimates shown earlier due to the larger sample size. The RPC reports that for fiscal year 2009 there were 18,838 refugees from Iraq compared to 21,760 Iraqi born in the ACS, which is more in-line with our expectations of the ACS number being a ceiling for "potential" refugees from Iraq.

---

[15] Another issue is that there will likely be an underestimation of total numbers for 2009, due to the "heaping" of YOE responses around "2010," whereby respondents are more likely to provide "year" information in years ending in "10" or "5" (see Logan and Cruz 2009). This will be easier to determine with the release of the 2011-2015 ACS 5-year file later this year

Table 2a

Refugees that entered the US in 2009 from selected Countries of Chargeability/Birth

| Country | Refugee Processing Center | 5-Year ACS | | | |
		Estimate	Margin of Error	Percent Difference	Margin of Error
Iraq	18,838	21,760	2,146	15.5	11.4
Burma	18,202	12,470	1,186	-31.5	6.5
Bhutan	13,452	11,805	1,404	-12.2	10.4
Iran	5,381	12,930	1,362	140.3	25.3
Cuba	4,800	32,700	2,396	581.3	49.9
Somalia	4,189	3,215	687	-23.2	16.4
Eritrea	1,571	1,600	462	1.9	29.4
Vietnam	1,538	24,830	1,761	1514.4	114.5
Dem. Rep. Congo	1,135	1,500	402	32.1	35.4

Source: Refugee Processing Center and U.S. Census Bureau, Population Division, 2010-2014 5-Year ACS file. Notes: ACS data uses country of birth. Sudan includes both Sudan and South Sudan. 90% confidence interval shown. Burma is listed as Myanmar in the ACS. Country of chargeability is most often the country of birth. The ACS place of birth variable was used to approximate the country of chargeability. The ACS estimates presented in this paper are subject to sampling and non-sampling error. For more information, see: http://www.census.gov/programs-surveys/acs/technical-documentation/code-lists.html.

22.   ACS 5-year data figures for Iran, Cuba, and in this case Vietnam, are far higher than their RPC equivalents, which again illustrates the limitation of this method when faced with historical migration patterns and subsequent network migration related to refugee resettlement. The fact is most migrants from those countries are no longer refugees, though a significant number of refugees are still resettled from those countries. In addition, ACS 5-year estimates were still found to be lower than RPC figures for specific countries, such as Burma (-32 percent), Bhutan (-12 percent), and Somalia (-23 percent), though to a lesser extent than for the 1-year ACS data. In this case, looking at ancestry of Burmese (3,000 born in Thailand and Malaysia), Somalian (900 born in Kenya and Ethiopia), and Bhutanese (850 born in Nepal) foreign born has a much greater impact of bridging the gap between ACS and RPC data, as inclusion of these groups brings results within or closer to expected sampling variation. This suggests that country of chargeability could be playing an important role in differences between ACS and RPC data, but also suggests there could still be some systematic ACS measurement error (undercoverage or sampling), or RPC data accuracy issues in these cases as well. One would expect the use of a 5-year as opposed to a 1-year file to improve coverage for newly arrived refugees, who would have had more time to settle over a 5-year rather than a 1-year period.

VI.   Other attempts to identify refugees using Census data

23.   As mentioned earlier, comparing ACS year of entry data to measure refugees only provides a maximum ceiling for refugees in a given year. Several other organizations have attempted to measure refugees with ACS or CPS data, most notably the Urban Institute and the Migration Policy Institute (MPI), both think tanks in the Washington, DC area. The ability to impute refugee status to Census

data would be particularly helpful for measuring the characteristics or long-term integration of refugees resettled in the United States. Both methods use something of an "all or nothing" approach to measure refugees based on country of birth and year of entry, vis a vis administrative data for those years, though MPI makes some adjustments based on ancestry and the Urban Institute on public assistance use.

24.   The Urban Institute developed a method which has been adapted by other institutions (e.g. Pew Research Center) and Federal agencies (e.g. HHS) to measure refugees with Census data[16]. Initially developed to analyze data from the Current Population Survey, using country-specific administrative data from DHS, it first determines the number of migrants attaining legal permanent residence, resettled refugees, and asylum grants for a given year, and identifies countries for which refugees/asylum seekers make up the majority. Next, for a given year of entry (as far back as 1980), persons born in countries identified as "refugee sending" are imputed as refugees. These results are then compared to a demographic estimate of the total number of refugees/asylees residing in the United States (total number of OPR admissions since 1980 adjusted by deaths) and adjusted if needed, remaining consistent by region of origin and period of entry.

25.   Beginning in 2013, additional adjustments to recently arrived migrants (since 2009) receiving public benefits were made. For those arriving to the United States in 2009 or later, and reporting that a nuclear family member receives food stamps or cash aid, they are assigned refugee status if for prior periods of arrival their country of origin had previously been classified as refugee-sending countries or if the country currently sends a significant number of refugees, but not a majority of arrivals[17]. It should be noted that the number of persons granted asylum status or legal permanent residence are not necessarily equal to arrivals during a given year, such as for refugee resettlement, and it is unknown to what extent this is adjusted for, if at all[18].

26.   An added benefit of allocating refugee status with Census data is being able to learn more about their characteristics. Among other things, detailed information on sex, age, English language ability, educational attainment, employment status, income, health insurance, housing conditions, and public assistance use are available on various Census surveys. To measure these characteristics MPI (2015) estimated demographic and socioeconomic indicators for the U.S. refugee population resettled between 1980 and 2011 using ACS data, combining 2009-2011 data to increase sample size and precision of estimates (though it appears YOE bias towards 2009 was not taken into account). Refugee status was imputed from country of birth and year of arrival in the United States, comparing these characteristics to RPC administrative data on the number of arrivals by year and country of chargeability. Foreign born for whom a country/year combination from

---

[16] Urban Institute 2014. This method was developed by Jeffrey Passell and Rebecca Clark, currently with the Pew Research Center and NIH respectively.
[17] Administered through the Supplemental Nutrition Assistance Program (SNAP)
[18] It is also unclear how the issue of refugees whose status is changed to legal permanent residence are handled using this method. For example, if a large cohort of refugees from a specific country achieves LPR status in a given year, this will decrease the likelihood that new refugees from that same country will be classified as refugees that same year.

RPC data exceeded 40 percent of the equivalent estimated foreign-born population in the ACS were assigned refugee status. For example, if the ACS found 20,000 Iraqis moving to the United States in a given year, then if RPC admissions exceeded 8,000, then all ACS Iraqi migrants were assigned refugee status. Conversely, if the ACS found 30,000 Cuban-born immigrants in a given year, and RPC admissions were less than 12,000, then no Cuban born migrants were assigned refugee status.

27.   In addition, refugee status was also assigned to foreign born for whom a country/year combination from DOJ asylum grant statistics exceeded 20 percent of ACS foreign born population estimates. Further adjustments were based on inconsistencies between RPC country of birth and nationality data, in which case assignment of refugee status in ACS was based on ancestry. For example, for some years refugee admission from Kenya exceeded the 40 percent threshold for country of birth and not nationality, due to the high number of children born to Somalian refugees in Kenya. As a result, those who reported Kenya as their country of birth also needed to report Somali ancestry to be coded as refugees (MPI 2015).

28.   The biggest drawback to these methods is the assignment of all migrants from certain countries as either refugees or non-refugees, thus potentially missing refugees from countries with large migration flows. There appear to be network or chain migration patterns at work, whereby after settlement refugees become legal permanent residents (one year) or U.S. citizens (three years) and then extended family members move to the United States as part of family reunification immigration policy. Using these "all or nothing" methods, these network migrants would initially be counted as refugees, until refugee flows have ebbed to a point where network migration becomes larger, then would no longer be counted as refugees. A more precise method that could impute refugee status based on other individual characteristics would be of interest to researchers, as will be addressed later in the paper.

VII.   The Survey of Income and Program Participation (SIPP)

29.   The U.S. Census Bureau does conduct one survey that directly measures refugee status. The Survey of Income and Program Participation (SIPP) is a longitudinal household survey conducted by the U.S. Census Bureau since 1983. The SIPP survey design is a continuous series of national panels, with sample size ranging from approximately 14,000 to 52,000 interviewed households. The duration of each panel ranges from 2 ½ years to 5 years, with each "wave" of interviews conducted every four months[19]. The sample size is organic, in that as households split or gain new members, the sample increases to accommodate these changes. As with all longitudinal surveys, later waves suffer from respondent attrition, as people stop participating in the survey.

---

[19] The SIPP incorporated a new design in 2014 whereby interviews are now only conducted every 12 months.

30.   The primary purpose of the SIPP is to monitor economic progress and participation in government public assistance programs over time, but it also asks a multitude of other related questions on topical modules. Prior to 2014, during the 2nd wave of the survey (4th month) a Migration History Topical Module is conducted. This module includes information on place of birth, date of move (specific month and year), and immigration status when they moved to the United States. Among the six coded responses is the category "granted refugee status or granted asylum." [20] This allows one to directly measure refugees in the sample, and thus their characteristics, and given the longitudinal nature of the survey, track their progress over time.

31.   Though the most recent SIPP panel started in 2014, the data are still not available, thus we were limited to looking at the 2008 SIPP panel, which had a sample of about 50,000 households or 78,000 individuals during the second wave. Looking at data from the Migration History Module, 14 percent of the sample reported having moved to the United States and reported an immigration status. Among these 34.1 million persons, only 1.7 million had been granted refugee or asylum status upon entry (see Table 3), or about 5 percent of international migrants. Refugees on the SIPP come from a wide variety of countries, but the largest proportion (24 percent) come from Cuba. However, the relatively small sample sizes from these countries make analysis by individual countries difficult.

Table 3
Reported Refugee or Asylum Status in the SIPP: 2008

Country	Total	Percent
Cuba	421,898	24.1
Vietnam	154,384	8.8
Ukraine	89,060	5.1
Russia	87,741	5.0
Iran	68,304	3.9
China	54,127	3.1
Laos	52,218	3.0
Somalia	45,591	2.6
El Salvador	43,816	2.5
Cambodia	39,303	2.2
Bosnia and Herzegovina	35,315	2.0
All Other Countries	655,633	38.0
Grand Total	1,747,391	100.0

Source: U.S. Census Bureau, 2008 SIPP – Migration History Topical Module

32.   These numbers represent refugee stocks, but when looking at year of entry the numbers dwindle dramatically. For example, the annual number of refugees who entered between 2000 and 2008 ranged between 70,000 and 105,000. The low

[20] This question wording could prove difficult to measure asylum seekers, as asylum status is often not granted upon entry to the United States, but rather after they have already been living there for some time.

numbers of persons in the sample necessitate limiting analysis of characteristics to the refugee stock population, though they could still be divided into broader year of entry cohorts (e.g. before 2000, 2000 and after). Though we did not do so in this paper, it is also possible to link refugees with other characteristics on the SIPP, such as age, sex, occupation and industry, poverty status, etc.

33.   Given the longitudinal nature of the SIPP, it would be possible to track refugees to see how their family formation, income, employment history, and program participation have changed over time. This is critical for monitoring the long-term integration of refugees. The small sample would necessitate treating all refugees and asylum seekers as a whole, which could potentially miss the specific situation experienced by refugees from different regions of the world.

34.   Longitudinal surveys such as the SIPP are challenging to analyze due to their complex structures. However, the SIPP currently represents the only Census Bureau source to directly measure refugees, and it seems to be an underutilized resource for the investigation of refugees, particularly their characteristics.

VIII.   Data comparability issues

35.   As mentioned earlier, comparing ACS survey (or other Census data) to RPC administrative data comes with a number of limitations. First, administrative data on refugees are collected from a variety of sources and application processes, whose primary purpose is not necessarily the production of refugee statistics. To what extent data from different administrative sources are comparable and accurate is not known, particularly with regard to characteristics collected in the application process. It is possible that refugee data are overstated for specific countries, since there is an incentive for refugees to claim to have come from one country rather than another in terms of being granted refugee status. It is assumed vetting is thoroughly conducted by organizations like UNHCR or the International Organization for Migration (IOM) or by DHS prior to resettlement in the United States, but this might not be possible under certain adverse working conditions.

36.   Census survey data and RPC data are collected for slightly different time periods. Census uses a calendar year, while the State Department uses the fiscal year for reporting purposes. This could cause slight variations between numbers, especially given the rapidly changing origins of refugee countries over time. In addition, the concepts of country of chargeability and country of birth differ slightly, in that some children of refugees are assigned the country of their parent, even if born outside that country. While these differences are likely to be small, cumulatively they could have an impact on comparability of numbers, particularly flows.

37.   Census data are faced with their own set of limitations, not only in terms of sample size as discussed earlier, but also in terms of limitations to the use of questions like "Year of Entry." Research on the YOE question has shown that there is a high degree of measurement error when using this question, due to issues like recall bias, proxy respondents, and misunderstanding of the question (Cruz and Logan 2009). For example, response to year of entry questions is often "heaped" on years ending in 5 or 10, meaning respondents are inexact in their responses. Further, year of entry could refer to either year of first or most recent entry, which could be problematic if a respondent has moved to the United States more than

once[21]. This could also be particularly salient for asylum seekers already in the country who do not necessarily report year of entry as the year they received official asylum status, though this should be less of an issue for resettled refugees.

38.   The discussed methodologies all suffer from limitations, in that the Census analysis only provides an "upper" estimate of potential refugees from a particular country, while the Urban Institute and MPI methods use an "all or nothing" approach, which assigns refugee status based on countries of origin identified as refugee sending. This obviously includes a number of non-refugees from these specific countries, but also more importantly, misses refugees from countries with high non-refugee flows (e.g. Cuba). The UI and MPI methods do attempt to adjust their imputations based on welfare usage and ancestry, but neither is the ideal solution, which would assign refugee status to specific individuals, rather than all persons coming from specific countries.

IX.   A better way to simulate refugee status?

39.   The Census Bureau has begun to consider options for measuring refugee status from its data, which is an issue receiving attention by international bodies like the UN Statistical Commission. There are also occasionally requests to disaggregate the refugee component from subnational Net International Migration (NIM) estimates, which is not possible with the current methodology. Given the United States does not maintain a national population register, which would facilitate the linking of refugee status to other characteristics, other possibilities must be considered.

40.   One possible option would be to add a new question on "reason for move" to the ACS, similar to what is asked on the Current Population Survey, which would specify "humanitarian reasons" as a response category, or alternatively add a question similar to what is asked on the SIPP regarding immigration status. However, with current concerns about respondent burden and question sensitivity it would be difficult to have such a new question added to the ACS, unless required by either Federal or legislative needs. Given this is unlikely to happen, another method would be to find a way to link RPC administrative data with Census and survey data.

41.   As discussed earlier, the State Department collects and maintains administrative data on refugees and asylum seekers in the United States through a shared system called the Worldwide Refugee Admissions Processing System (WRAPS), while DOJ compiles its own records on asylum cases. If the internal records maintained by these agencies contain personally identifiable fields, they could then be linked to the ACS. For example, refugees are eligible to apply for social security numbers immediately upon arrival to the United States, as they are expected to begin participating in the labor force. Thus a direct-match method using a Personal Identification Key (PIK), if social security numbers or Individual Taxpayer Information Number (ITIN) are available, should be possible for

---

[21] The 2015 ACS has addressed this by adding instructions on the survey instrument to collect most recent year of entry if multiple moves to the U.S.

refugees and asylees with administrative records, or else a probabilistic matching method using other fields could be used.

42.   While it could prove difficult to match a sufficient number of refugees/asylum seekers to ACS data, another possible method would be to model detailed characteristics of refugees from administrative data to ACS data.   For example, administrative data include detailed information about refugee age, sex, country of origin, country of nationality, and other variables (MPI 2015). Using this information, it should be possible to find persons on the ACS with a high probability of sharing these same characteristics, similar to what is done during hot deck imputation procedures for survey respondents with missing data. While this work would still need to be developed, there is great potential for this sort of modeling exercise, assuming access to individualized administrative data.

43.   Linking administrative records of refugees and asylum seekers to census and survey data has several advantages over other approaches to measuring the characteristics of this population. First, by not adding a new question to the ACS questionnaire it would not increase overall respondent burden. Next, administrative records could be immediately linked to past census and survey data creating a times series rather than waiting several years for data to be collected. Finally, the validity of a refugee status measure is potentially higher in administrative records than a self-administered census or survey. Since a migrant must be registered in administrative records to obtain refugee or asylum status, the potential for undercoverage in these data is very low, which also increases the validity of the linked data.   While there are many administrative hurdles to overcome to gain access to detailed administrative data, this method could provide better information on the refugee population and a greater understanding as to their characteristics and adaption to life in the United States over time.

X. Discussion

44.   As the refugee population continues to grow, so does the necessity to accurately measure their size and monitor their adjustment in destination countries. This topic has achieved prominence and high visibility at recent UN Statistical Commission meetings in New York, which approved the formation of a UN expert working group on refugee statistics, which will develop a handbook on statistics on refugees and internally displaced persons. These sorts of internationally coordinated activities should serve to improve the comparability, quality, and timeliness of refugee statistics.

45.   U.S. refugee and asylum statistics are derived from administrative data collected by different Federal agencies. While the U.S. Census Bureau does not collect or directly measure refugees, we have begun looking at ways to disaggregate refugees from our data sources. If we were to use figures provided by the State Department, given the refugee population is already included in the national and subnational estimates produced by the Net International Migration Branch (derived from ACS data), it would be necessary to be able to measure refugees so as to not double count them. A number of data comparability issues exist between administrative refugee and U.S. Census data, including variations in time period under consideration, countries of origin, and possible undercoverage for some specific refugee populations in the ACS.

46.   Given that the United States lacks a national population register, there is no simple way of linking data between sources, in order to measure both the characteristics and track the long-term integration of individual refugees. It would be advantageous to be able to measure specific refugees on Census data, rather than imputing refugee status to all migrants coming from specific countries in a given year. This paper briefly discussed several potential methods of integrating data to accomplish this, including linking of administrative data to individual ACS records or using detailed characteristics from administrative data to impute refugee status on ACS records. Both of these methods need much more development before future implementation. Armed with this knowledge, it would then be possible to look at the specific characteristics of refugees, disaggregated by period of arrival and country of origin, in order to track changes in their outcomes or geographic distribution over time, which are all important from a policy perspective. Whatever approach is taken, it will be critical to keep abreast of international developments in the statistical field, in order to maintain international comparability with our statistics.

References

Capps, Randy and Kathleen Newland. 2015. The Integration Outcomes of U.S. Refugees: Successes and Challenges. Migration Policy Institute.

Office of Immigration Statistics. 2014. 2013 Yearbook of Immigration Statistics. Department of Homeland Security.

de la Cruz, Patricia and Cassandra Logan. 2009. "August 2008 CPS Migration Supplement: APreliminary Look at the Citizenship and Year of Entry Sections." Paper Presented at theAnnual Conference of the Southern Demographic Association, October, Galveston, TX.

Greico, Elizabeth M., Luke J. Larsen, and Howard Hogan. 2015. "As Time Goes By: How Period Data Influence the Estimates of Recently Arrived Immigrants in the American Community Survey." U.S. Census Bureau working paper.

Jensen, Eric B., Renuka Bhaskar, and Melissa Scopilliti. 2015a. "Demographic Analysis 2010: Estimates of Coverage of the Foreign-Born Population in the American Community Survey." Population Division Working Paper Series No. 103, U.S. Census Bureau, Washington, D.C.

Jensen, Eric B. 2015b. "Refugees and Asylees in the U.S. Census Bureau's Population Estimates." Internal Memorandum for Population Division Management. U.S. Census Bureau

Jensen, Eric B. 2015c. "Linking Administrative Records and Census Bureau Data to Measure the Characteristics of Refugees and Asylum Seekers in the United States." Internal Memorandum for Population Division Management. U.S. Census Bureau

Office of Refugee Resettlement. 2014. Statistical Abstract for Refugee Resettlement Stakeholders. Department of Health and Human Services.

Urban Institute, internal memo. 2014. "Imputation of legal status of the foreign-born, and eligibility-related characteristics of noncitizens, on the March 2014 CPS."