APRIL 2022

EMERGING TRENDS IN CENSUS APPROACHES IN ASIA AND THE PACIFIC

with country examples

ESCAP 75
MOVING FORWARD TOGETHER

This report is issued without formal editing. Views expressed herein do not necessarily reflect that of Economic and Social Commission for Asia and the Pacific (ESCAP) or any United Nations agencies.

For more information, please contact:

Statistics Division
Economic and Social Commission for Asia and the Pacific (ESCAP)
United Nations Building, Rajadamnern Nok Avenue, Bangkok 10200, Thailand
Email: stat.unescap@un.org
Website: https://www.unescap.org/our-work/statistics

ACKNOWLEDGEMENTS

This report has been prepared by Stein Terje Vikan, an ESCAP Statistics Division consultant on population and housing censuses. The report is part of the Bloomberg Philanthropies Data for Health Initiative (D4H) Project: “Supporting improvements to civil registration and vital statistics (CRVS) systems in Asia and the Pacific (Phase III)”.

The work has been overseen by Petra Nahmias and Afsaneh Yazdani, from ESCAP Statistics Division. The author is grateful for their invaluable input and comments throughout the work of preparing this report. Panita Rattanakittiaporn and Nasikarn Nitiprapathananun from ESCAP have also provided excellent administrative support.

This report would not have been possible without invaluable contributions and generous support from Seiffe Tadesse, Meryem Demirci and Srdjan Mrkic from United Nations Statistics Division, Alison Culpin and Michael Sharp from the Pacific Community, Gloria Mathenge, currently at the Economic Commission for Africa but formerly the Pacific Community, as well as National Statistical Offices of Armenia (Karine Kuyumjyan), Australia (Ross Watmuff, Gemma Van Halderen, and Georgia Chapman), Indonesia (Ali Said, Nashrul Wajdi and Diah Ikawati), Mongolia (Batchuluun Dorj), Republic of Korea (Kang Jeong-hee, and Byungsik Lee), Thailand (Hataichanok Puckcharern), and Turkey (Mehmet Doğu Karakaya, Metin Aytaç, Mehmet Şaban Ucari, Alper Acar, Önder Değirmenci, and Osman Orhanli).

Whilst the support from these colleagues is recognized, any inaccuracies in the description of census operations in these countries remain those of the author.

CONTENTS

ABBREVIATIONS AND ACRONYMS

ABPRS	Address Based Population Registration System
ABS	Australian Bureau of Statistics
Armstat	Statistical Committee of the Republic of Armenia
BMIS	Border Management Information System
CAPI	Computer-Assisted Personal Interview
CATI	Computer-Assisted Telephone Interview
CAWI	Computer-Assisted Web Interview
COVID-19	coronavirus disease 2019
DOPU	Drop-Off/Pick-Up
ENEA	East and North-East Asia subregion
ESCAP	Economic and Social Commission of Asia and the Pacific
GPS	Global Positioning System
KOSTAT	Statistics Korea
NAD	National Address Database
NCA	North and Central Asia subregion
NSO	National Statistical Office
PAC	Pacific subregion
PAPI	Paper Assisted Personal Interview
PHRD	Population and Household Registration Database
RRR	Resident Registration Register
SEA	South-East Asia subregion
SPC	Pacific Community
SSWA	South and South-West Asia subregion
Statistics Indonesia	Badan Pusat Statistik (BPS - Statistics Indonesia)
Stats NZ	Statistics New Zealand
TurkStat	Turkish Statistical Institute
UNECE	United Nations Economic Commission for Europe
UNSD	United Nations Statistical Division

CHAPTER 1

INTRODUCTION – GLOBAL CENSUS TRENDS

A population census has been a long-standing central responsibility of a National Statistical Office (NSO). Its scope and objectives have remained very similar over time; however, aspects of its methodology have evolved as technology has developed and demands have changed.

The main aim of a population and housing census is to fully enumerate a country's population and housing conditions to provide detailed demographic and socio-economic data at highly disaggregated geographical levels and for subpopulations. They provide a primary source of data needed for formulating, implementing, and monitoring a range of policies and programmes within a country, and comparing between countries.[1] In recent decades, both national and international frameworks for development, such as the Millennium Development Goals and the subsequent Sustainable Development Goals, have all increased the need for more comprehensive and granular demographic and socio-economic data to monitor the progress towards these goals. In this context, population and housing censuses play a crucial role in providing necessary disaggregated data for assessing the situation of people on a range of indicators. The level of disaggregation, especially geographically but also for smaller population groups, distinguish a census from what is possible through a survey.

1.1 Traditional censuses

A traditional census, in principle, entails canvassing the entire country, reaching every single household, and collecting information on all individuals within a brief stipulated period of time. Previously, this entailed visiting each household in the country and posing questions to the respondents from a census questionnaire. Respondents would sometimes be given the option of returning the questionnaire through mail. The information collected would be captured on paper questionnaires for later processing by the census team. Often, the census team would have to conduct a mapping exercise prior to the data collection phase in order to demarcate enumeration areas on maps. For both the mapping and data collection phases, NSO conducting the census normally must hire a large number of extra staff. While the data collection operation would normally be restricted to a few weeks, with a clear reference date for the information, the actual data processing could take years. Therefore, the final publication of the census results could be relatively late compared to the reference date.

Several trends over the last decades have impacted how censuses are conducted. While many countries still conduct what would be labelled a "traditional" census, advances in technology

[1] See https://unstats.un.org/unsd/demographic-social/census.

have enabled accelerating the collection and processing of census data whilst reducing the significant costs of the census operation.[2]

One such change was the introduction of Optical Character Recognition, which enabled paper questionnaires to be read by a scanner. While this reduced the data entry time significantly, it still required data to be captured on paper (Paper Assisted Personal Interview or PAPI) and some, albeit fewer, manual check-ups. With the ever-evolving development of computers, especially the introduction of laptops and hand-held devices, the possibility to capture data electronically emerged in the late 1900s and early 2000s. Later on, the introduction of smart phones and tablets escalated the possibilities for rapid data collection. These technologies significantly changed data collection by enabling electronic data entry directly in the field and eliminating the stage of entering data from paper questionnaires into a computer for processing. They also reduced (but did not eliminate) the need for manual check-ups. This approach was labelled Computer-Assisted Personal Interview (CAPI). Another version of this was Computer-Assisted Telephone Interview (CATI). Using telephone interviews to collect data from households having a telephone, limited the need for field visits to those who could not be reached by phone.

The spread of the internet made it possible to administer questionnaires to households and individuals without physically visiting them. While not everyone has access to the internet or is able or willing for various reasons to submit a questionnaire through the internet, this approach could significantly reduce the need for field visits or the duration of field operation. It also reduced the need to hire enumeration staff, as the respondents filled in the questionnaire themselves. This approach is labelled Computer-Assisted Web Interviewing (CAWI).

All these technologies or approaches made carrying out a census simpler, more cost-effective, and timelier. However, they did not alter a fundamental aspect of a census that everyone in a population had to provide information through a census questionnaire. This means that such data collection approaches are still considered “traditional”, even if they have vastly improved the efficiency of carrying out a census.

1.2 Emerging challenges for traditional censuses

In the last decades, several trends have put pressure on NSOs to modernise censuses, both in terms of conducting more effective traditional censuses and looking for alternative approaches such as register-based censuses. A key development has been the ever-increasing cost of conducting a census. Although a census is typically carried out every five or ten years, it comes at a considerable cost to the NSOs. Given that it is by far the largest operation an NSO conducts, it requires special funding and the hiring of a large number of

[2] For a brief overview of how censuses have evolved globally and especially in Asia and the Pacific, see https://unescap.org/kp/2021/stats-brief-august-2021-issue-no-30-are-population-censuses-asia-and-pacific-evolving.

temporary staff to carry out field enumeration and supervision, in addition to data processing and dissemination of results. This has prompted governments to look for ways to reduce these costs.

Further, many countries are facing a reduced willingness amongst the population to respond to census questionnaires, leaving NSOs with increasing challenges of non-response. This means that NSOs are constantly looking for ways to reduce the response burden or find alternative ways of producing statistical information. For instance, the lower-than-expected response rate of the 2018 census in New Zealand prompted Statistics New Zealand (Stats NZ) to introduce new methods using alternative data sources to produce census information for individuals who had not provided a response through a census questionnaire.[3]

And lastly, there is pressure to process census data faster in order to release census results in a timelier fashion. Census data plays an essential role in government planning. The sooner the results can be released, the more relevant the data is for the planning and policy-making processes. A shorter census period will also reduce the strain on the NSO, who may have to dedicate a large number of staff to this work.

1.3 Alternatives to the traditional census approach

In the last decades of the twentieth century, several countries started using administrative data to produce census data. Denmark was the first country to conduct a complete population and housing census based on administrative data in the 1980 census round, followed by other Scandinavian countries.[4] By the 2000 round, this approach was gaining ground.

1.3.1 Register-based census

A full register-based census collects all information at the individual level from registers or administrative data sources, with no field enumeration. This assumes all of the variables required at unit level for a full census are available in up-to-date registers or administrative sources in the country. The Guidelines on the Use of Registers and Administrative Data for Population and Housing Censuses[5] of United Nations Economic Commission for Europe (UNECE) differentiates between registers, administrative data sources and statistical registers in the following way:

A register is defined as a systematic collection of unit-level data organized in such a way that updating is possible. Updating is the processing of identifiable information with the purpose of establishing, bringing up to date, correcting, or extending the register, that is, keeping track of any changes in the data describing the units and their attributes. Administrative data sources, on the

[3] See https://stats.govt.nz/2018-census/data-quality-for-2018-census.
[4] See https://unece.org/fileadmin/DAM/stats/publications/Register_based_statistics_in_Nordic_countries.pdf.
[5] See https://unece.org/fileadmin/DAM/stats/publications/2018/ECECESSTAT20184.pdf.

other hand, are data holdings that contain information collected primarily for administrative purposes. This type of data is collected by government departments and other organizations for the purposes of registration, transaction and record keeping, usually during the delivery of a service. They include administrative registers (with a unique identifier) and possibly other administrative data without a unique identifier. Statistical registers are registers created for statistical purposes. They are typically created by transforming data from registers and/or other administrative data sources.

Typically, the following four key registers are used in a register-based census:[6]

- *Population register* – A register of all residents in the country.
- *Address/Dwelling register* – A register of all (whether occupied or not) addresses or dwellings in the country.
- *Business register* – A register of all businesses or legal entities in the country.
- *Activity register* – A register of activities that residents are engaged in, such as (un)employment/work, benefits and pensions or studies.

These registers are often referred to as *base registers*. In addition, countries may have several other *supplementary registers* that provide additional information at the unit level that can be used to generate census variables.

As mentioned earlier, it is important to note that these registers have been originally established and maintained by other government agencies than the NSO for their administrative purposes. When considering moving towards a register-based or combined census, it is critical to assess the availability of high-quality administrative data sources that are appropriate for the census purpose.

The UNECE lists six key questions that should be answered when considering the suitability of any administrative source:[7]

- How are data compiled and for what reason?
- Is there a legal obligation?
- What is the target population?
- Are data regularly updated?
- Are there plausibility checks?
- How are variables defined and are the definitions comparable with statistical concepts?

For a register-based census to take place, it is therefore necessary for the NSO to establish a good working relationship with the custodians of the registers to access the required data and metadata. It is often also necessary to ensure legal access for the NSO to these registers

[6] Ibid.
[7] Ibid.

through a statistical or census law. Further, the data will have to be transformed from its original administrative state to a statistical register.

In all register-based censuses, it is necessary to combine information, at the unit level, from different registers. This implies that somehow it must be possible to ensure that information on Person A from one register is connected to information on the same individual (and not Person B) from another register. The simplest way to ensure this is to assign each individual (usually at birth) a unique number or identifier that follows them throughout their lives. If this unique identifier is used to identify an individual in every administrative register, then it becomes fairly straightforward to combine information from different administrative sources for each individual. Many countries have put in place such a system, for instance the Scandinavian countries, Lithuania, and Turkey.

However, many countries do not have such a unique identifier, either because they have not developed it yet, their civil registration system is not well matured, or they have chosen not to have a unique identifier for their residents. Examples of the latter are Australia and New Zealand. In these countries, matching information from different registers becomes more difficult but not impossible. Other characteristics such as name, sex, date of birth, and address will then be used to match individuals.[8] This can be accomplished either through a deterministic (where the match has to be exact) or a probabilistic (where matching is based on a similarity score, often referred to as fuzzy matching) method.

1.3.2 The combined census

For some countries, such as the Scandinavian ones, migrating from traditional to the register-based census was a planned transition over several decades, where key features were purposefully put in place. The key register infrastructure was already developed and the legal basis for register-based censuses and statistics in general was amended or established. The NSOs would run trials with register data alongside the traditional census until they were confident that the quality would not be compromised by making the transition. As a result, when they made the switch, they did it so completely.[9]

However, some countries may not have register data available for all census variables, but they still want to use currently available register data to support their census. Some of these countries have opted for a *Combined Census*, which uses register data for some census requirements, and traditional enumeration is used for the remainder.

There is no clear definition of the combined approach, as there are many ways register data can be used to assist the implementation of a census. Different countries have chosen different strategies based on their specific circumstances. However, the main point is that

[8] Ibid.
[9] See https://unece.org/fileadmin/DAM/stats/publications/Register_based_statistics_in_Nordic_countries.pdf.

register data is being used to provide as much census information as possible, complemented by data either from a full enumeration or one or several surveys.

We can thus broadly divide the combined approaches into three categories:

- Register data combined with full enumeration
- Register data combined with survey data
- Full enumeration supported with register data

Register data combined with full enumeration

In these instances, register data may be utilized to pre-fill questionnaires, which respondents can then update or verify. This reduces the response burden and may assist in checking the quality of register data in preparation for relying solely on register data for certain census variables. Some countries may already have some census variables available from registers of sufficient quality and do not need to ask this information through the questionnaire. Reducing the questionnaire size, increases the likelihood that respondents be willing to participate in the census.

Register data combined with survey data

Another approach is to rely on register data for key census variables (similarly to a "short form" in some traditional census approaches), supplemented by survey data for more detailed information not available in registers. This model can take on a variety of forms. Some may rely on existing surveys, while others would have a census survey carried out specifically to produce census data. Often, a large sample size is required for such a survey to allow for providing detailed statistics for small geographical areas and subpopulations. However, even a large-scale survey is less costly than a full enumeration operation, and it also reduces the response burden for the population.

Full enumeration supported with register data

Some countries may not have a population register or alternative registers that can be combined to make a full count of the population, or they may not want to move away from a full enumeration for their census. However, they may still want to utilize registers to improve the quality and the efficiency of their census. For other countries, the long-term goal may be to transition to a register-based census or a combined approach where register data provides a significant portion of the census information.

There are several ways register data can support a census without being the sole provider of a set of census variables. For instance, information on addresses/dwellings, possibly with geographical coordinates, in one or several registers may assist in developing a list of addresses for the distribution of census questionnaires. It may also help fill data gaps for hard-to-reach population groups or non-respondents. It can also be used for quality assurance purposes for collected census information.

Australia is a country that utilizes register information to improve the quality of its census but has no stated aim of moving towards a register-based census. Australia does not have personal identification (ID) numbers or a population register. The Australian Bureau of Statistics lists five ways they used administrative data for their 2021 census:[10]

- *Help communities during the census*. Administrative data is used to help identify communities and/or population groups that may need special assistance in filling in the census questionnaire. This could for instance identify the need to deploy census staff that speak a certain language to a specific area.
- *Help improve the census count*. Administrative data can assist in identifying dwellings that are unoccupied on the census night.
- *Preparation for unexpected events*. Events such as bushfires (or other natural disasters) or the coronavirus disease (COVID-19) pandemic, may suddenly make data collection very challenging in the short time span of a census. Administrative data can assist in filling gaps in the data collection caused by such events.
- *Update register of addresses* for distribution of census questionnaires.
- *Count people in hard-to-find places.* This could be in institutions such as prisons, and, for Australia, for people residing in their Antarctic Base.[11]

[10] See the country's case study in chapter 3 for more details.
[11] See https://abs.gov.au/census/planning-2021-census/using-administrative-data-2021-census.

CHAPTER 2

CENSUS IN THE ASIA-PACIFIC REGION

2.1 Trends in the Asia-Pacific region

Every census cycle, the United Nations Statistical Division (UNSD) conducts a survey to map various aspects of approaches used by countries.[12,13] In terms of data collection approaches, several important trends emerged from these surveys for the 2010 and 2020 cycles.[14]

In 2010, traditional censuses were the most common globally, with 85 per cent of countries reporting this as their main census method (see table 1). For Asia and the Pacific, the corresponding figure was slightly higher at 94 per cent. For the 2020 round, the proportion of countries reporting traditional censuses as their main method had decreased globally to 71 per cent, a continuation of the previous trend albeit at a slower pace. In 2020, about 78 per cent of responding countries in Asia and the Pacific conducted a traditional census.

The reduction in the use of traditional approaches to censuses naturally means the introduction of alternative approaches. The key alternative approach is a register-based census, but it is also common to use a combined approach. This trend started in Europe in the late 1900s, gained momentum in the 2000 and 2010 rounds and has strengthened in the 2020 round. In 2010, some 10 per cent of responding countries reported using registers as their main census approach, whereas 5 per cent used some combination of methodologies. The corresponding figures for Asia and the Pacific were 0 and 6 per cent, respectively. These 6 per cent represented two countries, Singapore and Turkey. For the 2020 round this had developed further, so that of all responding countries globally 10 per cent used registers only, 13 per cent used a combination of registers and full enumeration, and a further 5 per cent used registers in combination with sample surveys. In Asia and the Pacific, among responding countries, 2 per cent conducted (or planned to conduct) a full register-based census (which represented one country, Turkey[15]), 12 per cent combined registers with full enumeration and 7 per cent combined registers with sample surveys. The countries and territories reporting the combined use of full enumeration and registers were American Samoa, Indonesia, Iran, Malaysia and New Zealand, and the countries that reported the combined use of sample surveys with registers were Armenia, Mongolia, and Singapore.

[12] See https://unstats.un.org/unsd/statcom/51st-session/documents/BG-Item3j-Survey-E.pdf.
[13] See https://unstats.un.org/unsd/demographic-social/census/documents/UNSD/overview.pdf.
[14] It should be noted when looking at global and regional findings from these surveys that not all countries provided responses. This means that strict comparison between the two points in time, and comparisons with other sources of information on similar topics, should be done with caution. But it is still possible to discern some overall trends.
[15] Turkey actually conducted a combined census in 2021, but with an increasing reliance on registers.

Table 1: Proportion of countries by reported main census methodology, 2010 and 2020 census rounds, percentage

Census approach	*World*		*Asia and the Pacific*	
	2010	*2020*	*2010*	*2020*
Traditional	85	71	94	78
Register	10	10	0	2
Combined	5	18	6	20

Sources: UNSD, "Overview of national experiences for population and housing censuses of the 2010 round", June 2013 and UNSD, "Report on the results of the UNSD survey on 2020 round population and housing censuses", March 2020.

Note: Number of responding countries differs between surveys (for World, 2010: N=123, 2020: N=158, and for Asia and the Pacific, 2010: N=33, 2020: N=41).

There is thus a slow, but steady uptake of the use of registers, either alone or in combination with full enumeration or sample surveys. This applies in Asia and the Pacific as well as globally. It should be noted that transitioning from a traditional census to either a combined or register-based census takes careful planning that may span several census rounds. It is therefore not likely to change very rapidly.

The UNSD surveys also looked at what kind of enumeration was applied for field-based data collection. However, the questions on this topic changed between 2010 and 2020 and are thus not comparable. This is partly due to the introduction of new approaches such as CAPI and CAWI that were not widespread in 2010. In 2010, UNSD[16] reported that PAPI was the most commonly used enumeration method globally, with 75 per cent. The CAPI was still not widely used, with 11 per cent of responding countries reporting that this method was used in the 2010 census round. The use of online questionnaires (CAWI) was becoming more common, though, with 26 per cent. Table 2 shows the enumeration methods for the 2020 round, globally and in Asia and the Pacific. In 2020, CAPI had taken over as the most common enumeration method, with 72 per cent of responding countries globally using this method and 63 per cent in Asia and the Pacific. The CAWI was also becoming more popular, with 35 per cent of countries globally and 48 per cent amongst responding Asia-Pacific countries. The PAPI was still frequently used, with 43 per cent globally. As many as 60 per cent of countries in Asia and the Pacific still used PAPI in the 2020 round. These figures point to a shift towards electronic data collection, either via internet self-response or face-to-face interviews. This shift has three important implications. Firstly, it speeds up data processing and therefore, the timeliness of census reports, as data entry from paper questionnaires (either manually or by scanning) is no longer necessary. Secondly, it greatly improves quality by reducing data entry errors, as several checks are built into the applications (for instance by prohibiting out-of-range entries). And thirdly, it has the potential to significantly reduce the overall cost of the census by reducing the staffing needs.

[16] See https://unstats.un.org/unsd/demographic-social/census/documents/UNSD/overview.pdf.

Table 2: Method used (or planned to use) for field-based enumeration, 2020, percentage

Method	*World*	*Asia and the Pacific*
PAPI	43	60
CAPI	72	63
CAWI	35	48
CATI	11	13

Source: UNSD, "Report on the results of the UNSD survey on 2020 round population and housing censuses", March 2020.

Note: Number of responding countries differs between surveys (for World, N=142 and for Asia and the Pacific, N=40). Countries could select more than one method.

These UNSD surveys are thus indicating that two major shifts in census taking are underway globally as well as in Asia and the Pacific. The first is the rapid uptake of electronic devices and the use of the internet for data collection. This has been motivated by the need to reduce costs and been facilitated by easy access to new technology, leading to gains in terms of quality and timeliness. The second shift is a slower one, has only recently begun and may take several census cycles to complete. This change involves using administrative data or registers, either to replace field enumeration entirely or in conjunction with some field enumeration. In both cases, the motivation is to reduce costs and improve timeliness and frequency.

2.1.1 Impact of COVID-19 on the 2020 census round

The COVID-19 pandemic coincided with the main implementation period of the 2020 census round. With severe restrictions around the world focusing on reducing mobility and interaction among people in order to reduce the spread of the virus, this was destined to have significant impact on the implementation of censuses in many countries.

The UNSD conducted two special surveys among member States, the first in March/April 2020 and the second in December 2020/January 2021. These two surveys sought to map the impact of the pandemic on the implementation of censuses around the world. They have also convened expert meetings[17] to share lessons learned and provide guidance to NSOs on planning for and carrying out censuses during the pandemic.

Several trends were discernible from the surveys.[18] A key finding was the majority of countries that responded to the survey had to postpone the implementation of the census to a later stage. Other countries, who were able to keep their original census reference date, had to extend the data collection period in order to ensure that everyone was enumerated. It is therefore safe to say that the COVID-19 pandemic has had a massive impact on the 2020

[17] See https://unstats.un.org/unsd/demographic-social/meetings/2022/ws-census-20220127.
[18] See https://unstats.un.org/unsd/demographic-social/meetings/2021/egm-COVID19-census-20210209/concept-note.pdf.

population and housing census round. A key challenge was, unsurprisingly, to deal with the severe risk of spreading the virus from face-to-face interviewing of large amounts of people.

The UNSD[19] notes three key modifications that NSOs explored in order to meet this challenge:

- Introducing or extending data collection methods that do not require face-to-face interviewing. These include internet-based self-enumeration (respondents respond to a questionnaire online), paper-based self-enumeration (where questionnaires are dropped off and picked up again without physical contact), and telephone interviews.
- Changing or reducing the questionnaire size to ease the burden on respondents.
- Exploring the use of administrative registers to improve the census coverage or provide some variables from these registers. Linked to this is also attempts at imputations for non-responding households by using proxy respondents (i.e., transfer values from households with similar characteristics).

With regards to Asia and the Pacific, the first UNSD survey on the impact of the COVID-19 pandemic on the 2020 population and housing round[20] found that 12 out of 16 responding countries in this region had to postpone all or parts of the census from the original reference date.[21] This clearly shows the impact that the pandemic has had on the 2020 census round in the Asia-Pacific region. The four countries that reported the pandemic had not impacted their census implementation plans had either managed to conduct the census data collection in early 2020 before the pandemic reached their country or had not yet reached a decision regarding data collection scheduled for later in the year.

The second UNSD survey was also sent out to countries that had originally planned for their census to take place in 2021. Of the 23 responding countries from Asia and the Pacific, 18 reported having to postpone their censuses, one had already completed their census, three had not been affected by COVID-19 at the time of reporting, and one planned to conduct a register-based census and therefore would be unaffected by COVID-19. These results are in line with the findings of the first survey, indicating that very few countries were able to carry out their census operations as planned.

Of the countries that reported COVID-19 affecting or postponing their census operations, 12 reported making changes to their mode of data collection.

[19] Ibid.
[20] See https://unstats.un.org/unsd/demographic-social/census/COVID-19-SurveyT1.
[21] The first survey only went out to countries originally scheduled to conduct their population and housing census in 2020.

Table 3: Number of the Asia-Pacific countries introducing or increasing various data collection methods due to limitations from COVID-19

Data collection method introduced or increased	*Number of countries*
Use administrative data	3
Internet (CAWI) data collection	6
Telephone (CATI) data collection	8
Self-enumeration with paper questionnaire (mail-out/mail-back, drop-off/pick-up)	7
Other changes/adaptations	8

Note: Countries could select more than one method. The number of countries that responded was 23.

As Table 3 indicates, the main changes implemented are the ones that involve moving away from face-to-face interviewing. Different approaches require varying degrees of technological inputs. Changing to a Drop-Off/Pick-Up (DOPU) approach requires logistical planning, but very low technological input. Setting up facilities for interviewing by telephones also requires fairly low technological inputs, whereas providing an online questionnaire requires a higher level of technological sophistication. However, this would then depend on whether the online facility was already in place and merely extended to more respondents, or whether this facility would have to be prepared from scratch. Several of the responses in the "other" category indicate that a number of countries, for instance, Singapore, Sri Lanka, and the Philippines, have extended the use of already existing internet and/or telephone interview facilities.

The table also shows that three countries had used or expanded the use of administrative data to counter the challenges due to COVID-19, these include the Republic of Korea, Malaysia, and the Federated States of Micronesia. The fact that fewer countries resorted to this strategy is perhaps not surprising. Using administrative data requires first and foremost that this source of data is available. Although more and more countries will have government administrative systems that collect and store such data, it may not yet be suitable for statistical purposes. Furthermore, the NSO needs to legally be able to access and use such data and have the technological know-how to utilize it. This may take longer to implement than, for instance, setting up a facility for telephone interviewing. Although a crisis such as the COVID-19 pandemic may make a shift toward using administrative data more appealing to a country, this shift may be more suitable as part of a long-term plan rather than as a short-term solution. We may thus see more countries exploring the use of administrative data for their census in the next census round as a consequence of the challenges faced by the COVID-19 pandemic.

Countries were also asked to list the three main challenges in conducting their census linked to the COVID-19 pandemic. The most common factor, listed by 18 countries, was the need to reduce face-to-face interaction. This was followed by mobility restrictions/problems with transportation, listed by nine countries, and procurement difficulties. Another challenge also

mentioned was that personnel were not available. A common thread in these challenges is the need to reduce interaction between people (especially enumerators and respondents) as well as the need to reduce the reliance on staff, as in a pandemic their availability will be precarious. The obvious solution to these challenges for a census operation, as we have seen, would be to rely as much as possible on online self-response and administrative data. However, for many countries these solutions are not always ready at hand, especially the use of administrative data, and they may only be a partial solution, at least in the short term.

2.2 Census approaches in Asia and the Pacific for the 2020 round

As part of the preparation of this report, the author has mapped the planned and/or implemented census approaches in 58 of ESCAP's members (including associated members) for the 2020 census round. Four of ESCAP members, France, the Netherlands, United Kingdom, and the United States of America, were excluded since they are not located in the Asia-Pacific region. This mapping was completed by reviewing the websites of NSOs. Additional information was required for some members to determine their census methodology. The UNSD[22] and SPC[23] websites were also consulted for census years.

The main methodological difference between this mapping exercise and the UNSD survey mentioned earlier is that all members are included here whereas the UNSD survey was voluntary and did not include all members. Another difference is that UNSD surveyed planned approaches, whereas this mapping exercise looks at the actual approach for the countries where the census was conducted. The numbers therefore differ somewhat, but the overall trends are the same.

Table 4 indicates that 47 of the 58 member States and associate members included in this review (representing 81 per cent) have used or are planning to use a traditional approach in the 2020 census round. A total of ten countries (representing 17 per cent) are planning to or have carried out a combined census, with five countries combining with full enumeration and five with sample survey data. In addition, one country, Afghanistan, has carried out a hybrid census, which is a combination of a sample survey with satellite images. This was due to long-standing security and accessibility issues in the country (see the country brief in chapter 4).

[22] See https://unstats.un.org/unsd/demographic-social/index.cshtml.
[23] See https://sdd.spc.int.

Table 4: Number of members in the Asia-Pacific region by census type, 2020 round

Type of census	*Number of member States*	*Percentage*
Traditional	47	81
Combined - Register and full enumeration	5	9
Combined - Register and sample survey	5	9
Hybrid	1	2

Source: Author's research

We can thus see that currently, all Asia-Pacific countries that have taken steps away from a traditional census methodology have done so by combining register data with either a full enumeration or sample surveys. In order to combine register data with sample surveys, it needs to be possible to establish a statistical population register based on either a population register or by a combination of administrative data sources. Republic of Korea, Singapore and Turkey are good examples of this category, as they base a key set of census indicators on register data, and then complement this with a large-scale survey for the remainder of the indicators. This reduces both the cost of the operation and response burden.

For countries that are either not yet able to do so, or choose not to, the available option is to combine the register data with full enumeration, as five countries in the region have done.

Such a combination of register data and full enumeration can take many forms, as discussed earlier. The Australian case study shows one such variation of this approach, where data from the full census enumeration is the preferred choice, but gaps in the census data will be filled with administrative data. For more information on which census approach each country or territory in Asia and the Pacific are implementing for the 2020 round, please see Annexure.

CHAPTER 3

COUNTRY CASE STUDIES

In the following we will present four country case studies, showcasing how some countries in the region have approached exploiting administrative data in their censuses. In some cases, administrative data is eliminating the need for full enumeration of the population, whereas in others, it is utilized to support and improve the implementation of a full enumeration census. These cases show there are many ways that administrative data can be meaningfully used to improve the quality and efficiency of census-taking. They also highlight the different types and quality of administrative data that countries may have access to.

The four case studies are followed by a presentation of a few brief examples from other countries in the region that to a varying degree are using administrative data in their census or are exploring ways to do so. These brief examples underline the overall picture of a region undergoing significant change when it comes to approaches to census-taking.

3.1 Australia – Using administrative data to support census operation

The Australian Census is listed in our overview as a Combined Census. This is primarily because administrative data is used to support traditional enumeration to carry out the census. The Australian Bureau of Statistics (ABS) is currently not aiming to produce any census variables from administrative data only, but rather use administrative data in innovative ways to improve the efficiency, cost-effectiveness and quality of the census operation itself.

3.1.1 Introduction of the use of administrative data for the 2016 Census

The year 2011 marked a major change in the way ABS carries out its Population and Housing Census. Up until this time ABS has mostly carried out a traditional census. A large temporary workforce was recruited to firstly go out in the field and list every household across the country (the mapping or listing exercise), and then make contact with these households to deliver either paper questionnaires or online Census login codes to them (the data collection stage). This exercise, especially in a country as large and sparsely populated as Australia, was costly and time-consuming. ABS noted that during the 2011 Census, less than half of the doors were answered when field enumerators tried to visit the households.[24] This meant that enumerators had to visit the households multiple times, and it was difficult to recruit sufficient

[24] See https://abs.gov.au/ausstats/abs@.nsf/Lookup/by%20Subject/2900.0~2016~Main%20Features~How%20we%20collected%20your%20information~2.

enumerators to carry out this massive operation. It led to high costs and still resulted in dwindling response rates.

This experience led to an overhaul of the strategy for the 2016 Census. Three main changes were introduced:

1. The development of a national Address Register to support mailing out of materials to households across Australia. The Address Register was formed using the Geocoded National Address File as its base and then built on using information gained through visiting every address through a large canvassing exercise, in addition to analysing other available data. This register formed the basis of addresses to which information was mailed out.

2. The postal delivery of an instruction letter detailing how to complete the Census online or how to request a paper form. Paper forms were then dispatched and mailed back via Australia Post so that Census field officers only visited a minority of homes, where the use of the mail service was not considered viable or where a Census form had not been returned.

3. The use of a smarter online form. Many enhancements were made to the online form to improve quality and make it easier for respondents to complete.[25]

The combined effect of these changes was that it was possible to hire significantly fewer field staff while still maintaining the high standard of the census operation. According to ABS, this was one of the major contributors to the savings of over AU$ 100 million in the 2016 Census. It is worth noting that change was facilitated by two important aspects. Firstly, the use of administrative data on buildings or addresses, combined with other data collection methods and verification processes to establish a national Address Register for use in subsequent censuses. Secondly, the scale-up of the use of online Census forms as the default approach for the majority of Australian households.

We can already here see the shape of a strategy where administrative data is used primarily to support the modernization of the Census operation based on full enumeration, rather than aiming to produce some census variables from register data. Coupled with the increased use of online-based data collection, the goal is to simultaneously reduce the cost of the census operation and the response burden, while maintaining the quality of the census outputs.

3.1.2 Expansion of the use of administrative data for the 2021 Census

This strategy was further developed for the 2021 Census. In addition to the already mentioned use of administrative data to establish and maintain the national Address Register and also to provide information for people living in institutions such as prisons and on the Australian

[25] Ibid.

Antarctic Base, ABS noted that three additional uses of administrative data were introduced in this census round.[26]

i. Help communities during the Census

ABS is using administrative data, often in combination with previous census data, to better identify and understand communities that in the past have had lower than average response rates. These communities or groups can be Aboriginal and Torres Strait Islanders groups, young adults, elderly or migrant populations where English skills may be low.

Once the groups have been identified, ABS then works closely with organisations that target these groups to reach out with important information about the Census and how to fill in the questionnaire. This may involve information sessions or provision of information in relevant languages. By identifying areas where, for instance, there are many elderlies who may struggle to fill in questionnaires online, ABS is able to dispatch staff to these areas who may assist the target group in completing the census form.

In doing this, ABS uses administrative data from these sources:

- health register data
- social security register data
- migration information
- taxation register data

ii. Help improve the Census count

The Australian Census is a *de facto* census, that means it counts where people are on Census night (and not their usual place of residence (*de jure*)). A key question for the Australian Census is therefore if a house (or dwelling) without a response is a house that was unoccupied on the night, or if someone was, in fact, home, but no response was obtained. If no data is obtained, ABS will then attempt to impute information for this household from a similar household to adjust the count. Research[27] conducted by ABS showed these imputations could sometimes lead to overcounting certain demographics.

In order to address these challenges, administrative data was used to improve the census count in two ways.

Firstly, administrative data was used to assess which houses were empty on Census night. When no information has been obtained for occupants of a dwelling, ABS used various administrative data sources to determine if this was in fact an occupied dwelling. Data used included Government service data from sectors such as health, social services and taxation. In addition, administrative data from private actors such as data on electricity use from

[26] See https://abs.gov.au/census/planning-2021-census/using-administrative-data-2021-census.
[27] See https://abs.gov.au/websitedbs/D3310114.nsf/Home/Can+administrative+data+help+to+improve+the+Census+count.

electricity distributors were used at area level (i.e., not individual level). These data sources assisted in assessing if anyone from the address in question had used any of these services lately.

Secondly, administrative data was used to adjust the census count. Once a house had been determined to be occupied, and no data was obtained, administrative data was used to choose how best to determine the number of people living in this house. This was used by copying census data from a house that had common features of administrative data with the house in question.

iii. Being prepared for unexpected events

There are many events that could potentially disrupt data collection in the small window that a census provides. A census is attempting to give a snapshot picture of the total population at a particular time and is therefore dependent on simultaneous data collection across the whole country. Australia in the past has had experience with natural disasters such as bushfires that can severely disrupt data collection in parts of the country during the census. Some areas may for whatever reason turn out a lower response rate than the average, or what is deemed acceptable. More recently, countries across the world have had their census plans severely impacted by the COVID-19 pandemic, as especially face-to-face data collection became difficult.

In such circumstance, ABS may use administrative data from aforementioned sources such as taxation register data, health register data and social security register data to fill gaps in the census data.[28] There are two main ways that gaps can be filled by administrative data.

Firstly, administrative data could be used when census information (from regular data collection) seemed to be missing for households or groups of households from the area where the unexpected event occurred. Basic demographic information may then be transferred from the administrative data source. Secondly, more detailed census information for these households or groups of households may be filled either by administrative sources or by census data from the previous census. Date of birth, sex, country of birth are examples of variables that would not change over time. ABS is also looking into using administrative data to fill in indigenous status, if this information is missing. This will then be done in consultation with and with support from the Aboriginal and Torres Strait Islander community.

3.1.3 Data privacy

Concerns around data privacy are pivotal in the discussions around the use of administrative data for statistical purposes in Australia. This is one of the reasons that ABS will only use administrative data to produce census variables as a last resort, and rather use administrative

[28] See https://abs.gov.au/statistics/research/using-administrative-data-fill-possible-data-gaps-census.

data to improve the census operation based on full enumeration. Keeping the public's trust that data gathered for statistical purposes is only used for this and not by ABS or others for compliance or other checks is very important to ABS.[29] Thus, ABS has developed the following principles for when administrative data can be used to fill gaps in the census:

1. *Census data quality is significantly affected*: There must be a large enough impact to the quality of Census data for ABS to use administrative data. For example:
    - ABS can't get data for a town or local region
    - The gaps for a particular population are large enough that Census can't give accurate information for planning and policy decisions.
2. *Administrative data is of high enough quality*: Administrative data must be of high enough quality to fix the impact on Census data quality.
3. *Delays to Census results are acceptable*: The benefits to data quality must clearly outweigh the costs, particularly any delays to Census results.
4. *Owners of the administrative data are supportive*: Before using administrative data to fill gaps in the Census, the owners of the administrative data (like the Australian Taxation Office) must agree to ABS using their data for this reason.
5. *Transparency and keeping the public's trust*: ABS is transparent about how they might use administrative data in the Census and the benefits. ABS must assess and limit any impacts on privacy before using administrative data.

3.1.4 Comparing census data with administrative data

ABS is also conducting research or quality control projects where the census results are compared to administrative data. Such a comparison was done for the 2016 Census and a similar exercise is planned for the 2021 Census results. A central motivation for this exercise is that in order to use administrative data to support the census, ABS needs to know how well the administrative data represents the Australian population. Research from 2016 shows that the difference between administrative data and the official ABS population estimate for Australia is only about 8000 people or 0.03 per cent of the population.[30] There are some more significant differences for small population groups, but overall, the administrative count of the Australian population is very accurate.

To be able to conduct such comparisons and use administrative data to support the census and other projects, ABS maintains what is referred to as Person Linkage Spine (or the Spine).[31] The Spine is based on three datasets with health register data, social security data and taxation register data, and it is updated annually. This Spine enables ABS to make linkages with other datasets and improve the accuracy of research.

---

[29] Ibid.
[30] See https://abs.gov.au/statistics/research/assessing-administrative-data-quality-enhance-2021-census.
[31] See https://abs.gov.au/websitedbs/d3310114.nsf/home/Person+Linkage+Spine.

3.1.5 Conclusions

As can be seen from this case study, labelling the Australian Census as a Combined Census may be slightly inaccurate. It may be more accurate to say that the Australian Census is a 'Census supported by administrative data'. We have seen that Australia does not, at this stage, aim to move away from a census based on the principles of full enumeration, but rather to make this enumeration as effective and high quality as possible. Two key features in their modernization process are to use administrative data to improve several census processes and to use online data collection whenever possible to reduce the need for field visits.

Even though there is no plan to move towards a register-based data collection system, administrative data is being used in several ways to support the census operation. The establishment and maintenance of the national Address Register, supported by administrative data, has reduced the need for an expensive mapping exercise at the outset of the census work, and facilitated distributing census forms through the mail rather than through dispatching enumerators. And even though administrative data is seen as a 'last resort' when it comes to filling in actual data for household, consideration is being given to using such data when response rates are particularly low and other means of data collection have been exhausted.

The Australian census may be a useful example for countries who are not aiming to establish a population register or collect census information from administrative data but would like to improve the efficiency of their traditional census using already existing administrative data. This will of course depend on what kind of administrative data is available in any given country.

3.2 Indonesia – Using administrative data for pre-listing households[32]

The 2020 Population and Housing Census in Indonesia was the seventh population census in the country, but the first that used a combined approach. Up until the 2010 Census, the approach had been traditional. However, Indonesia decided to leverage the existence of population data from their Population Registration System to modernize their census approach.

3.2.1 Legislative changes

In order to facilitate this transition towards a combined census, with increasing use of register data, several legislative instruments had to either be introduced or amended. Law Number 16 of 1997 on Statistics was already assigning Badan Pusat Statistik (BPS or Statistics Indonesia) the task of providing statistics, including censuses, for Indonesia. It also gives a generic reference to the use of administrative data for statistical purposes. In 2006, the Law on Population Administration (Number 23) was established, providing a mandate for Population Registration and Civil Registration. Every resident is provided a unique identifier (the Indonesian abbreviation is NIK) through this provision. This law was further elaborated in 2013. Several other laws also touch upon the administrative collection of population data, especially in relation to registration of civil events.

Then, in 2019 two Presidential Decrees were introduced to further prepare for the combined census and the use of administrative data. First, Presidential Decree Number 39 on One Data Indonesia provided regulation to standardize data collected by various government agencies. Such standardization is important as it greatly facilitates the linking of information from various administrative sources at unit level. Secondly, Presidential Decree Number 62, of the same year, introduced National Strategies for the Acceleration of Population Administration for the Development of Vital Statistics. This decree specifically granted Statistics Indonesia the right to use administrative data from the Ministry of Home Affairs, the custodian of the Population Registration System, for the 2020 Population and Housing Census.

3.2.2 The Population Registration System

The Population Registration System is managed by the Directorate General of Population and Civil Registry within the Ministry of Home Affairs. Civil registration is an integral part of the Indonesian Government's poverty-reduction policy, with a focus on providing identity documents to all citizens.[33] Having a legal identity is key to access several services such as education and health, and civil registration is central to produce vital statistics that inform many important policies at both national and local levels. Several of the legal changes

[32] Source was from communication with NSO representatives.
[33] See https://getinthepicture.org/country/indonesia.

mentioned above are also very relevant to the development of the civil registration and vital statistics system.

However, there are still many challenges when it comes to registration of vital events in Indonesia.[34] For instance, Indonesia Socioeconomic Survey[35] indicates that 34 per cent, or 27 million children, were still unregistered in 2016. Both marriages and divorces were also often not registered. The same applied to death registration. Another important event, domestic change of residence, is also highly under-reported. Put together, this means that the population data from the civil register is not yet of a quality that can be used to provide census variables in a combined census. Even the basic population variables, the ones normally collected in a short form, do not have sufficient quality to provide an accurate picture of the population size and structure in Indonesia.

It should be noted that the Government of Indonesia is placing heavy emphasis on improving the registration of civil events and providing legal identity to all. This is illustrated by Statistics Indonesia's strategy to move towards a register-based census in 2030. While the Population Register System is not capable to produce the short form variables for the 2020 Census, the aim is to do so in 2030.

3.2.3 The 2020 Census – introducing administrative data and new technologies

The shift in Indonesia towards using administrative data for their census should be seen as part of a wider shift in taking advantage of new technologies in data collection. In parallel with leveraging administrative data, Statistics Indonesia also introduced at large-scale the use of online data collection together with expansion of electronic data collection in the field using tablets (CAPI). All these new technologies and approaches can be seen at various stages on the census implementation.

i. Civil registration data as input in the pre-listing phase

Even if the quality of the population data from the civil registration system was insufficient to be used for creating census variables, the data was considered very useful for the census operation. The main use of administrative data was in the pre-listing phase. Following the Presidential Decree Number 62, Statistics Indonesia entered into a Memorandum of Understanding (MoU) with the Ministry of Home Affairs to access civil registration data for use in the 2020 Census. Once the population data from the civil registry was obtained by Statistics Indonesia, this file was matched with Statistics Indonesia's own population file, which is based on the previous census. Even though all citizens are supposed to have a unique ID (NIK), the low quality of this variable meant that Statistics Indonesia had to conduct various matching exercises. The first step in this matching exercise was a deterministic approach

[34] See https://getinthepicture.org/resource/civil-registration-and-vital-statistics-crvs-indonesia.
[35] Ibid.

where NIK and name of individuals from the two sources were identical and a match could be performed. The second approach was a probabilistic approach where more variables were used for the matching, but the match did not have to be 100 per cent. In addition, matching was conducted for administrative areas to facilitate the preparation of the population listing with locations. This exercise was carried out in collaboration with the Ministry of Home Affairs with several validation rounds for both source data sets.

ii. First stage data collection – Online census

The information collected in the pre-listing phase resulted in a temporary Statistical Population Register, with basic information at individual level. In the second stage, an online census was carried out. During this stage people were encouraged to fill in their census information via an online questionnaire. This online questionnaire was pre-filled with the information from the Statistical Population Register, and respondents were asked to verify or update the information (or add if information was missing). Based on the online census an updated Statistical Population Register was produced.

iii. Second stage data collection – Verification at local administrative level

In this stage the updated Statistical Population Register was verified at local level by Statistics Indonesia staff in collaboration with local leaders. The local leaders would be asked to assist with verifying the list of households for their administrative area. Any households that the local leaders were not sure existed would be flagged. These households would then, in turn, be identified in the field and visited by field enumerators. After confirming the presence of household members, a paper census questionnaire would be dropped off for later picked up (DOPU).

iv. Third stage – Field enumeration

The remaining households from the verified Statistical Population Register, who had not responded to the online questionnaire or been identified in the second stage as doubtful and visited for a drop-off questionnaire, were then visited in the final stage of the short form data collection. These households were then interviewed by field enumerators using either CAPI or PAPI.

A post enumeration survey was also conducted after the completion of the short form data collection. A key objective with this survey was to assess the quality of the field enumeration and the administrative registers.

v. Long form – Sample survey

While the short form enumeration entailed canvassing the whole population and producing basic population data, the long form, providing more detailed information such as fertility,

mortality and migration, was conducted as a sample survey. About 5 per cent of the population, or 3.05 million households, were included in this exercise. This exercise, undertaken in 2021, provided representative estimates at subregional level. The data collection exercise used both CAPI and PAPI methods.

3.2.4 Towards a combined approach

As we have seen, Indonesia has been able to take advantage of the fairly new existence of administrative population data to reduce costs and improve the effectiveness of the census operation. While the administrative data was not considered of sufficiently high quality to be able to provide census variables directly, it was used to improve and speed up the household listing. In combination with online census, it significantly reduced the need for field enumeration.

As Indonesia prepares to rely more heavily on administrative data for the 2030 Census, more efforts are needed to improve the coverage and quality of the civil registration system. For Indonesia to rely on this data source to produce the variables from the short form, and not only use it for household listing, the coverage of the population must approach as close as possible to 100 per cent. In addition, procedures must be improved to capture all vital events, including deaths and domestic relocations, in a timely fashion.

Despite these challenges, Indonesia is an example of a country that is using administrative data to support census operations and reduce the need for field enumeration significantly.

3.3 Republic of Korea – Introducing a combined census approach[36]

The Republic of Korea conducted its first Population Census in 1925 and included a Housing component for the first time in 1960. The approach taken up until the early 2000s was that of a traditional census with a short form collected through a full field enumeration and a long form with more detailed questions administered through a sample survey of approximately 10 per cent of the households. After the census in 2000 the Government of the Republic of Korea started raising issues around the census methodology. Motivating factors for exploring this shift were the quality, the cost-efficiency and the timeliness of the census. In 2008, the Republic of Korea officially announced plans to conduct research and prepare for a new approach to their census.

The aim of Statistics Korea (KOSTAT) was to establish a combined census approach, where data for the short form was collected through administrative data, thus eliminating the necessity of a full field enumeration. The long form was to still be collected through a large-scale sample survey. A starting point for KOSTAT was that the Republic of Korea had several well-established administrative data sources, including a civil registration, namely the Resident Registration Register (RRR), and vital statistics system, that could potentially be used to provide census information. As the Republic of Korea conducts their census every five years, 2010 was the next potential census for introducing the use of administrative sources. But the Republic of Korea chose to conduct a traditional census in 2010 and use the result of this census to carry out comparisons to data from administrative sources. This work, in turn, prepared them for moving to a combined census in 2015.

3.3.1 Existence of administrative data sources

As mentioned, the Republic of Korea had several administrative data sources that contained information relevant to a census. In the end, KOSTAT established a statistical register based on 25 administrative registers from 14 public institutions. This data was combined to create a microdata set for a register-based census. The administrative registers used for census are listed in Table 5.

[36] Sources were from communication with NSO representatives and the meeting presentations at the Data Integration Community of Practice (https://stat-confluence. escap.un.org/pages/viewpage.action?pageId=23790042) and the side event on Register Based Censuses at the Second Ministerial Conference on Civil Registration and Vital Statistics in Asia and the Pacific (https://getinthepicture.org/system/files/sites/default/files/5.%20UNFPA_Republic%20of%20Korea%20Presentation.pdf).

Table 5: Administrative registers used for the Korean census

Category	*Administrative Data*
Population (Koreans)	Resident Registration Register (RRR) which is the population register
Population (Foreigners)	Foreign register: Domestic residence register of Koreans having foreign nationality, Register of undocumented (short stay) immigrant, Nationality acquisition register
Supplemental actual residence	College register, Social facility and user list, Dormitory (college) user list, Entry and departure register
Household	Family relations register, Military register, Expatriates in the diplomatic service register, Conscripted police register, Conscripted maritime police register, Conscripted firefighter register, Reformatory inmate register, Juvenile correctional institution detainee register
Housing	Building and dwelling registers, officially announced housing price data, Building energy use data
Others	Employment insurance data of the insured, the Register of the insured in workplace, Nursing home user register, Presumed dead person data

All these sources were used to provide the various census variables that were s necessary to produce the short form. To do so, many different registers had to be linked at unit level. Personal ID numbers in registers were transformed into virtual numbers (10 digits) that functioned as the matching key. This matching key was created only for statistical purposes. In addition, any administrative data containing address information were subject to standardization, which was mostly carried out automatically. Approximately 8 per cent of addresses could not be standardized in an automated way and had to be standardized manually.

3.3.2 Comparative study of administrative data and census data

In order to assess the readiness of the administrative data sources to replace full enumeration for the 2015 Census, a large study was carried out after the completion of the 2010 Census. The basis for this study was an administrative population register based on the RRR that was compared with the actual 2010 enumerated Census Population.

Table 6: Population of the Republic of Korea by administrative and census data *(as of 1 November 2010)*

Population	*Administrative data*	*Census data*
Korean	50,470,000	47,990,000
Foreigner	980,000	590,000
Total	51,450,000	48,580,000

As shown in Table 6, there are some discrepancies between the two sources, prompting KOSTAT to identify the key sources of such inconsistencies. The RRR included individuals who may have moved abroad to work, study or for other reasons. These individuals had to be removed using data on immigration (entry and exit), in order to better reflect people currently present in the Republic of Korea. In addition, there were some delays in reporting of births and deaths that also caused some discrepancies. Some of these errors could, however, be rectified using late corrections in the register.

Another comparison was made for the number of households. Again, there was a discrepancy, as the RRR database reported 19.84 million households compared to 17.34 million households from the census. The difference was thus 2.5 million households more from the administrative data than from the census data. This difference was because some members of these households resided in a different address from what was reported in the administrative registers for reasons of schools and housing.

KOSTAT also investigated the match between the actual (as enumerated in the census) and the registered (in the administrative registers) place of residence. Looking at the lowest level of administrative boundaries, 92.1 per cent of the population's registered place of residence was located in the same village/town/community (Eup/Myeon/Dong) as their enumerated place of residence in the census. Several reasons for this were identified along with possible solutions. For instance, the RRR did sometimes not reflect enrolments in dormitories or admissions to social facilities. This could be addressed by obtaining lists of residents in the boarding facilities at schools and universities and at social facilities. Further, several individuals who had changed workplace had not had this change reflected in the RRR. KOSTAT used national pension and employment insurance data to correct these. In addition, potential omissions, delays and errors in administrative data were verified by using historical data.

3.3.3 Matching between administrative data sources and between administrative and census data

In these studies, a key activity was to find a way to link, at individual unit level, different sets of data from various sources. For instance, sometimes it was necessary to match several administrative registers to fill data gaps at unit level. In other instances, it was necessary to match two registers to add variables from one register to the main register. In all cases, it was necessary to ensure that an individual unit, whether a person or an address or dwelling, was linked with the same unit in the second register. As mentioned earlier, KOSTAT created virtual ID numbers for persons based on personal ID numbers. When administrative data such as the RRR were matched with survey data where such personal ID numbers were not included, names, date of birth, gender and address were used to aid the matching process. KOSTAT also standardized the addresses using a mix of automated and manual procedures.

The studies mentioned above are just some examples of the many checks KOSTAT conducted to ensure the best possible matching between administrative registers and census data. It was vital for KOSTAT to reach an acceptable level not only in terms of matching the size of the population from both sources but also the many population characteristics that a census describes.

3.3.4 Personal data protection is key

KOSTAT cites ensuring personal data protection as key in facilitating the shift towards a combined census and by doing so also makes it easier to win public approval for this shift. Official statistics in the Republic of Korea, including the census, are governed by the Statistics Act. This law was already in place before the shift to a combined census, but KOSTAT made sure that all changes in terms of access to and use of administrative data and exemptions from personal information protection were firmly based on the Statistics Act. An example of such a way to protect privacy and minimize the danger of misuse of data is the transformation into virtual ID numbers (these do not contain any real information) before inclusion of individuals in the statistical registers.

3.3.5 The 2020 Census and beyond

A result of the shift to a combined census is that KOSTAT is now able to release 13 census items on an annual basis. This means that reliable and up-to-date population statistics can be used for planning purposes more frequently. The linking of new administrative registers means that new statistics can also be produced, for instance concerning multicultural households and single parents. KOSTAT is also looking into expanding the number of census items that are provided from register data, within both population and housing statistics.

The sample survey carried out under the combined census in the Republic of Korea collects responses from 20 per cent of the population. It is important to note that this sample survey collects data for more detailed and additional items in the census as well as general items that are already being produced based on register data. The purpose is enabling quality control and identify ways to improve the part of the census produced from registers. It also aims to provide information that are difficult to grasp with administrative data.

KOSTAT lists five key challenges as they continue to move towards a timely and reliable register-based census:

- Institutionalization of secure access to administrative data for KOSTAT from the owners of the administrative data
- Review of the quality of the administrative data
- Standardization of administrative data across various sources
- Improvement of timeliness of administrative data (i.e., updating routines)
- Continuous protection of personal information

As can be seen, the Republic of Korea has implemented a transition towards a combined census where a core set of census items is produced from register data on an annual basis. This has been made possible due to the presence of strong administrative registers that KOSTAT could access. In 2010, the Republic of Korea implemented their last fully traditional census, but in parallel they collected data from registers for their short form. This enabled them to compare their register data with enumerated census data, and, in turn, better measure and understand the quality of their register data and identify areas for improvement. Such an approach is a good example for countries with well-established sources of administrative data who are looking to move towards a combined census.

3.4 Turkey – A planned transformation towards a register-based census[37]

The last traditional census in Turkey was carried out in 2000. Up until then, censuses were *de facto*, that is mapping where people were on census night. Censuses were usually carried out during a census day, with a national curfew ordering every citizen to stay at their usual residence while census enumerators visited each household to complete a paper questionnaire during a face-to-face interview. The Turkish Statistical Institute (TurkStat) indicates this exercise required around 1 million enumerators and it took about 3.5 years to process all the data for dissemination, even though Optical Mark Recognition, Optical Character Recognition and Intelligent Character Recognition were introduced in the 2000 Census.

3.4.1 The establishment of the address-based population registration system

In the early 2000s the Turkish Ministry of Interior established the Central Civil Registration System, called MERNIS. During this project all citizens, including some not alive, but registered in the past, were given a unique 11-digit ID number. However, a key piece of information was lacking in this system in order to be used for census purposes, and that was a place of usual residence for each citizen. In 2006, TurkStat was thus tasked with establishing what was called the Address Based Population Registration System (ABPRS). As the first part of this work, TurkStat established a National Address Database (NAD), which is a database with all addresses in the country with some basic information to describe each address or dwelling. TurkStat then conducted a full field enumeration to establish each person's usual place of residence together with the ID number from MERNIS. Combining the results from this enumeration with NAD resulted in the ABPRS. The ABPRS is not only a statistical database, it is also an administrative data source used by other government agencies. For instance, it is used by the Ministry of Education to enrol pupils in schools and by the Supreme Election Committee to create voter lists. The system is kept up to date as by law any changes of address have to be notified to the register within 20 days.

It is worth noting that legal issues pertaining to collecting and maintaining data had to be addressed. Firstly, the Civil Registration Services Law (Number 5490) was introduced in 2006. This regulation charged TurkStat with establishing the ABPRS and it gave the Ministry of Interior the mandate to maintain and update the system. The exercise to conduct the full field enumeration to establish people's usual place of residence was mandated in this law. In

[37] Sources were from communication with NSO representatives, the meeting presentation at the Data Integration Community of Practice (https://stat-confluence.escap.un.org/pages/viewpage.action?pageId=13336726) and Census methodology in Turkey: Transformation from traditional to register-based administrative data (https://content.iospress.com/articles/statistical-journal-of-the-iaos/sji190561)

addition, the Statistical Law (Number 5429), which was already in place, mandated TurkStat's activities pertaining to field application and data confidentiality.

3.4.2 The 2011 Census – Introducing the combined approach

The 2011 Census was the first opportunity to use the newly established ABPRS for census operations. The system was able to produce approximately 20 per cent of the census variables directly from the register, the remaining 80 per cent had to be collected using a large-scale sample survey. Approximately 2.2 million households, or 11.7 per cent of the population, were sampled. In addition, full enumeration was implemented in institutions, such as universities, nursing homes, prisons and military barracks. This is a large number, but significantly less than a full enumeration, which meant that fewer field staff were required. The number of staff was reduced from approximately 1 million to only 4,500. This, in turn, led to lower costs, better quality of staff and improved or simplified logistical operations. TurkStat indicates the total cost of the 2000 Census, conducted the traditional way, was 48.3 million United States dollars whereas total cost of the 2011 Census was 13.9 million United States dollars.

In addition to using register data from ABPRS to produce 20 per cent of the census variables, register data from NAD was also used to create the household listing and the enumeration areas for the field operations. Using NAD data, enumeration areas of approximately 100 occupied dwellings each were created to help facilitate the data collection in the field.

Another change was that the one-day curfew was abandoned and the data collection period extended to three months. This caused some challenges with recalling the situation on census day for the respondents. Tablets were also introduced for the first time in 2011, when approximately 60 per cent of the interviews were conducted using a tablet (CAPI) and the remaining 40 per cent using paper questionnaires (PAPI). A consequence of the combined use of register data and the introduction of tablets during field data collection was the reduced time gap between data collection and the release of the census results. This time lag was reduced from three and a half years in 2000 to one and a half years in 2011.

Even though ABPRS was the main source of administrative data for the 2011 Census, TurkStat also used some other administrative sources to support the census operation. For instance, Social Security Institutions registers, Turkish Employment Agency registers and municipalities building permits registers were all used for quality assurance purposes with regards to data collected in the field as well as to impute missing observations.

3.4.3 Key drivers in the use of administrative data for census in Turkey

TurkStat mentions several key drivers that are behind this shift towards a register-based census. The need for a large number of temporary staff and the associated costs have already been mentioned. In addition, there is a need to produce timelier, more relevant and more

frequent population statistics. Also, TurkStat, as the NSO of a candidate country for membership in the European Union, relates to Eurostat and the European Statistical System. Eurostat has a stated goal of producing annual census statistics from 2024 onwards. It is worth mentioning that an advantage of a full register-based census is that it can be conducted much more frequently at very low cost.

Another important aspect driving the shift towards a register-based census is the increasing accessibility of registers in Turkey. These registers are there primarily for administrative purposes and not statistical purposes, but their existence is a prerequisite for the shift towards a register-based census.

3.4.4 Expanding the use of register data for the 2021 Census

Between the 2011 Census and start of the planning for the 2021 Census, several important events took place. In 2012, the Institutional Living Quarters Address Database was established, making it easier to gather data on population groups living in these institutions. Several other registers were also established or further developed. The NAD was also transformed to a Spatial Address Registration System by adding coordinate information to the addresses.

In order to facilitate the use of these additional administrative sources for statistical purposes, TurkStat is developing the Population Characteristics Database. This database brings together data from several administrative sources which enables TurkStat to prepare an increasing number of census variables directly from these registers. The Population Characteristics Database includes information from these registers:

- Social Security Register (contains the insured's earnings, premium amount, etc.)
- Income and Tax Register (contains annual income tax returns and earnings records)
- Employment and Unemployment Register (contains job seeker records)
- Disability Register (contains type and rate of disability)
- Education and Exams Register (contains active and passive student records, higher education exams, etc.)
- Agricultural Register (contains agricultural activities)
- Tradesman Information System (contains tradesman records)
- Road Motor Vehicle Register (contains motor vehicle registrations, etc.)
- Retirement Register (contains retirement records)
- Social Assistance Register (contains pension records, social assistance records, etc.)

Key census information that can be produced from this database include age, sex, nationality, disability, education, employment status, property, migration and family structure. The key that makes it possible to combine all this data is the National ID number. Since all these underlying registers use the same unique ID number for every person, all information relating

to one person can easily be linked to their identity. By combining all these supplementary registers in the Population Characteristics Database with the ABPBS, TurkStat aims to produce approximately 70 per cent of the census variables for 2021 census from register data.

For the remaining 30 per cent of census variables on building and dwellings characteristics, TurkStat has carried out an ad-hoc survey, with an estimated sample size of 306,000 households, between October 2021 and March 2022. They used CATI technique for this survey. The estimated survey duration for each household is six minutes, which indicates the reduced response burden resulting from collecting most information directly from registers.

3.4.5 Quality control an important aspect in introducing new register data

For all administrative sources that have been included as part of the Population Characteristics Database, their quality has been thoroughly assessed prior to inclusion. Missing observations is a constant challenge. In addition, various checks including content control, target population control, definitions of variables, metadata control, missing data control, and plausibility control are carried out. Such quality checks were also conducted on various administrative data sources in the fields of labour force and building and dwelling statistics. For both of these topics, quality checks revealed that they were not yet of sufficient quality to be used as sole source for these census variables.

Indeed, TurkStat lists the following challenges with the use of administrative data in the production of statistics, including censuses:

- Lack of metadata
- Different classifications by different data owners
- Missing values in the registers
- Differences of definitions
- Discrepancies in reference dates between data sources
- The existence of multiple records (duplications)
- The lack of adequate IT infrastructure

The insights gained from the above-mentioned studies will be fed back to the register owners in order to improve their data for the future. TurkStat thus plays an important and active role in improving and sometimes establishing new registers in Turkey. The aim for the near future is to be able to produce all census variables from register data, thus eliminating the need for costly data collection in the field. Given that the registers are continuously updated, it is also the aim to produce detailed population statistics on an annual basis.

3.4.6 Conclusion

Turkey is an example of a country that is in the midst of a planned transition from a traditional census to a fully register-based census, via the combined approach. They have systematically

put in place the legal framework, built organisational capacity and the necessary IT and other infrastructure to take advantage of existing administrative data sources in the country. Further, they have worked closely with the owners of the various administrative data sources to both gain access to these sources and, when necessary, assist the data owners in improving the quality of the registers. Turkey still implements what TurkStat refers to as a “register dominant” combined census in 2021 but will in all likelihood move towards a full register-based census in the next round.

CHAPTER 4

BRIEF COUNTRY EXAMPLES

4.1 Afghanistan – A hybrid census to meet the need for population data

Afghanistan has seen many decades with conflict and instability, and as a result has not conducted a full census since 1979. According to Bahadur Hellali, Deputy Director from the Afghanistan National Statistics and Information Authority,[38] the census in 1979 only covered 67 per cent of the districts in the country. The main reason for this was security issues. More recent attempts to conduct a full census also stranded due to the same reasons. Although population projections based on the latest census provide some population estimates for Afghanistan, the lack of reliable population data, especially at lower geographical levels, is a major hindrance for effective government planning. The need to provide reliable estimates for the Sustainable Development Goals indicators was also a driver, as population figures are essential to calculate many of these.

In this situation, the Government of the Islamic Republic of Afghanistan collaborated with the United Nations and the University of Southampton and others to come up with a new approach for providing population data, given that the possibility of conducting a full census remained remote. The result of this collaboration was the implementation of a "Hybrid Census" in Afghanistan in 2016.[39] The starting point for this exercise was the premise that it was impossible to collect field data in large parts of the country. At the same time, the availability of detailed satellite imagery, the ability to collect geo-referenced codes for respondents in a survey, paired with sophisticated statistical models enabled analysis not previously possible.[40] Micro-census surveys were then carried out in areas where it was safe to do so, and this survey data was then linked using statistical models to the spatial data. This allowed the researchers to predict population totals and age/sex structures at small areas. The researchers could also produce estimates of uncertainty for the figures.

It should be noted that such a hybrid census is an outlier, in the sense that very few countries implement them. It is not an alternative to a regular census as it provides very few variables and comes with uncertainty. But when the alternative is to not conduct a census at all, for security or other reasons, it can provide very valuable, albeit basic, essential information for planning purposes.

[38] See https://unstats.un.org/unsd/undataforum/blog/hybrid-census-to-generate-spatially-disaggregated-population-estimates.
[39] See https://unfpa.org/sites/default/files/resource-pdf/Hybrid_Census_Brief_v9.pdf.
[40] Ibid.

4.2 Armenia – Transition from traditional to combined census[41]

Armenia is conducting a combined census for the first time in 2022. The census was supposed to take place at the end of 2020. But due to the COVID-19 pandemic, it was postponed to 2022. Their last census, in 2011, used a traditional approach. However, for the upcoming census, the Statistical Committee of the Republic of Armenia (Armstat) is planning to use administrative data for a small set of variables available in the State Population Register (SPR), whereas a supplementary household sample survey of 25 per cent of addresses in the SPR will be used to improve the data from the register. The Republic of Armenia Statistics Act supports Armstat's access to the data and metadata hold by the administrative authorities for producing official statistics.

The Armstat is planning to use SPR combined with the Border Management Information System (BMIS) to count the population. The SPR, established in 2006, includes all citizens of the Republic of Armenia with their corresponding place of residence, according to the procedure established by the Law "On State Population Register". Using these two combined administrative data sources, Armstat can produce a full count of the population, including Armenians living abroad, foreigners living in Armenia, asylum seekers, and refugees. The SPR provides the following information:

- Personal ID number
- Name and surname
- Status (residence or refugee)
- Actual registration address
- Citizenship
- Date and place of birth
- Sex

The sample survey will collect all census variables and be used to verify and improve the SPR, in addition to being the primary source for many of the census variables, including variables related to assessing the coverage of the civil registration. The survey data will be collected by census enumerators using tablets (CAPI).

A census pilot was carried out in October 2019 in two regions. A key lesson learned from this pilot was the need to improve the system for electronic transfers of data between SPR and BMIS. This is being implemented for the main census.

[41] Source was from communication with NSO representatives.

4.3 Mongolia – Transitioning towards a combined census[42]

Mongolia's census in 2020 was the first combined census where administrative data was used for significant parts of the census. In 2014, the National Statistics Office of Mongolia set up the National Population and Household Registration Database (PHRD) as a repository for information stemming from various government administrative systems. The PHRD is a statistical database that is only used for statistical purposes. It is an online system with continuous updating where regional Government staff are tasked with updating the database.

As the Statistic Law in Mongolia currently does not grant the NSO legal access to all administrative data in the country, the NSO had to establish data exchange agreements with over 40 Government agencies in order to populate the database. For instance, prior to 2017, the NSO did not have access to the Civil Registration Database. However, since then, NSO has received regular updates from the Civil Registration Database to update its population data. According to the principle of One Citizen One Registration, the PHRD receives information on births from the General Authority of State Registration on a monthly basis. Several other agencies also provide data to the PHRD on various topics such as education and employment. All in all, the PHRD currently contains about 20 variables at the household level and 20 variables at the individual level. However, the PHRD still encounters some issues. The country does not have a standardized address system yet, and standardization of classifications for many topics across Government agencies is a challenge.

For the 2020 Census a core set of census variables were collected directly from the PHRD, and was also compared and validated by data from Government agencies such as the Department of Labor and Welfare, the Ministry of Labor and Social Welfare and the Ministry of Health. For the remainder of the census variables a more detailed questionnaire was administered to 10 per cent of the households, selected from the PHRD.

4.4 New Zealand – Census transformation programme[43]

In 2012 Stats NZ launched a new programme called the Census Transformation Programme. A key reason behind this new initiative was the rising cost of running a traditional census. There were also challenges with bias caused by non-response. In addition, opportunities from new technologies and increasing availability of relevant administrative data sources in New Zealand, made it an especially opportune time to explore new ways of conducting the census.

The Census 2018 was the first census where this new programme was introduced. Administrative data was used to provide information on individuals not reached through traditional enumeration. The Stats NZ notes that approximately 89 per cent of the population

[42] Sources were from communication with NSO representatives and the meeting presentation at the Data Integration Community of Practice (https://stat-confluence.escap.un.org/pages/viewpage.action?pageId=13337660).

[43] See https://stats.govt.nz/topics/census.

was counted through census responses and 11 per cent through administrative data. It is worth noting that Stats NZ made the decision to not impute units for their census count. However, imputations were used to provide missing census characteristics when these could not be produced from either enumeration or administrative data.

The Stats NZ are continuing their census transformation as they plan for the next census in 2023. The plan is still to conduct a full enumeration, but with further support from administrative data sources. The Stats NZ has conducted numerous studies to support their census transformation programme, many of them are available on their website[44].

4.5 Thailand – Piloting the use of register data to support the census[45]

The National Statistical Office of Thailand had to postpone their planned census for 2020 due to the COVID-19 pandemic. Their last census prior to that was in 2010. The 2010 Census was a traditional census, but in this round new data collection approaches were introduced for the first time. In addition to regular in-person interviews and Drop-Off/Pick-Up (DOPU) questionnaires, the NSO also introduced the options of telephone interviews and filling in the questionnaire online.

The NSO of Thailand planned to further modernize their data collection for their 2020 Census, and prior to the postponement they implemented two pilots to assess how administrative data potentially could support the census implementation. The first pilot, conducted in 2018, used register data to pre-fill some information including address information prior to data collection by enumerators using tablets. A key finding from this pilot was that for some individuals, the registered addresses did not correspond to their usual residence. This complicated the process of identifying the interviewees. Another important finding was the increased interview time because a considerable amount of the pre-filled information (about 30 per cent) differed from the information collected in the field.

The second pilot, conducted in 2019, aimed at finding the most effective way of collecting census information from high-rise buildings. Information about accessing census questionnaires online was sent out to respondents, but the response rate was less than 10 per cent. A conclusion from this pilot was the need for better public relations campaigns to improve public awareness about the census and to boost the online response rate.

Given the ongoing challenges with COVID-19 and the necessity to reduce the cost and to address declining response rates, the NSO of Thailand are continuing to explore ways to use register data to support their census. Several administrative data sources are being explored

[44] Ibid.
[45] See https://unstats.un.org/unsd/demographic-social/meetings/2021/egm-COVID19-census-20210209/docs/s05-06-THA.pdf.

and further pilots are planned. This will include efforts to combine various administrative sources to improve the quality of the register data.

4.6 The Pacific Island Countries and Territories – Increasing interest in administrative data

The Pacific Island Countries and Territories (excluding New Zealand and Australia) still employ a traditional approach to their censuses, although there is growing interest in applying new technologies and approaches. Whereas the 2010 census round in this region was dominated by paper forms (either self-enumeration where questionnaires are dropped off and picked up or filled in by enumerators based on a personal interview), the 2020 census round saw increasing use of tablets in the data collection process. This transition contributed to faster data collection, but it still required enumerators to prepare a household list and then physically visit all households to collect the information.

There are, however, signs that this region may also start the introduction of various forms of combined censuses in the future, perhaps in the next census round. For instance, Cook Islands are currently exploring using Government administrative data to support income and education information from the census. Both Government payroll data and university education data have been accessed for the current round, with the aim of imputing figures for gaps in the census data. The Cook Islands do not currently have a national ID system or a population register or a national address register, but discussions on their development are ongoing.[46] Such a system would greatly increase the possibility of moving towards a combined census where some of the census variables are based on administrative data only.

Samoa is another country in the Pacific region that is introducing infrastructure that will enable a shift towards a combined census. The Government of Samoa is working towards establishing a country wide identity management system and deploying a civil registration system.[47] They are thus aiming to establish a comprehensive civil registration system to give every citizen a unique identity, and this will also form the basis for a population register. This infrastructure could, in turn, be used as the basis for parts of the census questionnaire, thus reducing the burden of census data collection. These examples show that there is every chance that the next census round in the Pacific will look quite different from what it has been in the past.

[46] See https://sdd.spc.int/events/2020/09/webinar6-pacific-island-use-admin-data.
[47] See https://sbs.gov.ws/nationalid.

CHAPTER 5

CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions

Over the last decades there have been, globally, increasing challenges to the traditional census. Collecting information from every person in a country using traditional methods is a massive and costly exercise and thus a key concern. Further, reduced willingness amongst the population to respond to the census questionnaires makes it increasingly difficult for NSOs to produce reliable figures with the necessary geographical and substantive detail. And lastly, census data is crucial in various government decision-making processes, that increases pressure to produce more granular and timelier census results.

On the other hand, developments of new technologies and approaches to data collection mean that there are also emerging opportunities. These can be divided, broadly speaking, into two sets. The first is the development of new technologies such as the internet and tablets and mobile phones with Global Positioning System (GPS) technology and applications for electronic data collection and processing. While these have been around for two decades or so, there are still countries who have not fully utilized their potential and are still relying partly on paper questionnaires for their census data collection. The second is the increasing introduction and the quality of electronic administrative record keeping in government systems across the world. This introduction has happened at different times and at different pace and with varying quality in different countries, but the overall trend is that these administrative data sources are more prevalent now than they were a few decades ago.

These global trends are also present in Asia and the Pacific. While there is no country in the region that has yet conducted a fully register-based census (i.e., a census involving no census-specific data collection being conducted), several countries have started to use administrative data in their census operations in various ways, and some countries are very close to a fully register-based census (e.g., the Republic of Korea, Singapore and Turkey).

Amongst the countries in Asia and the Pacific that have used administrative data in the 2020 census round or have started exploring the use of administrative data for future censuses, there is much variation in their approaches. As mentioned, some countries have advanced far towards a register-based census such as Turkey and the Republic of Korea. These countries are producing a significant and increasing proportion of the census solely from register data and collecting the remaining information from a sample survey. They rely on well-established population registers and are aiming to integrate more registers in their census work in the future. Other countries, such as Mongolia and Armenia, rely on data from their population

register to produce a small amount of census variables, whereas the remainder is collected in a long-form sample survey. Furthermore, other countries, such as Australia, are not aiming to develop a population register, but are using various administrative sources to greatly increase the efficiency of their census in innovative ways. Australia and New Zealand are also examples of countries where register data is used to fill gaps at unit level when enumerated data cannot be obtained. Some countries, such as Indonesia, are using population register data in their listing exercise to reduce the need for a large-scale mapping exercise. Even if the data emanating from their population register is not yet of a sufficient quality to produce census variables, they are using it to pre-fill online questionnaires for respondents to verify. This way they can reduce the response burden and increase the quality of the population register. We also see that some countries, such as Thailand, are exploring ways to take advantage of administrative data sources by assessing their quality in various pilot exercises. It is worth mentioning that all these countries are also taking advantage of other technological advances such as online data collection and the use of electronic data collection devices. This also applies to many countries still conducting a traditional census.

The countries mentioned above, while being at very different stages and employing different strategies, are all listed as conducting a "Combined Census". While this is, perhaps, technically correct, as all countries are using a combination of administrative data and field-based data collection, it is not very useful in describing their individual approach to their censuses. There is a need to better describe and categorize countries in their various stages of their census development, perhaps by creating several categories for a combined approach. At the very least, distinction should be made between countries producing some census variables solely from administrative data (whether in combination with a full enumeration or with a sample survey) and countries using administrative data to support the census operation in various ways but without relying on it for census variables.

This study has shown that countries are currently using administrative data in a myriad of ways to support their census. Some are relying on well-functioning population registers with unique ID numbers that can be used as key to link to information from other registers. Others may have a population register that does not yet have sufficient quality in terms of coverage and completeness but are using it to support the listing/mapping phase of the census. Others may not have a population register but are using a combination of other administrative data sources such as health, tax or social security registers, to support their census in various ways. This could for instance be to provide information for people who did not return a census questionnaire, or it could be to conduct post-enumeration quality control of the census results. It would be very useful to collate all these different approaches to using administrative data in census work, as assistance and guide for countries contemplating the use of or expansion of their use of administrative data for censuses.

The various approaches taken by countries in Asia and the Pacific region illustrate that there are different pathways to different goals. Not all countries may want to or, indeed, should

establish a population register as such but may still want to take advantage of growing access to administrative data. Similarly, there can be no one-size-fits-all approach towards register-based censuses, but rather country-specific strategies depending on data availability and future potential. It should also be noted that replacing enumeration with administrative data without proper quality control procedures can lead to reduced quality of the census outputs. Some caution is warranted when it comes to implementing such a transition too quickly.

Given that administrative data is not generated for the sole purpose of a census, NSOs may want to collaborate closely with other initiatives that support developing and improving administrative data in the country, such as the push for improving civil registration systems through the Get Everyone In The Picture campaign.[48] As seen in the country study of Indonesia, the drive towards improving civil registration may open up possibilities for taking advantage of administrative data in a census. While civil registration and vital statistics are especially important for Population and Housing Censuses, better administrative data on addresses, dwellings/buildings, health, education, taxation and social security, to name a few, would also be extremely beneficial.

5.2 Recommendations

- Further work on detailing the potential ways that various administrative data sources may assist a census operation, from providing information for the mapping/listing exercise to producing census variables and quality control would be useful in order to provide examples, lessons learned and insights for countries considering the use of administrative data sources for census purposes.
- Countries considering using administrative data for their future census should first undertake a full mapping exercise of current available and relevant administrative sources in the country. Such an exercise may also look at current legal framework relevant for the use of administrative data for censuses as well as relevant strategies for the development of administrative data sources in the country. Such a mapping exercise could be the starting point for a strategy for moving towards a combined or register-based census.
- NSOs may use the planning process of the National Statistical System to align and establish good working relations with potential custodians of relevant administrative data.
- Development partners may consider providing technical support to countries wishing to undertake such a mapping exercise.

[48] See https://getinthepicture.org.

- There is a need for better guidance on quality control of administrative data by NSOs. Many NSOs in Asia and the Pacific region may have no or limited experience in using administrative data for statistical purposes. Thus, there may be a need for the training of staff in assessing whether any given administrative data source is fit for statistical purposes. Guidelines on the use of registers and administrative data exist, such as the ones from UNECE,[49] and these could be the used as a basis for training.

- Countries should be encouraged to prepare and disseminate methodological notes on how their censuses were conducted. These should be made available to other countries so that knowledge and experience can be shared, and learning facilitated.

- Countries should be cautious when transitioning away from full enumeration and certainly before moving to a full register-based census. It may be useful to run a short form full enumeration alongside a register-based short form for one census round. This will allow for comparison and proper quality assessment of the administrative data used.

- Countries should keep in mind that administrative data can support a census in many ways, as shown in this report. Even if the administrative data sources in a country is still not capable of producing census variables, they may be very useful for the census operation.

- Countries can benefit greatly from available resources such as those developed by UNECE[50] and UNSD.[51] Similar to the findings of this report, the UNECE's guidelines recommends to: a) identify administrative sources against specific use, b) build and strengthen relationships between NSOs and data suppliers, with a legal basis and collaborative mechanisms, c) harness supplier relationships to ensure a comprehensive understanding of source metadata, d) assess the coherence and compatibility of the administrative source to the census, to understand differences, e) understand restrictions and challenges to acquiring an administrative source and integrating it into a census, f) assess and manage the risk implied by use of an administrative source, g) be transparent in communication with data users and with the public, h) undertake feasibility research and test runs prior to including administrative data in census production, i) make use of expert review, j) record and publish results of quality assessment at all stages, and k) develop an NSO-specific quality assurance framework and strategy.

[49] See https://unece.org/fileadmin/DAM/stats/publications/2018/ECECESSTAT20184.pdf.
[50] See https://unece.org/sites/default/files/2021-10/ECECESSTAT20214_WEB.pdf.
[51] See https://unstats.un.org/unsd/statcom/53rd-session/documents/BG-3e-Handbook-E.pdf.

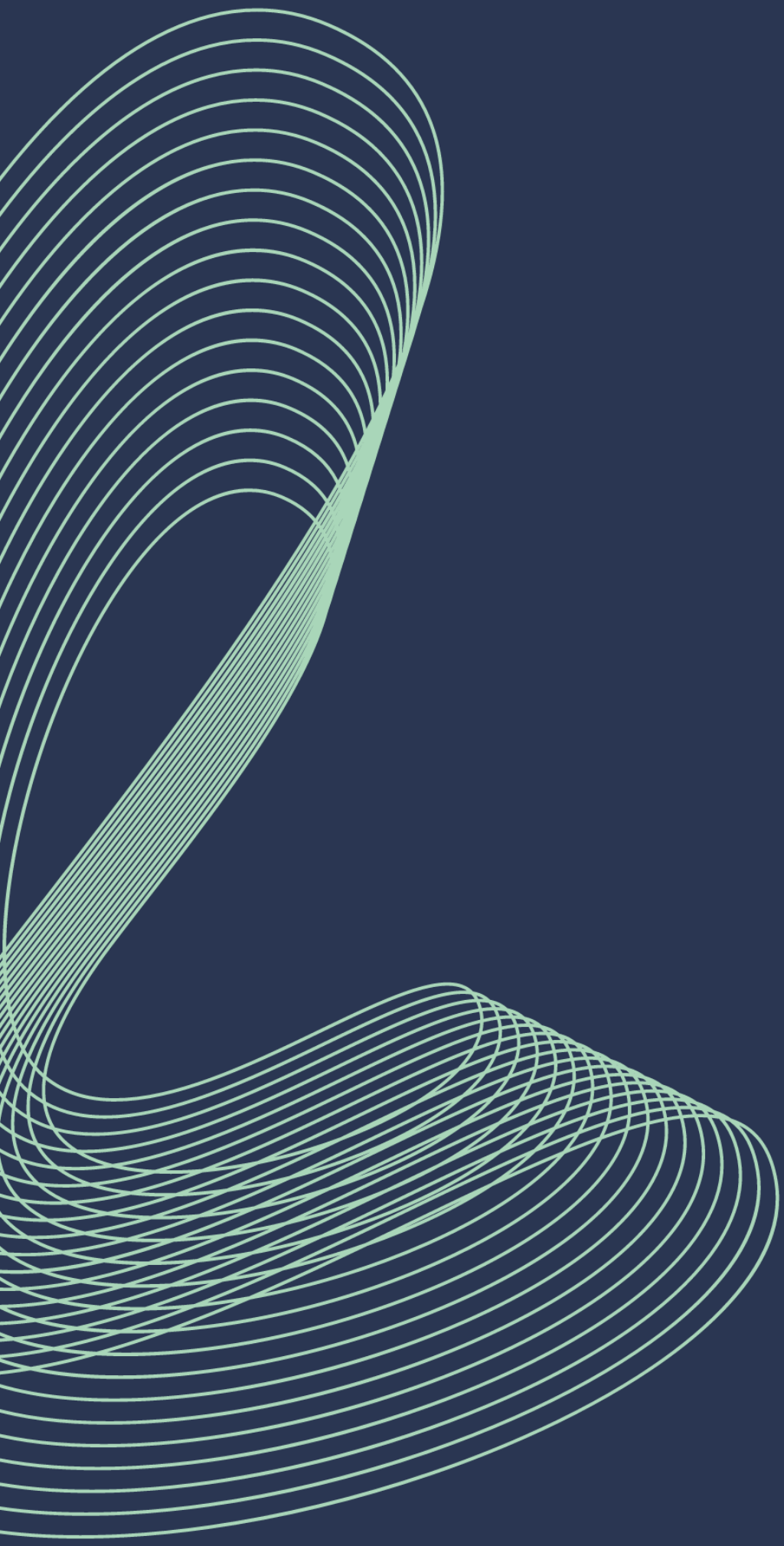
ANNEXURE

Table 7: Type of census in the Asia-Pacific region for the 2020 round

Country	ESCAP Subregion	Type of census	Census year
Afghanistan	SSWA	Hybrid	2016
American Samoa	PAC	Traditional	2020
Armenia	NCA	Combined - Register and sample survey	2022
Australia	PAC	Combined - Register and full enumeration	2021
Azerbaijan	NCA	Traditional	2019
Bangladesh	SSWA	Traditional	2021
Bhutan	SSWA	Traditional	2017
Brunei Darussalam	SEA	Traditional	2021
Cambodia	SEA	Traditional	2019
China	ENEA	Traditional	2020
Hong Kong, China	ENEA	Traditional	2021
Macao, China	ENEA	Traditional	2021
Cook Islands	PAC	Traditional	2021
Democratic People's Republic of Korea	ENEA	Traditional	2019
Fiji	PAC	Traditional	2017
French Polynesia	PAC	Traditional	2022
Georgia	NCA	Traditional	2024
Guam	PAC	Traditional	2020
India	SSWA	Traditional	2021
Indonesia	SEA	Combined - Register and full enumeration	2020
Iran (Islamic Republic of)	SSWA	Combined - Register and full enumeration	2021
Japan	ENEA	Traditional	2020
Kazakhstan	NCA	Traditional	2021
Kiribati	PAC	Traditional	2020
Kyrgyzstan	NCA	Traditional	2021
Lao People's Democratic Republic	SEA	Traditional	2015
Malaysia	SEA	Combined - Register and full enumeration	2020
Maldives	SSWA	Traditional	2022
Marshall Islands	PAC	Traditional	2021
Micronesia (Federated States of)	PAC	Traditional	2020
Mongolia	ENEA	Combined - Register and sample survey	2020
Myanmar	SEA	Traditional	2024
Nauru	PAC	Traditional	2021
Nepal	SSWA	Traditional	2021
New Caledonia	PAC	Traditional	2019
New Zealand	PAC	Combined - Register and full enumeration	2023
Niue	PAC	Traditional	2021
Northern Mariana Islands	PAC	Traditional	2020
Pakistan	SSWA	Traditional	2017
Palau	PAC	Traditional	2020
Papua New Guinea	PAC	Traditional	2024
Philippines	SEA	Traditional	2020
Republic of Korea	ENEA	Combined - Register and sample survey	2020

Country	ESCAP Subregion	Type of census	Census year
Russian Federation	NCA (and ENEA)	Traditional	2021
Samoa	PAC	Traditional	2021
Singapore	SEA	Combined - Register and sample survey	2020
Solomon Islands	PAC	Traditional	2019
Sri Lanka	SSWA	Traditional	2021
Tajikistan	NCA	Traditional	2020
Thailand	SEA	Traditional	2022
Timor-Leste	SEA	Traditional	2015
Tonga	PAC	Traditional	2021
Turkey	SSWA	Combined - Register and sample survey	2021
Turkmenistan	NCA	Traditional	2022
Tuvalu	PAC	Traditional	2022
Uzbekistan	NCA	Traditional	2023
Vanuatu	PAC	Traditional	2020
Viet Nam	SEA	Traditional	2019

Sources: a) Census approach — Author's compilation from websites of NSOs, UNSD and SPC.
b) Census year for Pacific countries — https://sdd.spc.int/census-and-survey-calendar?from_year=2021.
c) Census year for Asian countries — https://unstats.un.org/unsd/demographic-social/census/censusdates.



ESCAP 75
MOVING FORWARD TOGETHER