The Challenges of Maximizing the use of Microdata in PICTs

Professor Wadan Narsey

[Report for PACSTAT and PSMB (SPC) 26 June 2022]

Contents

Acknowledgements

I thank the staff of SDD SPC for their assistance for this study:

Michael Sharp (Economic Statistics and Microdata Specialist)

David Abbott (Manager, Data Analysis and Dissemination)

Andrea Borlizzi (Statistics Advisor)

Scott Pontifex (Microdata Curator and Dissemination Adviser)

Olivier Menaouer (Microdata Officer)

I thank especially Michael Sharp, Andrea Borlizzi and David Abbott for their meticulous corrections and suggestions for improvements.

I also thank all the Reviewers for their advice and especially Len Cook and Vince Gavin..

I thank the PICT Government Statisticians and heads of NSOs who responded to the Questionnaire.

Professor Wadan Narsey

wadan.narsey@gmail.com

26 June 2022

Abbreviations and Acronyms and Glossary

ADB	Asian Development Bank
CROP	Council of Regional Organizations in the Pacific
DHS	Demographic Health Survey
EUS	Employment and Unemployment Survey
FPOS	Fundamental Principles of Official Statistics
HIES	Household Income and Expenditure Survey
ICPSR	Interuniversity Consortium for Social Research.
IHSN	International Household Survey Network
IMF	International Monetary Fund
LFS	Labour Force Survey
MICS	Multiple Indicator Cluster Survey
NSO	National Statistics Office.
PACSTAT	Statistical Innovation and Capacity Building in the Pacific, IDA funded project to improve quality of welfare data collections.
PDH-ML	Pacific Data Hub Microdata Library
PHC	Population and Housing Census
PICTS	Pacific Islands Countries and Territories
PSGF	Pacific Statistics Governance Framework
PSMB	Pacific Statistics Methods Board
PUF	Public Use Files
SDD	Statistics for Development Division (SPC)
SDG	Sustainable Development Goal (of UN)
SIDS	Small Islands Developing States
SPC	The Pacific Community
SUF	Scientific Use Files
TYPSS	Ten Year Pacific Statistics Strategy
WB	World Bank
WBML	World Bank Microdata Library

1. Introduction

Organization context for this study

The Pacific Community (SPC) is the principal scientific and technical organization supporting development in the Pacific region. It is an international organization established by treaty (the Canberra Agreement) in 1947 and is owned and governed by its 26 members, including all 22 Pacific Island Countries and Territories (PICTs).

The Statistics for Development Division (SDD) of SPC is the main stakeholder in the implementation of the Ten-Year Pacific Statistics Strategy (TYPSS), which was adopted by PICT leaders in 2009 in recognition of the need for a comprehensive plan to drive improvement and development of statistics in the Pacific region.

The SDD is currently exploring issues around microdata dissemination, including the costs and benefits, challenges of anonymizing the microdata, and possible inadequacies in statistics legislation. The immediate objective of this project is to help the Pacific Statistics Methods Board (PSMB), and other Pacific statistics decision makers, with information to guide decision making in relation to the use of microdata, especially focusing initially on HIES.

To fully understand the context for this research project and to guide its content and recommendations, it is necessary to understand the rationale for the establishment of the PSMB, and the PACSTAT project.

The PSMB was established by Heads of Planning and Statistics of the member PICTs in 2017 with the objective of ensuring best practice methods and standards in the collection, compilation, analysis and dissemination of all Official Statistics, and other statistical indicators arising from censuses and surveys. These core collections include housing and population census, household income and expenditure surveys, demographic and health surveys, multiple indicator cluster surveys, agriculture census, labour market surveys, price indices, trade statistics and disability surveys. The PSMB is expected to work closely with the PICT National Statistics Offices (NSO) and also liaise with the relevant global and regional agencies.[1]

The Pacific region conducts numerous censuses and surveys and has numerous collections of administrative data. However, the data remain underutilised in most cases, except for the publication of very basic reports. This underutilisation of microdata is now generally recognized to be missed opportunities hindering evidence-based policy derivation and monitoring in the PICTs.

The PACSTAT project is an International Development Association (IDA)[2] funded project that is to be implemented over the period of June 2020 to June 2025, with the overall objective of improving the quality of welfare data collection and accessibility to comparable welfare data in the Pacific. It has the following components and sub-components, all of specific relevance to the contents of this Report:

[1] These include the UN Statistics Division (UNSD), World Bank (WB), Asian Development Bank (ADB), UN Economic Commission for Asia and the Pacific (UNESCAP), Pacific Islands Forum Secretariat (PIFS), Pacific Financial Technical Assistance Programme (PFTAC), United Nations Population Fund (UNFPA), the United Nations Children's Fund (UNICEF), the International Labour Organisation (ILO), the Food and Agriculture Organisation (FAO), Australian Bureau of Statistics (ABS) and Statistics New Zealand.

[2] The IDA is a section of the WB Group providing zero or low interest loans for development.

i. To support the administration of PSMB:

   a. including overseeing PSMB commissioned statistical development research
   b. Dissemination of best practice recommendations to regional NSOs
   c. Improve the analysis and publication of gender statistics in the region.

ii. To strengthen SPC-SDD's ability to play a "statistical system leader" role and provide technical assistance on data collection methods:

   a. in welfare data collection, analysis, and dissemination
   b. create comparable indicators, including gender statistics, and micro-datasets to promote data use in the region
   c. promote statistical innovation and experimentation in the Pacific.

Terms of Reference

The Contract specifies a range of general and specific objectives of the assignment, some direct and some indirect resulting from this report. The general are to

1. provide information to help the Pacific Statistics Methods Board (PSMB), and PICT statistics decision makers, to guide decision making in relation to access to, and use of, their development microdata (or unit record data) sourced from population-based censuses and household surveys, particularly HIES.

2. to disseminate best practice recommendations from the PSMB to regional statistics stakeholders.

3. provide capacity building opportunities for Pacific Statistics Officers, SDD and other stakeholders.

4. make recommendation to PSMB in relation to technical matters associated with microdata anonymization, and more general matters in relation to increasing access to Pacific microdata; and

5. synthesise recommendation in a Policy Brief for PICT statistics decision makers that will lead to the increasing use of their microdata.

Very specifically, the contract advises the Consultant to

6. Consult stakeholders, conduct a situation analysis to understand the "blockers" for microdata dissemination, and to review, evaluate and critique the current situation and current procedures and protocols in relation to microdata dissemination and anonymization.

7. Report on issues relating to: microdata sharing and anonymization; the current situation in the Pacific; practices and policy of other organisations (the World Bank, UNICEF, DHS Programme, WHO, ABS ('5 safes') and Statistics New Zealand;

8. critique the practices and policy of Pacific Data Hub – Microdata Library;

9. Prepare a policy brief to support Pacific statistics stakeholders to make decisions in relation to microdata dissemination.

Then section 3 of the contract titled "Scope of Work and Description of Tasks" states "The Project is specifically interested in access to data used in the measurement of welfare, the focus is on issues relevant to HIES microdata dissemination ... for poverty, food security, labour and gender analysis".

Given the objectives specified by the Contract, this Report attempts to

(a) first outline international best practices and policies on "Open Data" in general as the overall context for discussion of openness of microdata, given that many of the indicators for the open data in general, are the result of analyses of microdata (which is the subject of the next chapter).

(b) outline microdata repository policies of global organizations like International Household Survey Network (IHSN), World Bank (WB), Australian Bureau of Statistics and Stats New Zealand.

(c) examine the experiences of the Pacific Data Hub – Microdata Library for the PICTs.

(d) consult stakeholders (via a questionnaire to PICT NSOs) to understand the political economy of "blockers" for microdata use and dissemination in the Pacific.[3]

(e) Outline the benefits possible through freer access to microdata for independent academic researchers, as well as to researchers associated with multilateral organizations.

(f) outline a tiered approach so that PICT microdata can be preserved for posterity (Tier 1); be used in a manner co-ordinated and harmonised by SDD/PDH-ML for priority research into poverty, food security and gender inequalities (Tier 2); access by application and strict conditions (Tier 3); and Free Downloadable Access (Tier 4).

(f) the importance of formal agreements and appropriate legislation to facilitate access.

(g) prepare a policy brief for Pacific statistics stakeholders (PSMB, PICT NSOs and Governments) that outline pathways to increase the use of microdata in the Pacific, with the assistance of SPC-SDD, PICT donors and multilateral organizations.

[3] Unfortunately, the response to the Questionnaire was not as good as hoped for.

2. The Record of PICTs in Open Data

It is globally recognized that "open data" is a facilitator of development, as well as of governments' accountability to their citizens. A most useful organization is Open Data Watch (ODW) an international non-profit organization of experts aiming at improving the coverage and openness of official statistic. ODW is dedicated to promote data *"as a vital input for policies that seek to eradicate poverty, fight inequality, and promote inclusive, sustainable growth…. advocates for increased data investments, promotes data use and impact, improves statistical capacity building efforts, and encourages strong political leadership for data. …. [helping] to produce better data for better policies and better lives".*[4]

A central pillar of the work of ODW is through its Open Data Inventory (ODIN) whose *International Open Data Charter*[5], states the following objectives and principles:

> " Building a more prosperous, equitable, and just society requires that governments are transparent and accountable" (Paragraph 2)
>
> " [enabling] governments, citizens, civil society and private sector organizations to make better informed decisions." (Paragraph 3)
>
> "help highlight trends, identify social and economic challenges and inequities, and benchmark progress in public programs and services." (Paragraph 5)
>
> " presents opportunities to provide innovative, evidence-based policy solutions and support economic benefits and social development for all members of society." (Paragraph 8)

The Open Data Charter not only sets out all the benefits to the signatories, but also commits them to the following principles: that data should be

> (a) open by default[6]
>
> (b) timely and comprehensive
>
> (c) accessible and usable; and
>
> (d) comparable and interoperable.

The signatories to the Open Data Charter are national, sub-national and local governments, and multilateral organizations.[7] They are required to develop action plans for the implementation of the Charter principles. The present study will try to examine how far these principles can be applied to PICT microdata.

---

[4] https://opendatawatch.com/our-work/

[5] This document can be downloaded from Opendatacharter.net.

[6] This is an extremely powerful principle which implies that access should not require an approval process.

[7] None of the PICT countries or the LDC Comparators are signatories to the ODC. Australia and NZ are.

The Global Ranking of PICTS in Open-ness of Data

The Open Data Inventory is a valuable resource for not only for all development practitioners but also for PICT National Statistics Offices (NSO)to assess the openness of their countries' statistical datasets by all the key variables, and to set targets for their own improvement. This resource can also be of great use in understanding where the country's data outputs are in relation to other countries in the global rankings.

The following tables are derived from the ODIN rankings database for 2020.

Table 1 ODIN Global Rankings (2020)

	Average Rank	Number of Countries
Northern Europe	28	10
Eastern Europe	30	10
Australia and New Zealand	35	2
North America	39	3
Western Europe	47	8
Eastern Asia	55	7
Southern Europe	72	15
Western Asia	77	18
South-Eastern Asia	87	11
South America	93	11
Central Asia	102	5
Central America	114	8
Northern Africa	116	5
Southern Asia	117	9
Southern Africa	121	5
Caribbean	124	12
Eastern Africa	126	16
Western Africa	126	16
Middle Africa	148	8
Pacific Islands	**152**	**8**
All	94	187

Source: ODIN country and regions rankings database.

The ODIN database for 2020 suggests that among the 20 world regions, the Pacific region (covering 8 countries) is ranked last, with an average rank of 152, i.e. a lower rank compared to poorer Asian and African regions (see Table 1). The Caribbean (12 countries) was 16th, with an average rank of 124, while Australia and NZ were 3rd, with an Average Rank of 35.

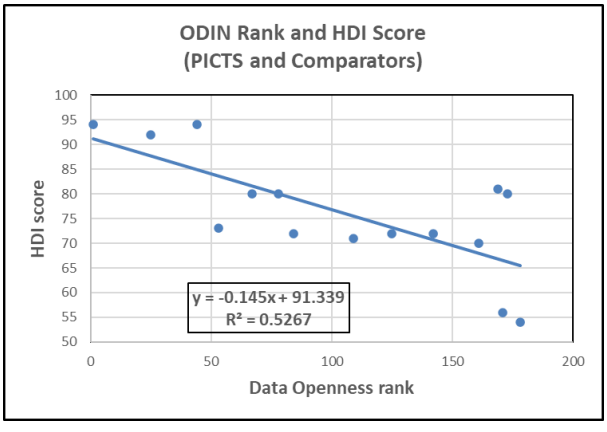


Table 2a describes the ODIN ranks of the PICTs (bottom block) with some useful comparators: Less Developed Country (LDC) Island Comparators are in the middle block, and Australia, NZ and Singapore are in the top block. Table 2 also displays their Human Development Indicator

(HDI) score, while Figure XX shows the direct correlation between the two scores: the better the HDI score (higher numerically), the higher the ranking (lower numerical score) (and vice versa).[8]

Table 2 ODIN ranks of PICTs and comparators		
Countries	ODIN Rank	HDI
Developed country comparators		
Singapore	1	94
New Zealand	25	92
Australia	44	94
LDC Island Comparators		
Jamaica	53	73
Mauritius	67	80
Malaysia	78	80
Maldives	84	72
Trinidad and Tobago	173	80
PICTSs on ODIN database		
Samoa	109	71
Tonga	125	72
Fiji	142	72
Marshall Islands	161	70
Palau	169	81
Solomon Islands	171	56
Papua New Guinea	178	54
Source: ODIN database		

The Country Profiles given by ODIN can also be used as "Report Cards" for the NSOs in assessing their own performance by the specified constituent development criteria which are used to estimate the overall Open Data "Rank.

Annex 1 provides a summary of the detailed assessments made by ODIN of each country's performance by the various social, economic and environmental indicators in terms of coverage and openness.[9] The PICTs by and large score poorly relative to the other island comparators.[10]

Annex 3 gives the World Bank Statistical Capacity (Performance) Indicators for the PICTs and the Comparator countries, while the graph depicts their strong correlation to the ODIN ranking, as expected, with a correlation coefficient of 0.815.

The above discussion on the importance of the openness of data in general is intended to set the context for a similar discussion on the importance of microdata to the development strategies of PICTs, with a particular focus on HIES data. In fact, within the framework of the international movement towards open data, there have also been parallel efforts towards making microdata more available internationally.

It may be useful to keep the perspectives of multilateral organizations - for whom openness of national microdata systems is a "bread and butter" issue which facilitates their operations - separate from the perspectives of large nation states, developed and developing

[8] Of course, "correlation" should not be taken as "causation".

[9] Most of these criteria are also present in the UN Sustainable Development Goals. PICT NSOs ought to be able to present more up-to-date assessments for every indicator.

[10] Trinidad and Tobago, from the Caribbean region, is an exception with poorer scores than the PICTS.

whose interests, in terms of both benefits and costs, may be more akin to the PICTs' NSOs highlighted in this Report.

It is also important to recognize that the emphasis given to open data in a given country is largely influenced by the scale, depth and interests of its research capability in government, academia and community organizations.

The Pacific is Lagging Behind on Microdata Analysis

While the previous section has indicated that the Pacific has been lagging behind on openness for development indicators in general, this section shows that they are also lagging behind on the indicators that require microdata analysis.

The PICTs are generally lagging behind their other SIDS comparators on availability of data on poverty/income and food security/nutrition. While the ODIN tables in Chapter 2 were on development data in general, Table 3 gives the ODIN scores for Coverage and Openness on Food Security and Nutrition and Poverty and Income, which are specifically in this Report's Terms of Reference. Although only 7 PICTs are on their database, most appear to be lagging behind on the Coverage and Openness for both sets of indicators of Poverty and Income, and Food Security and Nutrition. The two graphs[11] suggest the close correlation of these ODIN scores with GDP pc.

Annex 2a presents data on ODIN Aggregate Scores for PICTs and the comparators for all the indicators (social, economic and environmental) and their positive correlation with GDP per capita (current. US$) derived from the WB Indicators database.

Annex 2b also gives a summary table derived from the ODIN Country Profiles of the selected countries, including PICTs, of the overall status of the national data systems with:

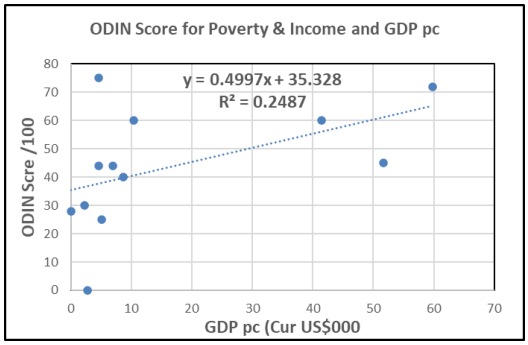
(a) Percentages of indicators not present (most PICTs are in the higher ranges).

(b) Percentages o indicators lacking gender disaggregation (a mixed performance of the PICTs).

(c) The availability of the Websites: whether Open or Available or Not Available (no PICTS, but Singapore, Australia and NZ were classified as being Open).

(d) The existence of national data portals (Singapore, Australia, NZ and Jamaica. No PICTs).

[11] The graphs have excluded the countries with data points showing zeros to reduce distortion of the correlation coefficient.



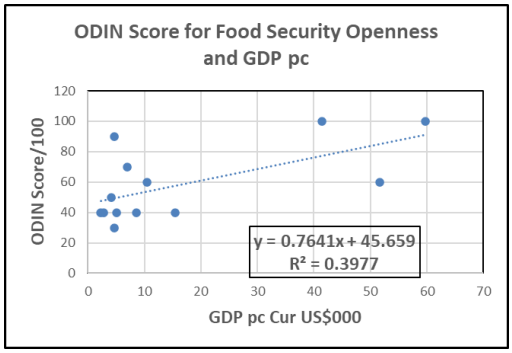


Table 3 ODIN Scores for Food Security and Nutrition, and, Poverty and Income

	Food security and nutrition (ODIN Score /100)			Poverty and Income (ODIN Score/100)		
	Coverage	Openness	All	Coverage.	Openness	All
Singapore	50	100	81	38	100	72
New Zealand	38	100	72	40	80	60
Australia	25	60	44	30	60	45
Jamaica	50	90	75	60	90	75
Mauritius	25	40	33	30	50	40
Malaysia	50	60	56	60	60	60
Maldives	33	70	56	38	50	44
Trinidad and Tobago	25	40	33	0	0	0
Samoa	0	0	0	20	40	30
Tonga	50	30	38	50	40	44
Fiji	13	40	28	10	40	25
Marshall Islands	50	50	50	0	0	0
Palau	0	0	0	13	40	28
Solomon Islands	50	40	44	20	40	30
Papua New Guinea	25	40	33	0	0	0

Source: ODIN Country Profiles Database 2020.

3. The global movements to Open Microdata Access

It is useful for PICTs to trace how the global movement towards "open microdata" have evolved from efforts by both developed nation states like the United Kingdom and international organizations like ICPSR, IHSN, Open Data Watch and WB to not just preserve but also encourage the greater open-ness and use of microdata by researchers and policy makers.

While it might be questioned why PICTs should be mindful of the practices of global multilateral organizations, the reality is that that some international organizations, like IMF, World Bank and WTO have extremely powerful mandates in PICTs for influencing, if not determining, their monetary, economic and trade policies, and in many instances may be seen almost as "binding", especially on developing countries.

Many multilateral organizations like the UN have enormous influence over development strategies of developing countries, and indeed have become the guiding institutions on setting Sustainable Development Goal targets, which have now become the international yardsticks for development. FAO also fulfils a similar role on food security, and ILO on labour standards. All are involved in fostering the creation, storage and dissemination of official statistics.

Even though the priorities of these multilaterals and their dominant stakeholders are much broader than those of individual nation states (including PICTs), it is useful to outline in detail their efforts in microdata preservation and dissemination simply because many of the issues and challenges they have faced are the same issues and challenges PICTs currently face, except that the latter's challenges are more severe given the general paucity of their resources, both economic and skilled manpower.

While there is some duplication of international views in this Report, their views are given as coherent packages rather than just as a summary of key points from different sources.[12] But their general agreement with each other should reassure PICTs that the challenges they face have been faced by many other developed countries and multilateral organizations as well.

PICTs and the PSMB should note that the issues they are grappling with, as with policies for the Pacific Data Hub Microdata Library, have been faced by other developed and developing countries and international organizations, and resolved to some extent. Their solutions may not necessarily be applicable to the smaller and under-resourced PICTs, but are nevertheless useful for PICTs to consider and modify if necessary to suit their own interests and national context.

Clearly, national and global multilateral interests are not identical. PICT NSOs, like those of any country, are funded by their national taxpayers, who are their primary stakeholders in term of outputs. International research organizations and multilaterals on the other hand, have global stakeholders and their effectiveness depend critically on the openness of individual countries' microdata.

[12] One Reviewer suggested that that the views of the different multilaterals ought to be summarized. I humbly disagree.

What is Microdata

Microdata are "unit-level information on people or entities such as individuals, households, business enterprises, farms, or even geographic areas" usually gathered through censuses and household surveys.[13]

Censuses attempt to gather information on every household and person in the country, whether in urban areas or distant rural areas, on some particular day, usually once every ten (or five) years. By definition, censuses are extremely expensive and difficult exercises to mount, especially in PICTs like PNG, Solomon Islands and others where people are truly scattered in remote areas.

Household surveys, on the other hand, attempt to gather information on a small sample of representative households, from which national estimates may be derived, without incurring the high costs of censuses. If well planned and designed, household sample surveys can give extremely accurate national estimates of a variable of interest, which are just as useful as those resulting from censuses. For the PICTs, sampling fractions are often of an order of magnitude greater than those seen in countries like Australia and New Zealand; involve very high response rates, and can provide as reliable small group analyses compared to those larger countries.

Because of its granularity and unit-level information, microdata can "facilitate the investigation of the unique ways a certain phenomenon may affect different groups or sub-populations". Official statisticians and researchers can then obtain national estimates to gain scientific insights made available on an impartial basis, to reflect Principle 1 of the Fundamental Principles of Official Statistics (FPOS) which is "to honour citizens' entitlement to public information".

All analysis and reporting must nevertheless abide by Principle 6 of FPOS that "Individual data collected ... whether they refer to natural or legal persons, are to be strictly confidential and used exclusively for statistical purposes.... This places limitations on what the data can be used for and who can use it" so that trust of the units who have volunteered their data, is maintained.

PICTs also need to take into account the powerful international movement to encourage the greater use of microdata in addition to the *preservation* of microdata for posterity. These two ought also to be important objectives for PICTs especially given the under-resourcing of NSOs, allied with paucity of technical personnel to undertake the safe archiving of national data, including microdata. There is evidence that some PICTs' microdata may have been permanently lost.

The perspectives of global research and multilateral organizations like ISPSR, IHSN and the World Bank all show to PICTs different perspectives that can better inform their own decisions regarding the challenges, the benefits and costs of opening access to PICTs' microdata. It should also be kept in mind that the thought-leaders of these organizations all were part of collaborative efforts.

The experience of UK

This Report begins by summarizing the experience of UK, a developed country which quite early on displayed similar concerns to those being expressed by PICT NSOs. The UK and

[13] Swanson et al (2021). This section relies on this excellent paper.

PICTs share a common concern that data integration for government's purposes are distinct from statistical and research use. Moreover, "confidentiality" and "privacy" can clearly have different meanings for an NSO.[14]

The problems faced by PICTs have also been faced by large, developed countries like the UK. Naturally, the problems of maintaining confidentiality are far more pronounced in small PICTs with tiny populations, where individual households and legal entities in survey responses may be more easily identified, than in large countries.

Len Cook (2003) and the UK Statistical Office recognized early on that not making full use of microdata resulted in an inadequate identification and analysis of policy options. He articulated[15] the many benefits of improving global access to British microdata:

* To build trust in official statistics

* To allow better research into more complex questions.

* To encourage scientific safeguard and verification of common results by official statisticians and independent researchers

* To assist in improving survey design through researchers' feedback about data quality

* To help develop a scientific community exploring common issues using the statistics.

Cook noted that while there were risks to allowing freer access to microdata, the resulting public damage was much less than that caused by other issues such as: oversensitivity to health ethics questions limiting the quality of statistics being collected; difficulties in accessing households; alienated communities reluctant to collaborate with surveys; and increasing survey burden on households. These observations are quite pertinent to many small PICTs with small populations, often on small islands where statistical accuracy may require proportionately large sample sizes.

On the other hand, Cook pointed out that broadening access to microdata conveyed widespread benefits:

* It enlarged the credibility of the research community;

* It facilitates evidence-based policy initiatives of governments, beyond the civil service research personnel;

* It strengthened the ethical base for ensuring confidentiality;

* It improves the quality of public debate through objective non-partisan research findings and facts;

* It strengthens data protection legislation to preserve confidentiality of individual responding units.

[14] I am grateful to Len Cook for these observations.

[15] Len Cook also acknowledged the contribution of Julia Lane at the Conference of European Statisticians.

These are of ever greater relevance to small PICTs with small civil service departments and small numbers of research personnel who are nevertheless expected to cover the full gamut of development problems faced by larger developing countries and multilateral development organizations like the UN or the World Bank.

Finally, Cook introduced the challenge of ensuring equity of access to the *research results* of increased microdata access.

ICPSR (2009)

The Interuniversity Consortium for Political and Social Research (ICPSR) is an international consortium of about 700 academic institutions and research organizations. It "provides leadership and training in data access, curation, and methods of analysis for the social science research community. It maintains a data archive of more than 500,000 files of research in the social sciences, and hosts 16 specialized collections of data in education, aging, criminal justice, substance abuse, terrorism, and other fields".

The ICPSR has published the "Principles and Good Practice for Preserving Data" (IHSN 2009).[16] This document aims to assist official data producers in defining and meeting their digital core preservation requirements and obligations, identifies steps for developing an effective and suitable preservation approach, while also trying to comply with standards of good practice within the digital preservation community. I return to this document when discussing the operations of the PDH-ML.

Dupriez and Boyko (2010)

Olivier Dupriez and Ernie Boyko (2010) set out the "Principles, Procedures and Practices of Dissemination of Microdata Files".[17] This document address the following key questions:

Why should data producers disseminate microdata?
What are the costs and risks, and how can they be addressed?
To whom should microdata be made available?
Under what conditions should microdata be provided?
What is meant by microdata anonymization
should microdata be sold or provided free of charge?
What are the technical infrastructure requirements for disseminating microdata?
What are the institutional requirements for disseminating microdata?
How to promote use of microdata?

These questions are all of relevance to the managers of the PDH-ML and to this Report.

[16] ICPSR (2009). The document was funded by IHSN and financial support from WB. Copyright shared with OECD, WB and University of Michigan.

[17] The paper was authored by Olivier Dupriez (Manager of IHSN and Senior Economist-Statistician with the WB Development Data Group) and Ernie Boyko (former staff member of Statistics Canada).

World Bank

World Bank (2014) [18] noted the benefits of sharing microdata: (It) "fosters diversity of research, increases transparency and accountability, and can mitigate duplication of data collection work and increase the quality of data through feedback received from data users".[19]

The WB also noted that while they had not noted any incidents, there was a "reputational risk in case of violation of privacy protection rules and regulations" which led many data producers and depositors to adopt "a conservative approach by severely limiting or excluding access to their microdata". Nevertheless, the WB felt that "Adopting international standards and good practices of microdata dissemination offers greater guarantees that the various technical, ethical and legal issues will be properly addressed, and in so doing mitigate risk".

The WB felt that the objective of maintaining confidentiality could be achieved by ensuring that:

(a) the microdata released does not identify individuals;

(b) the microdata is only used to derive statistics that refer to a group of persons or legal entities, and not to specific individuals (anonymization);

(c) there are legal and other necessary arrangements to protect confidentiality; and

(d) the procedures for researchers' access to microdata should be transparent, and publicly available to increase public confidence in the process of microdata release.

WB (2014) supported an OECD (2007) enunciation of "Principles and Guidelines for Access to Research Data from Public Funding" summarized here:

(a) Openness[20] meaning access on equal terms for the international research community at the lowest possible cost, preferably at no more than the marginal cost of dissemination.

(b) Transparency: Detailed metadata must be provided, and conditions must be transparent. Different multilateral organizations had different conditions: Eurostat's scientific use files are only released to researchers from registered approved institutions; UNICEF, WHO and the World Bank disseminate microdata as Public Use Files or Licensed Files, accessible to registered users.

(c) Legal conformity and protection of privacy. The microdata on households, individuals and firms are only to be released after the data are anonymized with adequate safeguards. Factors entering the decision include: sensitivity of the

[18] "Microdata dissemination best practices". While the note was initially drafted by the World Bank and presented at the 22nd CCSA session (Ankara, 2013) it was improved after inputs received from Eurostat, FAO, UNECE, UNICEF.

[19] Ironically, Annex 12 on this author's analyses of HIES data happens to disagree with WB findings on poverty in Fiji. It is reassuring that the WB acknowledges the importance of diversity of views.

[20] The OECD view excluded "Open Data" or "unrestricted access" from this definition.

data, legislation, reputational risk, potential consequence for the re-identified respondents, and political context.

(d) Protection of intellectual property, with consideration of copyright or other intellectual property laws. It is important that any microdata obtained through international organizations should specify ownership rights and obligations.

(e) Interoperability, to ensure technological and semantic interoperability to facilitate research and use.

(f) Quality Standards are explicitly followed.

(g) Security of data: Specific attention should be devoted to supporting the use of techniques and instruments to guarantee the integrity and security of data.

(h) Accountability: The performance of data access arrangements should be subject to periodic evaluation by user groups, responsible institutions and funding agencies.

The WB is also working with other international organizations to articulate best practice standards for data storage and microdata dissemination.

Swanson et al (2021)[21]

An excellent discussion of most of the relevant issues on access to microdata in national and international repositories is provided in Swanson et al (2021) which was a draft paper for ODW. The following sections rely heavily on this paper.

Swanson et al point out the importance of microdata from censuses, surveys and administrative records in providing *"the core information needed to address critical development challenges such as poverty and food insecurity, access to health care and education, and gender inequality. They are the basis for the SDG indicators and many other statistics used by governments and businesses and cited in public discussion... [and] they can provide a more nuanced, multidimensional record of the needs of vulnerable people that is essential for validating previous analyses, testing new hypotheses, and designing programs".*

The development benefits for statistical data systems are also quite clear to them: *"Making microdata more widely available can also enhance the credibility and reputation of the national statistical system and place it at the center of the effort to achieve the Sustainable Development Goals and leave no one behind."*

Unfortunately, while microdata had the attributes of a "public good"[22] in that they were virtually costless to disseminate once they have been produced and they were not diminished in value upon use by one party, they were "excludable" thereby becoming a "club good" where access could be restricted by technical and legal requirements.

[21] In the earlier Draft Report, I had referenced this Report as ODW (2021).

[22] In economic theory, goods and services can be "non-rival" in consumption i.e. consumption by one does not reduce consumption by another (as in the use of microdata), but may be excludable by some means (e.g. refusing access or download or setting a price). Pure public goods would be those where there is non-rival consumption AND excludability cannot be practiced (e.g. clean air).

Swanson et al (2021) documented how and where low- and middle-income countries stored public data from censuses and surveys and what were the barriers to access them, or the "blockers" that this Report is tasked to investigate. They found that for the majority of the developing countries they surveyed, it was not known exactly what surveys had been conducted or where their datasets were stored - or even what had been lost. Their Table 1 is provided as Annex 4 in this Report. It summarizes the microdata repositories provided globally by multilateral organizations and nation states.

Swanson et al (2021) noted that the International Household Survey Network (IHSN) Microdata Catalog (which was housed at the World Bank) and the World Bank Microdata Library (WBML) were the largest repositories of censuses and surveys serving as models for other organizations and countries. Between them they contained records of roughly 10,000 datasets. The IHSN however contained only metadata records and access to the underlying microdata files have to be obtained from external repositories, a large proportion of which are listed as unavailable. The WBML on the other hand listed only surveys for which microdata were publicly available. Unfortunately, the two repositories were not consistent with each other even in identification numbers.[23] Both repositories provide options for filtering their holdings by year, country, access level and by key words contained in the study description.

The repositories of other international agencies are more specialized, usually limited by topic or survey program. The World Bank Microdata Library and the IHSN Microdata Catalog redirect users to many of these repositories to download microdata files not kept at the World Bank. The IHSN Catalog and the WBML , housing about 8,500 references or links to datasets (or 11,581 if we also include those at the Pacific Data Hub). Other organizations with sizeable microdata sets are:

FAO	Agriculture, food security and many others
UNHCR	Refugee populations and impacts
WHO	Noncommunicable Diseases (NCDs)
WHO	Health Surveys
ILO	Labour Force, child labour, income and expenditure surveys
DHS (USAID)	Demographic and Health Surveys
UNICEF	Multiple Indicator Cluster Surveys
IPUMS Int.	Census microdata

Except for the field of refugee populations and impact, the PSMB might wish to discuss the extent to which the Pacific Data Hub needs to try to cover most if not all these fields and be compatible with international repositories

Swanson et al recognized that the catalogs of both IHSN and WBML may be incomplete partly or largely because some nation states chose to be exclusive with their datasets, just as some PICTs are with the PDH-ML. Moreover, the countries and regions are not represented evenly in all the repositories listed above, and Table 2 of Swanson et al (2021) gives the distribution of the regions in their datasets. The Pacific is represented in all the major repositories with the exception of UNHCR and WHO multi-country Studies.

Of the 83 Low and Lower Middle Income Countries LLMCs) assessed by Swanson et al (2021), only 27 managed their own microdata repositories. Similar to the international

[23] It is surprising that some surveys listed in the WBML do not appear in the IHSN, and some surveys shown as unavailable in the IHSN are listed as having data available through the WBML. Both are incomplete.

repositories, the coverage of indicators was unclear with doubts about the consistency of the naming of surveys among countries and international organizations. Many catalogs in national repositories did not appear to overlap with the international repositories as they should, and were clearly under-represented in the WBML and IHSN catalogs. Some LLMCs had severe problems of lags in updating.

The Pacific Data Hub might wish to ensure that its holdings fill all the Pacific gaps in the holdings of the major repositories, if necessary with the financial and technical assistance from IHSN and WBML.

The confidentiality issue

Principle 6 of the Fundamental Principles of Official Statistics (FPOS) refers specifically to restrictions that must be considered before the release of microdata: "Individual data collected by statistical agencies for statistical compilation, whether they refer to natural or legal persons, are to be strictly confidential and used exclusively for statistical purposes".

To ensure these two conditions were followed, Swanson et al (2021) refer to the Five Safes framework which was developed originally by the UK Office of National Statistics and used by other international microdata repositories like IPUMS International. The ideas behind this framework also apply to the management of the PDH-ML and advise future reforms.

1. Safe people: i.e. safe researchers and organizations
2. Safe projects: approved uses of the microdata
3. Safe outputs
4. Safe Data: security of the data
5. Safe settings: security of access systems.

Unfortunately, taking the Five Safes to extremes can also act as effective barriers to access and the tradeoffs need to be carefully assessed. Subjective judgements may need to be made about the potential damage to individuals if they and their characteristics are identified. Normally in a large sample, age, sex, education or employment may have little consequence. In very small samples however, such as in PICT surveys, high levels of education and employment in upper income brackets could easily identify individuals. Individuals may also be sensitive about health information, political or religious views.

Statistical Disclosure Control (SDC) or Anonymization

Swanson et al (2021) point out that there is growing literature as documented by Mathias Templ et al (2014) to facilitate SDC by

(a) eliminating direct identifiers such as names, addresses, and other unique attributes that might identify individual persons, households or firms.

(b) eliminating indirect identifiers (such as birth dates)

(c) modify some data to obscure individual records that could identify individuals.

(d) adding "statistical noise".

The challenge for small PICTs like Niue and Tokelau would be to modify individual data to obscure identities, while not significantly changing the "truth" of what the related data reveals. This might be particularly difficult if a sub-group (such as a village) had a very small number of observations. Another difficulty may be that changing the location of the individual in geographically small countries (like Niue and Tokelau) may be also a difficult exercise. It is extremely unlikely that (d) above will need to be used in PICTs.

Access Control and Licensing

Table 4 in Swanson et al (2021) outline the various levels of access used by the WBML (presented here as Annex 5), which are all of relevance for consideration by the PSMB and PICTs. Similarly, this Report suggests a simplified tiered structure of access and control which tries to reduce the administrative burden on the Pacific Data Hub (see Section 9 and the Policy Brief).

The WBML found it useful to categorize the microdata as "Public Use Files" (PUFs) and Scientific Use Files (SUFs) given by Table 5 in Swanson et al (2021) and presented here as Annex 6.

Annex 7, then, describes the profile of access to WBML microdata for different regions, including the Pacific. The majority of microdata files (60% to 80%) are available as PUFs, around 20% in external repositories, and less than 5% as SUFs. To be noted is the low percentage of PICTs' surveys available in external repositories.

Table 7 in Swanson et al (2021) gives a profile of the access levels of other multilateral organizations such as FAO, UNHCR, WHO, DHS, UNICEF and IPUMS International. For FAO for instance, 57% of the microdata sets are available as SUFs, 30% in an external repository, and 13% not available.

Other limitations on data access

There were other limitations on data access pointed out by Swanson et al (2021), perhaps more applicable to national repositories rather than international ones. Similarly, they may apply to PICT national repositories rather than the far more efficient PDH-ML.

(a) Faulty registration systems

To obtain access to the microdata, users typically complete online forms with a user name, email, and country, with the email verified before receiving access to the repository. This may be automated in international repositories but not in national ones where the approval may require action by a staff member.

(b) Unclear redirect links

International repositories may not store the microdata themselves and hence need to redirect to where the data is stored. That redirection may not reach the data immediately but the user may have to go through various hoops, some practically insurmountable.

(c) Faulty downloading of data

For a variety of technical reasons, data on many national repositories are not easily downloadable "at a click" as one may with the WB database.

(d) Limited file formats

Many national data repositories do not have user friendly nonproprietary formats (such as ASCII) or Excel, but formats not familiar or manageable to the user.

(e) Repositories not always online

For many national data repositories, especially in poorer developing countries, the data server may simply not be on line for a variety of reasons, such as electricity outages or break-downs in the computer services of that country.

It is axiomatic that PICT NSOs with poor national infrastructure such as reliable electricity supplies and stable internet access would face these difficulties even more. Most if not all of these difficulties could be minimized by the NSOs using the PDH-ML as the central storing and disseminating agency

This very practical aspect of ensuring that approved international users have complete and efficient and prompt access to the PICTs' microdata, is a sound justification for encouraging PICTs to store their microdata with the PDH-ML (at least for those datasets that they could grant public access, with or without conditions as discussed below in the various Tiers).

General recommendations by Swanson et al (2021)

Swanson et al (2021) recommend that

1. Countries should take the lead in publishing their data, preparing a complete catalog of surveys and censuses, maintain repositories that include metadata and direct users to where the microdata are available. They expected that external assistance would be needed in resource poor countries.

2. Access controls should be proportionate to disclosure risk, which should be managed with the Five Safes principles considering "both the risk of reidentification of natural or legal persons and the potential harm done by a disclosure breach".

3. the same microdata should be made available at different access levels corresponding to the risks and dangers of disclosure: "A file with stronger disclosure controls can be made available with fewer or no restrictions for public use or direct access, while a version with fewer controls can be licensed with stricter controls or made available only through a secure facility for users requiring more granular data". These are all of relevance to the PICTs and the PDH-ML.

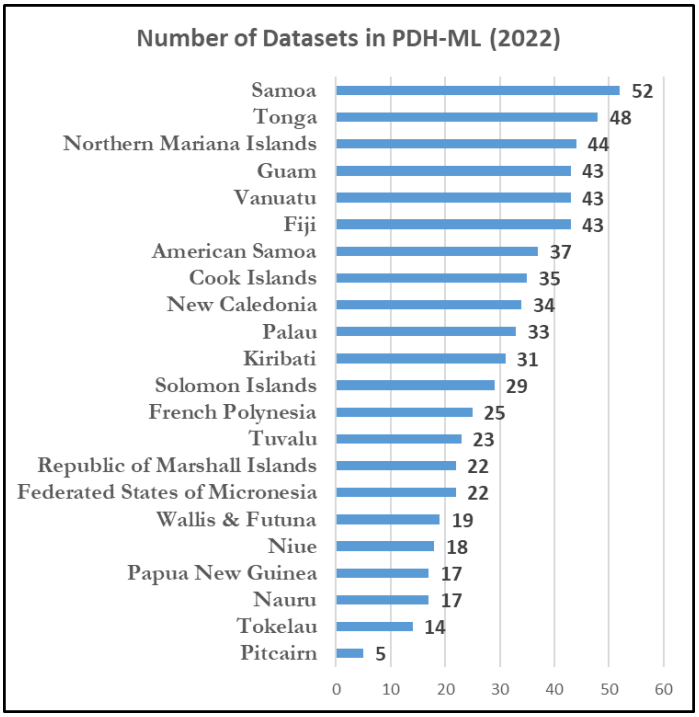
The importance of formal agreements

To ensure that the principles enunciated above are followed where appropriate, it is strongly advisable that PICTs enter into formal agreements with researchers, including international multilateral organizations.

The next section outlines the performance of the PDH-ML while the section after than outlines existing agreements and the template that can be followed by PICTs.

4. The Performance of the Pacific Data Hub - Microdata Library

As pointed out in the Introduction to this Report the Pacific Community (SPC) is recognized[24] for its scientific and technical expertise in addressing some of the region's most complex development challenges, including climate change, disasters, non-communicable diseases, gender equality, youth employment, food and water security, and biosecurity for trade", as well as poverty identification and alleviation.



Given the broad coverage of advisory activities, SPC (and SDD in particular) also considers itself responsible for preparing and publishing any statistical aggregates and indicators, both for the SPC region and the individual member countries where necessary. It also sees as part of its broader responsibilities, the support of the national efforts to collect and share microdata where individual PICTs gave granted approval.

To this end, SPC houses the Pacific Data Hub- Microdata Library (PDH-ML), which, as of February 2022, contained 654 datasets from all the SPC Member countries.

The PDH-ML Collection

While individual countries have different time trends in censuses and surveys ever conducted (see Table 4), overall the aggregate number of datasets added per decade has risen steadily for each decade, from the 1980s (53) to the 2010s (195).

[24] Explicitly stated in the SPC Data License Agreement (DLA) with NSOs.

Table 4 Censuses and Surveys ever conducted (by decade)							
Country	A Pre-1980	B 1980-1989	C 1990-1999	D 2000-2009	E 2010-2019	F 2020+	All
American Samoa	11	4	7	8	7		37
Cook Islands	2	4	3	9	10	7	35
Federated States of Micronesia	4	3	2	5	4	4	22
Fiji	19	2	3	8	9	2	43
French Polynesia	5	3	1	4	7	5	25
Guam	14	3	5	9	11	1	43
Kiribati	6	1	3	7	7	7	31
Nauru	1	1	1	5	6	3	17
New Caledonia	4	4	3	6	11	6	34
Niue	2	3	2	4	6	1	18
Northern Mariana Islands	4	3	6	12	16	3	44
Palau		3	5	14	9	2	33
Papua New Guinea		2	3	4	4	4	17
Pitcairn		1	1	2	1		5
Republic of Marshall Islands	8	2	1	6	4	1	22
Samoa	4	3	3	15	23	4	52
Solomon Islands	3	2	3	8	9	4	29
Tokelau		1	2	4	6	1	14
Tonga	3	2	4	16	17	6	48
Tuvalu	7		2	4	6	4	23
Vanuatu	4	4	5	10	16	4	43
Wallis & Futuna	2	2	2	6	6	1	19
Grand Total	103	53	67	166	195	70	654

Source: From data provided by PDH-ML catalogue.

Microdata availability

Table 5 indicates that the majority of the datasets of the surveys ever conducted are not available to the PDH-ML, although the percentage available has risen steadily from 8% in the 1980s to 38% in the decade (2010 to 2019).

Some individual countries have an excellent record over the last decade (2010-2019), like Tuvalu (100%), Kiribati (86%), Tokelau (83%) and Cook Islands (60%).

Table 5 Percentage of Datasets Available

Country	A Pre-1980	B 1980-1989	C 1990-1999	D 2000-2009	E 2010-2019	F 2020 +	All
American Samoa	0	0	0	13	14		5
Cook Islands	0	0	33	44	60	43	40
Federated States of Micronesia	0	0	0	40	50	0	18
Fiji	11	50	33	38	44	0	26
French Polynesia	0	0	0	75	0	0	12
Guam	0	0	0	33	0	0	7
Kiribati	33	100	67	71	86	14	55
Nauru	0	0	0	80	67	0	47
New Caledonia	0	0	0	50	18	17	18
Niue	0	0	0	50	67	0	33
Northern Mariana Islands	0	0	0	0	0	0	0
Palau		0	0	21	33	0	18
Papua New Guinea		0	33	25	25	50	29
Pitcairn		0	0	0	0		0
Republic of Marshall Isl.	25	50	100	33	50	100	41
Samoa	0	0	0	13	22	0	13
Solomon Islands	67	50	33	50	22	25	38
Tokelau		0	50	50	83	0	57
Tonga	0	0	25	13	65	0	29
Tuvalu	0		0	50	100	0	35
Vanuatu	0	0	20	30	38	25	26
Wallis & Futuna	0	0	0	67	83	0	47
Grand Total	8	8	15	33	38	14	25

Source: From PDH-ML catalogue.

It is important for the PSMB encourage its members to deposit all available microdata sets with the PDH-ML at least for Tier 1 initially. This can be done with a signed Data License Agreement (DLA) for each dataset initially for archiving in the Tier 1 category recommended below, with no access to outside agents without the express permission of the NSOs who have sovereignty over the datasets. The ultimate goal is to protect this data for posterity should national repositories are affected by natural disasters.

NSOs may then consider making some of those in Tier 1 available to the higher Tiers upon satisfaction of the specific requirements set by each NSO, including suitable

anonymization, and other conditions they would like to impose, some of which have been expressed in the Questionnaire responses (see Annex 11).

Table 6 indicates the number of datasets by category and decade, with the most common and persistent over time being censuses of population, censuses of agriculture and HIES.

The step-up in aggregate numbers between the 1990s and 2000s occurred for HIES, Business/Economics/Labour Force. Health in particular saw large increases because of the increased focus by multilateral organizations on SDGs including family violence and health issues such as NCDs.

Table 6 PDH datasets by category and decade (Ever Done, Available at PDH-ML, and Perc. of All Done)

Category	A Pre-1980	B 1980-1989	C 1990-1999	D 2000-2009	E 2010-2019	F 2020 +	All
	Numbers of surveys and censuses ever done						
A Census Population	73	30	29	32	36	19	219
B Census Agriculture	17	12	12	18	17	6	82
C HIES	9	10	12	23	26	10	90
E Ec., Bus. and LF *	2	1	7	29	49	16	104
H Health	1		7	62	61	14	145
X General	1			2	6	5	14
All	103	53	67	166	195	70	654
	Number of Microdata sets Available						
A Census Population	8	4	9	28	24	3	76
B Census Agriculture				1	2		3
C HIES			1	15	17		33
E Ec. Bus. and LF				2	8		10
H Health				8	20	7	35
X General				1	4		5
Grand Total	8	4	10	55	75	10	162
	Percentage of datasets available in PDH-ML						
A Census Population	11	13	31	88	67	16	35
B Census Agriculture	0	0	0	6	12	0	4
C HIES	0	0	8	65	65	0	37
E Ec., Bus. and LF.	0	0	0	7	16	0	10
H Health	0		0	13	33	50	24
X General	0			50	67	0	36
All	8	8	15	33	38	14	25

Source: From PDH-ML catalogue.

* Ec., Bus. and LF: Economy, Business and Labour Force.

The Censuses of Population have had the highest proportion deposited in the PDH-ML in the 2010s, some 67%, although this is lower than the 88% observed in the decade 2000-2010. It seems to be absolutely vital that the PSMB and the PDH-ML make a special effort to complete their archive of this most precious of microdata sets which describe the entire population of each country, person by person. There cannot be any more basic microdata set than the census of population of any country at the census date.

For the HIES which is the special focus for this Report, only 37% in aggregate have been stored at the PDH-ML, although 67% were stored in the decade 2010-19.

The category E (Economy, Business and Labour Force) had a great variety of surveys with no great consistent trend over time for any of the components, except for Labour Force Surveys (see Annex 8).

Similarly, category H (Health) had a great variety of surveys with no great consistency over time of the components, except DHS (Annex 9).

PICT NSOs might like to compare the content of the collections in Category E and Category H of the PDH-ML with the content of the international collections like the World Bank Microdata Library to assess whether the PDH-ML is generally in step with the global microdata collections, although PICTs ought naturally to have more unique ones focused on their needs, like the impact of climate change on atoll countries.

What the spread of Item Types in Annexes 8 and 9 indicate is that for these two general category types, the PDH-ML microdata collections have been ad hoc over time, probably responding to the changing focus of international organizations in the Pacific and resource availability, rather than well-though out needs of the PICTs themselves according to their development priorities which ought not to be changing dramatically from decade to decade. The latter of course presupposes that the data needs of the PICTs are consistent over time and across countries. But one would at least expect that the larger more populous Melanesian PICTs (PNG, Fiji, Solomon Islands, New Caledonia and Vanuatu) have consistent needs for microdata sets over time, and very similar development priorities.

It is recommended that PDH-ML and the PSMB have on their policy agenda a discussion about PICT priorities for microdata sets in economics, business, labour force and health to guide the PICTs on the important priorities for microdata collection exercises, so that the holdings can be consistent over time.

These priorities will not be uniform across all PICTs: atoll PICTs will probably have climate change impact at the top of their list; high population growth countries will have employment creation and labour force changes as the top priority; most PICTs have urgent health (NCDs) issues, which have been present for decades.

It is extremely important that adequate resources are made available to the PDH-ML to enable a special initiative to collect all backlogs of microdata sets, so that they are not lost to posterity.

The HIES microdata collections

Table 7 provides the number of HIESs which have been conducted by each PICT by decade. The total number have increased steadily from in the eighties decade, to 26 in the decade 2010 to 2019. Most PICTs have conducted at least 1 HIES per decade, with some of them (Fiji, Samoa, Tuvalu, Vanuatu, and RMI) undertaking 2 surveys in the 2010-2019 decade.[25]

[25] Only Pitcairn Islands did not have any HIES.

Table 7 Number of HIES conducted, by decade							
Country	A Pre-1980	B 1980-1989	C 1990-1999	D 2000-2009	E 2010-2019	F 2020 +	All
American Samoa		2	1	1	1		5
Cook Islands		1	1	1	1		4
Federated States of Micronesia		1	1	1	1	1	5
Fiji	8	1	1	2	2		14
French Polynesia		1		1	1		3
Guam				1	1		2
Kiribati			1	1	1	1	4
Nauru				1	1	2	4
New Caledonia	1	1	1	1	1		5
Niue				1	1		2
Northern Mariana Islands			1	1	1	1	4
Palau				2	1	1	4
Papua New Guinea		1	1		1	1	4
Pitcairn							
Republic of Marshall Islands				1	2		3
Samoa			1	2	2	1	6
Solomon Islands			1	1	1		3
Tokelau					1		1
Tonga				2	1	1	4
Tuvalu			1	1	2	1	5
Vanuatu		1	1	1	2		5
Wallis & Futuna		1		1	1		3
Grand Total	9	10	12	23	26	10	90
Source: Author's elaboration from PDH-ML data							

Table 8 gives the percentage of HIES microdata sets actually held by PDH-ML and potentially available for analysis.

The blank cells mean that there were no HES done in that period in that country. The "0" indicates that none of the HIES done had their datasets deposited with the PDH-ML. The Averages in the last column represent the total HIES sets deposited with the PDH-ML as a percentage of the total HIES ever done. American Samoa deposited 1 (20%) out of 5 ever done; Fiji deposited 1 out of 14 ever done.[26]

Unfortunately, only 37% of the HIES microdata sets have been deposited with the PDH-ML in total, although the share in the last two decades is higher (65%). In fact, 13 out of the 22 PICTs have all their HIES datasets available in the PDH-ML for the decade 2010-2019, and 2 PICTs have 50% of the datasets in the database. On the other hand, 6 PICTS did not deposit any data set in the decade 2010-2019. These PICTs may wish to do so, especially if they would like to be part of the HIES data analysis exercise that the PDH-ML may co-ordinate in the near future.

[26] It is totally unclear what was the quality of the pre-1980 HIES for Fiji.

Table 8 Percent. of microdata sets available							
Country	A. Pre-1980	B. 1980 – 1989	C. 1990 - 1999	D. 2000 - 2009	E. 2010 - 2019	F. 2020 +	Average
American Samoa		0	0	100	0		20
Cook Islands		0	100	100	100		75
Federated States of Micronesia (FSM)		0	0	100	100	0	40
Fiji	0	0	0	50	0		7
French Polynesia		0		100	0		33
Guam				100	0		50
Kiribati			0	100	100	0	50
Nauru				100	100	0	50
New Caledonia	0	0	0	100	0		20
Niue				0	100		50
Northern Mariana Islands			0	0	0	0	0
Palau				50	100	0	50
Papua New Guinea		0	0		100	0	25
Pitcairn							
Republic of Marshall Islands				100	50		67
Samoa			0	0	50	0	17
Solomon Islands			0	100	100		67
Tokelau					100		100
Tonga				50	100	0	50
Tuvalu			0	0	100	0	40
Vanuatu		0	0	100	100		60
Wallis & Futuna		0		100	100		67
All	0	0	8	65	65	0	37
Source: derived from PDH-ML data							

HIES Microdata Requests and PDH-ML Responses

Table 9 gives the results of requests for HIES Microdata over the years. Out of 119 requests, only 34 were approved and 85 denied.[27]

Table 9 HIES Requests and Approvals (numbers and percent.)				
	Numbers			Percent.
	No	Yes	All	Perc. Approved
Cook Islands	7	4	11	36
Federated States of Micronesia	9		9	0
Kiribati	9	15	24	63
Niue	2	1	3	33
Nauru	13	3	16	19
Papua New Guinea		4	4	100
Solomon Islands	13	3	16	19
Tokelau		3	3	100
Tonga	6		6	0
Tuvalu	16		16	0
Vanuatu	10		10	0
Wallis and Futuna		1	1	100
All	85	34	119	29
Source: From PDH-ML spreadsheet of requests.				

Some preliminary observations may be made. Compared to the number of HIES microdata sets held in the PDH Library, the number of requests is extremely small. Kiribati and Nauru stand out, with a high number of requests for Kiribati and a high percentage of approvals. Staff of the SDD feel that demand for access to microdata may perhaps be inversely related to the data/indicators that are made available through NSO websites and to the approvals: the more indicators available the higher the demand for access; and the higher the approvals, the more the demand.

Table 10 gives the sources of the requests by the type of requestor. Of serious concern is that only 29% were granted access despite the undoubted reputations of most of the research organizations.

Table 10 HIES Requests (Approved, Denied)				
	Numbers			Percent.
Type of Requesting organizations	No	Yes	All	Perc. Approved
Dept. of Foreign Affairs and Trade (Aus.)		2	2	100
Individual	2		2	0
SPC	1		1	0
UN generic	46	13	59	22
University	21	11	32	34
WB/IMF/ADB	15	8	23	35
All	85	34	119	29

[27] An earlier table was incorrect and its source could not be found.

There were very few requests from regional donors DFAT (Australia) and MFAT (NZ) and very few (only 32) requests from universities, evenly split between Australia and NZ universities and elsewhere.

The largest number of requests (59) were from UN generic organizations (including WHO, FAO etc.) of which only 22% were granted access. Only 34% of university applications were approved, and only 35% of requests from international finance organizations like IMF, WB and ADB were approved.

Given that the PDH-ML administrators would have had little problem with the track record of UN organizations or WB/IMF/ADB requests in terms of keeping the confidentiality of individual records and not attempting to profit commercially from the microdata, then it is quite likely that the NSOs - with ultimate sovereignty over their country data sets - would have made the final decision to refuse access.

What should be of greatest concern is that there were no requests from any regional university, including USP. This ought to be a serious policy issue for all regional universities, given that all their Mission and Vision statements declare that they are there to serve the Pacific peoples not just in teaching but also research. This Report makes a strong recommendation for a concerted effort to involve academics from the regional universities (including those in Australia and NZ) in research using the microdata sets available at the PDH-ML.

It would be important for the PSMB to ascertain what were the real national NSO blockers to these requests for access and to address their fears and worries if access to PICTs microdata is to be enhanced. The Questionnaire responses do give some indication but the response rate was unfortunately too low to draw strong conclusions. Discussions between the PSMB and PDH-ML administrators may be needed.

PDH-ML Protocols and Procedures for public access

The SPC is currently reviewing its "Protocols/procedures for licensed data access requests to the Pacific Data Hub – Microdata Library". When approved, users agree to comply with all of the conditions (terms of use) of the respective licenses.

1. Researcher(s) must be affiliated with a credible research and/or teaching institution, such as an accredited university or recognised research organisation or an NGO.

2. Researcher(s) must either have a proven track record of analysing large datasets or, in the case of students, must have a supervisor or thesis adviser who could oversee her/his work

3. Researcher(s) must accept and adhere to the SPC terms of use and any other specific conditions outlined in a signed Data License Agreement (DLA) between SPC and data owner and stipulated in Data Access Agreement (DAA) between the researcher and SPC, or under a 3rd party license between data owner and the researcher.

4. Researcher(s) must agree to provide the final output of his/her work as a report, paper or otherwise. All data sources must be cited in the produced work.

The draft document also has additional selection criteria for the quality of the proposal, which may be administratively burdensome. The administrators of the PDH-ML may wish to consider whether they would wish to judge the quality of the research proposal, a function more appropriate for the research supervisor. [28]

In the draft document there are also rigorous recommendations for the establishment of a Review committee to deal with Microdata License Access via the Pacific Data Hub Microdata Library. The committee could be composed of four or five persons, including members of the SPC Statistics for Development Division and data owners, and possibly include legal representation via SPC's Legal and Risk Office. The committee will provide advice to the data owner (such as the Government Statistician or NSO) on the eligibility or criteria they have used to determine whether or not access should be given. But the data owner is the final decision maker in determining data access requests.

Also contentious and needing further discussion is the recommendation that members of the Review Committee will also agree to review final reports and draft manuscripts for publications from the selected researchers once the research is complete. It is unclear why the SPC PDH-ML should have this role at all.

The draft document has appropriate sections on legal requirements for transfer of microdata, follow up, data deemed too sensitive or damaged, licensing external depositories, transparency of decision making by NSOs on the application for release of microdata.

While one could discuss the above processes which are in train at the PDH-ML, it would be useful for the PSMB and the managers of the PDH-ML to place their own efforts in the context of the efforts and experience of the ICPSR in data preservation.[29]

ICSPR (2009) Principles and Good Practice of Data Preservation

The Report by this author does not discuss the many complex technical aspects of the integrity features[30] of data preservation and curation in an era of continuously evolving information technologies. Nor do I discuss the seven principles of a "Trusted Digital Repository" (TDR): administrative responsibility, organisational viability, financial sustainability, technological and procedural suitability, system security, procedural accountability and Open Archival Information System (OAIS) compliance. The PDH-ML administrators would no doubt be thoroughly familiar with both these sets of technical issues associated with data preservation and archiving.

What is of great relevance for this Report to the PSMB is the comprehensive Section 5: "Evaluation and Audit for Data Preservation", bringing together the experience of a number of global exercises for assessing the soundness of digital collections in serving their community needs and expectations. This section discusses two options: Option 1 is *Data Seal*

[28] Some of the conditions proposed were: justification for and significance of chosen topic, how the data will be used and research compiled, not replicating analysis already conducted or ongoing, appropriateness of the methods and analysis, dissemination approach or plan of results, the planned sharing of work between team members. These would be appropriate for the Academic Supervisor, not the source of the microdata.

[29] ICPSR (2009) can be considered a solid "must read" manual for the PDH-ML administrators. It had been reviewed by an excellent team of reputable experts in the field.

[30] ICPSR (2009) discusses content, fixity (record of all changes), reference to related content, traceability of the data to its source.

of Approval[31] and Option 2 is *Ten Characteristics of a Digital Preservation Repository,* with clearly common elements to both.[32]

The *Data Seal of Approval* (Option 1) consists of a concise set of higher-level principles that defines expectations for the three sets of parties: *producers* of the stored digital content, for the *repositories* that manage digital content and for the *consumers* of digital content. I repeat them here as all are of relevance to the operations of the PDH-ML and the DLAs under which the datasets made available to users.

1. The data-producer: provides the research data in formats recommended by the data repository and the metadata requested by the data repository.

2. The data repository:

   (a) has an explicit mission in the area of digital archiving and promulgates it;
   (b) ensures compliance with legal regulations and contracts including regulations governing the protection of people;
   (c) applies documented processes and procedures to managing data storage;
   (d) has a plan for long-term preservation of its digital assets;
   (e) archiving takes place according to explicit workflows across the data life-cycle;
   (f) assumes responsibility from the data producers for access and availability of the digital objects;
   (g) enables the users to utilise the research data and refer to them;
   (h) ensures the integrity of the digital objects and the metadata;
   (i) ensures the authenticity of the digital objects and the metadata;
   (j) follows internationally-accepted archival standards .

3. The data consumer:

   (a) complies with access regulations set by the data repository;
   (b) conforms to and agrees with any codes of conduct that are generally accepted in higher education and research for the exchange and proper use of knowledge and information;
   (c) respects the applicable licences of the data repository regarding the use of the research data.

[31] Details may be read at http://www.datasealofapproval.org/.
[32] I give both authoritative sets rather than summarizing elements from each.

Option 2 has the following basic principles, requiring the Repository to

1. commit to continuing maintenance of digital objects for identified community/communities.

2. demonstrate organisational fitness (including financial, staffing structure and processes) to fulfil its commitment.

3. acquires and maintains requisite contractual and legal rights and fulfils responsibilities.

4. has an effective and efficient policy framework.

5. acquires the digital objects that satisfy the stated criteria

6. accurately defines and maintains the scope of preservation

7. maintains/ensures the integrity, authenticity and usability of digital objects it holds over time.

8. have control over its digital content (know where it is and who has or is able to gain access to it)

9. associate minimal metadata data with the content and create and control more than one copy of the digital content.

10. fulfils requisite dissemination requirements.

11. has a strategic programme for preservation planning.

12. has technical infrastructure adequate to continuing maintenance and security of its digital objects.

These principles are easier stated than implemented or complied with especially when you have such a large number of diverse and differently resourced PICTs, as well as an under-resourced PDH-ML.

It is extremely important for this Report to give full weight to the views of the PICT NSOs and to assess where their thinking may be in relation to that of good practice in the global digital community. The SDD conducted a Questionnaire Survey towards this end.

5. The NSO Responses to the Questionnaire

To try and ensure that this Report takes full account of the PICT NSOs, a Questionnaire Survey was sent to all the 22 PICTs, with only 3 responding initially, and another 2 after a follow-up reminder. Thus this section is based on only 5 out of 22 PICTs responding, not really enough to make strong conclusions from.[33]

The following section gives the summary of the 5 Questionnaire responses, with Annex 11 giving the more detailed range of responses.

HIES Outputs

Individual PICTs had a great variety of outputs from their HIES microdata sets. While the most common was the basic HIES Report, there were also several on poverty, food security, hardship, well-being and labour markets. There were several multilateral organizations (WB, FAO, ILO) who facilitated the reports in addition to the technical survey support provided by SPC:

All NSOs stated they would have liked to produce more reports, with suggested topics being: gender, health, multidimensional poverty index, unemployment and marine food consumption.

Virtually all NSOs reported: (a) a lack of demand from government (b) a lack of demand from the public (c) lack of funds (d) lack of technical staff (e) lack of outside researchers.

Given that there were only 5 responses to the Questionnaire, there needs to be some discussion by the PSMB about their assessment of the general applicability of these responses. If they are confirmed, how these adverse factors may be addressed. In particular, whether (a) and (b) can be improved by making governments and the public more aware through public workshops, of the potential for HIES analyses to assist with evidence-based policy in the very areas that they perceive the country to be facing development problems.

External demand for access:

There was a wide variety of responses on external demand for access to microdata sets, the conditions that NSOs laid down, the availability (or lack) of subsequent reports to NSOs and whether conditions were met or not.

Legislation

There were a wide variety of responses on what their legislation allowed or did not allow external access to their microdata sets. Most NSOs felt that their legislation was outdated and needed to be revised to clarify their ability to deposit their microdata sets with the PDH-ML and for them to be used by external researchers, under some unform regulation.

[33] I would suggest that this be confidential to SDD.

Data anonymisation

All NSOs saw no reason why their HIES microdata could not be anonymized.

All NSOs would welcome technical assistance on anonymization from PDH-ML.

Attitudes towards Open data

All NSOs saw multiple benefits of an open data philosophy such as better evidence-based development policy and monitoring of progress, improvement of research, increased public confidence and trust in good use of data and NSO.

One benefit which the PSMB might want to discuss is the possibility that external researchers into the HIES (and other) microdata may produce insights into development issues which are qualitatively different from (and perhaps beyond that) of NSO's decision makers, including any Statistics Committee they may have. They might also be different from the multilateral organizations (like WB) which are more easily granted access.[34] Should the concept of "national data sovereignty" preclude access to these external researchers?

Conditions of data access

Most NSOs would like to set conditions for the PDH-ML to release their microdata sets to external researchers. These included (but not necessarily for all NSOs) guarantees of data anonymization, permission from NSOs, availability of reports, national capacity building, penalties for breaches. These conditions would all need further discussion by the PSMB as some would be quite administratively burdensome and defeat the very purpose of freeing up access.

Most NSOs would like to have fairly standard conditions applied: Resultant reports must be submitted to the NSO, data to be used only as applied for, results to be politically neutral. Some advocated penalties for breaches of the conditions.

What might be more contentious is that some wanted greater involvement in the approval process, the report analysis and writing.

There seemed to be some evidence that NSOs declined external requests for access to microdata because the NSOs could not see how their country and people would benefit.

It would seem to be important that these conditions need to be discussed and some agreement reached which minimizes bureaucratic delays, while safeguarding the PICTs interests in preserving anonymity of the data and proper data use.

Risks of Open Data Philosophy

Perceived risks were breaching of confidentiality, loss of public trust, misuse of data, and possibility of legal actions against NSO.

[34] See the discussion below in the Annex on how the outputs of this academic researcher can provide an alternative view on that produced by multilateral organizations.

Safe Storage of Data

Many NSOs did not have any policy of secure external storage of their microdata, which would not be vulnerable to national disasters.

It would seem urgent for the PSMB to reach some consensus on the PDH-ML as a very desirable Tier 1 archive for all PICTs microdata sets.

PDH-ML would also need to consider its own backup in the Cloud to guard against catastrophic risk to the SPC servers in its headquarters in New Caledonia.

Storage with PDH-ML and PDH-ML Reporting to PICTs

There was a mixed set of views on possible storage of microdata with the PDH-ML, inadequate awareness of the extent to which their microdata was used by external researchers, and lack of deposit of research results with the NSO.

There seems to be a need to examine the efficacy and regularity of reporting by the PDH-ML to the NSOs who had deposited their microdata with the PDH-ML, on all aspects of their microdata usage.

Most of the NSOs were prepared to support any initiative by the PDH-ML which led to the production of more reports from their microdata sets through a collaboration between researchers and their own technical staff and staff from other government departments.

They were also confident that such initiatives would receive support from their public stakeholders.

Unusually, only one PICT NSOs reported that there might be some researchers at their national academic institution.

Legislative Blockers

Most NSOs did not perceive any legislative blocker to them depositing the microdata sets with the PDH-ML but also thought that their legislation needed to be updated and made more explicit in terms of their ability to make the datasets available to the PDH-ML and external researchers.

General challenges for NSOs

The NSOs by and large reported very similar challenges facing their offices with differing orders of importance: funds for more surveys, loss of technical staff, lack of replacement staff locally, and inadequate salaries. Most did not see government interference by ministers as blockers to their reports. Some saw lack of support from donors and multilaterals.

Role of SPC

There was a view that the SPC assistance had diminished in recent years compared to their previous more active role. This may be more the result of a lack of PICT-wide publicity by SPC of their many contributions to HIES and other data analysis in recent years. I would point to a whole series of work coordinated by SPC/FAO experts and funding on Food Security in Vanuatu, Kiribati, Marshall Islands, and Solomon Islands, based largely on HIES microdata sets.

Perceptions of declining contributions by SPC may largely also be the consequence of a lack of systemic funding and participation by some NSOs in such regional exercises that FAO has been conducting.

One of the strong recommendations of this Report will indeed be a widening and deepening of such systemic and harmonized analyses of HIES of all the PICTs where such work has not been already done, on the key areas of poverty, food security and hardship, in collaboration with multilaterals (like WB, FAO and ILO) and academic institutions.

While there have been several Australian academic institutions involved, there seems to be a lack of participation of regional universities who could ensure long term sustainability of such initiatives.[35]

[35] This presupposes that there are qualified regional analysts who are available and willing to participate.

6. Justifying freer access to HIES microdata

The Questionnaire responses have indicated that all the NSOs would like to have more analyses from their microdata sets they have sovereignty over. It would seem from Annex 11 that all the PICTS who have already produced the basic HIES Reports are also looking for more Reports that widen the range of policy perspectives for their own governments.

It is clear from the Questionnaire responses that virtually all the PICTS report not just a lack of technical staff, but also analytical skills available to them for data analysis. This wider range of analytical reports can be encouraged if the NSOs made their datasets available to external users like the international multilateral organizations and academic researchers focused on poverty, food security and gender issues.

It should be kept in mind that NSO technical staff involved in censuses and surveys specialize in data collection and data integrity, rather than data analysis which requires a totally different set of academic skills.[36]

It should also be emphasized that there are many stakeholders, other than governments, who have a legitimate interest in the outputs from HIES and EUS microdata sets. There are also NGOs, multilateral organizations and donors who work in the areas of poverty, food security, gender issues, employment, health, education, housing, social welfare and sport, to name just a few policy areas which can be informed by microdata analysis. PICTs' peoples will benefit from more policy-oriented studies and reports, if multilateral and academic researchers are given greater access to their microdata sets.

The multilateral researchers facilitated by WB, FAO and ILO do bring a depth of skills and expertise that is largely lacking in the Pacific, either in the academic institutions or the public services. The region well knows the many poverty and hardship analyses facilitated by WB and ADB. There has also been a recent excellent initiative by FAO in the area of food security, for several PICTs (documented in this Report).

This Report suggests that NSOs and more importantly, the decision-makers behind them, need to be convinced of freer access for their HIES microdata not just to influential and well-resourced multilateral international organizations (like WB, FAO and ILO) but also to other independent academic researchers whose content and methodological approaches may be different from that of multilaterals. It is also a reality that the resource requirements of multilateral organizations are quite large. What PICTS might like to investigate is the use of national and regional academic staff as collaborators on these major projects with the multilaterals to not just economize but also help to build some ongoing domestic capacity in the PICTS and the region.

As discussed later in this report, after each household survey NSOs need to "close the virtuous cycle" for statistics gathering by running workshops in which the stakeholders discuss the policy recommendations in Reports based on analysis of the microdata. NSOs need to keep in mind that while these non-government stakeholders are legitimately part of the "public" who need greater access to microdata, they rarely have any influence on internal policies of NSOs towards access to microdata, which are usually the domain of governments.

[36] This latter point was made by one of the reviewers to an earlier draft.

The cost effectiveness of household surveys

It should first be noted that statistics legislation of most PICTs stipulate that their NSOs will conduct censuses of population, at least every five years, and government budgets readily make financial provision for these relatively expensive exercises. On the other hand, household sample surveys (such as HIES, EUS and LFS) have often struggled to obtain budgetary support, despite their capacity to provide almost equivalent national level information at a fraction of the cost. They are quite often planned but postponed because of budget constraints.

The Questionnaire Survey has suggested a relative lack of support of government and the public for analysis of HIES microdata sets. This may be due to an inadequate appreciation of the enormous statistical power of relatively cheap well-run sample surveys like HIES, to reveal development characteristics of the population, way beyond just counting the number of persons and their demographic characteristics as evident from censuses. The increasing need for this far greater range of development indicators may be seen in the hundreds of SDG indicators for which all countries are now committed to gathering statistics on.

The PSMB might wish to consider a "marketing exercise" targeting the higher up decision-makers including the political heads and senior civil servants who may be in a position to facilitate more household surveys through funding and staffing of NSOs.

One possible way of reducing the problem is to prepare a summary of the types of analyses that can inform the huge range of key decisions that governments need to make through small sample household surveys.[37]

For instance, NSOs understand only too well that HIES or EUS microdata may be derived from only a fraction of the households in the country compared to a Census of all households. But if properly sampled, weighted[38] and analyzed, they can fairly accurately represent the entire country for the fraction of the cost of a census. From household surveys, governments can, for instance, obtain solid national estimates of contributions not being made by workers to national provident funds [39] or even national populations.

The value to policy makers of all the HIES information can be easily guessed at by noting that the microdata from even the small samples, gives comprehensive information on

- different regions of the country (urban, rural, squatter settlements)
- all the incomes of everyone in the household
- all the expenditures (including all the food items, imported and locally produced)
- the quality of the dwelling and all the services (rooms, water, sewerage, etc.)
- demographic data on all the occupants (gender, age, marital status, education, citizenship, employment status, etc.).

[37] I am grateful to Len Cook for this suggestion.

[38] The "weight" is the factor (the inverse of the probability of that household being selected) by which the household observations need to be multiplied to derive its contribution to the national estimate effectively.

[39] For example, this consultant found that a less than 3% EUS sample of Fijian households was able to give an extremely accurate estimate of what the Fiji National Provident Fund should have received as annual contributions, accurate to within half a percent of actual receipts. It also revealed the significant amounts not being contributed as was legally required.

These HIES microdata sets provide information on virtually all the major policy areas that have been the focus of not just PICT governments, but also multilateral organizations like the UN, FAO, WHO, ILO etc.[40]

It is often the case that PICTs' HIES data have historically been analyzed by experts from multilateral organizations like UNDP and WB (referred to below) usually as part of their international and regional initiative, for example in poverty analysis (WB and ADB) or food security issues (FAO). Unfortunately, by their very irregular nature, their analyses are not nationally sustainable over time, their analysts are not continuously accountable to local stakeholders and sometimes, and their methodology may be determined by their need to be internationally consistent with their analyses in other developing countries, some of which may not be suitable for middle income PICTs.

It would be useful for PICTs to have a wider range of analyses before them especially national and regional researchers and analysts who continue to be accountable after the production of reports to policy makers and provide continuity over time. It cannot be emphasized enough that there are many contentious issues in the methodology of analysis, of even one area such as poverty analysis. This is brought out in Annex 12.

Critical importance of data preservation

In many policy areas, stakeholders wish to know how particular development indicators have been changing over time. Is poverty increasing or decreasing? Is food security improving or regressing for PICT people? Is gender equality improving in employment, incomes, household work, leisure activities?

Answering such questions requires the analysis of microdata over time, in a fluid situation where methodologies of analysis are constantly improving. The results of one report done ten years before may not be strictly comparable with one performed ten years later, because of the changes in methodology.

Data analysis could be enriched considerably if analysts could analyze both the most current survey data and the raw data of previous surveys with best practice current methodology.

It is vital, therefore, that any survey data collected by PICTs be preserved for posterity, for later analysis if necessary. Only thus can there be accurate estimates of time trends in development indicators.

In modern science, it is important that scientific experiments can be truly replicated if the results are to constitute reliable "scientific knowledge". Similarly, in social sciences such as poverty analysis, it is crucial that the basic HIES microdata is preserved for posterity so that scientific replication of the analysis is possible.

Recent catastrophic events such as the volcanic eruption in Tonga or possibilities of tsunami damage to entire atoll countries, drive home the catastrophic risk to the security of local microdata repositories and also readily available microdata to guide catastrophe response. This is recognized in the legislation of Samoa, Tonga and Cook Islands.[41]

[40] New studies over this next decade may reveal the impact of "over-crowding" on the spread of COVID in countries like Fiji and Solomon Islands.

[41] I am grateful to Len Cook for this information.

Nonetheless, it might not be affordable for under-resourced PICTs to store microdata, whether "in the cloud" or elsewhere. The PDH-ML (Tier 1 as proposed in this Report) and other international repositories like the WB-ML, or both, are feasible possibilities.

As PICTs become more and more experienced in poverty analyses and perhaps with a greater range of international researchers, it is quite likely that they may wish to revisit previous studies, their methodologies, their analyses and their results.[42] This can only be done if the original microdata is safe and available to be analyzed.

Given the high costs of each PICT creating a safe repository, a collective repository with experienced and reliable staff, such as PDH-ML can provide this invaluable service for PICTs, should they wish.

Official HIES Reports

Annex 10 is a partial list of the HIES Reports issued by the NSOs of the PICTS. It should be noted that SPC/SDD technical staff often provided statistical support to facilitate the implementation of many of these HIES and HIES Reports.

The HIES and EUS had a wide range of different sponsors (WB, ADB, FAO, donors etc.) of the analysis and report writing. The content usually reflected the interests of the sponsoring agency and analysts. Thus WB, ADB and donors like AusAID were typically interested in poverty and income analyses, while FAO was more interested in food security and ILO in labour.

Some reports had just main tabulations, some also regional and provincial analysis in addition to the national analysis, some had poverty analysis by various methodologies, and some had analysis of food security issues.

Just a cursory glance at Annex 10 would suggest that the timing of many of the HIES and EUS were largely the result of the sponsors' interest at the time. It cannot be emphasized enough, and this is supported by the NSO responses to the Questionnaire, that for most PICTS technically capable staff is a major constraint to implementing long-term plans for HIES and other surveys. Implementing solid sample surveys requires sophisticated sampling techniques, interviewer training, and subsequent analyses, all drawing on previous experiences. Few PICTs will have senior survey staff of long enough tenure, as staff turnover is endemic, especially with grossly inadequate local civil service salaries, compared to what the regional and international labor markets can offer.

There would seem to be significant potential advantage if there were one regional organization which coordinated all such analyses and harmonized the methodology to be used. SPC's SDD already has a track record in assisting many PICTs with their HIES analysis and recent coordination with FAO on regional analyses of food security. This work could be extended to other surveys as the statistical skills would be largely the same, although the analytical content would differ.

Given that the SDD and PDH-ML are engaging in an exercise to harmonize all the data in the various HIES, there would seem to be considerable scope in a grand Pacific wide analysis using agreed upon methodology, to produce priority outputs. These would be part of

[42] This is a possibility with the Fiji Bureau of Statistics which has had its 2020-21 HIES analysed by WB with a quite different methodology from that used previously so that poverty results could not be compared with that derived from earlier HIES. The 2020-21 results were also significantly changed after initially publishing the results, with some political controversy.

the Tier 1 HIES microdata sets recommended in The Policy Brief allowing access to external researchers who satisfy the PICT conditions.

Part of the work schedules of PICT government ministers, NGOs and donors includes regular attendance at international stakeholder meetings, at which they make commitments on SDG indicators for monitoring poverty and gender inequality. It would be beneficial for their reputations, good will and future funding, if they were to produce real data demonstrating that progress was being made, however slowly or rapidly.

It should also be pointed out that assistance from multilateral organizations like WB is not guaranteed continuity over time, nor is there any guarantee of consistent methodology over time, for example on the criterion for measuring poverty. [43]

Annex 12 briefly outlines some of this consultant's publications whose content (briefly summarized) illustrates the greater choice that PICTs can have in terms of analyses of their HIE or EUS data, some different from external multilateral organizations and their perspectives. These are the kind of research outputs possible through Tier 2 microdata sets recommended in the Policy Brief.

Risks of Microdata Sharing

The experience of this author who has been given access to numerous PICT microdata sets is that there has been no identification of any individuals, households or legal entities, even if the risk does exist. The risks do not seem to have been materialized in any recorded instances. The consultant is also not aware of any studies in the Pacific that have focused on this issue.

While the risks to confidentiality was listed by virtually all the NSOs responding to the Questionnaire Survey there were no examples given of breaches. The PSMB with their personal knowledge of the PICTs might wish to consider the potential costs to PICTs populations, governments or entities.

The PSMB may wish to consider the recommendation by several NSOs responding to the Survey Questionnaire, that there be penalties specified in the legal agreements granting microdata access to external users. Unfortunately, such penalties may be difficult to enforce. As one Reviewer suggested, the possibility (not actually doing so) of "naming and shaming" individuals may be more than sufficient to address any possibility of microdata misuse.

In contrast, the benefits of microdata sharing with external users, and especially multilateral organizations like WB, FAO and ILO, have been immense, in terms of research outcomes and guidelines for evidence-based policy improvements, but also to guide ODI investment.

It may be also useful to explore whether donors to the PICTs may be interested in funding a grand project whereby the HIES (and other microdata sets such as EUS and LFS) may be analyzed across the Pacific systematically to produce harmonized Reports for policy use generally, involving expert external research organizations like WB, FAO, ILO and universities from the region and internationally.

[43] Many developing country HIES have weak data on income, so WB has usually preferred expenditure variables as indicators of welfare. There are some countries, however, (like Fiji) that have excellent data on incomes which was used as the criterion in their previous poverty analyses.

An ongoing Review of SPC Data Governance Policy

It was pointed out by Reviewer 4 that that this Report should cross-reference that SPC was currently conducting an overall review of SPC's Data Governance Policy for microdata management and dissemination in the Pacific.

Another Reviewer pointed out that the

(a) new Data Governance Policy should needed to place considerable emphasis on building and enhancing trust, an important part of maximising the use of microdata in PICTs

(b) SPC will need to both provide tangible benefits for members and communicate success widely.

(c) Data should be considered as an asset for the public good.

(d) SPC was starting from a relatively low and varied level of data governance maturity but there was broad support and willingness on governance reform.

(e) There was a need to build people capability, supporting standards and guidelines, and improve the tools.

These are all useful observations which are mirrored in various parts of this Report.

7. Legislation as Blockers and Facilitators

In the Pacific, the most sophisticated system of access to microdata may be seen in the operations of the Australian Bureau of Statistics and NZ Statistics. It is therefore useful to outline their policies and processes for safeguarding confidentiality as models that PICTS can follow or at least use as bases to develop their own models from.

Given the significant technical difficulties that are likely to be faced, PICTS may also wish to draw on technical assistance from ABS and NZ Stats to advance their cause.

Australian Legislation and Australian Bureau of Statistics

The philosophy behind the ABS approach to freeing up access to microdata may be seen clearly in the *Australian Government Public Data Policy Statement (2015)*. The Australian Government regards data as "a strategic national resource ... for growing the economy, improving service delivery and transforming policy outcomes ... for the benefit of the Australian people."

The Australian Government "commits to optimize the use and reuse of public data; to release non sensitive data as open by default[44]; and to collaborate with the private and research sectors to extend the value of public data for the benefit of the Australian public. A number of other very detailed commitments are made on how to make data available, all useful for PICTS to look at:

1. with free, easy to use, high quality and reliable Application Programming Interfaces (APIs);

2. only charge for specialised data services;

3. where possible, publish the resulting data open by default

4. build partnerships with the public, private and research sectors to build collective expertise and to find new ways to leverage public data for social and economic benefit

5. uphold the highest standards of security and privacy for the individual, national security and commercial confidentiality; and

6. the public has redress through an appeal process if access is denied.

The ABS website also has clearly published rules about how it maintains confidentiality, the "five safes" framework they use (as explained earlier in the Report), how they treat microdata to ensure confidentiality.

The ABS website sets out other conditions that PICTS might wish to consider for their own microdata sets either directly or through the PDH-M: the microdata user obligations, undertakings and declarations, secure storage of microdata.

[44] Note the term "open by default". Denying access would need to be justified.

The ABS Tiered Access.

The ABS sets out three basic levels of access (Tiers) to microdata:

1. TableBuilder: uses highly detailed underlying microdata to allow the public to produce their own tables, graphs and maps. The outputs are aggregated and anonymized automatically, so the user can access results immediately. As users cannot access the microdata records directly, some microdata obligations, such as undertakings, do not apply.

2. Basic Microdata is the least detailed type of microdata (usually at broader levels of their classification) which can be downloaded and analysed by the user.

3. Detailed Microdata : Detailed microdata is accessible in the secure DataLab. Records are de-identified and confidentialised appropriately within the context of the other security features of the DataLab, and contain highly detailed data items.

NZ Legislation and Stats NZ

Stats New Zealand's Integrated Data Infrastructure (IDI) is a large research database containing de-identified microdata about people and households, from original sources in Stats NZ and other government departments.

The broad components are health, education and training, benefits and social services, justice, people and communities, population, income and work and housing.

The income and work data includes data on tax and income, Household Economic Surveys, Household Labour Force Surveys, New Zealand Income Surveys, and the Survey of Family Income and Employment. Most have data going back to 2006 and before.

"Core data" are available to all IDI researchers with an approved research project, while for restricted data and data dictionaries access is granted on a case-by-case basis.

There is also a rich library of all research outputs, including conference papers, research and methodology, based on data derived from the IDI. This ought to be emulated by the PDH-ML.

Ambiguities in PICTS legislation

One of the "blockers" to freeing up access to microdata is seen by some NSO respondents to the Questionnaire as the inadequacies or ambiguities in national legislation relation to outside access to microdata.

The NSOs of some countries think that their Statistics Legislation does not allow them to share microdata with external agencies. One NSO in particular stated that they would like their legislation to have an "open data" philosophy which allowed researchers access to their microdata.

The PSMB might wish to canvas among its members whether there is any parallel in PICT indigenous communities to the reluctance that has been perceived among the Maori of NZ towards data specifically on them being made public to external researchers.[45]

[45] I thank Vince Gavin for this observation.

Statistical Councils

One of the important institutional features that can facilitate the freeing up of microdata for public use is the establishment of "National Statistical Councils" or "National Advisory Boards" that can guide NSOs with the development of policies for their offices.

An important prototype that the PSMB can consider is that set up by the Cook Islands Statistics Act 2015-16. Chapter 12 of this Act states that the Minister may establish a Statistical Council drawn from government, businesses, and the community to: advise on the range and scope of official statistics; review the priorities and programmes of work; oversee the maintenance of public trust in statistics - including the processes that ensure public confidence in the confidentiality provisions ; advise the Government Statisticians on priorities for future access needs, and the form and quality of available statistics; designate as official statistics any statistical measure not produced by a department that meets the standards of official statistics and provides significant information about a sector of official statistics that would otherwise not be available.

The Fiji Bureau of Statistics around 2008 to 2011 did have a Statistics Advisory Board.[46]

[46] This consultant was the Chair of that Board, which seems to have been inactive after the departure of the Government Statistician then.

8. Importance of Formal Agreements With the Pacific Data Hub - Microdata Library

Several NSO responses to the Questionnaire Survey, stressed that it was important that access to PICT microdata sets be controlled by formal legal agreements. Without revealing details of any DLA with any specific PICT, the following appear to be common key provisions of the Data License Agreements (DLAs), given in detail as they compare quite favourably with the PCRSP (2009) principles and good practice for data preservation as summarized earlier in the Report through Option 1 (Data Seal of Approval) and Option 2 (the 10 characteristics of a good Digital Preservation Repository).

Legal provisions:

The Licensor represents and warrants to SPC (and SPC enters this Agreement) in express reliance on such representations and warranties that:

a) it has the right to provide the Dataset(s) to SPC;

b) the data included in any Dataset(s) being provided to SPC has been collected and compiled in compliance with the Licensor' s legal or regulatory requirements;

c) the Licensor has ensured that SPC has knowledge of all legal obligations that require researchers to sign a declaration of secrecy of the Licensor before viewing any microdata;

d) the data in the Dataset(s) contains no Personally Identifiable Information (PII); and

e) the Licensor maintains all liability in relation to all third-party claims arising out of these representations and warranties.

Intellectual property

The Parties agree on the treatment of Intellectual Property rights in the data shared under this Agreement as given below.

a) The Parties recognise the importance of protecting and respecting intellectual property rights. Pre-existing intellectual property rights remain the property of their current owner. Any intellectual property rights arising out of joint collaboration will vest jointly in the Parties.

b) The Licensor shall retain all rights to the intellectual property of the Datasets provided to the SPC under any of the data access protocols described in the SPC Terms of Use.

c) All derivative works created by SPC, using the Dataset(s) and metadata provided by Licensor under this Agreement, may only be redistributed by SPC as specifically stipulated in the Data License Agreement of the Dataset(s).

d) Dissemination will commence after any consultation between the two parties normally required for derivative studies.

This Agreement does not grant on the part of either Party the right to use any other materials belonging to, or created by the other Party, except for the Dataset(s) specified in Article I ("Name and Description of Dataset(s)"), or as specifically stipulated otherwise in Article III ("Data Use and Sharing").

Confidentiality and protection of privacy

The Parties shall take appropriate measures to ensure confidentiality and protection of privacy relating to the datasets under this Agreement. The Parties shall also ensure that users who access the data will maintain the absolute confidentiality of persons and households. Any attempt to ascertain the identity of a person, family, household, dwelling, organisation, business or other entity from the microdata is strictly prohibited. This includes alleging that a person or any other entity has been identified in these data is also prohibited.

Resolution of disputes:

The Parties shall make their best efforts to amicably settle any dispute, controversy or claim arising out of this Agreement

a) In case of a dispute, controversy, or claim between the Parties arising out of or relating to this Agreement or an agreement for any activity or project undertaken pursuant thereto, the Parties shall attempt to reach an amicable resolution in good faith.

b) The SPC recognises that the head of the statistical office and the judicial authorities of the country of the Licensor are the sovereign authority on matters relating to compliance with the statistical legislation of that country by all parties to this agreement.

Any disputes that might arise from or in relation to this Agreement, if not settled by negotiation, shall be settled by arbitration in Auckland, New Zealand in accordance with the Arbitration Rules of the United Nations Commission on International Trade Law (UNCITRAL).

This Agreement may be modified or terminated at any time by mutual written agreement of the Parties. Further, this Agreement may be terminated by either Party at its sole discretion with a specified number of days' prior notice in writing to the other Party.

There are other MOUs, such as with UNICEF and IMF, that may be examined for their relevance to PICTS.

What ought to be added to the DLA is what most NSOs asked for in their Quetionnaire Responses: all research outputs resulting from the access to the microdata, (whether published or not), must be deposited with the PDH-ML.

9. Policy Brief

For ease of discussion by the PSMB and the NSOs, this Policy Brief is presented as a numbered list, which format I believe will be more convenient for discussion by PSMB and the PICT NSOs.[47]

The global movement towards Open Data

1. There are strong international movements towards the freeing up of publicly generated data for greater use in evidence-based policy making. Organizations such as Open Data Watch are dedicated to monitoring the extent of openness of countries' data systems and their critical development components, ranking them globally and fostering greater openness. The PSMB may wish to discuss whether there are benefits for PICTs to be part of this global movement and confirm this for the benefit of any PICT NSOs which may have any remaining doubts.

2. There is evidence that openness of data systems is conducive to strengthening strategies for development through evidence-based policies. One result of making microdata more open is the encouragement of a wider research community beyond government departments and statistics officers, working together with common ethics, principles, scientific methods and objectives. The resulting research findings have a significant capacity for placing public debate on solid evidence-based research which is politically neutral or non-partisan. There is little doubt of this benefit among PICTs.

3. This benefit is especially relevant for under-resourced small PICTs with small cadres of professional statisticians and research personnel in the civil service. Local academic institutions can also be similarly under-resourced in the relevant fields. Widening the research community for these small PICTs has the potential for many benefits which would be otherwise foregone. The possible benefits are illustrated by Annex 12, the record of just one Researcher.

The PICTs are lagging behind

4. By international standards and openness rankings, PICTs are generally lagging behind other comparable developing countries globally as indicated by Table 1 for development data in general and Table 2 for microdata relating to poverty and gender in particular.

5. This sentiment that there is a significant underutilization of the PICTS microdata is implicit in the common NSO response to the Questionnaire Survey that they wanted to see more analytical Reports resulting from their HIES microdata sets. It is also indirectly reflected in the instigation of the PACSTAT Project at SPC and the commissioning of this Report by the PSMB.

[47] One Reviewer was not favourably disposed to a numbered list.

The Performance of the PDH-ML

6. The SPC's PDH-ML is a valuable resource for resource-scarce PICTs not just in terms of data preservation and archiving, but also for facilitating a greater level of access to the microdata for global, regional, national and independent researchers.

7. The performance of the PDH-ML Repository may be assessed by

(a) the percentage of all generated micro datasets actually stored at the PDH-ML (25%)
(b) The historically small number of requests for usage of datasets and
(c) the low 35% of approvals of the requests.

All three of these Key Performance Indicators ought to be the subject of inquiry by SDD and PDH-ML.

8. The PSMB should note that the rate of approval for microdata access is low not only for independent researchers and regional academic institutions, but also for multilateral organizations such as WB and FAO, whose bona fide would not be in doubt. The PSMB might wish to initiate an inquiry with the PDH-ML and PICT NSs for more information on the process of granting access to the microdata sets and whether the approval rate can be improved.

The microdata collections in general

9. The PDH-ML can only store microdata sets if they have in the first instance been generated by the PICTs. Annexes 8 and 9 indicate the large diversity of Item Types and their irregularity over time. This is hardly surprising given the extreme paucity of taxpayer resources allocated to NSOs of most PICTs, even large ones like Fiji and PNG, to conduct surveys regularly, according to some long term program. Many microdata sets were generated as a result of a new focus on particular PICTs by international organizations with their relatively abundant resources.

10. It is recommended that PSMB place on their agenda, a discussion about PICT priorities for the conduct of regular national surveys and censuses which can result in microdata sets in priority areas (such as economics, business, labour force and health). Such programs ought to also guide the international donors and multilateral organizations in funding microdata collection exercises.

11. These priorities will not be uniform across all PICTs. For instance, atoll PICTs will probably have climate change effects near the top of their list; high population growth countries will have employment creation and labour force changes more important; etc.

The storing of the microdata sets at PDH-ML

12. Given the low percentage of existing microdata sets actually deposited at the PDH-ML, the PSMB consider a special project to encourage PICTs to deposit all missing microdata sets (including the HIES) conducted over the years as well as all the Publications, Questionnaires and Manuals.

Blockers for greater microdata use

13. Different PICTs have different dynamics at work to explain the relative under-utilization of their microdata sets and their lack of approvals to requests. Some saw ambiguous legislation not helping their decision-making; some made the judgement that applications were not backed by reputable institutions; some thought that the research was of no use to their country and people; risks to confidentiality were articulated but there was no evidence suggested that confidentiality had been compromised. The low questionnaire response rate precludes stronger statements.

14. The NSO responses to the Questionnaire Survey suggest

    (a) they would all like to have an Open Data philosophy

    (b) that they would like to free up access to their HIES microdata to external researchers

    (c) that they would like to have more analytical evidence-based Reports resulting from their HIES microdata

    (d) that most have ambiguous legislation about their ability to make their microdata available to external researchers, and some even to other Government departments.

    (e) that most have a scarcity of technical and analytical staff.

    (f) most would like assistance on data anonymization and data analysis

    (g) most do not see Government ministers interfering with their work.

Special Project to foster greater microdata use

15. The PSMB and PICTs consider a special initiative to increase the utilization of the PDH-ML microdata sets of those PICTs giving approval for their datasets and in their areas of priority.

16. Annex 12 shows the potential for generating evidence-based policy analysis by independent academic researchers analyzing a wide variety of HIES and EUS microdata sets. Such initiatives can be facilitated by a confluence of funds from government and donors and quality academic researchers willing to work in the designated areas and producing the policy oriented reports.

Completing the virtuous cycle: holding workshops

17. All too often, NSOs consider the job done when the Reports are produced. However, the PSMB should consider encouraging the PICT NSOs to complete the "virtuous cycle" whereby the NSO not only conducts the national household surveys whose microdata is used to produce policy-oriented reports, but

    (a) funds are also made available for national workshops and policy discussions among the relevant government ministries, NGOs and donor stakeholders using the survey results.

    (b) ensuring an adequate media discussion of these dialogues which improves the goodwill of the public who laboriously respond to the NSO household surveys, often revealing sensitive information about their households' to NSO interviewers. This goodwill cannot but make it easier for the NSO to implement the next survey.

18. This public perception of the benefits of this "virtuous cycle" also helps to build public trust in the PDH-ML as a central Repository of great benefit to PICT NSOs and PICT taxpayers who ultimately pay for the household surveys that lead to the generation of the microdata sets.

19. This "virtuous cycle" of data gathering and publicization of results through national workshops ought to be considered as a valuable "incentive" for allowing greater access to PICT microdata, the opposite of "blockers" which tend to dominate the discussion.

Critical pathways forward: The 4 Tier Structure

20. While the PSMB and PICTS may clearly benefit by being in step with the international trends towards open data, it is also important that the process be driven by the development needs of PICTs with concrete benefits being delivered at each step of the process towards the freeing up of their microdata sets for analysis, report writing and public policy dialogue.

21. This Report suggests that the PSMB can assist the PICT to be in control of the pace and modus operandi with which their datasets are freed up, through a tiered approach as follows:

    Tier 1: The secure archiving of all microdata with the PDH-ML, with no necessary access to the public (unless agreed to as for the next tiers). The target KPI for this would be 100% of all microdata historically created.

    Tier 2: To consist of microdata which PICTs are willing to make available to a short-term Special Project managed by SDD for analysis and writing of reports as prioritized by the NSOs, including poverty, food security and gender inequalities (and any other areas considered important to them).

For this exercise, SDD can call for expressions of interest from multilateral organizations, universities (international and local) and research organizations, and recognized expert independent researchers to collaborate with the specific NSOs and their designated civil servants for the analysis of the specific microdata and Report.

Following every publication of the Report, the SDD/NSO can facilitate a national workshop bringing together all the national and global stakeholders in the policy issues discussed.

Eventually there can then also be an over-arching regional conference/workshop which brings together these Pacific-wide findings which can be published in a monograph.

Tier 3 Microdata sets which PICT NSOs are prepared to make available *upon application* by external researchers with the PDH-ML co-ordinating the process of approval by the PICTs who have sovereignty over that particular data set.

Tier 4: Microdata sets which PICT NSOs are prepared to make freely available for download, with strict conditions being met through terms of use.

Transparency of Approvals/Rejection Process and Reporting

22. There should be a formal reporting process whereby annually, the PDH-ML informs the PSMB and PICT NSOs of all Requests, Approvals and Rejections (with clear explanations for rejections).

23. The PDH-ML should also post on its website all Reports and publications that have resulted from access to the microdata sets in the PDH-ML, and be freely downloadable, unless specific restrictions have been put in place by the originating authority and agreed to by the PDH-ML and the PICT concerned.

24. The experience of the PDH-ML over time should then inform further changes in the structure of the Tiers and the conditions under which their datasets may be made available to researchers.

Key Performance Indicators

25. Each PICT should also be able to monitor its own progress in opening up access to their data and microdata through their rankings on the ODIN database.

26. The PDH-ML should also be able to monitor the Pacific's progress through external international indicators such as those collected by ODIN for open data in general and microdata in particular addressing poverty, food security and gender issues.

10. Annexes

Annex 1 ODIN Ranks and Scores (Social, Economic and Environment)

	ODIN	All Indicators Score (/100)			Social Indicators Score (/100)			Economic Indicators Score (/100)			Environment Indic. Score (/100)		
Countries	Rank	Cov.	Op.	All	Cov.	Op.	All	Cov.	Op.	All	Cov.	Op.	All
Singapore	1	80	100	92	62	100	83	95	98	100	88	100	95
New Zealand	25	57	81	70	44	72	58	77	89	84	55	82	70
Australia	44	56	69	63	49	70	60	77	80	79	45	58	52
Jamaica	53	52	70	62	43	68	56	67	63	65	48	80	65
Mauritius	67	54	60	57	40	53	46	77	57	66	50	60	57
Malaysia	78	53	53	53	40	46	43	71	54	61	52	58	55
Maldives	84	50	53	52	53	56	54	64	56	59	3	48	43
Trinidad & T	173	22	32	27	15	30	23	27	29	28	24	38	32
Samoa	109	40	53	47	26	33	29	65	73	70	33	52	43
Tonga	125	44	40	42	37	39	38	68	46	54	29	40	41
Fiji	142	41	38	39	30	31	30	67	50	57	31	32	32
Marshall Is.	161	34	33	33	27	32	30	43	31	36	32	34	33
Palau	169	38	24	30	23	19	21	66	36	47	29	24	30
Solomon Is.	171	30	28	29	27	38	32	58	39	47	10	8	9
PNG	178	17	24	21	12	24	18	38	39	39	5	8	7

Coverage scores: Based on 5 elements Data available for last 5 years and 10 years..

Openness scores: Based on Data being non-proprietary, download options, availability of metadata, machine readability, terms of use.

Social indicators: Population and vital statistics, education facilities, education outcomes, health facilities, health outcomes, reproductive health, food security and nutrition, gender statistics, crime and justice, poverty and income.

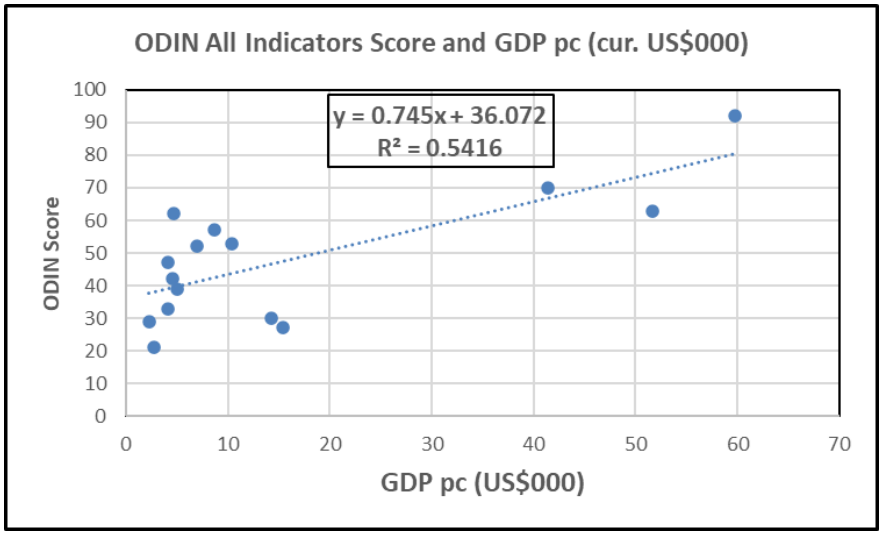
Economic indicators: National accounts, labour, price indexes, government finance money and banking, international trade, balance of payments.

Environment indicators: Agriculture and land use, resource use, energy pollution, built environment.

Source ODIN Country profiles, as of July 2021 . Link:

Annex 2a ODIN Scores (Coverage and Openness) and GDP pc (cur. US $)

Country	GDP pc Cur. US $	ODIN Score (Coverage and Openness) All indicators
Singapore	59798	92
New Zealand	41441	70
Australia	51693	63
Jamaica	4665	62
Mauritius	8628	57
Malaysia	10412	53
Maldives	6924	52
Trinidad and Tobago	15426	27
Samoa	4068	47
Tonga	4625	42
Fiji	5058	39
Marshall Islands	4130	33
Palau	14244	30
Solomon Islands	2251	29
Papua New Guinea	2757	21
Source: WB Indicators for GDP pc (cur. US $). Link:		



Annex 2b State of Indicators Not Published, Lacking Gender Disaggregation, and Availability.

Annex 2b				
	Indicators Not Published (%)	Indicators Lacking Gender Disaggregation (%)	WS A/NA	NSO Data portal
Singapore	12	29	O	O
New Zealand	26	43	O	O
Australia	26	33	O	O
Jamaica	15	33	O	O
Mauritius	14	48	SR	SR
Malaysia	26	52	NO	NA
Maldives	31	38	NA	NA
Trinidad and Tobago	51	24	NO	NO
Samoa	34	29	NA	
Tonga	34	48	NA	SR
Fiji	34	52	NO	
Marshall Islands	49	71	NA	
Palau	54	48	NA	
Solomon Islands	45	57	NO	
Papua New Guinea	62	48	NO	

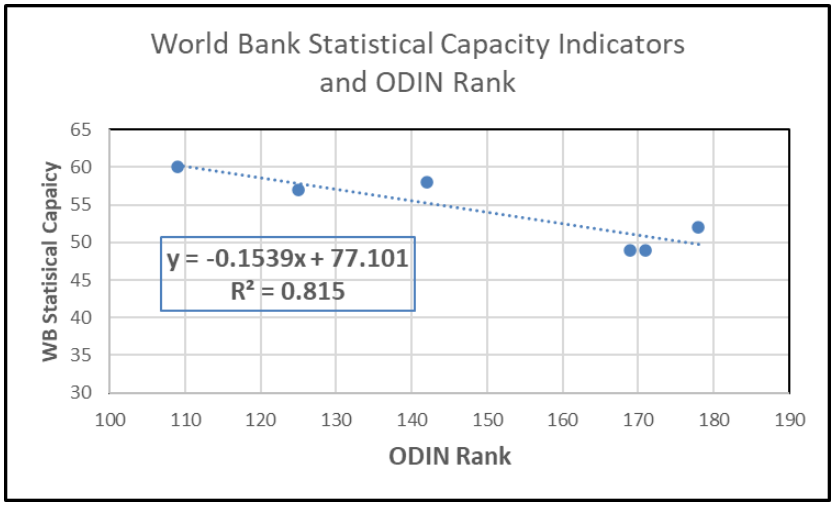
INP Indicators not published

ILG Indicators lacking gender disaggregation.

WSA/NA/NO Website Available/Not Available/ Not Open

Annex 3 World Bank Statistical Capacity Indicators and ODIN Rank[48]

	ODIN Rank	Overall Score (%)	Methodology Score (%)	Source Data Score (%)	Periodicity Score (%)
Samoa	109	60	60	50	70
Tonga	125	57	50	50	70
Fiji	142	58	30	80	63
Marshall Islands	161	32	20	20	57
Palau	169	49	30	60	57
Solomon Islands	171	49	40	40	67
Papua New Guinea	178	52	50	40	67
ODIN Website . Link:					



[48] The graph excludes the Marshall Islands outlier.

Annex 4 Principal International Microdata Repositories (2000-2022)

Repository	Survey Focus	URL	No. of countries included	No. of surveys included
World Bank Microdata Library***	General household surveys	https://microdata.worldbank.org/index.php	219	3,440
International Household Survey Network**	Household and institutional surveys and censuses	http://catalog.ihsn.org/catalog	218	5,130
Food and Agriculture Organization Microdata Catalog***	Surveys with agriculture and/or food security focus	https://microdata.fao.org	173	486
UNHCR Micro Data Library***	Surveys on refugee populations and impact evaluations	https://microdata.unhcr.org/index.php	69	242
World Health Organization NCD Microdata Repository*	Non-communicable disease surveys	https://extranet.who.int/ncdsmicrodata/index.php	186	374
WHO Multi-Country Studies Data Archive*	Health surveys	https://apps.who.int/healthinfo/systems/surveydata/index.php/catalog/central	101	46
ILO Survey Catalog**	Labor force, child labor, and income and expenditure surveys	https://www.ilo.org/surveyLib/index.php/home	171	564
The DHS Program (USAID) *	Demographic and Health Surveys	https://dhsprogram.com/	90	260
UNICEF MICS Surveys*	Multiple Indicator Cluster Surveys	https://mics.unicef.org/surveys	118	195
IPUMS International*	Census microdata	https://international.ipums.org	102	105
Pacific Data Hub - Microdata Library	Censuses and Surveys		22	739

Source: Table 1 of Swanson et al (2021), with last row added for PDH-ML.

\+ As of Oct 20, 2021.

* Allows users to download microdata directly

* * Redirects users to the other repositories to download microdata

*** Allows users to download some microdata directly and redirects for others.

Annex 5 Microdata Access Levels used by WBML.

Access Level	Description
Open Access	Data are made available to users without the need to login or register. Data are made available under an open license that permits commercial and non-commercial use.
Direct access	Data are made freely available under the following conditions: (a) Data may not be redistributed or sold (b) Data will be used for statistical and research purposes only (c) No attempt will be made to identify persons, facilities, or establishment. (d) Attribution of the data source will be provided in outputs using the data. (e) The repository and original source of the data bear no responsibility for their use or interpretations
Public Use	Data are made available to users *after registering and agreeing to a set of conditions that are the same as direct access.*
Licensed Data Files	Data are made available to registered users who have a legitimate need to access the data ("bona fide users"). Access is granted after a review of the user's registration information that includes a description of their affiliated organization who will take responsibility for use, the intended use of the data, expected outputs, and other researchers involved in the use of the data. After the application is reviewed and if access is granted, data are made available under a license that permits non-commercial use for statistical and scientific purposes only.
Data Enclave	Data are only made available to users in a secure location after being granted access through a process similar to licensed files.
Source: Table 4 of Swanson et al (2021).	

Annex 6 WB Microdata Library (WBML) Access Controls for Microdata Files.

	Public Use Files (PUFs)	Scientific Use Files (SUFs)
Authorized Users	Users agree to a basic set of conditions. Registration may be required. Users may also be asked for a brief description of how they intend to use the data, although this does not impact their access.	Users submit an application and receive authorization after signing an agreement governing the data's use. Often authorization is only given to those in relationships with a recognized sponsoring intuition, such as research institutions, universities, and development partners.
Access Process	Immediate access given.	Access granted upon successful review of application.
Data Content and Disclosure Controls	Any identifying information and indirect identifiers are removed from the file, often including geographical information below the subnational level. Sometimes, outliers or other records may be removed from the dataset as well. Disclosure control methods such as top and bottom coding, local suppression or data perturbation techniques are preferred over deleting entire records.	SUFs generally provide more granular information than PUFs. Direct identifiers are removed, but indirect identifiers may remain in the file. Limited recoding of the data may be used to disguise outliers.
Source: Table 5 in Swanson et al (2021)		

Annex 7 WBML access by regions – Low- and Lower Middle-Income Countries

Region	Surveys Available as PUFs (%)	Surveys Available as SUFs (%)	Surveys Available in an external repository (%)
Southern Asia	67.4	3.2	29.4
Southeastern Asia	73.9	3.4	22.7
Central Asia	64.8	3.5	31.7
Northern Africa	74.0	6.1	19.8
Sub-Saharan Africa	59.9	4.3	35.8
Central America	59.4	1.8	38.8
South America	79.0	1.2	19.8
Caribbean	69.4	8.3	22.2
Eastern Europe	71.4	0.0	28.6
Middle East	79.2	0.0	20.8
Pacific Islands	87.0	4.8	8.2
Source: Table 6 of Swanson et al (2021)			

Annex 8 PDH-ML Datasets in Category "Economy, Business and Labour Force"

	Decade						
Row Labels	A Pre-1980	B 1980-1989	C 1990-1999	D 2000-2009	E 2010-2019	F 2020 +	Total
Broadband Survey					2		2
Business Activity Survey (BAS)					1		1
Business Survey				1	2		3
Cattle survey	1						1
Census of Business Activities (CBA)				1			1
Community and Socio-Economic Survey				1			1
Community Perception Survey					2		2
Economic census	1	1	1	6	6		15
Education Experience Survey and Literacy Assessment					1		1
Electric Power Corporation survey (EPC)					1		1
Enquête Technologie de l'Information et Communication (TIC)					1		1
Enterprise study					1		1
Enterprise Survey				5	2		7
Hybrid survey					3		3
LFS			2	5	11	16	34
Local Market Survey (monthly)					7		7
Manufacturing survey			1	5	2		8
MCA Vanuatu Tourism Survey				1			1
Micro, Small & Medium Enterprises 2016 Survey					1		1
Price comparison with France					2		2
Public work study				1			1
Tourism Survey				1			1
Vanuatu International Visitor Survey					2		2
Village Resources Survey			1	1	1		3
Visitor Survey			2	1	1		4
Grand Total	2	1	7	29	49	16	104

Source: Derived from PDH-ML data.

Annex 9 PDH-ML Dataset types in Category "Health"

Row Labels	A Pre-1980	B 1980-1989	C 1990-1999	D 2000-2009	E 2010-2019	F 2020+	All
Adult Health Survey					1		1
DHS			1	7	5	2	15
Disability Survey				1	2	4	7
Family Health and Safety Study (FHSS)				3	5		8
Fertility Survey	1						1
Global School-based Student Health Survey (GSHS)					17		17
Global Youth Tobacco Survey				2			2
High Frequency Phone Survey on COVID-19 - W1						1	1
High Frequency Phone Survey on COVID-19 - W2						1	1
High Frequency Phone Survey on COVID-19 W1						1	1
Household Listing					1		1
Rapid Assessment Survey (RAS) - Round 1						1	1
Rapid Assessment Survey (RAS) - Round 2						1	1
Rapid Assessment Survey (RAS) - Round 3						1	1
Rapid Assessment Survey (RAS) - Round 4						1	1
Second Generation Surveillance survey (SGS)				19	1		20
Sexual and Reproductive Health Rights Needs Assessment (SRHR)					5		5
STEPS				13	7		20
Study of Risk Factors for Chronic Non-communicable Diseases				1			1
Territorial Testing of the Chronic Rheumatic Heart Disease in Primary School						1	1
VAW survey				1	3		4
Well-being Survey					1		1
Youth Health Survey					1		1
Youth Risk Behavior Surveillance (YRBS) survey			6	14	12		32
Youth Tobacco Survey				1			1
Grand Total	1		7	62	61	14	145

Source: Derived from PDH-ML data.

Annex 10 HIES Reports for PICTs

By alphabetical order of country name:

CNMI *HIES Report.* 2008. Department of Commerce.

Cook Islands. *HIES Report* 1998.

Cook Islands. *HIES Report* 2005-2006

Cook Islands *HIES* Report

Fiji *2002-03 HIES Report*. FBS.

Fiji *2008-09 HIES Report*. FBS.

Fiji *2013-14 HIES Report*. FBS.

Fiji 2019-20 HIES Report. FBS and World Bank.

Fiji Role of diets and food systems in the prevention of obesity and non-communicable diseases in Fiji. FAO. Sarah Burkart, Dana Craven, Bridget Horsey, Jenna Perry Tarli O'Connell and Steven Underhill.

FSM *1998 HIES Report.*

FSM *2005 HIES Report.*

FSM *2013-14 HIES Report.*

Kiribati *2006 HIES Report.*

Kiribati Food Consumption Report (2019-20 HIES) (Natalie Troubhat, Michael Sharp, FAO/SPC)

Marshall Is Food Consumption (Natalie Troubhat, Michael Sharp) (SPC/FAO)

Nauru *2006 HIES Report*

Nauru *2012-13 HIES Report.*

Niue *2002 HIES Report.*

Niue *2015-16 HIES Report*

Palau *2006 HIES Report.*

Palau *2014 HIES Report*

PNG *2009-10 HIES Report.*

Samoa 1998 HIES Report. SPC/UNDP.

Samoa 2002 HIES Report

Samoa 2008 HIES Tabulation

Samoa 2014 HIES Tabulation.

Solomon Islands 2005-06 HIES National Report

Solomon Islands 2005-06 HIES Province Report

Solomon Islands 2012-13 HIES National Report

Solomon Islands 2012-13 HIES Summary of Findings Poverty Report

Solomon Islands Food Security (Natalie Troubhat, Michael Sharp, Neil Andrew) (FAO) (2012-13 HIES)

Tokelau 2015-16 HIES Report

Tonga 2000-01 HIES Report.

Tuvalu 2004-05 HIES Report

Tuvalu 2010 HIES Report

Tuvalu 2015-16 HIES Report

Vanuatu 2006 HIES Report

Vanuatu Income and Expenditure Report. 2021 (ACIAR)

Vanuatu Hardship Report 2021 (WB Darian Naidoo and Jamie Tanquay)

Vanuatu Food Security Report 2021 (FAO, ACIAR, Nathalie Troubat)

Vanuatu Labour Market Report 2021 (ILO, Tite Habikare, Yamei Du Felix Weidennkaff Christian Viegelahn)

Vanuatu Well Being Report 2021.

Wallis and Futuna 2005-06 HIES Report.

Annex 11 Summary of Questionnaire Responses

The summary of the responses corresponding to the Questionnaire numbers is as follows, with most the responses being self-explanatory.

Scores out of 5 were averages for all the responses. The actual responses of each NSO are in a worksheet attached to the Report.

1. Respondent Name: .....

2. Position: .....

3. For the last HIES conducted by your Bureau, list the Reports and Publications your office has produced, with or without the help of outside agencies (name them):

   Reports: *The publications included mostly HIES Reports, but also food security, poverty and hardship, copra brief, labour market, well-being,*

   Agencies: *SPC, WB, ILO, FAO, ACIAR, Christensen Fund.*

4. List any other Reports or publications that external agencies - including academics and research institutions - have produced using your country's HIES data?

   Labour monograph, food security.

5. Would your office have liked to produce more Reports using your HIES ?

   All answered "YES"

6. If your answer to the above question above was "Yes" what topics or policy areas would such reports have focused on:

   Gender, disability, multidimensional poverty index, labour force, employment and unemployment, health and nutritional intake, marine food consumption, vulnerable groups.

7. How important were the following factors in your office not producing more Reports? *Highlighted are the averages of the responses:*

(a) No demand from Government:	Not Important	1	2	3	4	5	Important
(b) No demand from public:	Not Important	1	2	3	4	5	Important
(c) No funds available:	Not Important	1	2	3	4	5	Important
(d) No technical staff available:	Not Important	1	2	3	4	5	Important

(e) No outside researchers available:

Not Important	1	2	3	4	5	Important

Public access to microdata

8. Have any external researchers or research institutions requested access to your anonymized HIES microdata?

Yes: 4 *No 1*

9. If you answered "Yes" to Question 8, does your office keep a record of such requests and what your response was?

Yes: 3 *No 2*

10. Has your office approved any such requests?

Yes: 4 *No: 1*

11. If you answered "Yes" to Question 9 above, what conditions did your office impose?

MOU and PDH-ML conditions, Data Access Agreement, destruction of data after use, provide research products to NSO.

12. If you answered "No" to Question 9 above, why did your office refuse permission?

Not reputable. Unrelated to national needs.

13. If you answered "No" to Question 9, did you give that justification to the researcher?

Yes; several did not know.

Legislation

14. Can you please specify the exact legislation that regulates your Statistics Office?

All specified (widely varying years)

15. Does your Statistics Legislation allow your office to share anonymized microdata (such as HIES) with external researchers or research institutions?

(a) External researchers: *No: 4*

(b) Other Government Departments: *Yes: 3* *No: 2*

16. Do you see any technical reasons why your national HIES data records cannot be "anonymized"?

No: 5

17. If you feel that your Office does not have the technical capacity to anonymize your HIES data, would you like to receive such technical assistance from SPC's SDD?

Yes: 5

18. Does your Statistics Legislation have an "open data" philosophy that requires your office to share anonymized microdata (such as HIES) with external researchers or research institutions?

Yes: 1 *No: 4*

19. Would you like your Office to have an "open data" philosophy that legally requires your office to share anonymized microdata (such as HIES) with external researchers or research institutions upon certain conditions being met?

Yes: 4 *No: 1*

20. If you answered "Yes" to Question 18 above, what conditions would you like to impose on the researcher?

Data confidentiality
Use restricted to as applied
Submit reports to NSO
Acknowledge data sources
Politically neutral advice
Penalties if agreement breached.

21. If you answered "Yes" to Question 18 above, what possible benefits do you see for your country?

better development policy
monitoring development progress
public confidence in good use of data
improve research
build NSO credibility

22. If you answered "Yes" to Question 18 above, what possible costs or negative results do you fear?

breaching confidentiality
resulting in loss of public trust
resulting in legal actions
researchers uses data for other purpose not specified originally
commercialization of data
misuse of data

23. If you answered "No" to Question 18 above (i.e. you do not want to see a legislated "Open Data" philosophy for your Statistics Office, could you explain why?

N/A

SPC Datahub

24. What policy do you have in place for safely storing your microdata for posterity?

Only 1 had off-site storage
None

25. Has your Office stored any of your microdata with the SPC Datahub?

Yes: 3 *No: 1* *Don't know: 1*

26. If you answered "Yes" to Question 25 above, have you imposed any conditions for the use of that microdata by external researchers?

No: 1 *Not sure: 2*

27. Has your microdata stored with Datahub been used by external researchers?

Yes: 2 *Don't know: 3*

28. If you answered "Yes" have any of the results of the research been deposited with you?

No: 2 *Yes: 1* *Don't know:1*

29. What has been your experience so far with the microdata that you have deposited with the Datahub? (only 3 responses)

Not good	1	2	3	4	5	Good

30. If you circled "1" or "2" above, can you briefly explain why your experience has not been good?

N/A

31. If the SPC Data Hub wished to facilitate the production of more reports from your datasets, how much support would it receive from (circle the appropriate number).

(a) Your Office:	Not Important	1	2	3	4	5	Important
(b) Your government:	Not Important	1	2	3	4	5	Important
(c) Your public:	Not Important	1	2	3	4	5	Important

32 Are you aware of any legislation regarding your office that would prevent you from depositing your HIES dataset with the SPC Datahub? If yes, give details.

No: 3

33. If you answered "no" to question 32, would you like to see such legislation incorporated in your national statistics legislation?

Yes: 1 *No: 1* *Wait for legislative change*

34. If SPC was able to initiate a project to anonymize your Microdata and produce reports to be co-published by your office:

(a) how many technical staff could your office offer for this project?

(all positive)

(b) how many staff might other government departments (like planning) contribute?

(all positive)

35. If SPC were to suggest that your HIES microdata, after suitable anonymization, be made freely available for use by anyone anywhere in the world, what conditions would you like to see in place first?

First request permission from NSO and Statistical Advisory Committee
Report findings
Report findings before release approved
Use only for purposes specified
Destruction of data afterwards or return to source
Set penalties for breach of agreement
National capacity building
Input into final report

Most Important Challenges for Your Statistics Office

36. How important are the following challenges (in no order of importance) for your Statistics Office.

(a) Need funds for more surveys:	Not Important	1	2	3	4	**5**	Important
(b) Losing technical staff:	Not Important	1	2	3	4	**5**	Important
(c) No replacement staff available locally	Not Important	1	2	3	**4**	5	Important
(d) Inadequate salaries:	Not Important	1	2	3	**4**	5	Important
(e) Ministers not valuing office:	Not Important	1	2	**3**	4	5	Important
(f) Ministers restricting and/or regulating output of office:	Not Important	1	**2**	3	4	5	Important
(g) No demand from other departments for relevant statistics:	Not Important	1	2	**3**	4	5	Important
(h) No Statistics Advisory Committee:	Not Important	1	2	3	**4**	5	Important
(i) Lack of support from donors::	Not Important	1	2	**3**	4	5	Important
(j) Lack of support from regional organizations:	Not Important	1	2	**3**	4	5	Important
(k) Making microdata available to public researchers:	Not Important	1	2	3	**4**	5	Important

37. Are there any other challenges that you would like to add to the above list?

Lack of skilled people in data processing
lack of skilled people to do analysis
cyber security
lack of statistical leadership from top
lack of data forums
lack of outreach programs
linking data sources for "Big Picture".
SPC's role has been declining
Outdated legislation
Non-adherence to statistical protocols
minimum collaboration from donors

35 From the above list (and your additions), what do you consider the most important (in your order of importance): see attached worksheet.

Annex 12 Independent Academic Analyses of Fiji Microdata sets

This Annex of research publications by this author, illustrates the many "incentives" (the opposite of "blockers" for PICTs to free up access to their microdata to independent researchers who can work with NSOs and donors alike.

While focused largely on Fiji, they show that freeing up their microdata to independent analysts can give PICTs alternative interpretations and policy perspectives, from that of international multilateral organizations like WB. If these researchers are from a local academic institution, an additional benefit can be that they can be held accountable to local stakeholders for their analysis and findings, unlike the authors from the multilateral organizations who produce the reports and are difficult afterwards to have dialogue with over the methodology and results.[49]

Most of the Reports discussed here were focused on poverty and food security policy issues (using HIES) and gender (using EUS datasets), precisely the areas specified as important in the Terms of Reference for this Report.

This Annex also brings out the crucial importance of microdata preservation so that advances or changes in methodology can be applied to historical microdata to produce time trend analysis so important to assess whether public policies are working or not, and the rates of progress.

The PSMB may wish to facilitate a project with willing PICTs to conduct similar analyses in a large project co-ordinated perhaps by SDD and PDH-ML (as recommended in this Report). The SPC/FAO project with Vanuatu HIES is clearly along the same lines.

All publications mentioned here have been deposited with the PDH-ML and may be freely accessed. The author hopes that even if put into an Annex, it is important to put on record the views of one of the leading Pacific Island analysts of Pacific microdata.

Policy discussion workshops: closing the virtuous statistical cycle

One of the very positive developments following the analysis of 2002-03 HIES and the 2008-09 HIES were extremely useful national workshops launched by the donors assisting in the studies. These workshops were held not just in Suva but also in Nadi and Labasa.

They involved the full range of contributions and policy discussions by government departments and NGOs. An important result of these workshops was that the public (including households who took part in the surveys), were able to see the useful end results for policy application. This therefore provided strong goodwill for the FBS for their future household surveys and built "trust" considered so important by many international experts.

Fiji 1990-91 HIES

The FBS did a HIES in 1990-91, deemed by the FIBoS to be unreliable because of the political climate after the 1987 coup, and no official report was ever published. But in 1997, there was a UNDP Study which commissioned analyses of the 1990-91 HIES data by an external consultant Dennis Ahlburg who had the data "statistically adjusted" producing two reports

[49] This is not a trivial point as the recent controversy over the WB analysis of the Fiji 2020-21 HIES shows, with an initial report changing the methodology, reducing the Basic Needs Poverty Line, three months later changing the results again because they had apparently found errors in their initial calculations. Of course, they are not present to answer to very legitimate queries from local stakeholders.

(1995 and 1996). [50] The UNDP Study's central conclusion that the incidence of poverty was 25% was likely flawed, even if used uncritically[51] for decades afterwards to argue that poverty was rising compared to later estimates of poverty.[52]

If this 1991 HIES micro-data and all metadata had been preserved then today's analysts could use more current methodology to compare current rates of poverty and other development indicators such as food security, with that of the 1990-91 microdata using the later methodology.[53]

Fiji 2002-03 HIES

The 2002-03 HIES was believed by analysts to be solid and the original data has been preserved. There were several reports published out of this 2002-03 HIES, by the author together with FBS staff[54] having great national significance for political dialogue. Fiji has suffered from political strife (four political coups) with ethnicity an important factor in the political discourse on which groups need government support. This author presented his findings to the Fiji Cabinet which showed that going by the incidence of poverty, the most poor and not to be neglected for poverty alleviation were those of Indo-Fijian origin. However, the sharing guidelines for the *quantity* of poverty alleviation resources (based on the numbers of poor and their poverty gap), would fairly allocate more resources to indigenous Fijians than to other ethnic groups. This finding was clearly fair to all ethnic groups and may have defused political tension.

The resulting paper was published by ANU's *Pacific Economic Bulletin*.[55] The author later co-published a comprehensive academic monograph[56] with the Fiji Bureau of Statistics and School of Economics (USP), launched by the Fiji Government Statistician.

There are other PICTs (like Solomon Islands) where political tensions arising out of regional allocations of taxpayer funds can be reduced by solid poverty data derived from HIES.

Fiji 2008-09 HIES

The FBS conducted a solid 2008-09 HIES which ought to be a fascinating study for the PSMB and PICT NSOs, because this HIES microdata set was also analysed by a powerful WB team resulting in a WB publication. The FBS published a preliminary report.[57] This HIES resulted in a comprehensive FBS publication financed by AusAID, FBS (2012) *Poverty in Fiji: Changes 2002-03 to 2008-09*.

The Reports outlined changing household conditions between 2002-03 and 2008-09 in the key areas of poverty and income distribution, food security, health and NCDs risks

[50] There appears to be no internal FIBoS record of the statistical adjustments made nor the final data set used.
[51] It is natural that reports of international organizations are accepted more than that of independent researchers..
[52] Close reading of the original Ahlburg studies (in the author's possession) suggest that the incidence of poverty was at least 29% and possibly as high as 33%
[53] These issues are discussed in Narsey (2008) *The quantitative analysis of poverty*.
[54] Narsey, Wadan (2006) Report on the 2002-03 Household Income and Expenditure Survey. FIBOS.
[55] Narsey (2008) "The incidence of poverty and the poverty gap in Fiji: unpalatable facts for ethno-centric political parties". *Pacific Economic Bulletin.*
[56] Narsey (2008) The Quantitative Analysis of Poverty in Fiji. USP and FIBOS.
[57] Fiji Bureau of Statistics (2010) *Preliminary Report: Poverty and Household Incomes in Fiji in 2008-09*. FIBOS (Wadan Narsey, Toga Raikoti and Epeli Waqavonovono). Suva, Fiji. pp. 25. (Funded by AusAID).)

from narcotics consumption, private household contributions to education and health, and household assets and essential services.

Some *adverse developments* in food security over just this five year period were: reducing share of home production in food consumption; reducing share of local carbohydrates and increase in imported rice and noodles; reducing share of local marine foods; increase in sugar consumption for poorer urban families; complex changes in expenditure on junk food and narcotics all urgently needing further research. There was also evidence of a few positive developments, indicating the success of some government and NGO campaigns.

FBS (2012) has an Annex C of some seventy-three policy recommendations derived from the evidence in this study. FBS (2012) also has an Annex B which gives a comparison (and contrast) of the methodology and results, with the WB analysis of poverty in Fiji using the same 2002-03 and 2008-09 HIES data.[58] There is also an academic paper published on the important differences.[59]

Declining marine food consumption

One area on which all PICT HIES microdata can inform policy makers is in the consumption of marine foods. This author's analysis of the Fiji 2002-03 and 2008-09 HIES, published with a Japanese academic journal[60] found that just over this short five year period, local marine food consumption had declined as a proportion of total food consumption, while that of imported canned fish had risen. The proportion of population consuming local marine foods was also significantly declining; these trends were visible in both rural and urban areas. There was clear evidence of declining food security for Fiji.

Analysis of later HIES could show whether these trends have continued to current times. All PICTs policy makers would be interested in such analysis of their own HIES given that marine food consumption is an integral part of their diets.[61]

Fiji EUS 2004-05, 2010-11 and 2015-16

Given that the contract specified the importance of gender statistics to the PACSTAT Project, it is useful to examine the outputs possible from Employment and Unemployment Surveys three of which (2004-05, 2010-11 and 2015-16) which I have analysed for Fiji for the FBS, donors and NGOs.

The EUS collects household level and individual level data on incomes, employment, underemployment and unemployment, paid and unpaid work, and (for the later EUS) leisure activities (including sports, kava, religious activities, TV/videos). The EUS is therefore capable of showing a fascinating range of gender inequalities in the economy and society, and also by all the other variable such as age, education, marital status, location.

---

[58] World Bank (2011) *Poverty Trends, Profiles and Small Area Estimation (Poverty Maps) in Republic of Fiji (2003-2009).* Social Protection Unit, Human Development Group, East Asia and the Pacific Region, WB.

[59] Narsey (2015) "Poverty analysis for Fiji: the divide between World Bank theorists and World Bank practitioners in the field" *Journal of Pacific Studies.* Vol.35 No.3 pp. 108-1 26.

[60] (2012) "The Regression of Marine Foods Consumption in Fiji: changes 2002-03 to 2008-09". *South Pacific Studies*, Vol.32, No.2. 2012.

[61] Changes in the consumption of fatty imported meats also has serious consequences for NCDs.

The policy implications of the analyses of these EUS microdata may be seen in the following publications: FBS/Narsey (May 2007); Narsey (Nov. 2007); Narsey (2014) Unpub.); Narsey (2020); Narsey (2022) (awaiting publication by FWRM). [62]

The findings of all these reports are significant for policy dialogue for Fiji and global stakeholders in gender issues: the significant gender gaps in incomes, under-employment and unemployment, unpaid household work; total work; and leisure time activities.

Fiji 2013-14 HIES

The 2013-14 HIES was conducted and analyzed by FBS staff with preliminary findings released.[63] Poverty statistics were calculated with the 2008-09 BNPL adjusted by the change in the CPI so results could be compared with previous results for 2008-09 and 2002-03. There were no details given on the methodologies used and outside researchers were not given access to the data.

Fiji 2019-20 HIES

There was a 2019-20 HIES with a Report published with technical support from the World Bank and the University of Bristol (UK) focusing on poverty and relative deprivation. The Report stated that comparisons could not be made with previous estimates of poverty because of the changes in methodology (using expenditure/consumption as criterion instead of income, and significantly reduced values for FPL and BNPL.[64] The poverty results were changed a few months later because of errors discovered in the data modifications.

An innovation with Solomon Islands and Vanuatu HIES

The extremely conservative values used by international analysts while appropriate in poor developing countries like India or Bangladesh, are considered by local researchers and stakeholders to be insufficient for even the poorest of PICT peoples. How does one empower local stakeholders to use their own values for Basic Needs Poverty Line without having to go through the enormous work of analysing primary HIES data?

The author used the Solomon Is. HIES 2004-05 to publish an academic paper critiquing the methodology of a previous analysis done by the UNDP while presenting alternative methodology, analyses and results.[65] One innovation was data presented in a convenient tabular form, so that stakeholders could easily "read" the incidence of poverty in rural and urban Solomon Islands, using their own choice of values for the Food Poverty Line (FPL) and Basic Needs Poverty Line (BNPL).

The author also used the Vanuatu 2006 HIES microdata to publish an academic paper presenting alternative analyses to that used by the official report which was facilitated by

[62] This Report is the basis of a FWRM publication to be launched on International Women's Day on 8 March 2022.

[63] FIBOS Release No. 98, 2015.

[64] It may be an interesting exercise to see the extent to which new methodology can be applied to the microdata from the older HIES to ensure consistency..

[65] (2011) "The incidence of poverty in Solomon Islands: the importance of methodology". *Journal of Pacific Studies*. Vol.31, No.1, pp.31-58.

ADB.[66] This paper not only gave the alternative results (which were significantly different from that of the official report) but presented data in a convenient tabular form so that stakeholders could easily "read" the incidence of poverty in rural and urban Vanuatu using their own choice of values for the FPL and BNPL. Such a table[67] could have been extremely useful for the Fiji 2020-21 HIES if local stakeholders wished to use alternative (higher) values for the BNPL in order to compare with the estimates for previous years.

The FAO/SPC Initiative

An excellent example of what can be done in a concerted fashion is the recent initiative using the Vanuatu HIES where SPC-SDD staff provided technical support, staff from other government departments provided general support, and the data analysis and report writing was assisted by staff from multilateral organizations like WB, FAO, ILO and ACIAR. The outputs were an Income and Expenditure Report, a Hardship and Poverty Report, Food Security Report, a Labour Market Monograph, and probably a first for the Pacific, a Well-being Report.

The SPC/FAO initiative on food security was also extended to Solomon Islands, Marshall Islands and Kiribati. From their last HIES, the Kiribati NSO produced several reports: HIES Report, Food Security, Food consumption Report, Poverty Analysis, Copra Brief and a Labour monograph

This initiative can be emulated by individual PICTs but more easily through a project led by the SPC/SDD, across all willing PICTs, as recommended in this Report.

[66] Narsey (2012) "UNDP/ADB Poverty Results for Vanuatu: a critical review and alternative assessment". *South Pacific Studies*, Vol.33., No.1, 2012. Kagoshima University Research Center for Pacific Islands.
[67] There would need to be two tables, one using income and one using consumption expenditure.

References

ABS (2015) Public Data Policy Statement. Canberra

Cook, Len (2003) "Building trust in official statistics through microdata access". Office for National Statistics (UK) 30 June 2003.

Dupriez, Olivier and Ernie Boyko (2010) "Dissemination of Microdata Files. Formulating Policies and Procedures". IHSN Working Paper 005.

Fiji Bureau of Statistics and Wadan Narsey (2006) *Report on the 2002-03 Household Income and Expenditure Survey*. (funded by FIBoS).

Fiji Bureau of Statistics and Wadan Narsey (May 2007) *Report on the 2004-05 Employment and Unemployment Survey*. Fiji Islands Bureau of Statistics (funded by FIBoS)

Narsey, Wadan (November 2007) *Gender Issues in Employment, Under-Employment and Incomes in Fiji*. Fiji Islands Bureau of Statistics and Vanuavou Publications (November 2007). (Funded by AusAID).

Narsey (2008) "The incidence of poverty and the poverty gap in Fiji: unpalatable facts for ethno-centric political parties". *Pacific Economic Bulletin*.

Narsey, Wadan (2008) T*he Quantitative Analysis of Poverty in Fiji*. Fiji Bureau of Statistics and School of Economics, FBE USP and Vanuavou Publications.

Fiji Bureau of Statistics (2010) *Preliminary Report: Poverty and Household Incomes in Fiji in 2008-09*. Fiji Islands Bureau of Statistics (authored by Wadan Narsey, Toga Raikoti and Epeli Waqavonovono). Suva, Fiji. pp. 25. (Funded by AusAID).

Fiji Bureau of Statistics (2011) *Report on the 2008-09 Household Income and Expenditure Survey*. authored by Dr Wadan Narsey. Suva, Fiji. pp. 96. Funded by AusAID.

Fiji Bureau of Statistics (2012) *Poverty in Fiji: changes 2002-03 to 2008-09 and Policy Implications*. Authored by Dr Wadan Narsey. Suva, Fiji. 2012. (funded by AusAID).

Narsey, Wadan (2014) *Fiji Women and Men at work and leisure: changes 2004-05 to 2010-11* (unpublished report for Fiji Bureau of Statistics and AusAID).

Inter-university Consortium for Political and Social Research (ICPSR). 2009. "Principles and Good Practice for Preserving Data", International Household Survey Network, IHSN Working Paper No 003, December 2009.

IHSN (2019) Quick Reference Guide for Data Archivists. Olivier Dupriez, Diana Marcela Sanchez Castro, Mathew Welch (The World Bank).

ILO (2020) *Making Women's Work Visible*: The 19th ICLS Standards, Purpose and Progress.

Matthias Templ, Bernhard Meindl, Alexander Kowarik, Shuang Chen. 2014. *Introduction to Statistical Disclosure Control (SDC).* International Household Survey Network..

Narsey, Wadan (2022) *The Economic Empowerment of Fiji Women and Girls: beyond 33%*. Report for Fiji Women's Rights Movement (in process).

Narsey, Wadan (2020) *The Economic Vulnerability of Fiji Women to COVID-19*. Unpublished Consultancy Report for Pacific Women. Suva. August 2020.

Narsey, Wadan (2006) *Just Wages for Fiji: lifting workers out of poverty*. ECREA and Vanuavou Publications. (a publication used by the Fiji Government for its Wages Councils to determine sectoral wages).

Open Data Watch and FAO (May 2020) "Taskforce for household survey microdata dissemination".

Swanson, Eric Jamison Crowell, Taylor Hadnot and Elettra Baldi (2021) "Access to microdata in national and international repositories" Swanson, Eric, Jamison Crowell, Taylor Hadnot, Elettra Baldi. Draft Paper. ODW.

SPC (2022?) "Protocols/procedures for licensed data access requests to the Pacific Data Hub – Microdata Library" (Draft under review)

SPC "Review of obligations and condition in the legislation of Pacific Countries that enable release of microdata to third parties through the SPC Pacific Data Library" (Draft).

SPC (2021) "Options for change in Statistics Acts in the Pacific" (Draft Paper by SDD).

Tuvalu Government (2006) Household Income and Expenditure Survey 2004/2005 Final Report (with contributions by SPC consultants).

World Bank (2014) "Microdata dissemination best practices". 3 March 2014. Note presented at 22nd CCSA session (Ankara 2013) and revised with inputs from Eurostat, FAO, UNECE and UNICEF.