Public Disclosure Authorized

Public Disclosure Authorized

Policy Research Working Paper 9179

New Set of Indicators to Measure Population Remoteness and Dispersion

Estimates for 100 Countries, with Detailed Analysis of Pacific Island Countries

Imogen Halstead
Robert Utz
Boer Xia

WORLD BANK GROUP
Macroeconomics, Trade and Investment Global Practice
March 2020

Policy Research Working Paper 9179

Abstract

This paper defines three new indicators that capture the remoteness of local communities within a country context and the overall scale of population dispersion and settlement sparsity across a country. The paper also exploits the World Bank's subnational geography database to estimate the geographical scale of population dispersion and dispersion-adjusted population sparsity for 100 countries around the world. The new indicators are evaluated and explored for several Pacific Island countries, which are often characterized as being remote (in a global context) and highly dispersed. However, within each Pacific Island country, there is enormous variation in the remoteness of individual communities and the extent to which communities are clustered or dispersed from one another, and these conditions can be related to communities' socioeconomic characteristics. The results reflect this. The paper empirically contextualizes the settlement patterns evident in the Pacific Island countries within a broader global context, highlighting the extreme degree of population sparsity in the Pacific, relative to all the other countries that are assessed.

This paper is a product of the Macroeconomics, Trade and Investment Global Practice. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://www.worldbank.org/prwp. The authors may be contacted at rutz1@worldbank.org.

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

Produced by the Research Support Team

New Set of Indicators to Measure Population Remoteness and Dispersion: Estimates for 100 Countries, with Detailed Analysis of Pacific Island Countries[1]

Imogen Halstead
Robert Utz
Boer Xia

JEL Classification: R1, R12, R23

Keywords: Population dispersion, remoteness, Pacific Island Countries

[1] This paper was prepared as part of the "Spatial Development Issues of Pacific Island Countries" project.

Introduction

Pacific Island Countries (PICs) are often characterized as being both remote (in a global context) and highly dispersed, but within each country there is enormous variation in the remoteness of individual communities, and the extent to which communities are clustered or dispersed from one another. This paper seeks to better understand, and to empirically measure, the nature of dispersion and remoteness within individual PICs.

This paper defines a new measure of local-area population remoteness which is highly sensitive to variation in settlement patterns. Remoteness of a community is calculated as the population-weighted average distance between the community of interest, and all communities across the country. It is designed to be evaluated using commonly available population census data. A community is measured to be more remote if it is home to a smaller share of the population, relative to other communities further afield. It is also sensitive to the distance between the community of interest and any other community (in particular, a community is measured to be more remote, the more distant it is from large population centers). The remoteness measure for a particular community is unaffected by the overall population level of the country as a whole.

Our new measures of community-level remoteness can be aggregated to derive country-level indicators of (i) the geographical scale of population dispersion, and (ii) a dispersion-adjusted measure of population density, or inversely, sparsity. The geographical scale of dispersion is defined as the population-weighted average remoteness. This indicator is again highly sensitive to variation in settlement patterns (unlike more macro measures of geographical dispersion, such as the size of a country's Exclusive Economic Zone). The dispersion indicator will be higher when a larger share of the population live in remote communities. Like the community-level remoteness indicator, the dispersion indicator is population-neutral, and unaffected by the overall level of a country's population. The dispersion-adjusted measure of population density is defined as being proportional to the country's population divided by the square of the dispersion indicator (with the square operation applied to transform the distance-based dispersion measure to an area-equivalent). It's inverse captures population sparsity. This indicator is again highly sensitive to population settlement patterns (in contrast to calculations of the country's average population density).

In this paper, we calculate the local-area remoteness measures, and country-level dispersion and sparsity indicators for eight case-study PICs, relying on census data collected by country National Statistics Offices and made available through the Pacific Community (SPC) PopGIS platform. Our estimates of community remoteness are correlated with local area socio-economic characteristics (the link between remoteness and socio-economic characteristics is explored more comprehensively in a companion piece[2] to this work). At the country level, our indicators emphasize the enormity of the geographical scale of dispersion in Kiribati and the Federated States of Micronesia (FSM); residents live an average of around 650-700km from the rest of the population. The populations of Solomon Islands, Vanuatu and Republic of Marshall Islands (RMI) and are less widely dispersed, with people living an average of around 250km from the rest of the population. The populations of Fiji, Tonga and Tuvalu are more concentrated still, with people living on average around 100-150km from the rest of the population. This dispersion indicator is a population-

[2] World Bank, 2019. 'Socio-economic analysis: A spatial analysis of population and housing census data', a background paper prepared for *A Review of Spatial Development Issues in the Pacific Islands Countries* (P166346).

neutral measure of geographical distance, and obviously the lived experience of dispersion (or remoteness) is likely to be more impactful in Tuvalu than Fiji, given is vastly smaller population. Our dispersion-adjusted measures of population sparsity reflect this perspective. FSM and Kiribati stand out as being extremely sparsely distributed (with the sparsity of settlement patterns equivalent to 1 person per 20 km^2), as is Tuvalu (1 person per 5 km^2). Sparsity for the populations of Tonga, Solomon Islands and Vanuatu is somewhat more moderate (around 1-2 people per km^2), while the population of Fiji is by far the most highly concentrated across geographical space (16 people per km^2).

We also exploit the World Bank's subnational geography database to estimate the geographical scale of population dispersion and dispersion-adjusted population sparsity for 110 other countries around the world. This exercise highlights the extreme degree of population sparsity in the PICs, relative to all other countries assessed. Across the global database, the geographical scale of population dispersion ranges from a maximum of 3,600km in the Russian Federation, to a minimum of 40km in Trinidad and Tobago. None of the PICs lie at the extremes of this range, but dispersion for FSM and Kiribati is at the higher end of the distribution, while Fiji and Tonga are at the lower end. On the other hand, after accounting for population size, the PICs stand out as having extremely low levels of dispersion-adjusted population density compared to other countries, and equivalently, an extremely high degree of population sparsity**.**

This note proceeds as follows. In the first section we define our three key indicators which reflect, respectively: local-area community-level remoteness; the geographical scale of population dispersion across a country; and, country-wide, dispersion-adjusted population density, or inversely sparsity. The second section presents empirical estimates for each of the PICs and compares the results of our new measures to other spatial geography-related indicators. The third section contextualizes the results for the PICs in a broader global context, capitalizing on data available for countries included in the World Bank's sub-national geographic database. A final section describes some extensions to our benchmark work and concludes.

1. Indicator definitions

Population settlement patterns vary considerably across countries, and sometimes in nuanced ways which are difficult to capture in a simple, empirical indicator. This section presents three related indicators which reflect variation in population settlement patterns: (i) an indicator of community remoteness within a country context (a local-area-level indicator), (ii) an indicator of the geographical scale of population dispersion (a country-level indicator), and (iii) an indicator of dispersion-adjusted population density, or inversely, sparsity (a country-level indicator).

These indicators are designed with practical evaluation in mind. They can be calculated using commonly available geo-coded population census data, disaggregated at a local-area level; e.g. by enumeration area, or jurisdictional district areas. In theory, all three of our indicators can be rooted in a conceptualization of remoteness at the individual level and evaluating the indicators at the individual-level would yield the most accurate picture of the depth and severity of remoteness of settlements across the country. However, census data are rarely made available geo-coded at the individual level (i.e. indicating where each individual in the country lives) for obvious reasons. As such, we focus the definitions and discussion below on indicator definitions at the more applicable, community level (which effectively assumes that all individuals in a local area are located at the centroid of that area). For readability, we use the term ‘community’ to refer to local areas (though in practice, there are likely be communities which span data-defined areas, and areas which span communities).

Remoteness of a given community is defined as the population-weighted average distance between that community and all communities across the country. This is an adaptation of Horscroft’s (2014)[3] country-level measure of economic remoteness, which is calculated as the GDP-weighted average distance between a country of interest and all other countries in the world. In our context, we define the remoteness of a community i within a country with J communities (‘country J’, for simplicity), as the average distance of that community to all communities in the country, weighted by those communities’ respective population size:

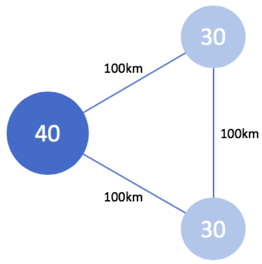
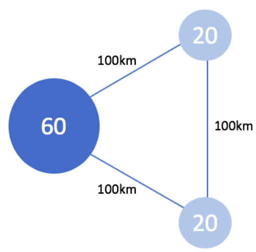
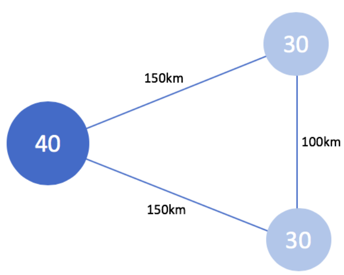
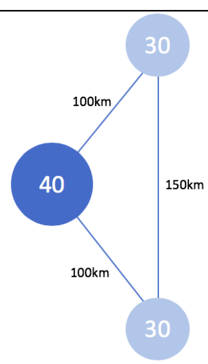
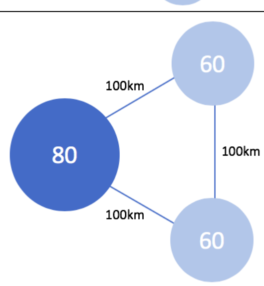
$$remoteness_i = \frac{1}{N}\sum_{j=1}^{J}\left(N_j \times distance_{(ij)}\right)$$

where N is the population of the country, $\sum_{j\epsilon J} N_j = N$, and $distance_{(ij)} = distance_{(ji)}$. Note that a district itself is included in its own the average (with distance zero), at $i = j$. The remoteness of area i will be increasing in its distance from other communities j, particularly those with large populations, and decreasing in the share of the population living in area i itself (in the formula, this increases the weight on the zero own-distance). The remoteness measure is unaffected by the total population for the country as a whole.

Table 1 provides a few illustrative examples. The nodes indicate communities, and they are annotated with the communities’ populations. Consider the remoteness of the dark blue node. In panel A, the remoteness indicator evaluates to 60km; on average, each person in the dark blue node is 60km away from every other person in the population. In panel B, the population is more heavily concentrated on the dark blue node (though the overall population is the same. In this context, the dark blue node is considered less remote; its remoteness indicator evaluates to 40km. In panel C, the dark blue node is more distant from the other light blue communities relative to panel A. Here, the dark blue node is considered more remote; its remoteness indicator evaluates to 120km. In panel D, the distance between the dark blue node and each of the two light blue nodes is unchanged from panel A, but the two light blue nodes themselves are further apart from one another. In this context, the remoteness of the dark blue node is unchanged from panel A, at 60km. In panel E, the population of each of the communities is doubled (though naturally their shares remain the same) relative to panel A; in this context again, remoteness of the dark blue node is unchanged at 60km.

[3] Horscroft, Virginia. 2014. *Public sectors in the pacific islands : are they 'too big' and do they 'crowd out' the private sector ? (English)*. Policy Research working paper ; no. WPS 7102. Washington, DC: World Bank Group. http://documents.worldbank.org/curated/en/986481468098052675/Public-sectors-in-the-pacific-islands-are-they-too-big-and-do-they-crowd-out-the-private-sector

Table 1: Remoteness, dispersion and sparsity calculations: A few examples

	Settlement pattern	Local-area level remoteness		Geographical scale of dispersion	Dispersion-adjusted density
	(Community-level nodes annotated with populations)	Dark blue node	Light blue nodes		
A		60km =(40*0 + 30*100 + 30*100) /(40+30+30)	70km =(30*0 +40*100 +30*100) /(30+40+30)	66km =((40*60) +2*(30*70)) /(40+30+30)	0.006 people/km^2 =K*(100)/66^2
B		40km =(60*0 + 20*100 + 20*100) /(60+20+20)	80km =(20*0 + 60*100 +20*100) /(20+60+20)	56km =((60*40) +2*(20*80)) /(60+20+20)	0.008 people/km^2 =K*(100)/56^2
C		90km =(40*0 + 30*150 + 30*150) /(40+30+30)	90km =(30*0 + 40*150 + 30*100) /(30+40+30)	90km =((40*90) +2*(30*90)) /(40+30+30)	0.003 people/km^2 =K*(100)/114^2
D		60km =(40*0 + 30*100 + 30*100) /(40+30+30)	85km =(30*0 + 40*100 + 30*150) /(30+40+30)	75km =((40*60) +2*(30*85)) /(40+30+30)	0.005 people/km^2 =K*(100)/75^2
E		60km =(80*0 +60*100 +60*100) /(80+60+60)	70km =(60*0 + 80*100 + 60*100) /(60+80+60)	66km =((80*60) +2*(60*70)) /(80+60+60)	0.012 people/km^2 =K*(200) /66^2

The geographical scale of population dispersion for a country is defined as the population-weighted average remoteness level. Intuitively, this measure captures the average distance between each two people in the population. For country *J*:

$$dispersion_J = \frac{1}{N} \sum_{i=1}^{J} (N_i \times remoteness_i) = \frac{1}{N^2} \sum_{i=1}^{J} \sum_{j=1}^{J} \left(N_i \times N_j \times distance_{(i\ to\ j)}\right)$$

Country-level dispersion is increasing in the distance between any two communities. It is particularly sensitive to the remoteness of (or distance between) communities with large shares of the population. Dispersion is also increasing in the share of the population living in more remote communities. The measure is population-neutral overall (if you were to double the population at each location throughout a country, both the local-area remoteness measures and the overall dispersion measure would remain unchanged).[4] The geographical scale of dispersion is expressed in units of distance.

Consider again the examples illustrated in Table 1. In the benchmark panel A, the dark blue node of population 40 has a remoteness of 60km, and the two light blue nodes each of population 30 have remoteness of 70km. The population-weighted average remoteness is therefore 66km. This is the geographical scale of population dispersion for the country as a whole. The population is more concentrated in panel B, where a greater share of the population reside in the largest community, with a dispersion indicator of 56km. Meanwhile, the scale of dispersion is greater in Panel C, where the largest community is more distant from the rest of the population, at 140km. In Panel D, the distance between the largest community and the rest of the population is unchanged, but the smaller communities are further from one another. Similar to panel C, the indicator shows greater population dispersion in panel D than panel A, however to a lesser extent.

Dispersion-adjusted population density is defined as being proportional to a country's population divided by the geographical scale of dispersion squared. This indicator is in effect, a population-weighted average measure of population density across a country, taking settlement patterns into account in their totality. Ordinarily, density is calculated as the population divided by land (or territory) area. In order to derive a dispersion-adjusted population density measure, instead of using the country's land/territory area in the denominator of the calculation, we use a dispersion-adjusted area-equivalent. To derive a dispersion-adjusted area-equivalent, we abstract away from the actual nuanced pattern of population settlement across a country (for which density calculations will vary at different points across the terrain, and for different definitional scopes), and concentrate solely on our summary indicator of the geographical scale of dispersion. We then derive the size of area that would be required to achieve this same geographical scale of dispersion, if the population were distributed evenly over a circular area. In this hypothetical situation, the formula for the geographical scale of dispersion would be: $(128/45\pi)r$ (where r is the radius of the circle); i.e. the average distance between any two points on a disc.[5] It follows that to achieve dispersion equivalent to the realized dispersion, $dispersion_J$, for a country J, the radius of the circular

[4] The implications of applying these calculations at different levels of geographical disaggregation are not immediately clear (it may be a case-specific empirical question). At one extreme, with only one geographical 'district' area, the entire population of a country would be at zero distance from itself; there would be zero estimated dispersion as a result. We did experiment with adding a 'radius'-equivalent additive scaling factor to circumvent this issue with the 'own distance' calculations, and we consider this a promising direction for future work to refine these measures.

[5] We are grateful to James Parkinson for this insight. See Appendix D for details.

area would need to be equal to $r_J = (45/128) * \pi * dispersion_J$, and the area of this circle will be $\pi {r_J}^2$. The dispersion-adjusted density measure will be:

$$\widehat{density}_J = \frac{N}{\pi\left(\frac{45}{128} * \pi * dispersion_J\right)^2} = k * \frac{N}{{dispersion_J}^2}$$

where the constant of proportionality $k = (128/45)^2(1/\pi^3) \approx 0.261$.

Since sparsity, rather than density, is the focus of the current body of work relating to the PICs, we focus on analyzing the inverse of this density measure: a dispersion-adjusted measure of population sparsity. It is defined as:

$$\widehat{sparsity}_J = \frac{\pi\left(\frac{45}{128} * \pi * dispersion_J\right)^2}{N} = \left(\frac{1}{k}\right) * \frac{{dispersion_J}^2}{N}$$

The examples in Table 1 demonstrate ways in which the value of this indicator varies in different empirical settlings. Dispersion-adjusted density is higher in panel B than panel A, reflecting the greater concentration of the population in the most highly populated (dark blue) node. In contrast, a straightforward density estimate would not differentiate between these two settings (since their overall population, and overall area of terrain, is the same). Consider now Panel A and Panel E. The pattern of population distribution is the same across these two settings, and this is reflected in their common geographical scale of dispersion. However, dispersion-adjusted density is twice as high in panel E as in panel A, reflecting panel E's larger population.

2. Evaluation for the PICs

In evaluating our indicators of local-area level remoteness, the geographical scale of dispersion, and dispersion-adjusted density for the PICs, we rely upon data from population and housing censuses, collected by country National Statistics Offices and made available through the Pacific Community (SPC) PopGIS platform (see Appendix A for detail). We access data for Fiji (2007,[6] by Tikina), Solomon Islands (2009, by Constituency), Vanuatu (2009, by Area Council), Kiribati (2010, by island/atoll grouping), FSM (2010, by Municipality), Tonga (2016, by District), RMI (2011, by atoll grouping) and Tuvalu (2012, by village). Note that some of these census data are quite dated, so the results (which reflect the distribution of the population at a particular point in time) presented in this paper should be interpreted accordingly.

The distribution of local-area remoteness measures, and the country-level geographical scale of dispersion are illustrated for each of our eight PIC case-study countries in Figure 1. Distances are indicated on the horizontal axis, and the population share is indicated on the vertical axis. Columns shows the share of the population living in local areas with remoteness measures within a particular distance interval. The geographical scale of dispersion—the population-weighted average remoteness—is marked with a vertical red line.

[6] Note that a more recent census was administered in Fiji in 2017, but the associated data were not available through PopGIS at the time this paper was prepared in March 2019. See Appendix A for more detail on the source data.

There is notable variation in the range of local-area remoteness levels, and the mean (the geographical scale of dispersion), but for all the countries, the distribution of local-area remoteness levels is skewed to the right. That is, populations are more heavily concentrated in areas with low levels of remoteness, with small shares of the population living in more remote parts of the country. This is not a feature of the structure of the remoteness index itself;[7] it reflects people's natural tendency to settle close to one another.

The geographical scale of dispersion is enormous in both FSM and Kiribati, with population-weighted average local-area remoteness levels of 714km and 651km, respectively. All local-areas across FSM have relatively high remoteness levels. This reflects the sizeable shares of the population in all four states, which themselves are very distant from one another. In contrast, in Kiribati, there is a concentration of the population on Tarawa, with somewhat lower remoteness levels (the minimum local-area remoteness is 374km, for South Tarawa). However, there are also several settlements on extremely remote islands; people living on Kiritimati (more than 5 percent of the population) are on average 2,974 km from other i-Kiribati.

The scale of population dispersion is more moderate for Solomon Islands (267km), RMI (258km) and Vanuatu (221km). In Solomon Islands, remoteness levels are lowest for people living in the vicinity of central Honiara at around 180km, and highest for people living in Temotu Vatu, at 813km. In Vanuatu, Port Vila has a similar remoteness to Honiara, around 180km, and Luganville has a remoteness of 208km (note that the two centers themselves are around 290km apart). Meanwhile, people living on the most remote islands, the Torres islands in Torba, are on average 445km away from each other, in ni-Vanuatu. In RMI, around half the population live in Majuro, which has a similar remoteness level to Honiara and Port Vila (182km). The country's secondary hub at Kwajalein has a remoteness level of 338km. The 664 people living in the remote settlement at Enewetak atoll are on average 943km away from other Marshallese people.

The geographical scale of population dispersion is smaller still in Tuvalu (153km), and smallest in Fiji (116km) and Tonga (113km). The population of Tuvalu is highly concentrated on Funafuti; the district of Senala is the most central, with a remoteness level of 103km, but in fact more than three-quarters of the population live in areas with remoteness levels between 100-150km. The remainder of the population are scattered across somewhat more remote islands, between 200-400km from the other Tuvaluans on average. In Fiji, only very small shares of the population live in remote areas (compared to the other case study PICs); 99 percent of the population live in areas with a remoteness measure of less than 250km (more than half in areas with remoteness levels below 100km). Suva has a remoteness indicator of 85km; Nadi, 146km; and Labasa, 196km. The most remote population in Fiji lives on Rotama Island, around 590km from other Fijians on average. In Tonga, the majority of the population (around three-quarters) live on the relatively small island of Tongatapu, and they are on average 70-80km from other Tongans. The remoteness indicator for other settlements increases with distance from Tongatapu. The small settlement at 'Eua has a remoteness level of around 95km; on Ha'apai, remoteness levels range from 130-

[7] To see this, consider a settlement pattern where one million people are disbursed around the circumference of a circle, and one person is in the center. The remoteness indicator for the person in the center will be close to r, the radius of the circle, while the remoteness indicator for the people on the circumference will be close to $4r/\pi \approx 1.27r$ (see Appendix 4). In this case, the distribution of remoteness levels will be skewed to the left; a greater share of the population are living in more remote areas. Of course, this settlement pattern is not common empirically; it is much more natural for communities to evolve with a density of people settling in central locations.

180km; on Vava'u they range from 260-280km; and on the distant Niuas, remoteness levels are approaching 600km.

Figure 2 presents estimates of population sparsity for the eight case study PICs. These estimates reflect both population size and the geographical scale of dispersion and provide an indicator of the lived experience of remoteness in the context of the geographical spread of a country's population. The population is extremely sparsely settled in FSM (at a scale of 19.0km^2 per person), and to a lesser extent Kiribati (with 15.8km^2 per person). Sparsity is also notable in Tuvalu (8.4 km^2 per person), due to its small population (its geographical scale of dispersion is relatively moderate), and in RMI (4.8km^2 per person). The populations of Vanuatu (0.8 km^2 per person), Solomon Islands (0.5 km^2 per person), and Tonga (0.5 km^2 per person) are much less sparsely dispersed. Lastly, given its relatively large population size, coupled with the relatively modest geographical scale of dispersion, the Fijian population is more concentrated (with a sparsity measure of less than 0.1 km^2 per person).

Figure 1: Distribution of local-area remoteness and the geographical scale of dispersion

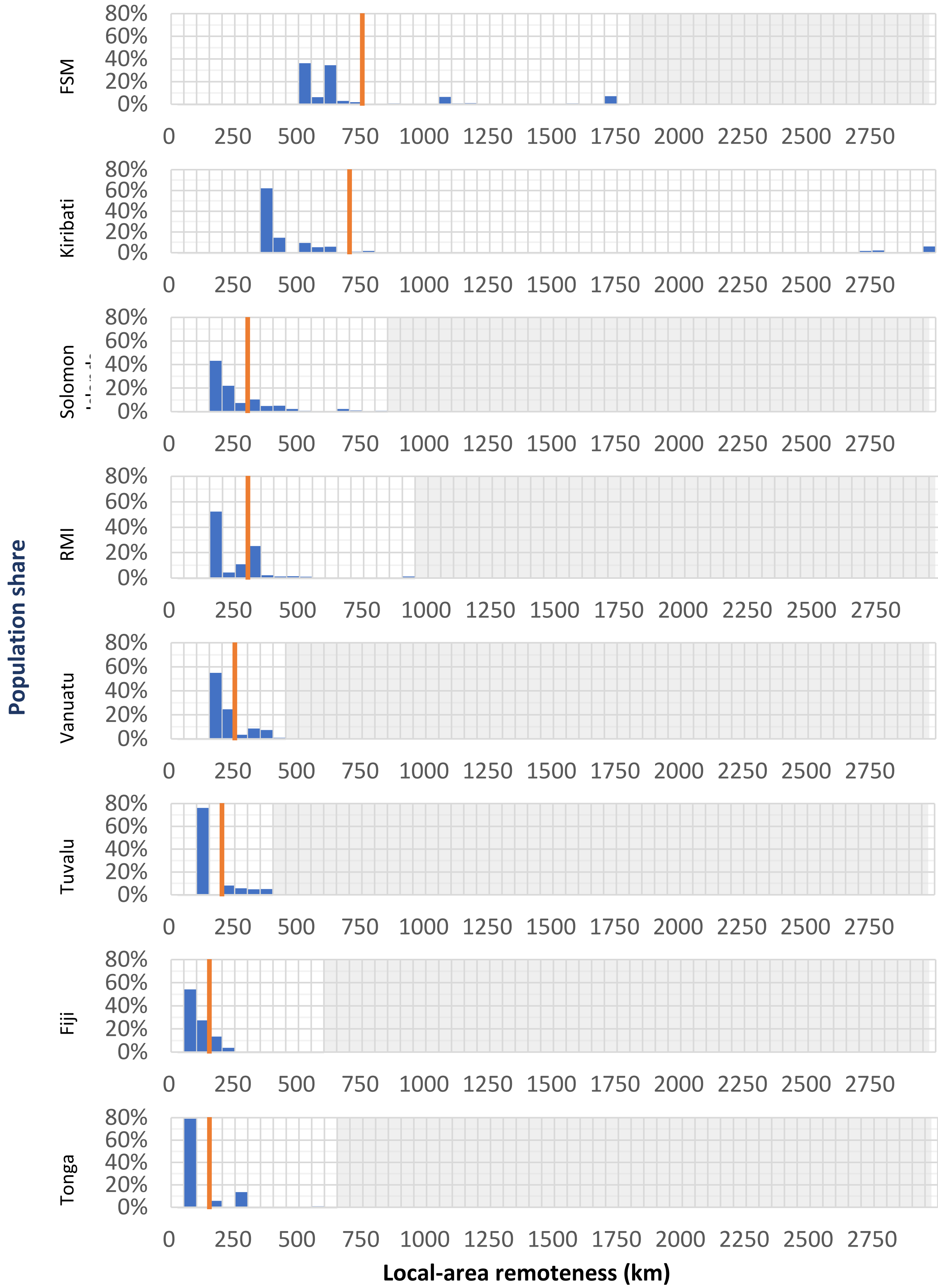
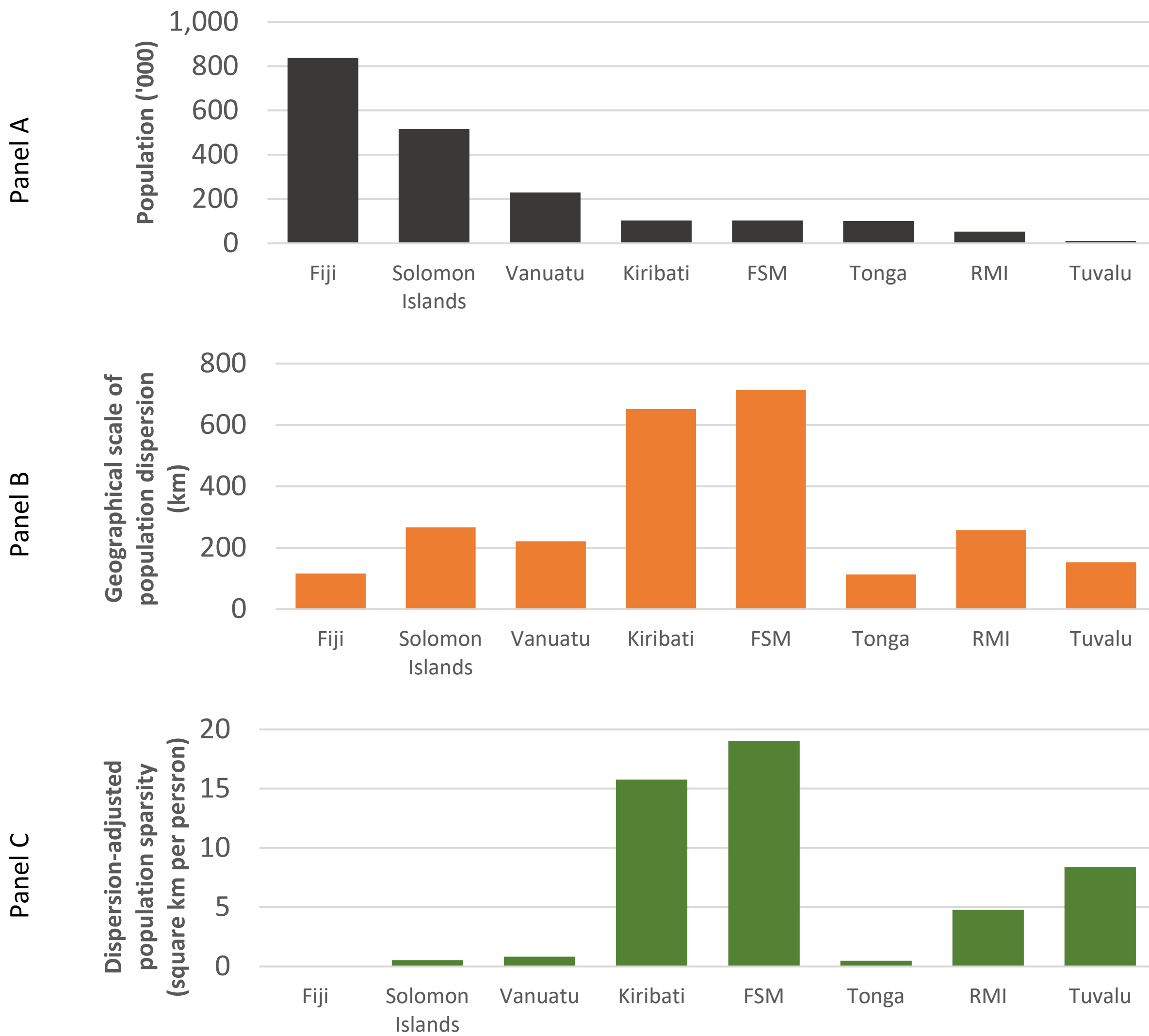


Figure 2: Population, geographical scale of population dispersion, and dispersion-adjusted population sparsity across case-study PICs



Comparison with related dispersion indicators

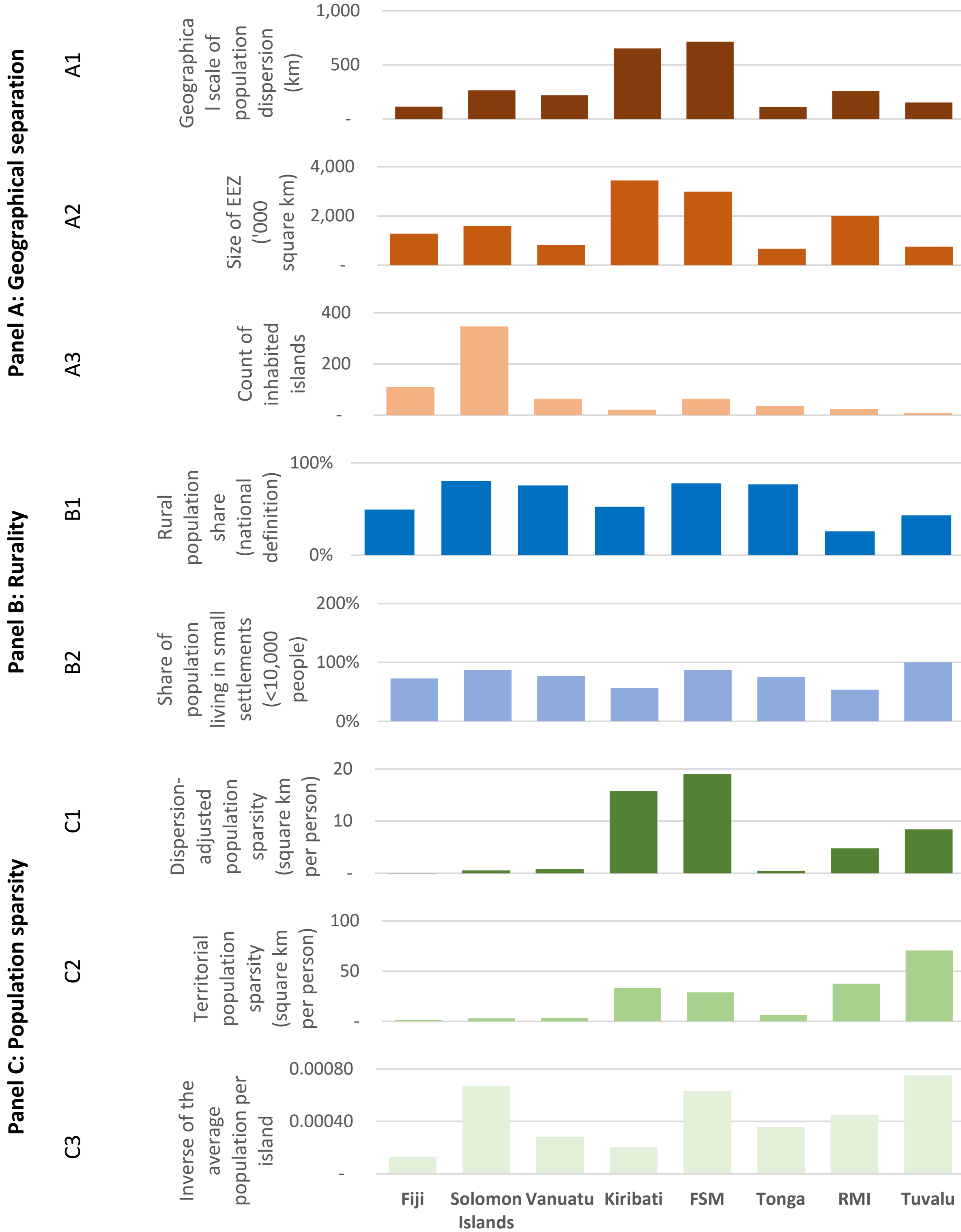
A range of existing indicators have been used in the past to describe the extent of population dispersion in the PICs (see, for example, the rich discussion in Horscroft, 2014). These indicators are often not directly comparable, but instead are complementary, reflecting different aspects of dispersion. Figure 3 presents a selection of these complementary indicators, including our new measures of the geographical scale of population dispersion, and the dispersion-adjusted population sparsity. The indicators are grouped into three panels, reflecting their different perspectives on dispersion.

Some existing indicators of dispersion reflect the geographical separation of settlements across the ocean (e.g. the count of inhabited islands, or the size of a country's EEZ). These are illustrated in Figure 3, Panel A, including our new measure of the geographical scale of population dispersion. Common patterns emerge with distance-based measures of dispersion across the PICs, with Kiribati and FSM more dispersed and Tonga less so. For the PICs, our new measure of the geographical scale of population dispersion is correlated with the size of a country's EEZ, but less closely related to the count of inhabited islands.

Other indicators of dispersion reflect the rural nature of population settlement patterns (the rural population share, or the share of the population living outside of settlements of a particular size). These are illustrated in Figure 3, Panel B. The two measures depicted—the rural share of the population, and the share living in settlements smaller than 10,000 people in size—appear quite closely correlated with one another. However, the rurality indicators reveal a slightly different picture to the indicators in Panel A. Solomon Islands has a relatively large share of the population living in rural areas or small settlements (implying a high level of dispersion, not evident in A1 and A2), and Kiribati has a relatively low share (implying a more concentrated population, not evident in A1 and A2). Fiji and FSM, for example, have quite similar levels of rurality for example (in Panel B), but the distance measures in Panel A (A1 and A2) suggest much more marked dispersion in FSM rather than Fiji.

Other socio-geographical indicators commonly reported attempt to capture population density, i.e. the concentration of people in space—a converse of dispersion. Figure 3, Panel C, inverts these measures to capture population sparsity. Panel C1 presents our dispersion-adjusted measure of population sparsity (the ratio of the dispersion-adjusted area equivalent to the population), Panel C2 presents the analogous measure of territorial sparsity (the ratio of a country's territorial land + ocean area to the population), and Panel C2 presents the inverse of the average population per (inhabited) island. The two explicit sparsity indicators are closely correlated, though notable differences exist. In particular, the territorial sparsity of Tuvalu's population is much more extreme than its dispersion-adjusted measure. Both these measures serve to highlight the sparsity of the populations of Kiribati, FSM, RMI and Tuvalu, relative to Fiji, Solomon Islands, Vanuatu and Tonga. Panel C3 presents a slightly different picture again, suggestive of a greater degree of sparsity in Solomon Islands, for example (as in Panel A3).

Figure 3: Complementary measures of PICs socio-geographical population dispersion



Local-area level remoteness and socio-economic characteristics

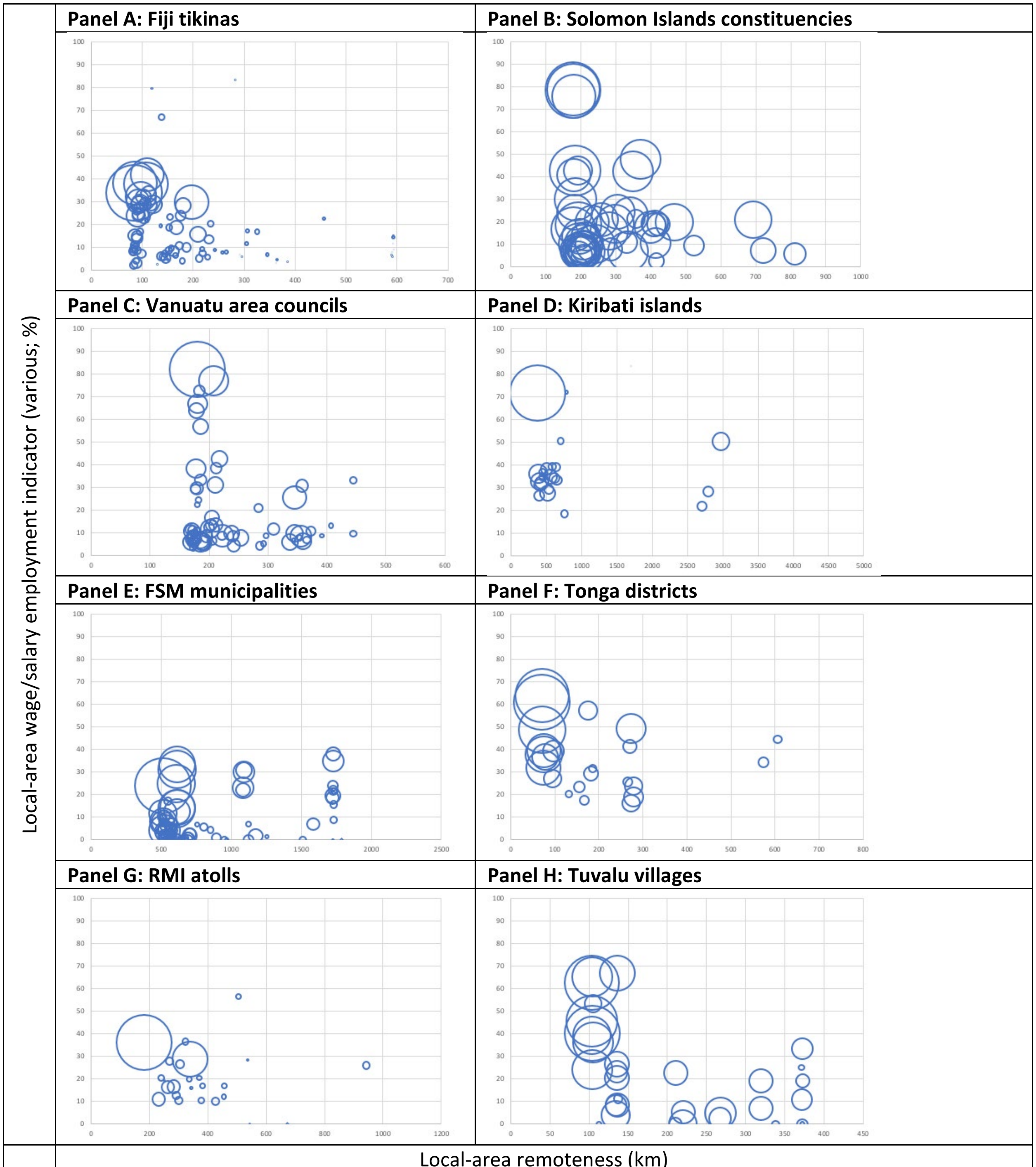
The degree of a community's remoteness within a country context is correlated with its socio-economic characteristics. For example, people living in more central parts of a country are more likely to be employed in paid work, or to have wages or salaries as a main source of income (Figure 4). In particular, the local areas with the highest proportion of people or households in paid work, or with wages or salaries as a main source of income, tend to be among the very least remote, according to our indicator.[8] However, there is variation around this general pattern across the case-study PICs, consistent with variation in the PICs' respective geographies. The relationship is less clear in FSM, where the population is particularly widely dispersed, and where there are several key population hubs some distance apart.

The link between remoteness and a variety of socio-economic characteristics is explored comprehensively in a companion piece[9] to this work, though that piece focuses on related but alternative definitions of remoteness defined more explicitly according to areas' distance from defined population hubs (e.g. the measures discussed in Appendix C).

[8] These same local areas also tend to be the most populous. For most of the case-study countries, the negative relationship between local-area remoteness and the respective employment indicator is only statistically significant when the data are population-weighted.

[9] World Bank, 2019. 'Socio-economic analysis: A spatial analysis of population and housing census data', a background paper prepared for *A Review of Spatial Development Issues in the Pacific Islands Countries* (P166346).

Figure 4: PIC local-area level remoteness and employment indicators



Notes:

1 Marker size reflects population (scaled separately for each country).

2 Wage/Salary employment indicators defined as follows:

- Fiji (2007): % population 15+ employees
- Solomon Islands (2009), Vanuatu (2010), Tonga (2016): % households with wages/salaries as the main source of income
- Kiribati (2010), Tuvalu (2012): % households with wages/salaries as a main source of income
- RMI (2011): % population 15+ with wages/salary income over past 12 months
- FSM (2010): % population 15+ paid work in past week

Source: Socio-economic indicators sourced from national census data available on the PopGIS platform via SPC; Authors' analysis.

3. Global estimates of population dispersion and sparsity for more than 100 countries

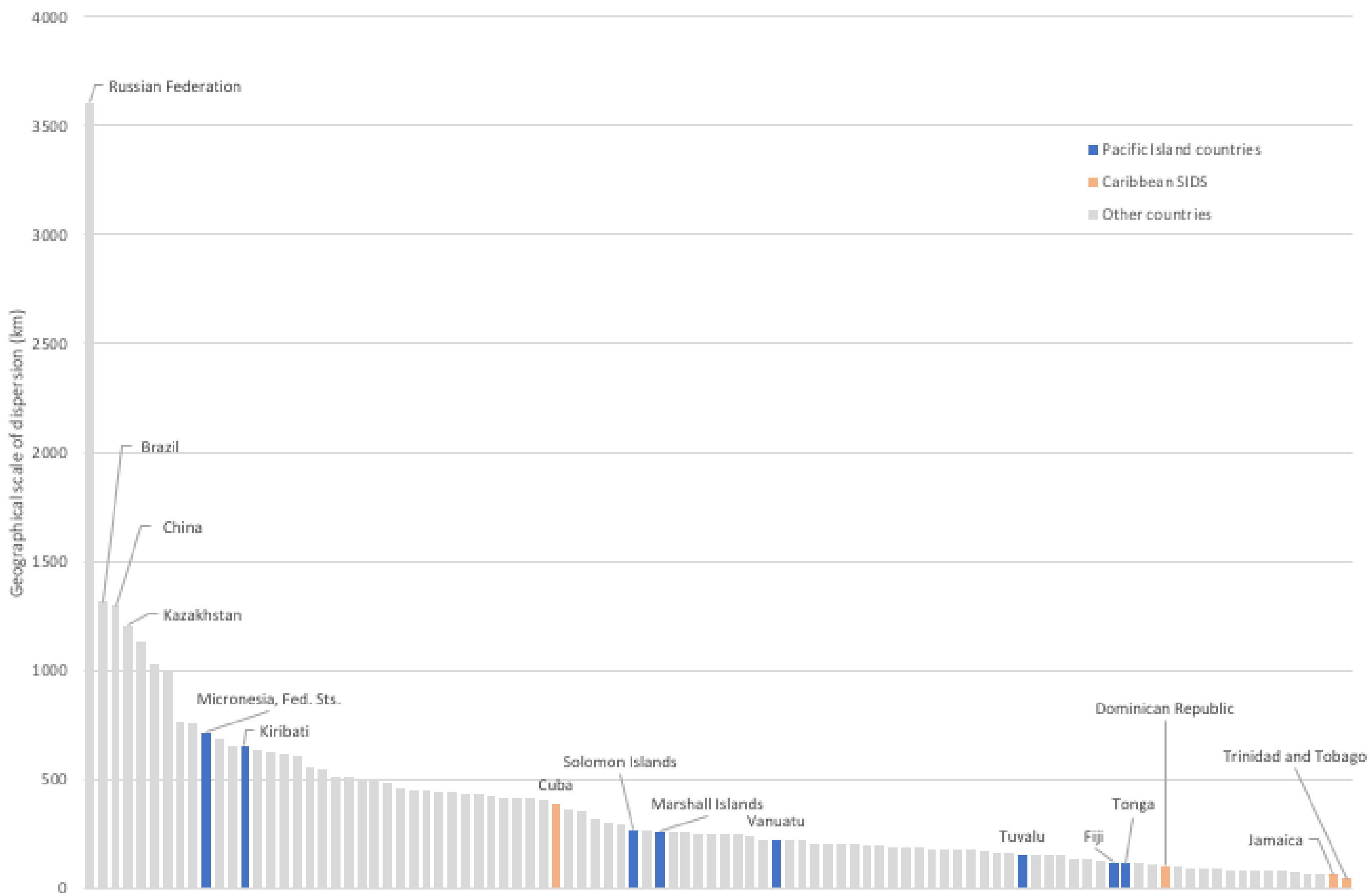
Estimates of the geographical scale of population dispersion and dispersion-adjusted population sparsity were calculated for each country with sufficiently detailed data (a minimum of five sub-national areas) in the World Bank's subnational geography database.

Across the global database, the geographical scale of population dispersion ranges from a maximum of 3,600km in Russia, to a minimum of 40km in Trinidad and Tobago (Figure 5, Appendix E shows the numerical values and country rankings according to the new dispersion measures). None of the PICs lies at the extremes of this range, but dispersion for FSM and Kiribati is at the higher end of the distribution, while Fiji and Tonga are at the lower end. The scale of dispersion in Russia is considerably larger than every other country across the database, with the estimated average distance between every two Russians more than 3,600km. The three countries with the next highest dispersion levels are Brazil (1,300km), China (1,300km) and Kazakhstan (1,200km). Dispersion is smallest in the Caribbean countries of Trinidad and Tobago (40 km) and Jamaica (60 km).

After accounting for population size, the PICs stand out as having extremely low levels of dispersion-adjusted population density compared to other countries (Figure 6), and equivalently, an extremely high degree of population sparsity (Figure 7). Figure 6 illustrates the cross-country distribution of dispersion-adjusted population density. This measure ranges from a maximum of 1,200 people per square kilometer in Bangladesh, to a minimum of 0.05 people per square kilometer in FSM. The inverse of this density measure, the dispersion-adjusted population sparsity, is illustrated in Figure 7. With the exception of Fiji, the degree of population sparsity in the PICs completely dwarfs that recorded for other countries in the database (the first panel shows a complete scale on the vertical axis; the second panel reduces the scale to better reveal the variation across the non-PIC countries). The seven case study countries of FSM, Kiribati, Tuvalu, RMI, Vanuatu, Tonga and Solomon Islands are the seven countries with most extreme levels of population sparsity across the global database.

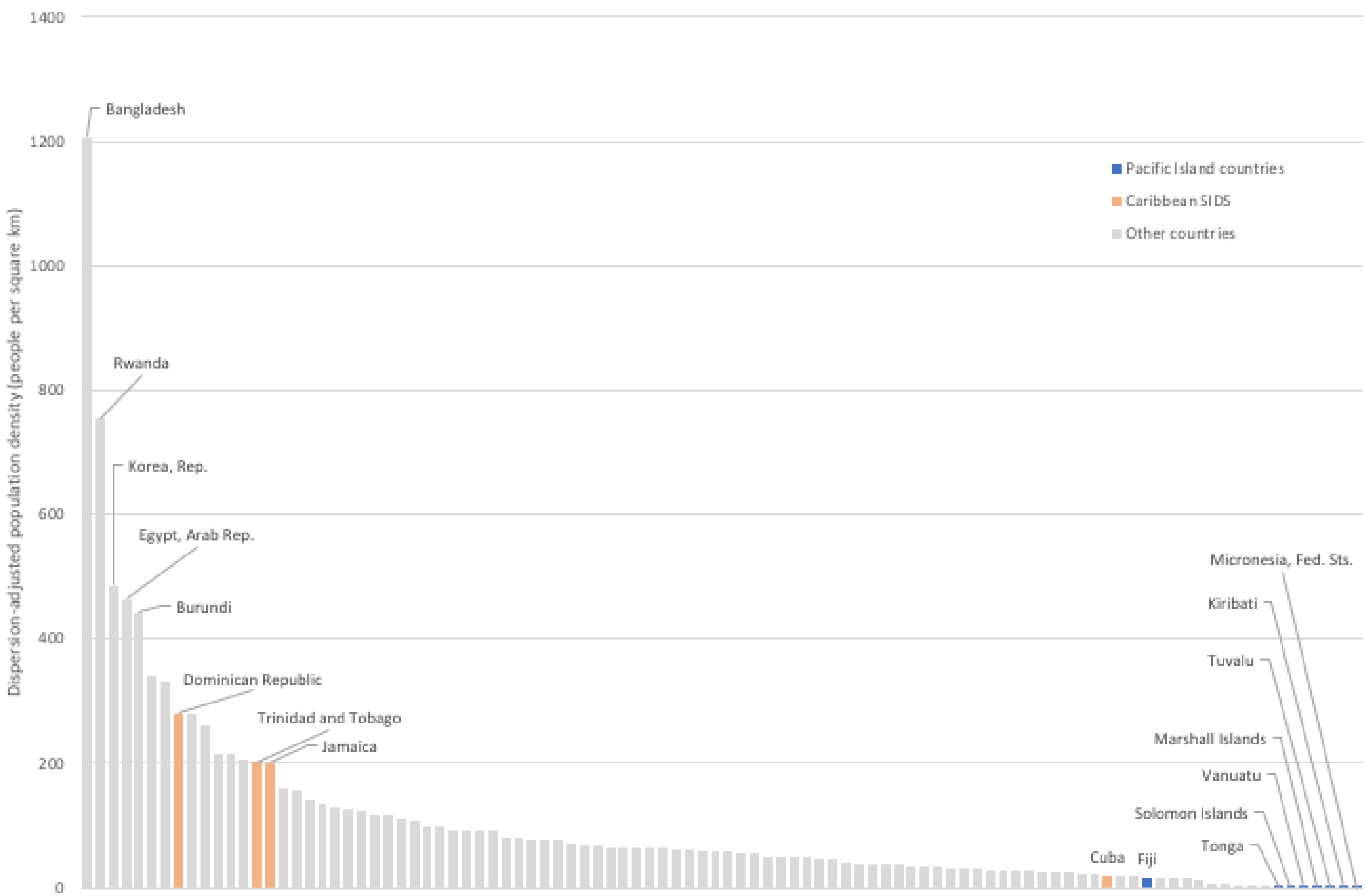
The pattern of population settlement in the PICs (which reflects the concentration of people on discrete islands) goes some way to reducing the impact of population sparsity. Figure 8 presents a cross-country scatter plot illustrating the correlation between a country's territorial population sparsity (the ratio of territorial area to population), and the dispersion-adjusted sparsity measure (the ratio of the dispersion-adjusted area-equivalent to population). The PICs stand out as having extreme levels of sparsity by both measures. Most countries lie below the diagonal line, which implies that the lived experience of population sparsity (i.e. after accounting for settlement patterns) is lower than implied by the simple ratio of territorial area to the population. The PICs lie noticeably far below the line. This is illuminated further in Figure 9, which shows the extent to which sparsity would be overestimated by looking solely at a country's territorial area, rather than adjusting for actual population settlement patterns. In Fiji, territorial sparsity is 25 times larger than the actual dispersion-adjusted sparsity estimate. Many of the island nations (the PICs and Caribbean countries) with large areas of ocean territory, but populations relatively concentrated on islands, are similarly at the upper end of this distribution.

Figure 5: Geographical scale of dispersion across countries



Source: Population data from national census statistics (PopGIS via SPC / World Bank subnational database); Authors' analysis.

Figure 6: Dispersion-adjusted population density across countries

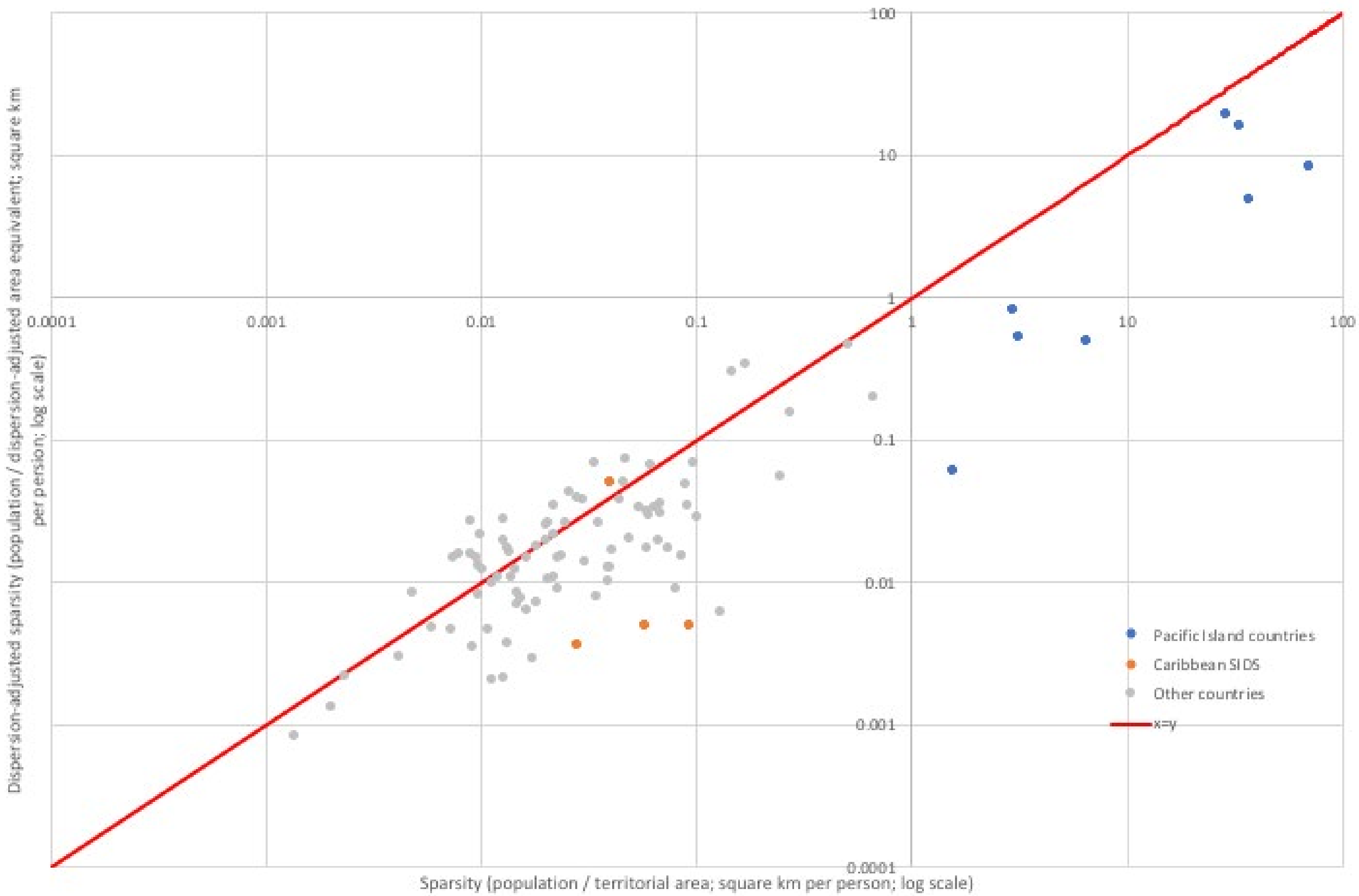


Source: Population data from national census statistics (PopGIS via SPC / World Bank subnational database); Authors' analysis.

Figure 7: Dispersion-adjusted population sparsity across countries

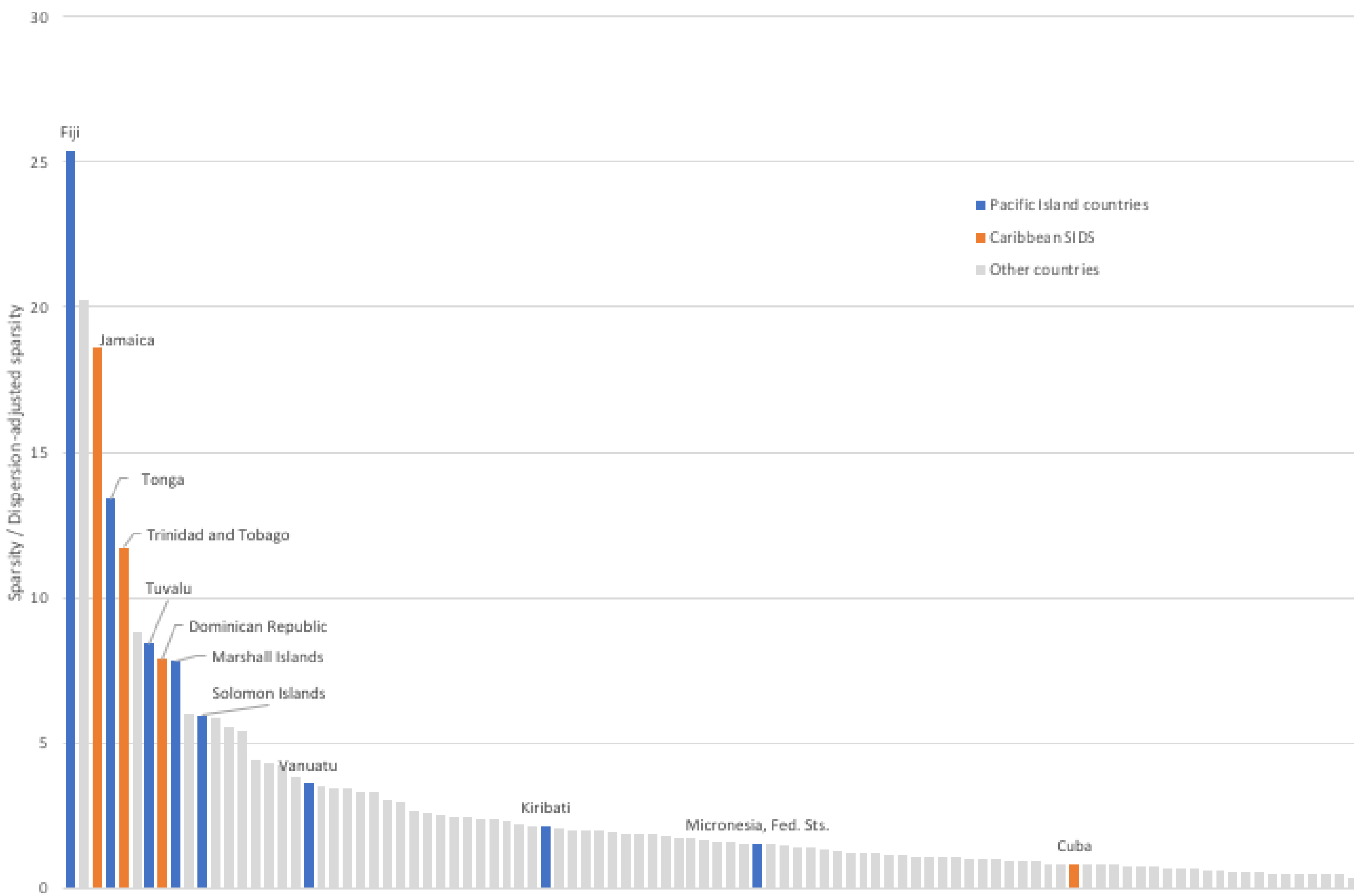
Source: Population data from national census statistics (PopGIS via SPC / World Bank subnational database); Authors' analysis.

Figure 8: Territorial sparsity and dispersion-adjusted sparsity



Source: Population data from national census statistics (PopGIS via SPC / World Bank subnational database); Authors' analysis.

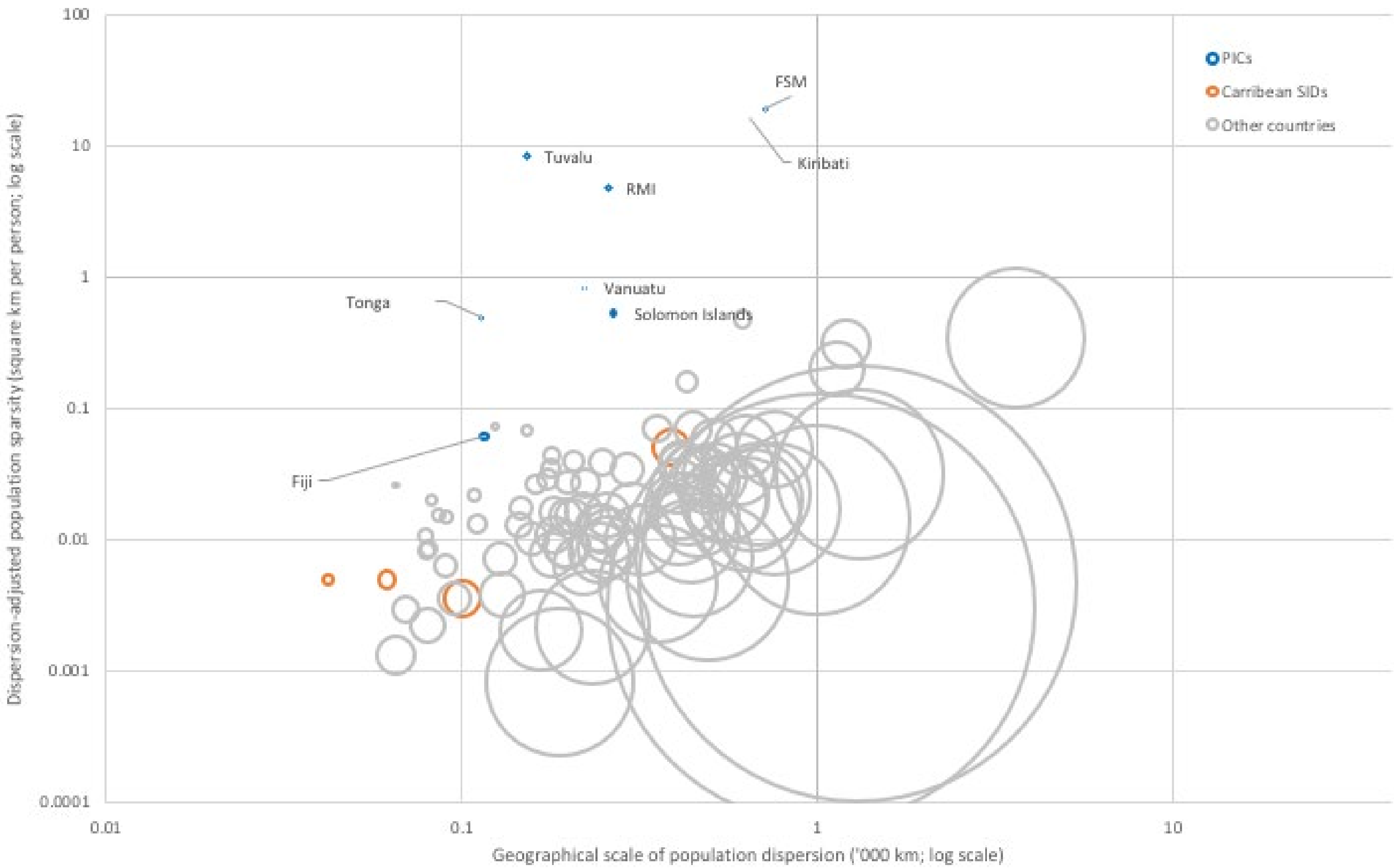
Figure 9: Impact of settlement patterns in reducing experience of sparsity



Source: Population data from national census statistics (PopGIS via SPC / World Bank subnational database); Authors' analysis.

In general, countries with smaller populations tend to be dispersed across smaller geographical distances, but the sparsity of population settlements is more pronounced. The intrinsic link between the geographical scale of population dispersion, dispersion-adjusted population sparsity, and population size is illustrated in Figure 10. Mechanically (that is, according to the formulation of the sparsity indicator), holding a country's population size constant, settlement sparsity will be increasing in the geographical scale of dispersion. This can be seen in the figure in the tendency for the countries with similar sized populations (indicated by markers of similar size) to be situated along isoquant-type lines running from the bottom-left to the top-right of the chart area. As population size decreases, these isoquants shift towards the top left. That is, although the geographical scale of population dispersion tends to be smaller for countries with smaller populations, the sparsity of population settlements is more acute.

Figure 10: Geographical scale of dispersion, dispersion-adjusted sparsity, and population levels



Note: Marker size reflects population size.

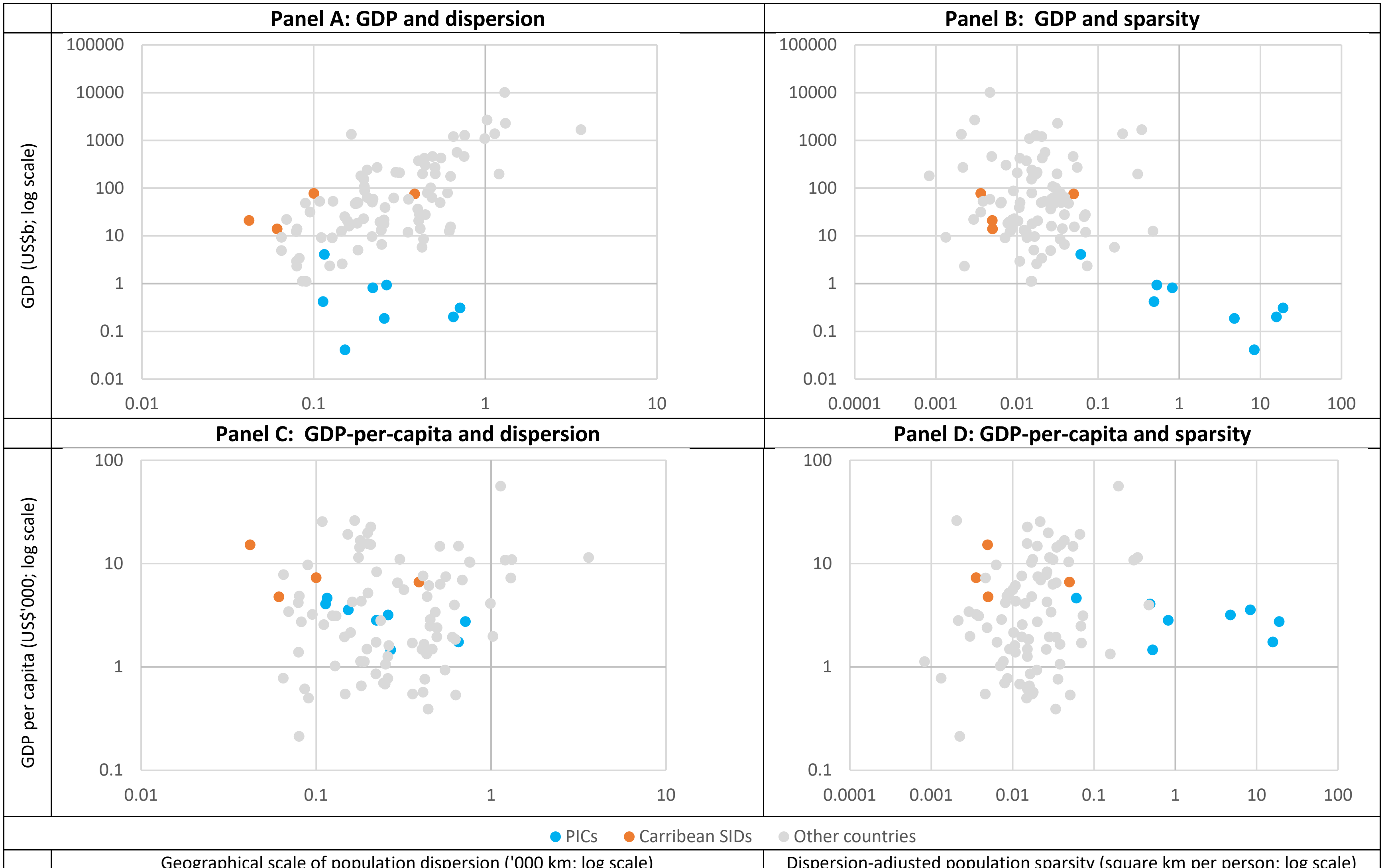
Source: Population data from national census statistics (PopGIS via SPC / World Bank subnational database); Authors' analysis.

Correlations with macro-variables

The new country-level indicators of population dispersion and dispersion-adjusted sparsity might prove to be useful in future cross-country analyses—for example, as variables to control for countries' socio-geographical characteristics (or settlement patterns). However, in a basic bivariate analysis with key macroeconomic variables, the new indicators are not in isolation particularly illuminating. Dispersion is correlated with the level of a country's gross domestic product (GDP) (Figure 11, Panel A), particularly abstracting away from the Pacific countries, but this reflects the more fundamental, general positive correlation between territory size (directly related to the distance measure of dispersion), and population (directly related to GDP). There is a much less clear relationship between dispersion and GDP-per-capita (a juxtaposition whereby population levels are less directly related to either variable) (Figure 11, Panel C). The Pacific countries drive what looks like a negative relationship between GDP and dispersion-adjusted sparsity (Figure 11, Panel B), but this relationship no longer holds if the Pacific countries are excluded. There is also no clear relationship between GDP per capita and sparsity (Figure 11, Panel D).[10]

[10] There is a statistically significant positive correlation between GDP per capita and sparsity (expressed in log form) if the Pacific Islands are excluded. However, this significance depends on the countries with the highest and lowest GDP per capita in the data set: Australia, which is particularly sparsely settled, Burundi, which is particularly densely settled, respectively.

Figure 11: Cross-country bivariate correlations with GDP and GDP-per-capita: Geographical scale of population dispersion and dispersion-adjusted sparsity



Source: National census statistics sourced from PopGIS via SPC for Pacific Island countries and World Bank subnational database otherwise; World Bank Development Indicators database; Authors' analysis.

Concluding remarks

This paper contributes a new approach to empirically measuring local area-level remoteness, and country-wide population dispersion and settlement sparsity. We apply this approach to study eight case-study PICs, in an attempt to better characterize their spatial geography. The new indicators reveal significant variation in the geographical remoteness of individual communities within each case-study PIC, and the country-level geographical scale of population dispersion across the PICs. Relative to other countries across the world, all case-study PICs display an extreme degree of sparsity in the pattern of population settlements, partly reflecting the countries' small populations.

The benchmark approach we have explored in this paper can also be extended to consider spatial geographic characteristics from slightly varied perspectives, and we offer two such extensions in Appendices to this paper.

1. Our benchmark approach conceptualizes remoteness (and relatedly dispersion and sparsity) within a country context. However, it is also possible to apply the same approach to sub-national regions; for example, to examine population dispersion within a provincial context. In Appendix B, we conceptualize remoteness and dispersion relative to a nation's capital. For each case-study PIC, we calculate alternative indicators of remoteness as communities' distance from the national capital—either directly, or via local regional hubs. The associated dispersion indicators follow, as the population-weighted average community-level remoteness. These alternative measures of remoteness (and relatedly dispersion) are less sensitive to changes in settlement patterns, but potentially more transparent.

2. The remoteness, and related dispersion and sparsity measures presented in this paper are somewhat abstract in that they do not explicitly reflect remoteness from (or dispersion around) any particular center. In an extension of this work, we show that an analogous mathematical approach can be applied to derive more concrete indicators. In Appendix C, we conceptualize remoteness and dispersion relative to a nation's capital. For each case-study PIC, we calculate alternative indicators of remoteness as communities' distance from the national capital—either directly, or via local regional hubs. The associated dispersion indicators follow, as the population-weighted average community-level remoteness. These alternative measures of remoteness (and relatedly dispersion) are less sensitive to changes in settlement patterns, but potentially more transparent.

The benchmark remoteness, dispersion and sparsity indicators we focus on in this paper have many appealing intuitive characteristics. In particular, they are highly sensitive to variation in population settlement patterns. Variation in these measures across different geographical settings broadly aligns with intuition around what it means for a community to be remote in a country-context, or for a population to be highly dispersed. For example, a greater level of dispersion will be recorded in a setting with less population clustering, e.g. where two inhabited islands are further apart, all else being equal.

Another appealing characteristic of our new measures is that there is a natural link between the micro-level geographic indicator (local-area remoteness) and macro measures of population dispersion and sparsity.[11] In empirical analyses in practice, local-area level remoteness is often characterized through the urban-rural dichotomy, which is rather blunt. Alternatively, the size or density of a particular settlement can potentially provide a more nuanced indication of a resident population's centrality or remoteness. However, this approach rests critically on the delineation of boundaries for any particular settlement, which is empirically fraught (for example, due to clustering of towns).

At the same time, one open question, which would be a fruitful direction for future work to refine our new indicators, relates to the ways in which the delineation of local areas affects the measures. Intuition suggests that the most precise and informative measures of remoteness and dispersion would be based on data geo-coded at the individual household level.[12] In practice, for the Pacific Island countries, our calculations center around data for local-area communities. These are based on exogenously defined local-area (administrative, or survey catchment area) boundaries. The implications of these boundary definitions on the remoteness and dispersion measures is, we suspect, an empirical question, dependent on interior settlement patterns, and the boundaries themselves. One immediate direction for further refinement would be to re-estimate and compare the indicators using a range of geographical areas—from the finer to the coarser.

As a final note, it is worth highlighting that none of the measures we present in this paper explicitly accounts for the island-nature of PIC's geography. However, they do reflect uninhabited areas of ocean in an implicit sense. Areas of uninhabited ocean lead to large measures of distance between communities (including between communities on outer islands and the national capital). This uninhabited distance results in high measures of remoteness, and relatedly, dispersion.

[11] Part of the motivation for this work was a desire, for the purpose of empirical analysis, to apply a consistent approach to identifying 'remote', outer islands in different PICs. It became clear that it is not immediately obvious how to empirically measure remoteness of a particular community, within the broader geographical context of a country as a whole. This was part of a broader task of analyzing socio-economic characteristics of Pacific Island communities from a spatial geography perspective, with a particular interest in describing living conditions for people living on more remote islands (see World Bank, 2019). World Bank, 2019. 'Socio-economic analysis: A spatial analysis of population and housing census data', a background paper prepared for *A Review of Spatial Development Issues in the Pacific Islands Countries* (P166346).

[12] At the other extreme—where a country has no disaggregated data at all (that is, data analogous to just one 'local area')—our indicators would be completely insensitive to its internal settlement patterns.

Appendix A: Source Data

Population census data

This paper relies on data collected by eight Pacific Island governments' National Statistics Offices (see Table A1) through their national population and housing censuses, and disseminated by the Statistics for Development Division (SDD) of the Pacific Community (SPC) through their PopGIS platform (https://prism.spc.int/regional-data-and-tools/popgis2).

Table A1: Pacific country census years and National Statistics Offices

Country	Census year	National Statistics Office
Fiji	2007	Fiji Bureau of Statistics
Solomon Islands	2009	Solomon Islands National Statistics Office
Vanuatu	2009	Vanuatu National Statistics Office
Kiribati	2010	Kiribati National Statistics Office
FSM	2010	FSM Division of Statistics
Tonga	2016	Tonga Statistics Department
RMI	2011	RMI Economic Policy and Planning Statistics Office
Tuvalu	2012	Tuvalu Central Statistics Divison

The data on the PopGIS site are encapsulated in maps for each PIC, illustrating variation in the results to select census questions across different parts of the country. Although the unit-record data for each individual and household are not available (and so it is not possible to consider cross-tabulations of responses across questions, for example), the geographical disaggregation is generally very detailed, and it is particularly helpful for this study focused on spatial aspects of development in the PICs.

Table A2 describes the geographical disaggregation of the data published on the PopGIS platform. Included in this table is a note on the nature of the overlap of geographical area boundaries with islands; with the exception of Vanuatu, which publishes a purposeful disaggregation of the census results by island on PopGIS, the data cannot be manipulated to present results at the island-level.

Geospatial data

The geospatial data we used were from several different sources. The PopGIS platform provides 'shape files' (the geo-coded outlines for specific regions) for each country, corresponding to the geographical disaggregation of the data. The locations of cities and towns of interest were pinpointed using OpenStreetMap and Google Maps; some coordinates were automatically generated using the name of the town, others were manually identified. Distances were calculated using QGIS software.

Table A2: Geographical disaggregation of available data

Country	Geographical level	Number of regions	Population			Land area (km^2)			Relation to islands
			Min	Max	Average	Min	Max	Average	
Fiji	Enumeration area (EA)	1602	27	1400	522.64	0.013	285.93	13.05	Islands include multiple Eas EA can include multiple Islands
	Tikina	86	104	144526	9735.71	3.17	884.33	243.01	Tikina can include multiple islands; Island can include multiple Tikinas
	Province	15	2002	231760	55818.07	44.15	3250.78	1393.28	Island can include multiple provinces Province can include multiple islands
Solomon Islands	Enumeration area (EA)	1344	0	2036	383.83	0.004	394.77	21.57	Islands include multiple Eas EA can include multiple Islands
	Ward	183	115	14806	2818.96	0.074	1082.98	158.43	Ward can include multiple islands, Island can include multiple Wards
	Constituency	50	2334	25018	10317.40	6.02	2781.11	579.87	Constituency can include multiple islands, Island can include multiple Constituencies
	Province	10	3006	136384	51587.00	23.52	5611.15	2899.33	Province include multiple islands
Vanuatu	Enumeration area (EA)	621	0	7111	368.57	0.029	331.97	21.47	Islands include multiple Eas EA can include multiple Islands
	Island	67	0	64327	3416.16	0.017	3996.23	199.04	Data by island
	Area Council	66	341	43276	3467.92	0.956	1035.4	202.06	AC can include multiple islands, Island can include multiple Acs
	Province	6	9189	77047	38147.17	879.35	4271.67	222.63	Province include multiple islands
Kiribati	Enumeration area (EA)	265	31	2311	388.90	0.01	4.59	0.27	Island include multiple EAs
	Village	180	31	15755	572.54	0.02	4.59	0.39	Island include multiple Villages
	Island (Atoll)	23	31	50182	4480.78	0.2	443.42	37.44[13]	Island(Atoll) include multiple islands

(continued over)

[13] Total area for Kiribati at EA and Village Level is 70.93 km^2, at Island(Atoll) Level is 861.192 km^2. Because the EA and Village Level map file does not cover the whole land area of the country.

Country	Geographical level	Number of regions	Population			Land area (km^2)			Relation to islands
			Min	Max	Average	Min	Max	Average	
FSM	Enumeration area (EA)	394	4	631	261.01	0.02	25.06	1.79	Islands include multiple Eas EA can include multiple Islands
	Municipality	75	6	13856	1371.19	0.11	9.86	9.38	Mu can include multiple islands, Island can include multiple Mu
	State	4	6616	48654	25709.75	110.27	372.18	175.87	State include multiple islands
Tonga	Census block	640	0	464	45.40	0.022	32.67	0.93	CB can include multiple islands, Island can include multiple CBS
	Village	166	0	8265	175.02	0.04	43.8	3.60	Village can include multiple islands, Island can include multiple Villages
	District	23	432	18064	1263.22	1.77	92.02	26.01	District can include multiple islands, Island can include multiple districts
	Division	5	1232	74611	5810.80	25.23	303.81	119.66	Division include multiple islands
RMI	Enumeration area (EA)	153	1	2346	347.44	0.005	3.46	0.72	Island include multiple Eas
	Atoll	25	9	27797	2126.32	0.5	15.53	4.38	Atoll include multiple islands
Tuvalu	Enumeration area (EA)	69	16	502	154.20	0.015	2.44	0.21	Island include multiple EAs
	Village	33	16	1207	322.42	0.042	2.43	0.44	Island include multiple Villages
	Island (Atoll)	9	27	6025	1182.22	0.059	5.39	1.62	Island(Atoll) include multiple islands

Appendix B: An extension: Remoteness and dispersion at the province level

The conceptual approach that we have outlined for measuring local-area level remoteness and country-wide population dispersion can also be applied at a sub-national regional level. For example, we can calculate the remoteness of each local area within a provincial context, and the associated dispersion of the population within each province. We can also derive a province-level indicator of the remoteness of each province within a country's broader context.

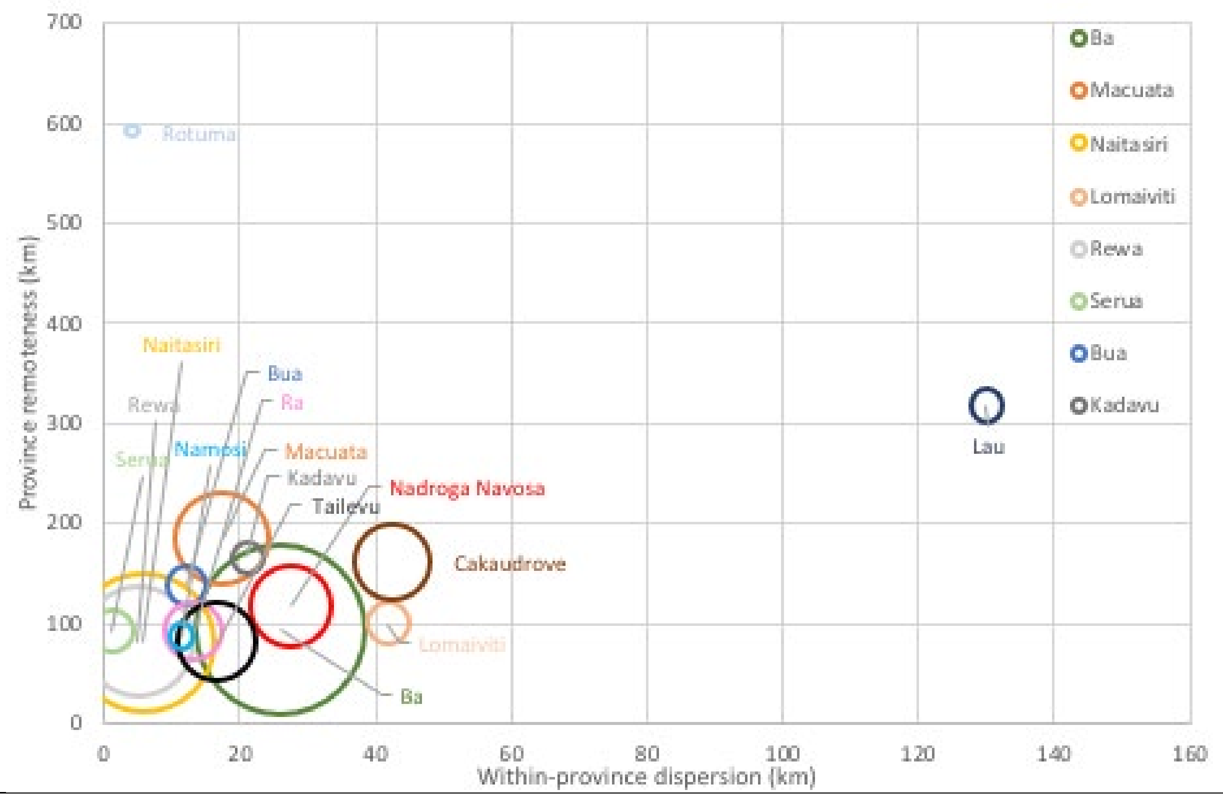
In the analysis that follows, we derive indicators of remoteness and dispersion at the sub-national level for seven of our case-study PICs (Figure B1). We do not complete the exercise for Tuvalu due to its small size. We reconceptualize community-level remoteness as remoteness within the provincial context. Dispersion of the province population is then calculated as the population-weighted average remoteness of each local area within the province; that is, the average distance from that local area, to all other local areas (within the province), weighted by those areas' populations. Remoteness of a province as a whole is conceptualized within the broader country context. For the calculations, we anchor the location of each province (or sub-national area) at the provincial capital (or other regional 'urban hub'[14]), and measure province remoteness as the (province) population-weighted average distance from that provincial capital to province capitals across the country.

For many countries, the results indicate that the more central provinces tend to have less dispersed (and often larger populations), while the most remote provinces often have widely dispersed (and small) populations. This likely reflects both geographical characteristics and (endogenous) settlement patterns, and the (endogenous) way in which administrative boundaries are determined. In Fiji, for example, Rewa Province (which captures the urban area of Suva, and had a population of around 101,000 at the time of the 2007 census) has a remoteness estimate (population-weighted average distance to all provinces) of 82km and an internal dispersion of 5km, while Rotuma Province (population 2,000) has a remoteness estimate of 318km and internal dispersion of 130km. Similarly, in Tonga, Tongatapu Province (which captures Nuku'alofa and had an estimated population of around 75,000 in 2016) has a remoteness estimate of 65km with internal dispersion of 9km, while Niuas Province (population 1,200) has a remoteness estimate of 569km with internal dispersion of 100km. The results hint at the extent to which agglomeration opportunities on the one hand, and service delivery challenges on the other, might vary across provinces within any given country.

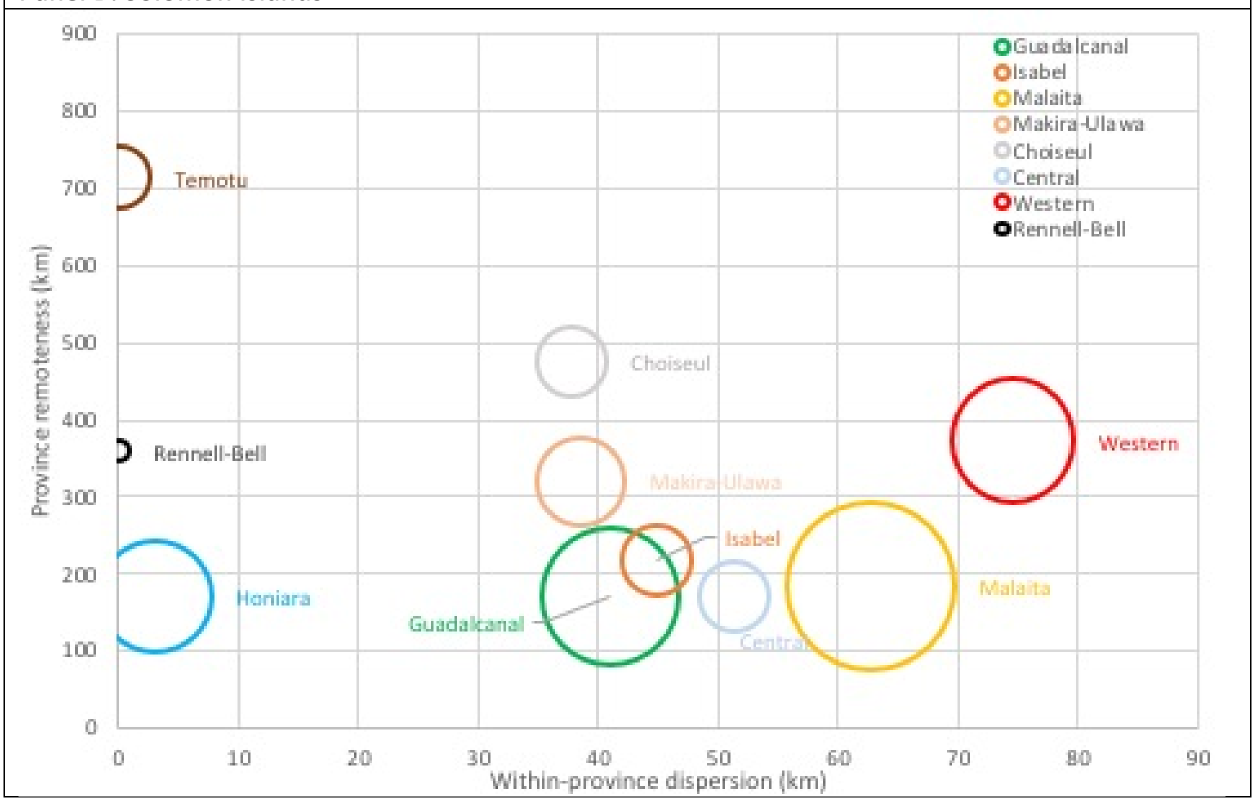
[14] See Appendix C for further details on the selection of regions and hubs.

Figure B1: Remoteness and dispersion of sub-national regional populations

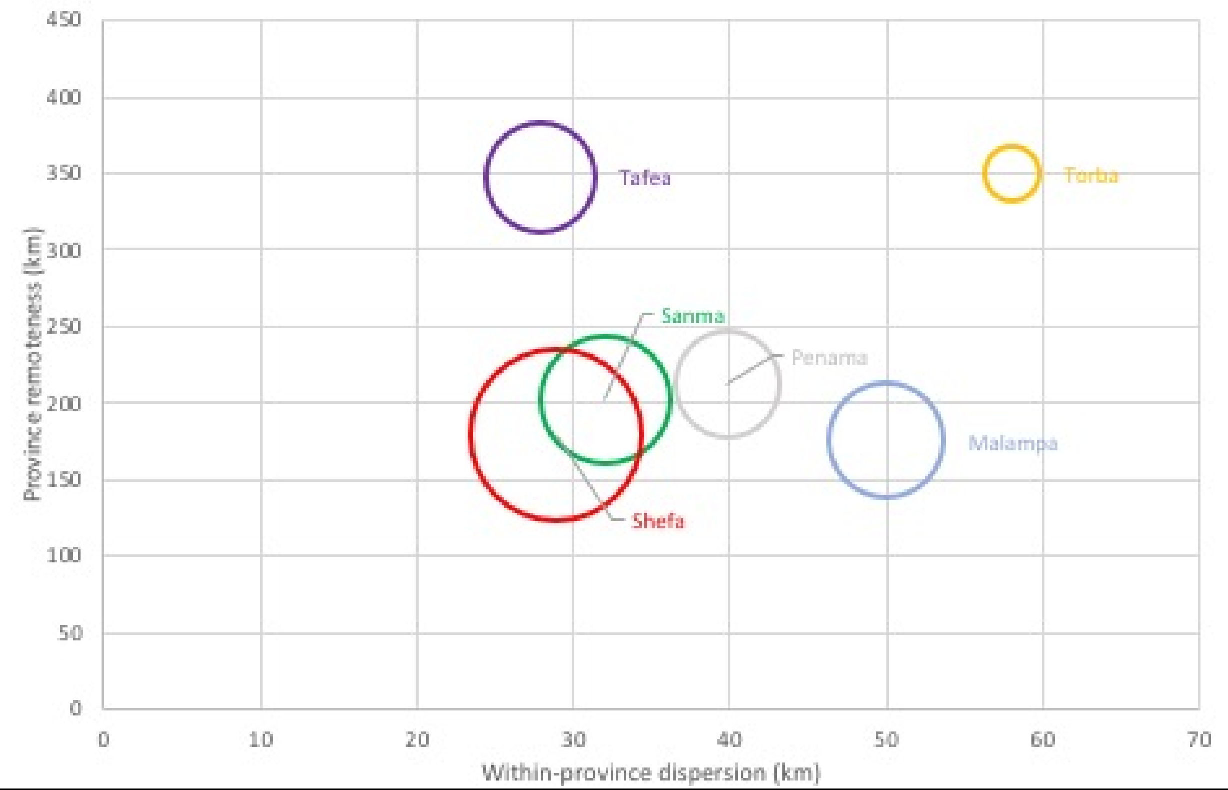
Panel A: Fiji



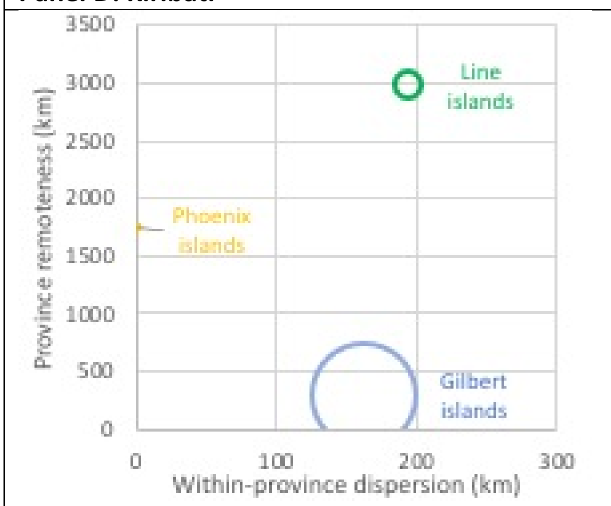
Panel B: Solomon Islands



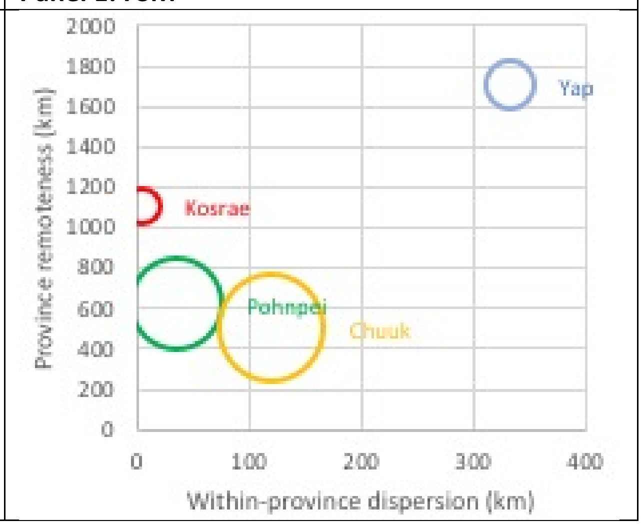
Panel C: Vanuatu



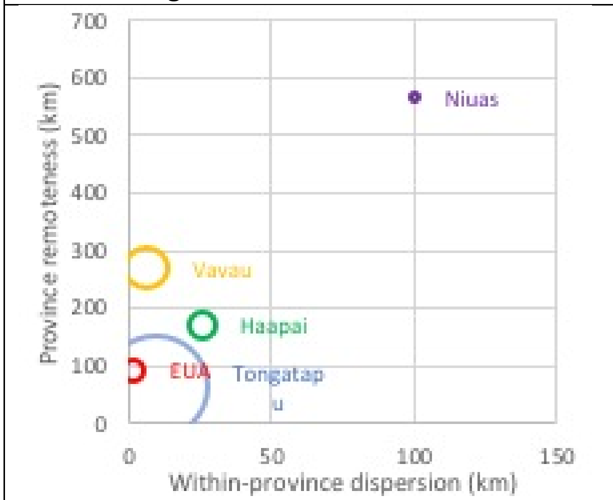
Panel D: Kiribati



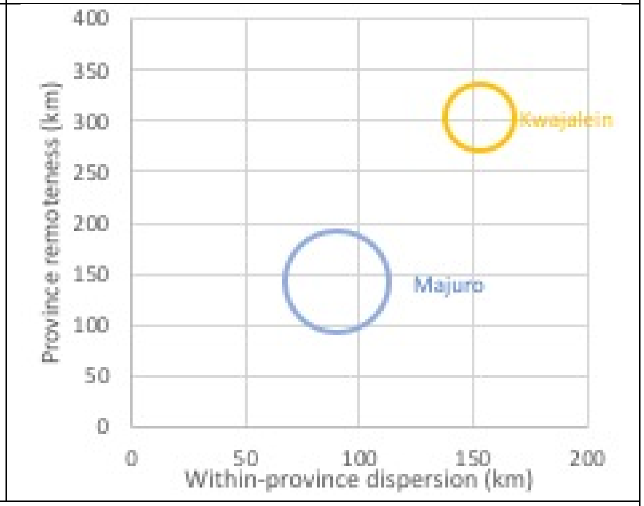
Panel E: FSM



Panel F: Tonga



Panel G: RMI



Note: Marker size reflects population (scaled separately for each country).

Appendix C: An extension: Remoteness and dispersion around key population hubs

In this section we present an alternative approach to measuring local-area remoteness and population dispersion for each country, anchored around a specific center—the national capital. We calculate two alternative versions. In the first version, we measured remoteness of a district by its direct (bird's eye) distance to the national capital (see Figure C1a). In the second version, we impose a hub-and-spoke model, nominating the national capital as the primary center, and local jurisdictional (e.g. provincial) capitals (as listed in Table C1 [15,16]) as secondary hubs. For each local area, we consider remoteness again as remoteness from the capital, but calculated this time as the sum of the distance from a local-area to the capital of its province, and the distance from that provincial capital to the national capital (see Figure C1b). In each case, dispersion of the population overall was calculated as the population-weighted average of these local-area remoteness measures. We refer to this approach as the 'distance-to-capitals' approach.

[15] In Fiji's case there were several levels of sub-national government which may have been useful for this analysis. We opted to look at provinces (in particular, rather than the larger divisions), because in our view, this provided a more different, and hence more valuable complement to the picture of remoteness presented through considering significant population/economic hubs as centers of activity.

[16] In RMI, in lieu of sub-national level jurisdictions, we nominated Ebeye as a hub; its population is significantly larger/more densely settled than any other part of the country (besides Majuro). We assigned sub-populations to either the Majuro of Kwajalein 'sub-region' based on whichever hub was closer. For Kiribati, we considered the Gilbert, Line, and Phoenix Island chains as separate sub-national areas. The Phoenix islands' population is extremely small (just 31 people at the 2010 census), but they are geographically remote from the Gilbert and Line Islands so we considered it appropriate to examine them as a separate group. In Tuvalu, due to the very small population overall, we did not apply a hub-and-spoke model and have no calculation of remoteness (or dispersion) by this approach.

Table C1: Case-study country sub-regions and centres[5,6]

Country	Jurisdiction/region	Local hub	Island
Fiji	Rewa	Suva*	Viti Levu
	Ba	Ba	
	Ra	Valeka	
	Naitasiri	Nasinu	
	Namosi	Namosi	
	Nadroga and Navosa	Sigtoka	
	Tailevu	Nausori	
	Serua	Navua	
	Macuata	Labasa	Vanua Levu
	Cakaudrove	Savusavu	
	Bua	Bua	
	Kadavu	Vunisea	
	Lau	Tabou	
	Rotuma	Ahau	
	Lomaiviti	Levuka	
Solomon Islands	Honiara/Guadalcanal	Honiara*	Guadalcanal
	Central	Tulagi	
	Temotu	Lata	
	Western	Gizo	
	Choiseul	Taro island	
	Isabel	Buala	
	Malaita	Auki	
	Makira-Ulawa	Kirakira	
	Rennell-Bellona	Tigoa	
Vanuatu	Shefa	Port Vila*	Efate
	Sanma	Luganvile	Espiritu Santo
	Malampa	Lakatoro	
	Penama	Saratamata	
	Tafea	Isangel	
	Torba	Sola	
Kiribati	Gilbert islands	Tarawa*	Tarawa atoll
	Line islands	Kiritimati	
	Phoenix islands	Kanton	
FSM	Pohnpei	Kolonia/Palikir*	Pohnpei island
	Yap	Colonia	Yap island
	Chuuk	Weno	Weno island
	Kosrae	Tofol	Kosrae island
Tonga	Tongatapu	Nukualofa*	Tongatapu
	Vava'u	Neiafu	
	Ha'apai	Pangai	
	Eua	Ohonua	
	Niuas	Hihifo	
RMI	Majuro ‘region’	Majuro*	Majuro atoll
	Kwajalein ‘region’	Ebeye	Kwajalein atoll
Tuvalu	--	Funafuti*	Funafuti atoll

* Denotes national capital

Figure C1a: Direct distance to capital

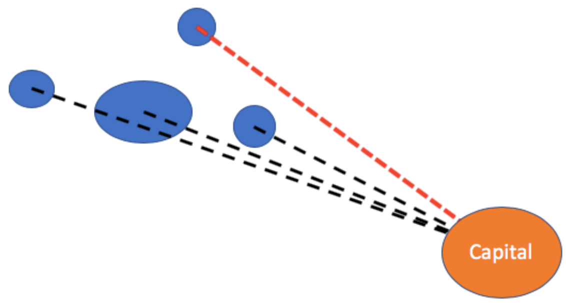
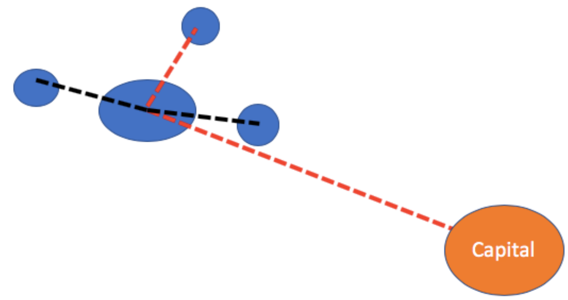
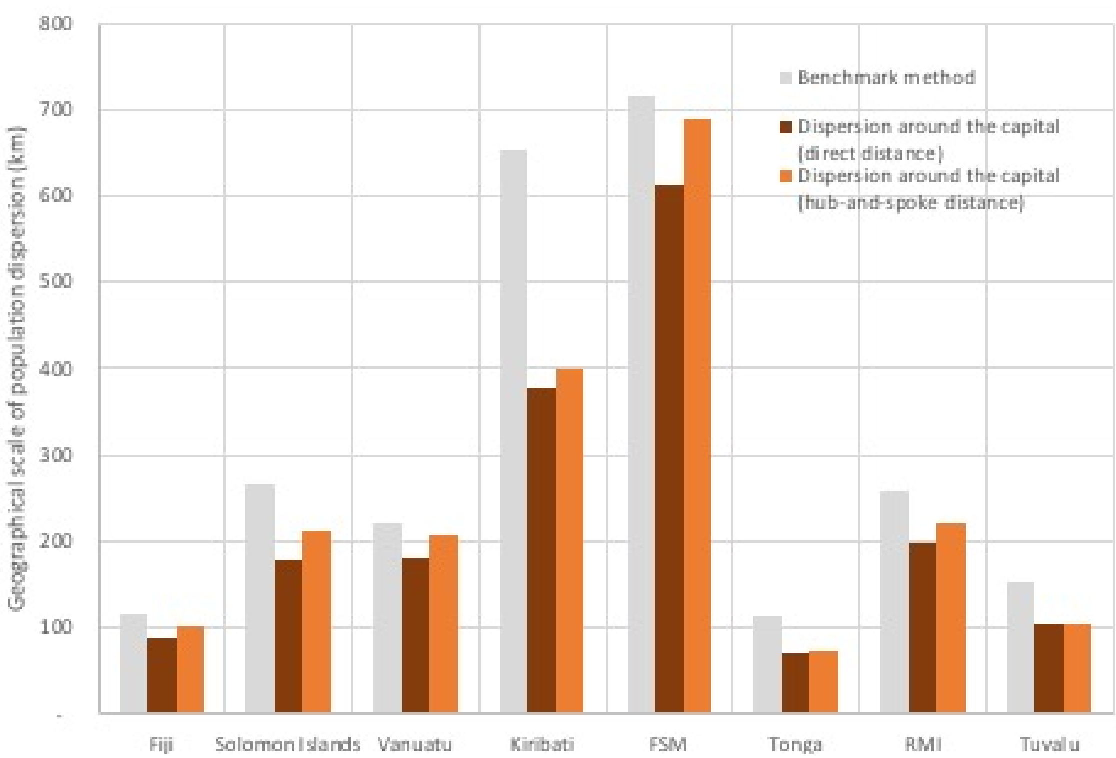


Figure C1b: Hub-and-spoke distance to capital



The estimated dispersion indicators by this distance-to-capitals approach are presented in Figure C2, juxtaposed with the results from the benchmark (more general) dispersion measure. Note that the cross-country pattern in the dispersion of the population is not markedly affected by the choice of methodology, with the exception of Kiribati. Kiribati has a far greater benchmark level of dispersion than dispersion around the capital specifically; this is due to the small settlement in the isolated Phoenix Islands from which all other communities are extremely remote (but which, for the vast majority of the population, has no impact on remoteness as measured by distance to the capital). The populations of Fiji, Tonga and Tuvalu are dispersed over relatively short distances (among the PICs) irrespective of the indicator used (though their very different populations mean this has different implications in each respective country context), while FSM repeatedly stands out as having the most widely dispersed population.

Figure C2: Geographical scale of dispersion: Benchmark measure and dispersion around the capital



By definition, the hub-and-spoke measure of dispersion around the capital increases the value of the dispersion indicator relative to the direct (bird's eye) measure of dispersion around the capital (i.e. it shifts each marker in the figure higher on the vertical axis). This is because traveling to the national capital via the provincial capital necessarily leads to a greater measure of distance than the direct line of sight. For each of the case-study countries, it is also true empirically that the measure of general dispersion results in a greater distance indicator than the dispersion measures centered around the capital. This is not a mathematical construct per say, but a likely empirical finding—indicative of the population density in (and around) the capital city of each country.[17]

The relative consistency in the overall picture painted by these three different aggregate dispersion measures masks noteworthy differences in precisely which districts are being identified as more or less remote under each alternative approach. That is:

- *Benchmark remoteness*: By this method (discussed in the main body of the paper), the remoteness of any particular district depends upon the settlement patterns across the country in their entirety. Generally speaking, larger areas of higher density populations will be measured as less remote than more sparsely settled, geographically isolated regions.

- *Remoteness from the capital, according to direct distance*: By this method, the identification of less and more remote communities is straightforward: those closer to the capital are less remote, those further afield are more remote.

- *Remoteness from the capital, according to the hub-and-spoke model*: By this method, communities closest to the capital, within the capital's own province, are again least remote; communities within other provinces are more or less remote from the capital depending on how far they are situated from their respective provincial hubs.

Figure C3a-c illustrates how the alternative approaches to measuring local-area remoteness play out in the Solomon Islands context. Areas shaded in a lighter hue are those classified as the least remote, and areas shaded in a darker hue the most remote. Note that in the Solomon Islands case, the direct distance-to-capital method and the general remoteness approach both lead to similar results in terms of which local areas are identified as more or less remote. This reflects the concentration of the population in Honiara, and the absence of any other substantial population hub. The distance-to-capital approach using the hub-and-spoke model results in a slightly different picture, with pockets of slightly less remote areas around provincial capitals (e.g. around Gizo, in Western Province).

These alternative measures are correlated with our benchmark measure of general remoteness and dispersion. However, they differ in a meaningful way, resulting in slightly different patterns of relative proximity and remoteness across different parts of each country. Areas in and around the capitals (and regional hubs) are by definition the least remote according to these alternative indicators. Communities distant from the capital and designated regional hubs will be considered increasingly remote, irrespective of their size.

[17] To see this, imagine an (unlikely) empirical setting where the reverse would hold: where the distance-to-capital measures would suggest a higher level of dispersion than the benchmark method. Suppose the capital of Solomon Islands were to be moved from Honiara to the most remote small island, where only one person lives. The distance-to-capital measures of remoteness, and relatedly dispersion, would now suggest a much higher level of dispersion (around the capital) because nearly everyone lives a long way from the small remote island. At the same time, the benchmark measure of dispersion is unchanged (since it is unaffected by the designation of the capital city). Note too that distances to the small, remote island would have a very low weight in the general remoteness and dispersion indicators, since the island has just one resident.

Figure C3a: Solomon Islands: Benchmark remoteness measure

Figure C3b: Solomon Islands: Remoteness from the capital: Direct distance measure

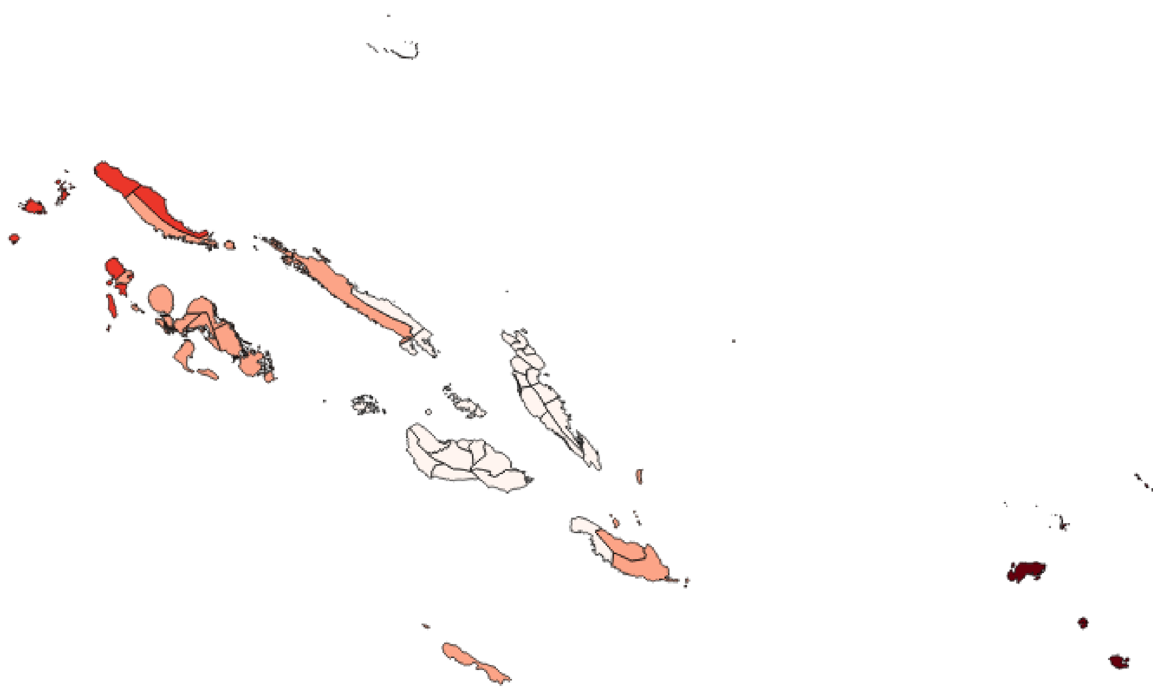


Figure C3c: Solomon Islands: Remoteness from the capital: Hub-and-spoke measure

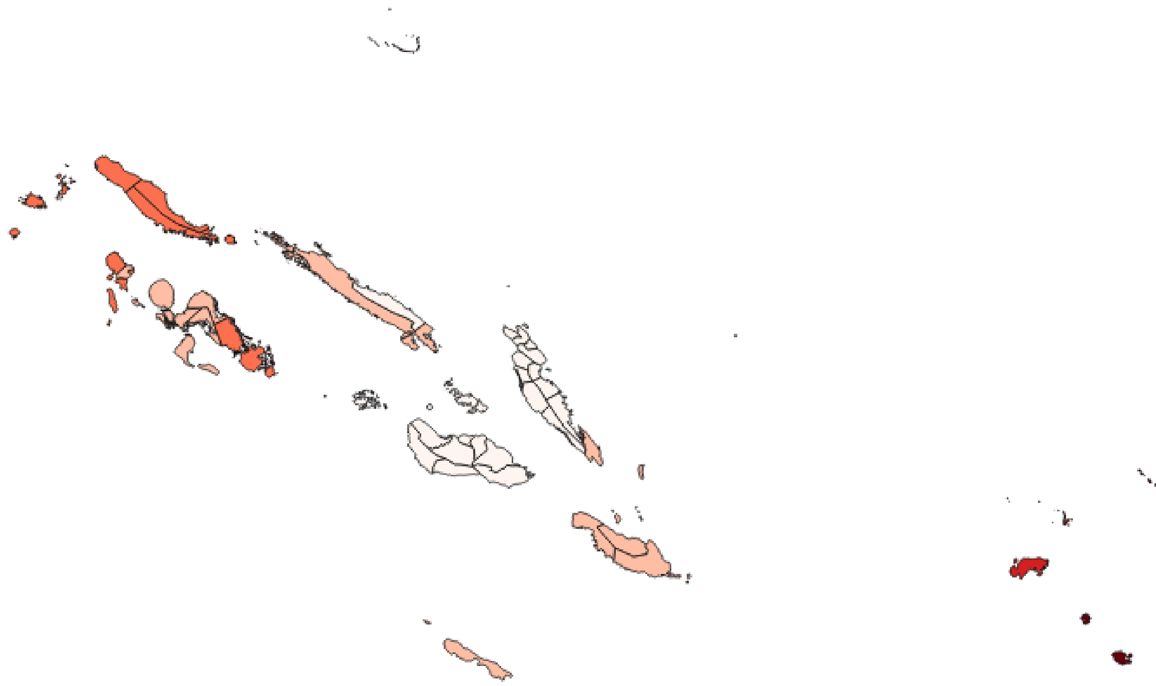
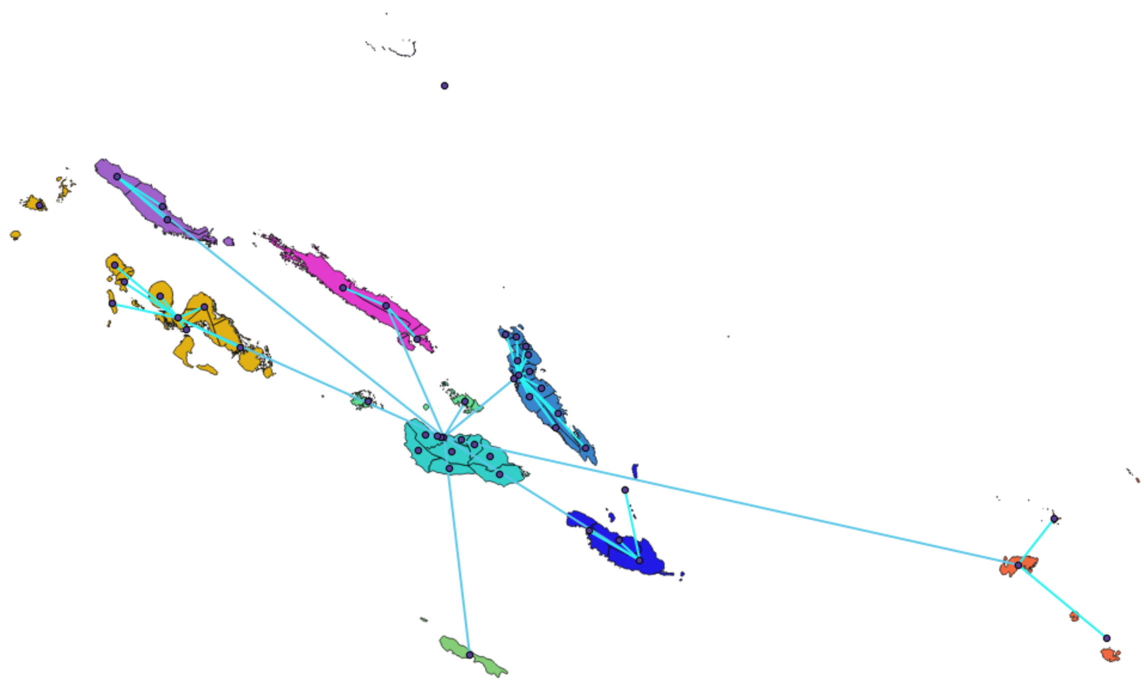
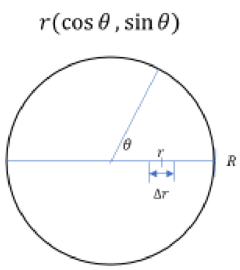


Figure C3c Supplement: Hub-and-spoke model as applied to Solomon Islands



Appendix D: Derivations using circles

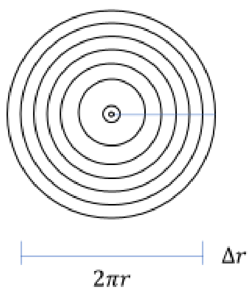
Expected distance between two points on a disc



$(\mathrm{r}, 0)$

$(\tau, \theta) \rightarrow (\tau\cos\theta, \tau\sin\theta)$

$$\mathrm{d}\big((\mathrm{r}, 0), (\tau\cos\theta, \tau\sin\theta)\big) = \sqrt{(\tau\cos\theta - r)^2 + \left(\tau^2 (\sin\theta)^2\right)}$$



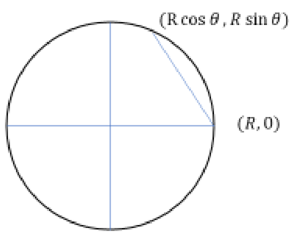
$$\frac{2\pi r \Delta r}{\pi R^2} = 2\frac{r}{R^2}dr$$

$$\mathbb{E}[\mathrm{x}] = \sum_x p(x)\mathrm{x}$$

$$\begin{aligned}
\mathrm{A} &= \int_0^R \int_0^R \int_O^\pi \sqrt{(\tau\cos\theta - r)^2 + \left(\tau^2 (\sin\theta)^2\right)}\,\frac{d\theta}{\pi}\frac{2rdr}{R^2}\frac{2\tau d\tau}{R^2} \\
&= \frac{4}{\pi R^4}\int_0^R \int_0^R \int_O^\pi r\tau\sqrt{(\tau\cos\theta - r)^2 + \tau^2 (\sin\theta)^2}\,d\theta d\tau dr \\
&= \frac{4}{\pi R^2}\int_0^1 \int_0^1 \int_O^\pi r\tau\sqrt{(\tau\cos\theta - r)^2 + \tau^2 (\sin\theta)^2}\,d\theta d\tau dr \\
&\triangleq \frac{128}{45\pi R^2}
\end{aligned}$$

Credit: James Parkinson

Expected distance between two points on the circumference of a circle



$$\text{Avg} = \int_0^\pi \sqrt{(R - R\cos\theta)^2 + R^2(\sin\theta)^2}\,\frac{d\theta}{\pi}$$

$$= \frac{R}{\pi}\int_0^\pi \sqrt{1 - 2\cos\theta + 1}\,d\theta$$

$$= \frac{R}{\pi}\int_0^\pi \sqrt{4\left(\sin\frac{\theta}{2}\right)^2}\,d\theta$$

$$= \frac{2R}{\pi}\int_0^\pi \sin\frac{\theta}{2}\,d\theta$$

$$= \frac{2R}{\pi}2\left(-\cos\frac{\theta}{2}\right)\Big|_0^\pi$$

$$= \frac{4R}{\pi}$$

$$(\cos\theta)^2 + (\sin\theta)^2 = 1$$

$$(\cos\theta)^2 - (\sin\theta)^2 = \cos 2\theta$$

$$1 - \cos 2\theta = 2(\sin\theta)^2$$

Credit: James Parkinson

APPENDIX E: Dispersion and dispersion-adjusted population density

Country	**Dispersion (km)**	Dispersion (Ranking)	**Dispersion adjusted pop density (population/km2)**	Dispersion adjusted pop density (Ranking)
Russian Federation	3609	1	2.90	9
Brazil	1313	2	31.68	31
China	1298	3	214.84	88
Kazakhstan	1205	4	3.24	10
Australia	1136	5	4.98	11
India	1024	6	333.14	93
Indonesia	993	7	69.91	61
Mexico	759	8	58.45	51
Argentina	754	9	20.31	20
Micronesia, Fed. Sts.	693	10	0.06	1
Iran, Islamic Rep.	685	11	45.13	41
Kiribati	669	12	0.07	2
Turkey	651	13	49.67	45
Mozambique	629	14	19.58	18
Vietnam	627	15	63.37	54
Mongolia	617	16	2.11	6
Sudan	602	17	29.19	29
South Africa	551	18	48.78	43
Tanzania	545	19	50.32	46
Peru	512	20	31.97	32
Chile	510	21	18.08	17
Nigeria	492	22	205.45	87
Uzbekistan	490	23	35.26	35
Angola	481	24	33.60	33
Myanmar	461	25	65.65	58
Philippines	448	26	136.14	80
Bolivia	448	27	14.38	15
Thailand	442	28	92.29	69
Niger	437	29	29.36	30
Algeria	432	30	57.84	50
Mauritania	428	31	6.28	12
Mali	419	32	27.58	26
Zambia	414	33	25.98	25
Afghanistan	409	34	55.44	47
Colombia	408	35	76.76	63

Country	**Dispersion (km)**	Dispersion (Ranking)	**Dispersion adjusted pop density (population/km2)**	Dispersion adjusted pop density (Ranking)
Cameroon	404	36	38.56	40
Cuba	387	37	19.97	19
Ethiopia	357	38	215.20	89
Lao PDR	355	39	14.23	14
Iraq	317	40	99.17	72
Romania	301	41	56.26	48
Belarus	293	42	28.96	28
Cote d'Ivoire	260	43	93.48	70
Nepal	258	44	115.28	75
Zimbabwe	257	45	65.38	55
Marshall Islands	254	46	0.22	4
Solomon Islands	253	47	2.50	8
Kyrgyz Republic	250	48	25.97	24
Burkina Faso	248	49	81.41	66
Syrian Arab Republic	244	50	80.41	65
Yemen, Rep.	243	51	124.73	78
Egypt, Arab Rep.	234	52	464.22	96
Bulgaria	222	53	37.55	38
Ghana	220	54	155.29	83
Benin	219	55	60.60	52
Croatia	205	56	25.54	23
Czech Republic	205	57	65.62	57
Ecuador	198	58	110.47	74
Slovak Republic	198	59	36.35	36
Hungary	196	60	66.23	59
Senegal	195	61	108.41	73
Vanuatu	195	62	1.89	5
Bangladesh	189	63	1207.23	100
Tunisia	181	64	91.41	67
Togo	181	65	62.01	53
Cambodia	179	66	129.71	79
Lithuania	179	67	23.08	22
Uruguay	177	68	28.66	27
Panama	175	69	34.99	34
Korea, Rep.	166	70	486.26	97
Georgia	161	71	37.51	37

Country	**Dispersion (km)**	Dispersion (Ranking)	**Dispersion adjusted pop density (population/km2)**	Dispersion adjusted pop density (Ranking)
Honduras	158	72	97.11	71
Estonia	152	73	14.88	16
Liberia	147	74	57.18	49
Nicaragua	145	75	76.86	64
Guatemala	130	76	261.25	90
Tajikistan	129	77	140.90	81
Bhutan	124	78	13.65	13
Tuvalu	124	79	0.19	3
Sierra Leone	116	80	147.48	82
Tonga	115	81	2.15	7
Moldova	111	82	75.56	62
Slovenia	109	83	45.77	42
Fiji	104	84	21.91	21
Dominican Republic	100	85	279.50	92
Jordan	95	86	278.96	91
Gambia, The	91	87	66.91	60
Costa Rica	90	88	158.85	84
Guinea-Bissau	86	89	65.58	56
Timor-Leste	83	90	49.63	44
Albania	80	91	116.72	76
Burundi	80	92	443.59	95
Lesotho	80	93	92.20	68
Armenia	79	94	121.84	77
El Salvador	70	95	342.29	94
Montenegro	65	96	38.38	39
Rwanda	65	97	755.23	98
Jamaica	61	98	200.12	85
Trinidad and Tobago	42	99	202.23	86
Bosnia and Herzegovina	33	100	859.01	99

Note: Cells in grey are small Pacific Island Countries