441 G St. N.W. Washington, DC 20548

July 3, 2024

The Honorable Harriet M. Hageman Chair Subcommittee on Indian and Insular Affairs Committee on Natural Resources House of Representatives

Subject: Responses to Questions for the Record; Hearing Entitled "Examining GAO's Findings to Address Data Gaps and Improve Data Collection in the Territories."

This letter responds to your June 18, 2024 request that I address questions for the record related to the Subcommittee's June 13 hearing. My responses to the questions are based on GAO's previous work and knowledge on the subjects raised by the questions.

If you have any questions about the responses to your questions or need additional information, please contact me at (202) 512-4409 or LoveGrayerL@gao.gov

Latesha Love-Grayer

Director, International Affairs and Trade

Laterha Lox-Aleagur

Enclosure

• "Examining GAO's Findings to Address Data Gaps and Improve Data Collection in the Territories."

Questions from Rep. Westerman for Ms. Latesha Love-Grayer, Director of International Affairs and Trade, Government Accountability Office

1. What role can interested outside groups play in helping to address territorial data gaps?

Groups like the Association of State and Territorial Health Officials (ASTHO) may play a role in bridging federal and territorial agency needs and interests. For example, during our audit work, our staff observed an ASTHO meeting that involved what appeared to be collaborative and constructive exchanges among participants from territorial agencies, academic institutions, and federal agencies. In addition, local universities may serve as the de facto statistical agencies in certain territories (e.g., Guam, U.S. Virgin Islands (USVI)) and can also help to mitigate territorial capacity limitations.

2. Why can't the federal statistical agencies /federal government already close the data gaps in the territories?

There are various challenges that federal statistical agencies face collecting data in the territories, including non-standardized address formats, confidentiality and precision concerns that arise when reporting data for small communities, and the cost and logistics of collecting data for the territories that are furthest away. Statutory differences among agencies can also affect whether they collect data in the territories. For example, the Census Bureau is required by law to include the Commonwealth of the Northern Mariana Islands (CNMI), Guam, Puerto Rico and USVI in the Island Areas Censuses, but must obtain the Secretary of State's approval to include American Samoa. It will take a concerted effort to work through such issues, which is why we recommended that the Office of Management and Budget (OMB) lead that effort in a coordinated way across the government.

¹13 U.S.C. § 191(a); 13 U.S.C. § 141(a).

- 3. Your report and written testimony list the various economic challenges the U.S. territories face.
 - a. Can you explain the connection between these challenges and the need for good data?

As we have previously reported, the territories face challenges to sustained economic growth, including undiversified economies based on few industries with limited job opportunities.² Consistent access to recent and accurate social and economic indicator data can help territories plan for how they will address these challenges and better demonstrate whether they qualify for federal assistance in addressing them.

4. What steps can federal statistical agencies take now, without waiting for additional funding or authorizations?

Agencies that have already taken steps to include the territories in their products can share lessons learned with agencies that haven't. Federal agencies can also engage in dialogue with the territories to identify which data gaps have the greatest impact on the territories and thus should be prioritized for addressing. OMB can ensure that agencies are using a consistent approach to assess the feasibility of including territories in their data collection and closing the gaps, as we recommended.

- 5. Your written statement notes that there have been few efforts by statistical agencies to research the costs, benefits, and feasibility of expanding statistical coverage in the U.S. territories.
 - a. For those few efforts that have been made, what were the results?

In 2018 the Census Bureau estimated that extending the annual American Community Survey to territories other than Puerto Rico would cost \$52 million per year. We are aware that the Bureau of Economic Analysis (BEA) also assessed and reported on the feasibility of expanding some statistics to include the territories. These BEA reports are not publicly available but were provided in December 2023 and January 2024 to the House and Senate Appropriations Subcommittees on Commerce, Justice, Science, and Related Agencies.

- 6. Your written statement notes that statistical agencies sometimes exclude territories from data collection because "they are accustomed to limiting the scope of their products to the 50 states and are not aware that they could also include the U.S. territories."
 - a. Can you elaborate on this point? Do you have any examples?

This was a statement made to us by officials from the Department of the Interior, so they would have insights into specific examples. Yet, implementing a coordinated, government-wide

²For more details, see GAO, *U.S. Territories: Public Debt Outlook – 2023 Update*, GAO-23-106045 (Washington, D.C.: June 29, 2023).

approach for examining and addressing data gaps as we recommended would help provide an informed and standardized basis for agencies' decisions on the scope of their data collection in the territories.

7. What did each territory identify as their specific need to improve their statistical operations?

Discussions with territory officials and our analysis of statistical products revealed that data gaps take several forms, including gaps in coverage, disparities and lags in reporting, and different mechanisms for measuring data quality. In these discussions, we learned that a range of factors at the federal and territorial government level contribute to these gaps, which we outlined in our report. Officials from the territories and the Department of the Interior (DOI) reported that DOI's Office of Insular Affairs has a technical assistance program that funds, among other things, efforts to improve data collection and statistical operations in American Samoa, the Commonwealth of the Northern Mariana Islands, Guam, and the U.S. Virgin Islands. (DOI does not have administrative responsibility for Puerto Rico).³ Assessing territory-specific data needs should be coordinated as we recommended. Such coordination would assist the territories to understand their highest priorities, their data needs, their data gaps and what's available, even locally, to assist with the data collection. Such a coordinated approach could include some of the parties that are stakeholders such as universities, which often are capable of producing some of the statistics that territorial agency officials may not have.

8. What did each territory identify as their most important existing gap in federal statistics coverage?

We reported on examples of gaps that territory officials identified as being important to their territories. Officials from American Samoa expressed concerns that some of their farmers may not be captured in federal statistics due to issues with the list that the Census of Agriculture uses. This omission may have the result that farmers who would otherwise be eligible for federal farm benefits may not be. CNMI officials emphasized the absence of reoccurring economic indicators such as employment data from the Bureau of Labor Statistics (BLS), which complicates their ability to understand their labor force. Guam officials reported that there is not enough information collected in the Island Areas Censuses on their smaller populations. This omission is a barrier to their ability to include these populations in other statistical products like the Social Vulnerability Index. USVI officials noted that they have had troubles assessing their recovery and their resilience from the pandemic because they do not have data that is collected by the Census Bureau in the current population survey. Puerto Rico also cannot obtain and use certain detailed information because they are also not included in the Current Population Survey. As we

Page 4

³See Exec. Order No. 12572, Relations with the Northern Mariana Islands, 51 Fed. Reg. 40401 (Nov. 7, 1986); Exec. Order No. 10264, Transfer of the Administration of American Samoa From the Secretary of the Navy to the Secretary of the Interior, 16 Fed. Reg. 6417 (July 3, 1951); Exec. Order No. 10077, Transfer of the Administration of the Island of Guam from the Secretary of the Navy to the Secretary of the Interior, 14 Fed. Reg. 5533 (Sept. 9, 1949); and Exec. Order No. 5566, Placing the Governance of the Virgin Islands under the Supervision of the Department of the Interior, (Feb. 27, 1931).

recommended, OMB should lead the coordination of federal agencies to work with territories to understand their most important data gaps.

- 9. Aside from the Office of the Chief Statistician, are the other federal statistical agencies aware of the gaps in data collection for the U.S. territories?
 - a. Do they agree that this is a serious issue?

To describe the scope of data gaps, we requested and received lists from the 13 principal statistical agencies in the fall of 2023 of what they considered to be their public-facing statistical products and an indication from those agencies of which products include the various territories. We also interviewed officials from principal statistical agencies, inquiring about the extent of agencies' and territories' efforts to mitigate federal data gaps and any challenges encountered in collecting and releasing data for federal statistical products or comparable local efforts. We provided the results of our analysis showing these gaps to all 13 principal statistical agencies, making them aware of our findings. Some agencies have engaged in discussions on this issue, as evidenced by ongoing efforts of the Census Bureau and Department of the Interior to form a working group on territorial data collection.

- 10. In the GAO report, it was stated that federal agencies attributed some of the data gaps to limitations in source data within the U.S. territories.
 - a. What limitations does each territory have?
 - b. Would addressing the data gaps at a federal level be sufficient or would this also require the U.S. territories to address some of these limitations? Why?

Our report provides examples of source data limitations within the territories that territorial and agency officials shared with us. These included the lack of robust and up-to-date address lists for the Island Areas Censuses, as well as the lack of unemployment insurance program data to allow participation in certain employment statistics in American Samoa, the Commonwealth of the Northern Mariana Islands, and Guam. A comprehensive assessment of source data limitations in the territories would be an important part of examining the costs, benefits, and feasibility of including territories in federal statistical products. We recommended that OMB develop a coordinated, government-wide approach to conducting such an assessment.

- 11. Do federal agencies have cost estimates and feasibility assessments for including the U.S. territories in their statistical products?
 - a. If so, what are they?

In 2018, the Census Bureau estimated that extending the annual American Community Survey to territories other than Puerto Rico would cost \$52 million per year. We are aware that the Bureau of Economic Analysis (BEA) also assessed and reported on the feasibility of expanding some statistics to include the territories. These BEA reports are not publicly available but were provided in December 2023 and January 2024 to the House and Senate Appropriations Subcommittees on Commerce, Justice, Science, and Related Agencies.

12. Why do some federal statistic products collect territory for one or some territories, but exclude the rest?

Various factors may contribute to why some federal statistical products collect data for one or some territories but exclude others. Statutes governing and funding principal statistical agencies may differ or be ambiguous about the requirement to collect data from the territories. For example, the Census Bureau is required by law to include the Commonwealth of the Northern Mariana Islands, Guam, Puerto Rico, and USVI in the Island Areas Censuses as part of the nationwide decennial count, as well as in the Economic Census.⁴ The Secretary of Commerce may also, with the Secretary of State's approval, include American Samoa. By contrast, the authorizing statute for BLS does not specify whether to include the territories, which can lead to variation in how territories are treated in BLS statistical products.⁵ In addition, distances from the U.S. mainland can also be a factor, and some statistical products have topics that may not be applicable to all the territories, such as those that measure geography-specific agricultural commodities or energy sources.

13. Why does Puerto Rico receive more coverage in federal data collection than the other territories?

Our report illustrates that there are a range of contributing factors to data gaps throughout the territories, including statutory, statistical, cost, geographic, and technical issues. As we note in our report, Puerto Rico and the U.S. Virgin Islands in the Caribbean Sea are geographically much closer than the Pacific territories to Washington, D.C.—the headquarters of most federal agencies. The distances of the Pacific territories from the mainland create time zone differences, increase travel costs, and contribute to other logistical considerations for federal agencies doing official business with them. Puerto Rico and U.S. Virgin Islands appeared as often as the Pacific territories—or more frequently—in all but one agency's lists of statistical products we reviewed.

14. How does addressing these data gaps sooner rather than later save American taxpayer dollars and improve fiscal responsibility in the long run?

As we reported, accurate, robust federal statistics could help territorial and federal officials better understand the implications of and potential solutions to various challenges that U.S. territories face. Improving statistics by addressing data gaps would better position the territories and federal agencies to manage and mitigate these challenges. Improved statistics would also better inform decision-makers about how to allocate resources to the territories and how to evaluate the effectiveness of those resources.

Page 6

⁴13 U.S.C. § 191(a); 13 U.S.C. § 141(a); 13 U.S.C. § 131.

⁵29 U.S.C. §§ 1-2.

• "Examining GAO's Findings to Address Data Gaps and Improve Data Collection in the Territories."

Questions from Rep. Jenniffer González-Colón of Puerto Rico for Ms. Latesha Love-Grayer, Director, International Affairs and Trade, U.S. Government Accountability Office (GAO)

1. In its report, GAO recommended that the Office of Management and Budget (OMB) should ensure the Chief Statistician develops a government-wide approach for federal statistical agencies to use to examine the costs, benefits, and feasibility of including territories in statistical products and, as appropriate, identify ways to address data gaps.

Could you discuss why, in GAO's view, a whole-of-government approach such as this one is necessary to understand and address federal data gaps impacting the territories? How would this help address the lack of coordination GAO identified in its report?

We reported that data gaps for territories are pervasive, but that agencies' efforts to understand and address territorial data gaps across the federal statistical system have been few and uncoordinated. Further, although OMB is responsible for coordinating the federal statistical system, it has not developed guidance or directives that explicitly address agency data collection efforts in the territories. Moreover, outreach to the Department of the Interior from federal statistical agencies looking to work with the territories tends to occur on a case-by-case basis, according to Interior officials. OMB, in its statutory role of coordinating the federal statistical system (through the Office of the Chief Statistician), is uniquely positioned to help develop a coordinated approach for federal statistical agencies to use to examine and address data gaps. Statistical agencies would then have to work in consultation with territories and stakeholders, as appropriate, to implement this approach. For example, such consultation could include working with the territories to understand their most important data needs and what options are available, including locally, to assist with data collection.

2. When GAO recommends that the Chief Statistician should develop a coordinated, government-wide approach, what do you envision that approach would look like in practice? Would it be OMB developing and publishing a Statistical Policy Directive,

⁶44 U.S.C. § 3504(e).

or distributing general guidance for all thirteen principal statistical agencies? Or a potential plan with short-, medium-, and long-term objectives?

We reported that OMB has not developed guidance or directives that explicitly address agency data collection efforts in the territories. Moreover, outreach to the Department of the Interior from federal statistical agencies looking to work with the territories tends to occur on a case-by-case basis, according to Department of the Interior officials. As a result, we recommended that OMB's Chief Statistician develop a coordinated, government-wide approach. The Chief Statistician has existing mechanisms, such as the Interagency Council on Statistical Policy and the Federal Committee on Statistical Methodology, that could help coordinate policy in this area. An important step in addressing this recommendation would be for OMB to leverage these mechanisms to lead a coordinated discussion with the territories, statistical agencies, and other stakeholders to identify the data gaps that are the highest priorities, the costs of addressing those gaps, and as appropriate, ways to address any data gaps. OMB will need to determine the most appropriate mechanism to use to disseminate the information resulting from that process. Over the coming years we will follow up with OMB to learn what steps it is taking to implement the recommendation and report this information on our website.

3. In my and my office's conversations with federal agencies like the Census Bureau, one of the major obstacles that is always raised when it comes to expanding statistical programs to include Puerto Rico is costs or the need for additional funding.

However, sometimes these statistical products or surveys are sponsored by two agencies—such as the Current Population Survey, which is jointly sponsored by the Census and the Bureau of Labor Statistics—or carried out by one agency but sponsored by another—such as the American Housing Survey, which is conducted by the Census but sponsored by HUD. This means that costs are shared between multiple agencies, and efforts to expand a particular survey or product to include Puerto Rico or the other territories would require interagency cooperation and agreements.

Could you discuss how OMB would be best positioned to address situations like this one?

OMB plays a key convening and sponsoring role with interagency bodies, such as the Interagency Council on Statistical Policy and the Federal Committee on Statistical Methodology. More broadly, OMB has a statutory role and responsibility to coordinate the federal statistical system.

4. As GAO notes in its report, OMB neither agreed nor disagreed with its recommendation on developing a coordinated, government-wide approach. Since the report's release, has GAO heard from OMB or had any follow-up conversations with them as to whether they are taking actions to implement the recommendation?

Is GAO concerned that OMB will not implement the recommendation, given they neither agreed nor disagreed with it?

Since the issuance of our May 2024 report, GAO has not heard from OMB regarding actions that it will take to address our recommendations. As part of our audit processes, we regularly follow up with agencies to determine what actions they have taken to address our recommendations, and we report that information publicly on our website. We plan to follow up with OMB using this process.

5. Now that the report is released, what comes next? How will GAO monitor compliance with its recommendation? For instance, how frequently do you intend to engage with OMB moving forward, if at all?

As part of our audit process, we will follow up regularly with OMB to learn what steps it is taking to implement the recommendation, and we will report that information publicly on our website.

6. In its report, GAO mentions that OMB officials shared that there are plans to open a new data center in Puerto Rico as part of the Federal Statistical Research Data Center Program. Would you be able to provide or share additional information on these plans?

According to the Census Bureau, these centers are partnerships between federal statistical agencies and leading research institutions. The centers provide secure environments supporting qualified researchers using restricted-access data while protecting respondent confidentiality. More information can be found here: https://www.census.gov/about/adrm/fsrdc.html.

7. One of the major concerns with federal data gaps is their potential impact on how federal funding is allocated to the territories. Could you discuss what GAO found regarding this issue? Were you able to identify instances of federal programs allocating funding to the territories differently because of the lack of data?

Data gaps can affect the mechanisms agencies use to allocate funds to the territories. For instance, federal Supplemental Nutrition Assistance Program payments to Guam and USVI are adjusted to reflect the cost of food in these territories based on federal poverty and consumer price data that are not collected the same way as for the 50 states. Without equivalent data, funds are estimated differently for the needs of residents in these jurisdictions than for the rest of the U.S. Additionally, Child Care and Development Block Grants to states and Puerto Rico are allotted based in part on estimates of the ratio of the number of children under 5 years old in a

⁷7 U.S.C. §§ 2012-2014; 42 U.S.C. § 9902(2).

• "Examining GAO's Findings to Address Data Gaps and Improve Data Collection in the Territories."

Questions from Rep. Sablan for Ms. Latesha Love-Grayer, Director, International Affairs and Trade, U.S. Government Accountability Office

1. In our discussion of sample sizes and associated costs, you stated that in smaller populations the cost per person is higher when collecting data. Could you tell us why the cost per person would be different? Could you tell us what the cost per person sampled would be in each of the U.S. territories for a representative sample? And could you tell us what the cost per person would be to collect that same sample in representatives states (Alaska, Hawai'i, Wyoming, Kentucky, California, for instance)?

Costs per person sampled are generally higher for areas that are more remote and where statistical agencies have less existing infrastructure at those locations. We describe in our report how estimation costs can vary based on sample sizes. However, we do not have estimates of these exact costs for specific territories or states. This information would be a valuable part of what can be further examined in a feasibility study, as we recommend.

It is more generally true that estimates tend to be more expensive for smaller populations relative to the size of the populations. This statement follows from statistical proofs that define the relationship between the size of a population and the size of a sample needed to achieve a given level of precision. Critically, the sample size needed for any given level of precision tends to be similar regardless of the size of the population unless the sampling proportion is high.

In the latter case, a high sampling proportion itself implies that the cost of the survey would be higher for the smaller location relative to the population of the location. In addition, statistical agencies often calculate national estimates in conjunction with more localized ones. The larger the sampling proportion required for an estimate, the more time and money would be required above and beyond what the agency would be spending already for the national estimate. Logistical issues can arise when sampling a large percentage of a population. Generally, a sample is much cheaper than a census, and the sample approaches a census as the sampling proportion increases.

To calculate the cost per person for producing an estimate for that population, one would need information on a large number of factors, including but not limited to: the target precision level, the target confidence level, the survey response rate, the survey design, and the planned estimates.

2. How is household or individual data typically collected by the federal government? Are there best practices or technologies—in use by the federal government, in use by other governments or institutions worldwide, or emerging—that can make such collection more accurate, more efficient, less costly? Are such technologies available for use in the U.S. territories?

There are a variety of survey methods used by federal statistical agencies. As we noted in our report, for example, the Census Bureau has evolved its practices to use automation and technology to help mitigate geographic challenges with data collection for the decennial census. More information on the Census Bureau's approaches can be found in 2020 Census: Innovations Helped with Implementation, but Bureau Can Do More to Realize Future Benefits (GAO-21-478). However, we note in our report on territorial data gaps that according to the Census Bureau the lack of a robust, up-to-date address list in some of the territories means that the Census Bureau has to rely on locally labor-intensive, paper-based door-to-door canvassing and enumerations. We did not assess the respective data collection methodologies of the other principal statistical agencies.

3. Could artificial intelligence play a role in extrapolating information about the U.S. territories that could be used as a reliable stand-in for data that may be collected directly or by other means in the rest of the United States?

We have not conducted the work necessary to answer this question.

4. Prior to statehood in 1959, what data collected by the federal government in the 48 states was not collected in Alaska and Hawai'i?

To describe the present scope of territorial data gaps in our report, we focused on information that the 13 principal statistical agencies provided to us in the fall of 2023. As such, we have not conducted the work necessary to answer this question.

5. For agencies and departments GAO surveyed, which have policies on record or are statutorily barred from collecting data in the U.S. territories? And what are those policies or statutes, if any?

We did not identify any relevant statutory language that barred the 13 principal statistical agencies from collecting data in the U.S. territories.

6. What data compiled by the federal government originates from state data collections and what data is collected directly by the federal government itself?

Given our focus on data gaps in the territories, we have not conducted the work necessary to answer this question on state data collections. However, our report includes examples of data gaps in federal statistical products that agencies attributed to limitations in source data within the territories. For example, according to officials with the Bureau of Labor Statistics, American Samoa, the Commonwealth of the Northern Mariana Islands, and Guam are excluded from

certain employment statistics, in part because none has a permanent unemployment insurance program yielding data that the Bureau would use to generate such statistics. Similarly, the Bureau of Economic Analysis produces annual gross domestic product estimates for the territories instead of quarterly estimates in part because the input data are primarily available annually.

• "Examining GAO's Findings to Address Data Gaps and Improve Data Collection in the Territories."

Questions from Rep. Velázquez for Latesha Love-Grayer, Director, International Affairs and Trade, Government Accountability Office (GAO)

1. Did your audit consider data gaps related to the use and impact of federal funding allocations to Puerto Rico, including funds for emergency response and/or postdisaster recovery and reconstruction? If so, what is GAO's evaluation of the available data? How does it differ from data gaps in tracking the use and impact of recurring federal funds?

Numerous factors make it difficult to quantify the effects of data gaps on federal funding to the territories. Principal statistical agencies we interviewed generally do not track the use of data from their products by federal programs or related decisions about allocations of federal funds, such as for emergency response post-disaster recovery and reconstruction. We have not conducted the work necessary to assess available data for specific funding decisions or for tracking the impact of those decisions.

2. Non-profit organizations, community-based entities, and scholars have contributed to highlighting Puerto Rico's primary needs through data collection, analysis, and publications. However, GAO's audit relies heavily on the insight of public officials for its data gaps assessment. To what extent do the gaps identified by GAO address the primary needs at the community level in Puerto Rico, and what can be done to ensure that the voices and needs of people on the ground are present in federal data?

Our report recommended that OMB work in consultation with the U.S. territories and other stakeholders to develop a coordinated approach to identify and examine ways to address data gaps. Our work that led to this recommendation included, among other efforts, interviews with selected stakeholder committees and groups, such as the Association of State and Territorial Health Officials, for information on the existence and effects of data gaps. We also reviewed our prior reporting on how the allocation of federal funds depends on decennial census data. Additionally, we conducted a literature search to identify relevant studies, agency documentation of any analyses conducted on the feasibility of expanding coverage of territories in statistical products, and any associated actions or proposals.

3. Resources are an important aspect when talking about collecting, analyzing, and disseminating data. In the case of Puerto Rico, recent austerity measures hinder these abilities. To that end, could you provide some information as to why the \$60 million allocated to other territories for improving their data were not available to Puerto Rico? What specific technologies are needed in Puerto Rico to address internal limitations?

We reported on a technical assistance grant program administered by the Department of the Interior's Office of Insular Affairs, which has administrative responsibility for coordinating federal policy for the territories, except for Puerto Rico. We did not review internal limitations or weaknesses of any territorial government.

4. According to your presentation, Department of the Interior (DOI) officials were unaware they could include U.S. territories in their data. What specific actions could be taken by the Office and Management and Budget (OMB) to address this situation?

Unlike most of the other agencies we reviewed for our report, DOI is not a principal federal statistical agency. Rather, DOI has administrative responsibility for coordinating federal policy for the territories (except for Puerto Rico). DOI officials told us that federal statistical products may sometimes exclude territories in cases where officials at statistical agencies are accustomed to limiting the scope of their products to the 50 states and are not aware that they could also include the territories. DOI reported that it is taking actions to help educate these agencies on the potential to include the territories in data collection and reporting. However, OMB is positioned to ensure that the federal statistical agencies take action to include the territories in their data collection efforts.

5. GAO stated in its report that a mix of cost, geographic, technical, and other issues have contributed to the exclusion of the U.S. territories from federal statistical products. Do you consider this lack of data a leading contributing factor to the neglect of federal agency attention to the needs of the territories?

We reported that data gaps in the territories can adversely affect territory and federal agencies' abilities to make informed decisions about the distribution of resources in the territories. Our report examined the extent and causes of these data gaps in publicly available federal statistical products, implications for federal funding to the territories resulting from data gaps, and examples of costs and benefits of addressing the data gaps. As we reported, there is not a coordinated, government-wide approach for agencies to use in deciding whether to collect data and report statistics from territories in federal statistical products.