DATA COALITION

Data Coalition | 🐦 @DataCoalition
1003 K Street NW, Suite 200 | ✉ info@datacoalition.org
Washington, D.C. 20001 | 🏠 datacoalition.org

March 23, 2020

RE: Strategies to Support Responsible Use of Government-Collected Data for Understanding Long-Term Impacts from the Coronavirus Pandemic on Public Health, Welfare, and the Economy

<u>**Delivered by Electronic Mail**</u>

Members of Congress –

During this unprecedented time in our country, **the need for valid, reliable data about the challenges facing Americans is clear**. The Data Coalition members strongly encourage our country's elected leaders to ensure we are collecting the data and developing the evidence necessary to understand these challenges, while also planning for our country's future policy needs. Enabling data access while protecting privacy is essential for accomplishing data-driven decision-making; fortunately, a set of bipartisan and well-vetted policy proposals can help address these challenges immediately.

In late-2017 a bipartisan commission created by Congress and the President offered 22 clear recommendations about how to improve the country's data infrastructure to support evidence-based policymaking. While the comprehensive Evidence Commission strategy was designed to work as an entire ecosystem, to date, Congress only taken action on half of those recommendations. **The coronavirus pandemic necessitates action on the remaining unanimous recommendations from the U.S. Commission on Evidence-Based Policymaking** to ensure elected leaders have critical information to understand not just the full impacts of the virus long-term on our economy and population, but also to study and learn from the policies being implemented today to attempt to mitigate the pandemic's effects. In addition to those suggestions, there are several common-sense, bipartisan policy ideas that can immediately improve our ability to measure long-term impacts.

The Data Coalition is America's premier voice on data policy, specifically advocating for data to be high-quality, accessible, and usable to improve our society. On behalf of the members of the Data Coalition – which include technology innovators, data analytics organizations, management consultancies, data vendors, and non-profit research organizations – we ask that Congress prioritize collection and management of high-quality data for addressing the pandemic. **Prioritizing high-quality data and a robust data infrastructure will bolster our country's capabilities to understand and address the current pandemic with sound policies**, while also preparing for future crises in the years ahead.

The Data Coalition's suggestions include ideas that explicitly align with or are complimentary to the Evidence Commission's unanimous recommendations. Each could also be considered to

temporarily address major data issues. Legislative specifications for each are included as an attachment:

1) **Rapidly Authorize and Launch a National Secure Data Service.** The Evidence Commission identified a strategy for quickly developing a shared service center for government data linkage and analytical capabilities that would vastly expand the research community's ability to produce statistical analyses for providing summary information about policy issues. While some work is underway to develop resource, Congress should act to establish the Data Service as quickly as possible enabling these data uses within a privacy-protective framework. The Data Service should be designated as a federal statistical agency and use the privacy framework reauthorized by Congress as part of the Confidential Information Protection and Statistical Efficiency Act of 2018 (CIPSEA; Title III of P.L. 115-435).

   Without this infrastructure in place, government will be unlikely to rapidly analyze and understand the data collected across federal agencies for assessing the full impact of government policies and interventions intended to address the coronavirus pandemic. A short-term authorization could get the infrastructure off the ground, and Congress could revisit the authorization following a "pilot" period to determine longer-term feasibility. Congress could even repurpose the long-delayed Advisory Committee on Data for Evidence-Building to provide transparency and advice as this work proceeds.

2) **Expand Access to Income and Earnings Data for Research Activities.** The Evidence Commission recognized a clear need for improved access to income and earnings data. This can immediately be accomplished in two ways using data already collected by government:

   - **Improve Data Quality for the National Directory of New Hires and Expand Access to Include CIPSEA Agencies.** With minor adjustments to federal law, approved researchers could have vastly improved, yet still restricted, access to existing data on wages and earnings. The Office of Child Support Enforcement at the department of Health and Human Services operates the National Directory of New Hires, created in the 1996 welfare reforms. This system is built on state-provided data. Enhancements to data access for research purposes within the CIPSEA privacy-framework paired with improvements to data quality, including increasing the periodicity of reporting and doubling the duration of data retention will vastly improve existing research capabilities. Notably, proposals to expand access to this system have previously been offered by the Barack Obama and Donald Trump administrations in annual budget proposals to Congress.

   - ***Pass the Measuring Real Income Growth Act of 2019.*** An existing legislative proposal would improve the Bureau of Economic Analysis' capabilities to use existing tax data for developing critical national economic indicators, such as the Gross Domestic Product. In addition to other useful provisions in the legislation, adjustments to 26 USC 6103(j) are especially timely for improving the quality of the country's economic

statistics using existing tax records.

3) **Fund the Evidence Act Implementation with $50 million the FY 2020 Supplemental.**
Congress and the President should provide federal agencies at least $50 million to support implementation of existing data analytic functions created as part of the bipartisan Evidence Act, in a targeted way for supporting the pandemic response. While the Data Coalition is advocating for $125 million in FY 2021, there are limited resources in the current fiscal year. Congress should include additional appropriations in further supplemental spending packages for FY 2020. These resources to implement the Evidence Act would specifically support agencies in rapidly building the capacity and infrastructure to support statistical activities, data governance, and program evaluation. These three functions must be successful across government to truly understand the impacts of the policy decisions that are being made today, as well as to prospectively plan for future crises. This funding should also be flexible enough to support federal statistical agencies in unique challenges associated with collecting critical and relevant data to assess long-term outcomes associated with the COVID-19 pandemic.

4) **Direct Agencies to Track Supplemental Spending through USAspending.gov.** With the influx of federal appropriations to support rapid response to the pandemic, federal agencies need clear direction from Congress about the intent to not only rapidly allocate appropriated funds, but to also ensure for future transparency and accountability of that spending. Federal agencies and the Office of Management and Budget should be advised by Congress to ensure the budget accounting framework enables tracking of federal funds allocated in supplementals similar to the strategy used for the American Reinvestment and Recovery Act of 2009, recognizing new requirements in place as part of the Digital Accountability and Transparency Act of 2014.

5) **Appropriate Funding to Support Development of a Household Survey on COVID-19 Impacts.** In support of efforts underway in the private sector, Congress should immediately appropriate at least $20 million to support initial development and launch of a large-scale household survey to daily monitor the impacts of COVID-19 on the American population at least over the next 6 months, though such a survey would likely be useful for policymakers if continued for a much longer time period.

6) **Improve System for Compiling National COVID-19 Test Data and Relevant Health Information.** The limitations of the country's existing health data infrastructure for monitoring COVID-19 are now abundantly clear, including the absence of a national health data exchange that encourages system interoperability and data sharing consistent with applicable federal laws. Using the existing infrastructure to rapidly shift for COVID-19 response, an improved reporting system with basic data standards could be implemented to ensure the capabilities exists for more timely reporting of test results and relevant vital records, along with capabilities to combine relevant information for improving predictive modeling and long-term analysis of impacts.

Reporting of the CDC's "Human Infection with 2019 Novel Coronavirus Person Under Investigation and Case Report Form" could be mandatory for states, rather than voluntary, with required data elements and electronic submission. These data could be made available in compliance with relevant federal privacy laws for restricted access research-purposes through existing privacy-protective protocols in place at the National Center for Health Statistics, a federal statistical agency, to support rapid modeling in partnership with experts outside government.

Further, mortality reporting at NCHS will be critical in coming months. NCHS must rapidly enhance its capabilities for geocoding, natural language searching, and artificial intelligence as part of the Vital Statistics Reporting System to ensure COVID-19 information is captured adequately, and reporting periodicity from states must become near real-time rather than monthly.

Building on existing partnerships with states, NCHS needs to improve capabilities to provide edits and data quality checks for the existing Electronic Death Reporting System in addition to developing a vastly improved capability for collecting electronic health records from medical facilities in the U.S. to support statistical activities and monitoring. Automating these reviews and providing the data within the CIPSEA privacy framework could rapidly bolster the capabilities. With improved data collection approaches, data scientists could more effectively develop artificial intelligence or machine learning algorithms to predict severity of the outbreak or estimate potential future risks. A supplemental appropriation of $20 million would support rapid development and launch of such a system dedicated to more robust and timely analyses.

7) **Temporarily Suspend the Student Unit Record Ban.** A provision in federal law currently limits the ability of the federal government to collect data that could be useful for analyzing student outcomes in the contemporaneous environment, particularly in higher education institutions. Lifting the ban for statistical purposes under the CIPSEA privacy framework, even temporarily, would ensure policymakers have information available to understand the shift to a virtual educational environment and potential effects on student outcomes for future policy consideration.

8) **Incorporate the Financial Transparency Act with 2020 Tax Credit for LEI Approval**. The Financial Transparency Act takes steps to provide insights across the financial regulatory community with regard to the information available to the public about firms and subsidiaries. This bipartisan legislative proposal would apply expectations to use global data standards for vastly improving financial services data reporting, access, and use. It would also enable long-term benefits for improving the validity and reliability of Gross Domestic Product estimates, among other national economic indicators. Paired with a tax credit for entities to cover the cost of registering for a Legal Entity Identifier and submitting required information, the cost of implementation would be relatively low while the benefits vast in

understanding our economy while reducing reporting burden on the private sector.

Thank you for your consideration of these critical data priorities as our country and partners around the world work to collectively address the coronavirus pandemic. We recognize the magnitude of the challenge before you and hope these suggestions will support your need for reliable information in the months and years ahead. For immediate assistance, please contact me or the Data Coalition's policy manager, Corinna Turbes (corinna.turbes@datacoalition.org).

Respectfully,

Nick Hart, Ph.D.
CEO, Data Coalition