THE NET UNDERCOUNT OF CHILDREN UNDER 5 YEARS OF AGE IN THE DECENNIAL CENSUS: AN ART OF THE POSSIBLE USE CASE

October 18, 2023

Joseph J. Salvo
Institute Fellow,
Social and Decision Analytics Division
http://orcid.org/0000-0002-1552-3387
ckp7gd@virginia.edu

Vicki Lancaster
Principal Scientist
Social and Decision Analytics Division
https://orcid.org/0000-0002-1517-3807
val7zv@virginia.edu

Stephanie Shipp
Acting Director and Professor,
Social and Decision Analytics Division
https://orcid.org/0000-0002-2142-2136
sss5sc@virginia.edu

UVA
BIOCOMPLEXITY INSTITUTE

Citation:
Salvo J, Lancaster V, Shipp S. The Net Undercount of Children Under 5 Years of Age in the Decennial Census: An Art of the Possible Use Case. Proceedings of the Biocomplexity Institute, Technical Report. TR# BI-2023-259. University of Virginia; 2023 October. DOI: https://doi.org/10.18130/nzyj-m621.

The Net Undercount of Children Under 5 Years of Age in the Decennial Census: *An Art of the Possible* Use Case

Joseph J. Salvo, https://orcid.org/0000-0002-1552-3387
Vicki Lancaster, https://orcid.org/0000-0002-1517-3807
Stephanie Shipp, https://orcid.org/0000-0002-2142-2136, Project Lead
Contact: sss5sc@virginia.edu

Social and Decision Analytics Division
Biocomplexity Institute & Initiative
University of Virginia

November 30, 2023

Funding: This research was sponsored by the U.S. Census Bureau Agreement No. 01-21-MOU-06 and Alfred P. Sloan Foundation grant, Grant No. # G-2022-19536

Acknowledgments: The authors would like to acknowledge the contribution of Ford Heilizer, who provided invaluable assistance with analyzing state administrative data on young children. Ford conducted this work while an intern at the Biocomplexity Institute and is now studying Computer Science at the University of Southern California. In addition, University of Virginia undergraduate fellows, Stephanie Zhang and Olivia Davis, provided valuable assistance with the manuscript.

We would also like to thank our colleagues: John Thompson, former Director of the U.S. Census Bureau; William O'Hare of O'Hare Data and Demographic Services; David Dolson, former Director of Statistics Canada; Matthew Snipp, Stanford University; Joan Naymark, Minnesotans for the ACS and 2030 Census (MACS) and Zhengyuan Zhu, University of Iowa, for their comments on earlier drafts of this paper. We would also like to thank our UC Census Bureau colleagues: Jennifer Ortman, Principal Demographer; Eric Jensen, Senior Advisor for Population Estimates and Coverage Measurement; and Sallie Keller, Chief Scientist and Associate Director Research and Methodology Division. (The views expressed in the report are those of the authors and not the Census Bureau.)

Citation: Salvo J, Lancaster V, Shipp S. The Net Undercount of Children Under 5 Years of Age in the Decennial Census: A Litmus Test for the Statistical Products First Approach, *Proceedings of the Biocomplexity Institute*, Technical Report. TR# 2023-259. University of Virginia; 2023 October. DOI: https://doi.org/10.18130/nzyj-m621.

Abstract

The undercount of young children, ages 0-4, is a persistent problem in the decennial census and surveys. It is the largest net undercount of any age group. This *Art of the Possible* Use Case identifies and proposes an approach to fill this data gap. Extending this to a demonstration use case would help to identify and build Curated Data Enterprise (CDE) capabilities to address this gap. The CDE approach goes beyond statistical products derived from single surveys, employing data re-use and integration to allow for accurate, timelier, and granular statistical products that help to fill data gaps and address pressing societal needs.

Since 1980, the deficit of young children has increased relative to adults, reaching an all-time high of 5.4 percent in 2020. Despite the best efforts of the Census Bureau and its local partners, this problem persists and is intractable using current methods. Finding ways to address this coverage error is important, given that funding for early childhood programs is linked to an accurate count of young children. Also, the undercount of young children affects the accuracy and timeliness of the decennial census overall. This paper proposes a new source of data to enhance the counts of children 0-4 years of age as part of the demographic frame being constructed by the Census Bureau. State-based Early Childhood Integrated Data Systems (ECIDS) collect, store, integrate, and maintain administrative data from early childhood programs across multiple agencies within a state. Analysis of these data informs policies, practices, and programs. ECIDS have become more prevalent as computational and data science skills facilitate the creation and use of these data systems. ECIDS offer a potential source of data for identifying young children who are missed in the census and surveys.

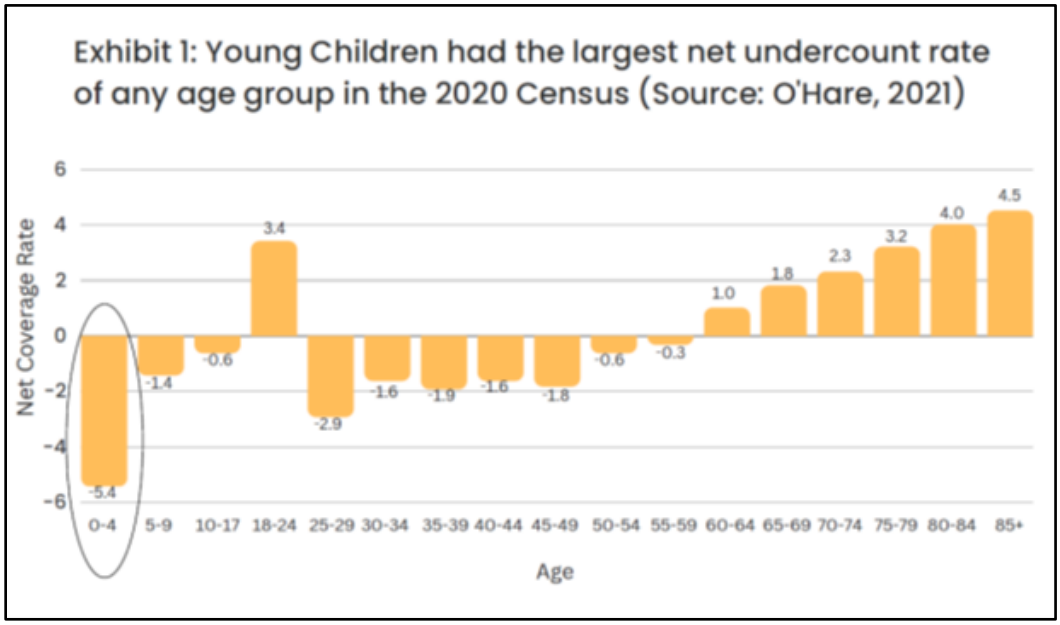
The Net Undercount of Children Under 5 Years of Age in the Decennial Census: A Litmus Test for the Statistical Products First Approach

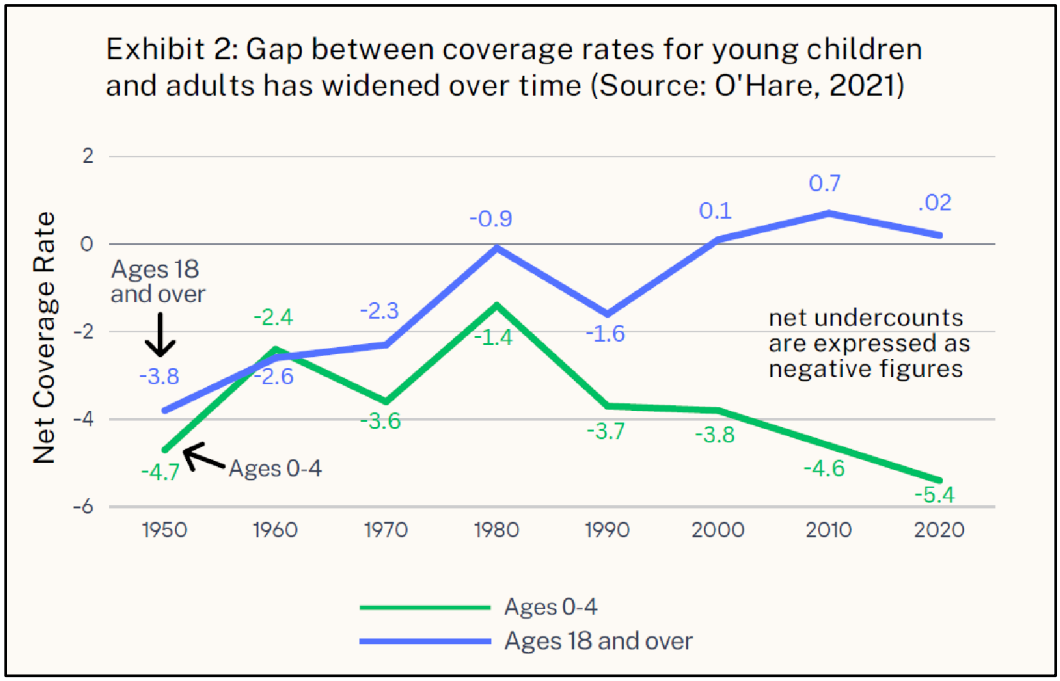
Overview of the Research Problem

The undercount of young children, ages 0-4, is a persistent problem in the decennial census and the largest net undercount of any age group (Exhibit 1). Not only was the undercount of young children very large, but the deficit of young children has increased relative to adults since 1980, reaching an all-time high in 2020 (Exhibit 2). Although precise reasons for this rise are the subject of debate, most research finds that increasingly complex living situations, the disruptions associated with the recent pandemic, and confusion over including those born around census day (April 1) all contribute to this problem. The net undercount of young children in the 2020 Census of 5.4 percent was determined using a technique referred to as Demographic Analysis (DA)[1], which provides a (mostly) independent estimate of the population. (See Box 1 for more detail).

[1] Demographic Analysis, or DA, can be distinguished from the generic use of "demographic analysis," which refers to general methods used by or analyses conducted by demographers. For more information, see Demographic Analysis (DA) (census.gov).

The deficit of young children in the census comes after years of effort to better understand and address the undercount, including campaigns to better inform families about counting their children on census forms. The 2020 Census results highlight the necessity of finding new ways to supplement the enumeration of young children and understand why historical undercount persists.





BOX 1 – Describing the Quality of a Census Count

"Net coverage" of a population (or subgroup) is a term used by demographers to describe the completeness of an enumerated population in a census. The omission of persons who should have been counted but were not is referred to as undercount. An overcount occurs when persons are included erroneously, mainly as a result of duplication or the double-counting of people because of address issues and confusion over where people are supposed to be enumerated (e.g., college students double-counted in their dorms and, incorrectly, at home). Thus, "net coverage" in the census results from offsetting errors – undercounts and overcounts.

Demographic Analysis or DA (capitalized) is a term coined by the U.S. Census Bureau to denote a method used to produce national estimates of the population that are mostly independent of the decennial census for evaluating net coverage of the population. (Sample data from the American Community Survey, which does incorporate some decennial census data for the preparation of estimates, are used for the foreign-born component of DA, hence the "mostly" qualifier.) For the major components of DA, the sources include National Center for Health Statistics birth and death vital records, the American Community Survey data about the stock of foreign-born persons in the U.S., and Medicare enrollment for persons 75 years and over. Net coverage error is calculated as the percent difference between the census count and the DA estimate and is considered the "gold standard" for estimating net coverage. For more information, see Jensen, 2022 and Mulry, 2011.

Existing Research on the Undercount of Young Children

The available evidence shows that the net undercount of young children is a product of omissions – children determined through the Bureau's coverage evaluation methods to be in the population but where the census failed to enumerate them. This was true in both 2020 and in 2010 (Khubba et al., 2022; Griffin and O'Hare, 2020). The undercount is disproportionately born by Black and Hispanic young children. Data from the 2010 Census coverage evaluation showed that estimated omission rates[2] for young Black (17.2) and Hispanic (13.1) children far exceeded those for whites (9.1) (U.S. Census Bureau, 2017: 15-16). Moreover, young children living in more complex arrangements where they were not the biological or adopted child of the householder, those born in the months just before April 1, young children who were grandchildren of the householder and in subfamilies, and those in large multi-unit rental buildings were all more at risk of being omitted (O'Hare et al., 2019; Griffin and Konicki, 2019: Fernandez et al., 2018).

[2] The rates cited are "non-match" rates between the census and the Post-Enumeration Survey (PES) sample. These non-matches are largely omissions but include situations where the lack of a match was due to incomplete information or geographic errors on the census records. In any case, higher levels of non-matches are associated with persons missed in the census and difficulties obtaining an accurate enumeration.

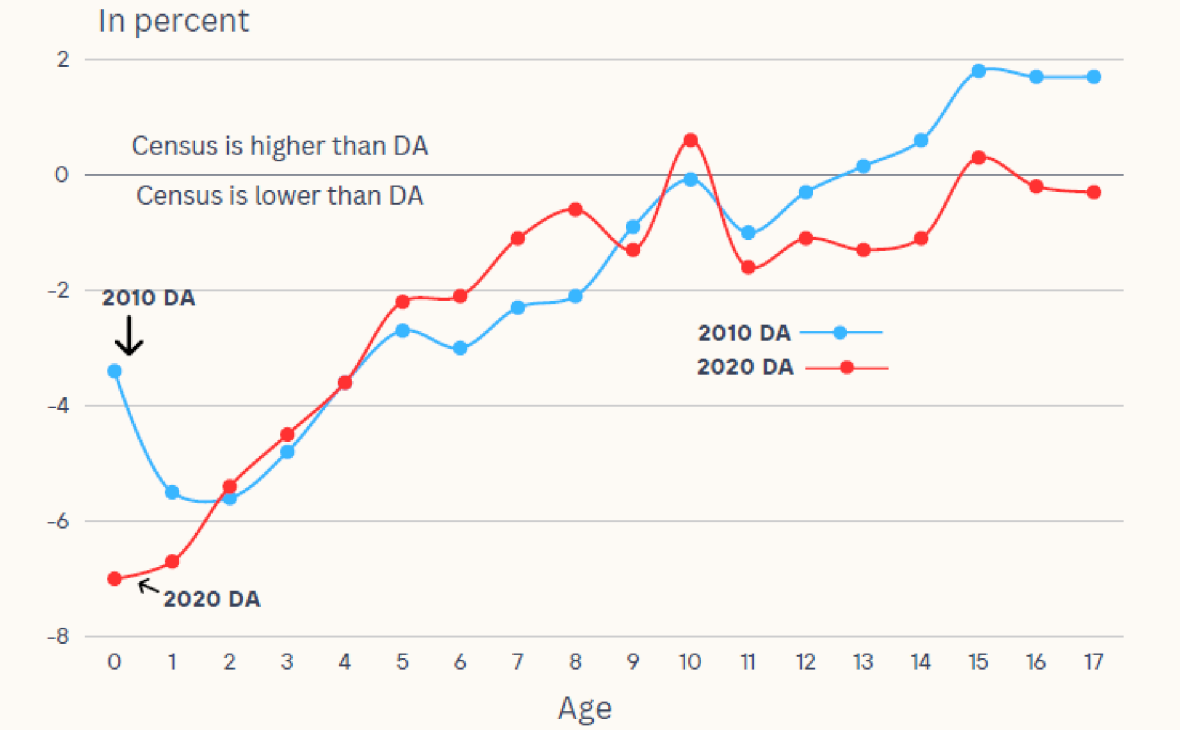
While information from proxy respondents was especially problematic, even in self-enumerated households, some parents mistakenly believed they did not need to include young children, leading to incomplete household rosters (Griffin and O'Hare, 2020). The highest net undercount among those 0-4 was for those under one year of age. The greater undercount for children under 1 in 2020 relative to 2010, shown in Exhibit 3, is difficult to interpret. On the one hand, it is possible that recall related to the composition and status of household members as of April 1 was more of an issue in 2020 compared with 2010, given the delays in the data collection associated with the pandemic. On the other hand, any increase in these recall errors may have been offset by the increased use of automated data collection instruments in 2020 relative to 2010, which may have acted to reduce erroneous enumerations.[3] Finally, Hogan (2019) has shown that the undercount of children under one year of age was likely too low in 2010, the result of coding children born in 2010 without exact birth date information to birthdays in the first quarter of the census year. Some number of these children should have been excluded from the census but were erroneously included, increasing the number of children under one year of age, resulting in an artificially low undercount for this age group.

Although population coverage estimates from the DA method were only available at the national level for 2020, Jensen and Johnson (2021) provided difference "benchmarks" at the county level for young children by comparing the population enumerated in the decennial census with the population estimates produced by the Census Bureau.[4] Using a similar approach, O'Hare (2022) showed that counties in the U.S. with higher estimated net undercounts of young children had greater Black, Hispanic, and American Indian populations, supporting findings noted above.

[3] The increased use of automated checks in the Internet data collection likely resulted in fewer births **after** census day being included erroneously in the 2020 count relative to the 2010 count. Thus, with **more** erroneous enumerations present in 2010 to offset omissions, the result was an artificially low net undercount in 2010 relative to 2020, and may help to explain the wide gap between the two censuses for those 0 to 1 year of age.

[4] It should be noted that the authors were emphatic that these comparisons do not represent coverage errors but provide some idea of the divergence between the estimates and the population count.

Exhibit 3: Demographic Analysis (DA) Estimates of Net Coverage for Children by Single Year of Age: 2010 and 2020 (Source: U.S. Census Bureau 2010 Demographic Analysis, revised, 2012 and 2020 Census Demographic Analysis)



Young children can be missed in several ways. Those answering on behalf of household members may inadvertently leave a young child off the roster. The entire household may be missed because of incorrect or omitted addresses on the Census Bureau's address list. Finally, a subfamily – a mother and child – may be omitted because of confusion over who should be included.

In the years leading up to the 2020 census, greater effort was made to encourage respondents to include young children and their families on their household rosters, in addition to ensuring that all households were enumerated (O'Hare et al., 2019). Through communication initiatives, the Census Bureau 2020 Census campaign with local stakeholders attempted to clarify the need to include young children at a given address.[5] Additionally, using prompts within the internet response option allowed larger and complex families to list all members at a given household more easily. Finally, Master Address File updates and the *Local Update of Census Addresses* (LUCA) program and *In-Office* and *In-Field Address Canvassing* helped to ensure that whole households were not missed.[6] Outside of the Census Bureau, the *Census Counts Campaign*,

[5] https://www.census.gov/about/cac/nac/wg-undercount-children.html
[6] https://www.census.gov/library/stories/2022/03/despite-efforts-census-undercount-of-young-childrenpersists.html

Leadership Conference Education Fund, *Children's Leadership Council*, *Count All Kids,* and the *Population Reference Bureau* all attempted to address issues of undercounting through external research and advocacy for stakeholders and public policy interests.

Why a Complete Count of Young Children is Important

As previously discussed, despite the best efforts of the Census Bureau and their stakeholder partners, the net undercount of young children increased – yet again – between 2010 and 2020. The importance of this problem is related to the purposes and uses to which the data are applied and the difficulties that ensue when the counts of young children are deficient. There are three main categories related to purpose and use:

1. *Acquisition and Distribution of funds* for childcare and education. These areas may draw on census estimates to determine counts of young children, necessary for determining the distribution of Childcare Development Block Grants (CCDBG) and Childcare Development Fund (CCDF) in low-income areas[7]. Furthermore, these funds are coordinated with other sources such as Temporary Assistance for Needy Families (TANF), which can be transferred to CCDBG. Beyond the effects of these programs, studies that focus on young children draw on census results to examine health outcomes, food insecurity, and childhood poverty, which provide a frame of reference for identifying and quantifying needs, and the design of local policies and strategies to address those needs. When children are missed, these efforts are impaired and policy responses are compromised.

   In a recent Census Bureau report, Villa Ross (2023) extended the work of Reamer (2020), to show that the federal government distributed more than $2.8 trillion in FY 2021 to states and localities using – at least in part -- data from the census and related programs (i.e., American Community Survey). The non-profit *First Focus on Children* has assessed the appropriation of federal dollars for programs aimed at providing services for all children, as well as those for those in early childhood (Exhibit 4). In FY2021, with the big increase in appropriations for COVID-19 relief, the total allocation for children was close to $800 billion, with $71 billion appropriated for programs aimed at young children. **Consequently, for local and state-level policymakers to argue for their fair share of what have been increases in federal dollars, they have a strong incentive to ensure that all children are accounted for in census products.**

[7] https://childcareta.acf.hhs.gov/ccdf-fundamentals/history-and-purposes-ccdbg-and-ccdf

Exhibit 4
Federal Spending on Children in the U.S.
FY 2016-2021 (Billions)

	2016	2017	2018	2019	2020	2021*
All Children	$391.44	$396.21	$410.60	$422.75	$500.17	$797.08
Early Childhood**	$16.33	$16.51	$19.53	$19.81	$24.44	$70.96

***The big increase in FY 2021 is related to the passage of COVID emergency relief funding passed in 2020 and 2021, which includes the CARES Act (March 2020), a COVID-19 package in December of 2020, and the American Rescue Plan (March, 2021).**

****This category includes: Head Start and Early Head Start; IDEA B-Preschool Grants; IDEA C-Grants for Infants and Families; Infant and Early Childhood Mental Health; Maternal, Infant, and Erly Childhood Home Visiting Program; National Early Childhood Collaboratives; Preschool Development Grants.**

Source: First Focus on Children, 2021: p 25, pp 43-50.

2. *The Calculation of Rates* for public health relies on denominators that define the population at-risk to illness and disease, as do determinations about the efficacy of an intervention to prevent and address such problems. An accurate number of young children is intrinsic to such evaluations; inaccurate denominators create the potential for rates that are poor representations of what is happening on the ground. This can lead to an inaccurate allocation of resources for healthcare and other critical services, such as pediatric care, immunizations, and early interventions for disease prevention. This is especially the case in low income minority communities, where the net undercount is highest and health care issues most pronounced. Finally, differential undercounts by race and Hispanic origin can hamper the ability of public health officials to accurately characterize disparities in disease incidence and the effectiveness of intervention strategies.

3. *Complete Representation in a Broad Range of Surveys.* The 2020 Census results are used as the basis for weighting and controls[8] for countless surveys – in the federal government, non-profit and private sectors. The numbers by age/sex and race/Hispanic origin are important in adjusting for survey non-response. Thus, deficits in counts regarded as "ground truth" for survey purposes can lead to distorted and unrepresentative results, with shortfalls in the number of young children reflected in key surveys, such as the American Community Survey (ACS). Responding to these deficits, the Census Bureau recently created the "blended base" as the launch point for postcensal population estimates by

---

incorporating adjustments for population coverage by age/sex derived from Demographic Analysis. This includes an upward adjustment for children 0-4 years of age (U.S. Census Bureau, 2022). However, this adjustment is the same for all areas of the nation and does not consider differential undercounts by race and Hispanic origin, or other attributes, such as income and household/family type (which may matter for funding and planning decision making, as discussed below).

Using the Statistical Products First approach for a Better Count of Young Children

The Curated Data Enterprise (CDE) provides a framework for guiding this research. It is the scaffold for the Census Bureau's *Statistical Product First approach (*Keller 2023). At the heart of the CDE framework are the purposes and uses that provide the context for the research (Exhibit 5). The research is iterative, with the inner gold loop primarily focused on the research activities and the outer blue loop primarily focused on engagement activities. The inner gray arrows portray the interactions between the two types of activities.

The CDE provides a framework for an end-to-end curation process for the data and methods needed to develop better counts of young children. End-to-end curation is a process that documents each step of the process, including decisions, tradeoffs, and challenges (Nusser et al., 2023 forthcoming). The process ensures transparency, replicability, and reusability of components built for a Use Case that addresses the undercount of young children.

---

[8] The term "controls" refers to the universe counts of population by age/sex and race/Hispanic origin, that are used as a basis for weighting cases and for adjusting survey estimates for known shortfalls in survey response.

Exhibit 5: Curated Data Enterprise Framework (Source: Keller et.al., 2022)

The CDE aims to produce better counts for decision-makers at all levels of government – federal, state, and local – and for those in the not-for-profit and business sectors. Counts of young children are a critical input into decisions on formulating policies, designing programs, and implementing strategies along with many other applications. This can take the form of:

- A local planning agency trying to get a handle on the cost of implementing a preK program today and moving forward.
- A local public health official trying to measure the success of an early childhood vaccination program by using the number of 0–4-year-olds in the denominator.
- A non-profit creating a plan on how to apply for and best allocate grant resources to support childcare for mothers of young children.
- An academic institution, as part of a cooperative extension program, helps a local government anticipate needed services for young children; by comparing the number of births in year X to the number of children in year X + 4, as means of estimating the net impact of migration (taking into account survival rates) on the number of young children. This is especially important for entry into Pre-K and Kindergarten programs.
- A decision to open a small business or franchise to provide products and services for the children of young families.

What is clear from the evaluation of data on young children is that more than one census or survey source of information will be necessary to create accurate estimates. Administrative records, in various forms, can provide crucial value-added information and help fill the missing survey data void. But administrative data have limitations too, such as missing data and incomplete coverage of the population. This does not preclude their usefulness, however, since all data have limitations.

As data users have made clear and the pandemic has exposed, increased access to timelier data on children needs to be a priority. Existing survey data are insufficient along those lines. Some data sources have the potential to measure children, as with enrollment records and other data on means-tested programs focused on children. However, major challenges include obtaining access to confidential data and the logistics of integrating data from several different state and local sources, especially regarding the identification of duplicate records.

The good news is that the Census Bureau has experience acquiring and integrating state files already has platforms in place, with the Longitudinal Employment Household Dynamics (LEHD) program being just one example. In addition, the recent development of the Demographic Frame as part of the larger *Frames* Enterprise is building in the CDE vision for integrating multiple data sources. The development of a database of persons with age information based on a variety of administrative data, such as IRS Tax Returns and the Social Security Numident file, has the potential to provide an independent estimate of young children (U.S. Census Bureau, 2023).

What is needed is a level of discovery, ingestion, and content curation involving alternative data sources in a model for estimating young children based on continuously updated data analysis. The CDE data curation process selects and leverages existing data, organizes the data and identifies gaps that inform new data collection strategies, all of which adds value to an understanding of what the data represent, and all in the interest of creating better statistical data products. Such an enterprise relies on expertise in understanding the needs of the data user community at the start and consensus on the data that are most useful.

Relevant and current data on young children that can be used to address the undercount are not likely to come from a single source but from a collection of integrated sources of various types, such as administrative data that are created as a result of service provision. These data may be reported for sub-annual periods and tracked to identify trends. Efforts are ongoing in several states and municipalities, made possible by recent advances in technology and data capture, allowing for the synthesis of large data sets. Moreover, organizations have emerged to develop methods and standards for compiling and analyzing these data in what has come to be known as Early Childhood Integrated Data Systems (ECIDS). Building on partnerships with state and local stakeholders, such data have the potential to become the next frontier for the construction of models aimed at more accurate and timelier estimates of young children.

Early Childhood Integrated Data Systems and Estimates of Young Children

The Early Childhood Integrated Data Systems (ECIDS) bring together data sources on young children across multiple state-level agencies that serve young children. These systems have allowed policymakers to respond more quickly and flexibly with up-to-date information about children and families, as exemplified in the COVID-19 pandemic with the closure of daycare and

childcare centers[9]. However, there needs to be a more robust disaggregation of ECIDS data to address questions of economic, racial, and gender disparities. This can be done through more extensive local, state, and stakeholder data sharing practices while working alongside community partners to highlight areas of need. Other data sources that could be included are Medicaid, TANF, SNAP, and school enrollment records, as state-level data could contain more granular information and resolve some of the limitations with federal data sources.

Early work on the development of ECIDS originated in Pennsylvania between 2003 and 2011, which led to federal investment in 2011 and 2013 through the Early Learning Challenge[10]. While the labels differ, ECIDS, in some form, have been implemented in approximately 20 states[11]. However, funding and technological limitations have slowed the ability of states to link data across multiple sources and with different kinds of identifiers. Additionally, data retrieval is also not always easy for specific user purposes. Further expanding ECIDS will require a way to navigate linkage issues while preserving data privacy and security, particularly around education and health data.

Stakeholders such as local organizations, advocacy groups, and welfare providers may also have a role in working with the data and determining purposes that ECIDS can serve[12]. For instance, Minnesota, North Carolina, and Pennsylvania worked with Head Start officials to determine what kinds of questions ECIDS could and could not address. Collaboration in North Carolina and Minnesota allowed stakeholders to get involved early. At the same time, Pennsylvania was able to mitigate concerns of stakeholders and dedicate resources to communicating the benefits of data sharing and privacy safeguards. In moving beyond preschool and K-12 linkages, data that can be included in ECIDS includes Head Start, childcare, early intervention and preschool services, homelessness, and public health data. These will require linkages that span agencies and communities, all of which will benefit from having a centralized data system.

Integrating early childhood data can help develop a coherent approach towards early childhood education and services[13] while revealing services that young children and their families access and addressing issues such as achievement gaps much earlier[13]. The degree of data integration

[9] https://www.childtrends.org/publications/integrated-early-childhood-data-is-an-essential-tool-for-advancingracial-and-ethnic-equity
[10] https://www.newamerica.org/education-policy/edcentral/modernizing-data-systems-matters-for-integratingearly-childhood-data/
[11] https://www.povertyactionlab.org/admindatacatalog/state-level-early-childhood-integrated-data-systems-ecids
[12] https://www.acf.hhs.gov/sites/default/files/documents/ecd/intergration_of_early_childhood_data_final.pdf
[13]https://www.flpadvisors.com/uploads/4/2/4/2/42429949/f_flp_importancemodernizingtechdevelopingecintegrateddatasystems_21june2022.pdf
[13] https://www.acf.hhs.gov/sites/default/files/documents/ecd/intergration_of_early_childhood_data_final.pdf

at the state level varies substantially for various ECIDS. Early Impact Virginia (EIV) brings together data through the Early Childhood Foundation's ECIDS project while working with other stakeholders and policy specialists[14]. Initial funding was granted through the 2009 Statewide Longitudinal Data Systems (SLDS) as an initial step towards ECIDS and towards a centralized model for data development and integration. A more centralized system allows for greater information linkages and accessibility, such as those found in North Carolina and Pennsylvania where all data are consolidated in one database, and linked as opposed to being deidentified beforehand. This model of ECIDS requires more time and poses more security concerns, but it only needs to be matched once and is much faster. Linkages and longitudinal data spanning multiple agencies are much harder to come by and have to be generated, but this poses fewer security concerns. A hybrid model exists where linkages are only done once in a centralized system, using a data feed from individual agencies. There may be similar challenges in ensuring data matches up and reporting challenges, but it can operate off much fewer resources[15].

Key Questions

The ultimate utility of ECIDS for the purpose of improving estimates of young children is a relatively unknown right now; however, this approach holds promise provided that the needed research takes place. In a system aimed at filling-in the gaps that occur in surveys through the preparation of model-based estimates, the data provided from each individual source can be flawed so no single source will suffice. The value added occurs when data elements are combined to produce better model-based estimates, as with programmatic data providing counts of minority children underrepresented in surveys. Thus, the strength of this curation

vision lies in the fact that different datasets contribute unique pieces to the overall estimation model. For this to happen areas of specific inquiry need to be addressed, which can potentially add value to models that estimate the number of young children:

1. **Are there more timely data available on births from localities as part of their ECIDS or related datasets?** (Currently, the Census Bureau experiences a lag in acquiring such data from the National Center for Health Statistics.)

2. **Are there data on program participation that can be used to better estimate those most likely to be undercounted, especially persons of Hispanic origin?** Can ECIDS provide more timely data below the state level on the number of young children at key points early in the lifecycle, where program participation is nearly universal? Some number of young children in minority groups who are underrepresented in surveys are likely to be captured in administrative records for programs that provide essential

[14] https://rga.lis.virginia.gov/Published/2019/RD243/PDF
[15] https://nces.ed.gov/programs/slds/pdf/ECIDS_System_Model.pdf

services, such as Head Start or Temporary Assistance for Needy Families (TANF). Right now, adjustments for overcounts and undercounts are applied using national rates, which can differ dramatically by geographic area and for specific subgroups known to experience undercounts (e.g., race/ethnic groups).

3. **Are there longitudinal data that can allow for "tracking" births from birth to pre-K and Kindergarten, using enrollment data?**

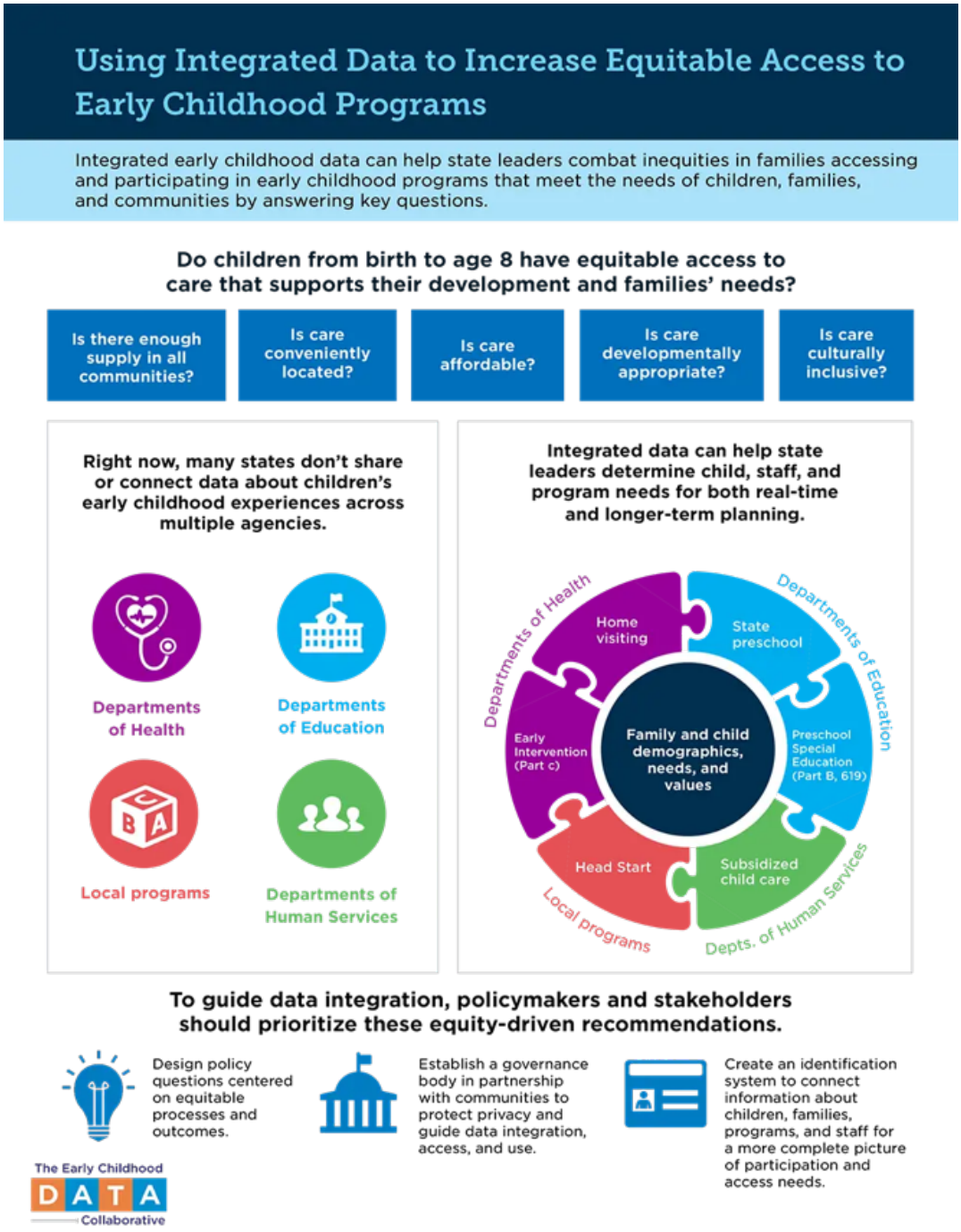
4. **Can the Census Bureau obtain data for individuals on a confidential basis through a Memorandum of Understanding (MOU), which can be linked to data from the decennial census and used to improve census and/or population estimates for young children?** The Census Bureau's *Frames* program utilizes multiple sources of administrative records to build the demographic frame and has made use of various types of state data. Given the undercount of young children, a logical extension of the demographic frames work would be to expand the use of local data to include information specific to young children. Can ECIDS-related systems provide actual records for children that are likely to be missed in the census and in sample surveys?

Collaborating with ECIDS-Related Networks

The growth of ECIDS and their potential to inform local public policy decisions has produced a great deal of interest in the methods, standards, and dissemination of the data that they collect. A number of groups concerned with issues affecting children, along with those interested in promoting data-driven public policy, have focused on providing expertise to states and localities on the ECIDS front. An example is *Child Trends,* a data collaborative that provides early care and education (ECE) data for use by state policymakers. Since 2009, the Early Childhood Data Collaborative (ECDC) provides tools and resources to encourage data-driven state policy and provides a national forum to support the development and use of coordinated state ECIDS[16]. This model is germane, given the disparities exacerbated by the pandemic and the need for the development of more equitable systems to guide the collection, integration, and sharing of data to support better outcomes for young children (Exhibit 6).

[16] https://www.childtrends.org/publications/integrated-early-childhood-data-is-an-essential-tool-for-advancingracial-and-ethnic-equity

Exhibit 6



Increasingly, data from local agencies are being integrated in an effort to address local needs, on this case access to Early Childhood Programs (Source: *Child Trends*[18]

Similarly, Actionable Intelligence for Social Policy (AISP) was born out of a desire to leverage data from a number of sources to improve the lives of children, through a network formed by two faculty members at the University of Pennsylvania, with an early focus on the city of Philadelphia in the early 2000s[19]. Since then, the goal has been to develop an inventory and connect with other similar data efforts nationally, with a focus on standards and best practices for data sharing. This has led to 36 sites in 23 states and municipalities, where information on the legal, administrative and technical requirements is exchanged on a regular basis. A 2018 survey by *Child Trends* found that the linkage of local data for young children presented formidable challenges among the states, but incremental improvements were being made[17].

Use Cases

Efforts to navigate the myriad requirements of acquiring data to improve the estimates of young children needs to start with Use Cases that employ data from systems where there is some indication that a readiness to do so exists. This includes states that have been at the forefront of developing the content and methods that undergird these systems. Research by the Social and Decision Analytics Division (SDAD) at the University of Virginia Biocomplexity Institute has sought to identify those states that can serve as Use Cases for testing the feasibility of using ECIDS information to model the enumeration of young children grounded in local data. An initial Census Bureau collaboration with just one or two states is the first step in the discovery of relevant data, accessibility, and fitness-for-use.

Luckily, there are at least two entities that have taken on the challenge of promoting data sharing related to educational program participation and outcomes. The Educaqon Data Warehouse Soluqon by *eScholar* is a longitudinal data warehouse that allows agencies to collect, store, and analyze data from a variety of sources. These data can be used to improve student achievement, idenqfy at-risk students, and make informed decisions about policies and programs. Several states make use of *eScholar* software to create and maintain the contents of their warehouses[18] and are part of the *eScholar* warehouse. The value of e*Scholar* lies in the way it creates and manages unique idenqfiers for students, staff, and other individuals. This informaqon can be used to improve data accuracy and security and make tracking individuals across different systems easier. For example, a student's unique idenqfier can be used to track their parqcipaqon in programs. These idenqfiers are unique to individuals, much like Personal

[18] https://www.childtrends.org/publications/using-integrated-data-to-increase-equitable-access-to-earlychildhood-programs) [19] https://aisp.upenn.edu/

[17] https://cms.childtrends.org/wp-content/uploads/2018/09/ECDC-50-state-survey-9.25.pdf

[18] eScholar is an educational technology company that offers a range of products and services aimed at helping educational institutions collect, store, and utilize data effectively. See: https://escholar.com/solutions/educationdata-warehouse/

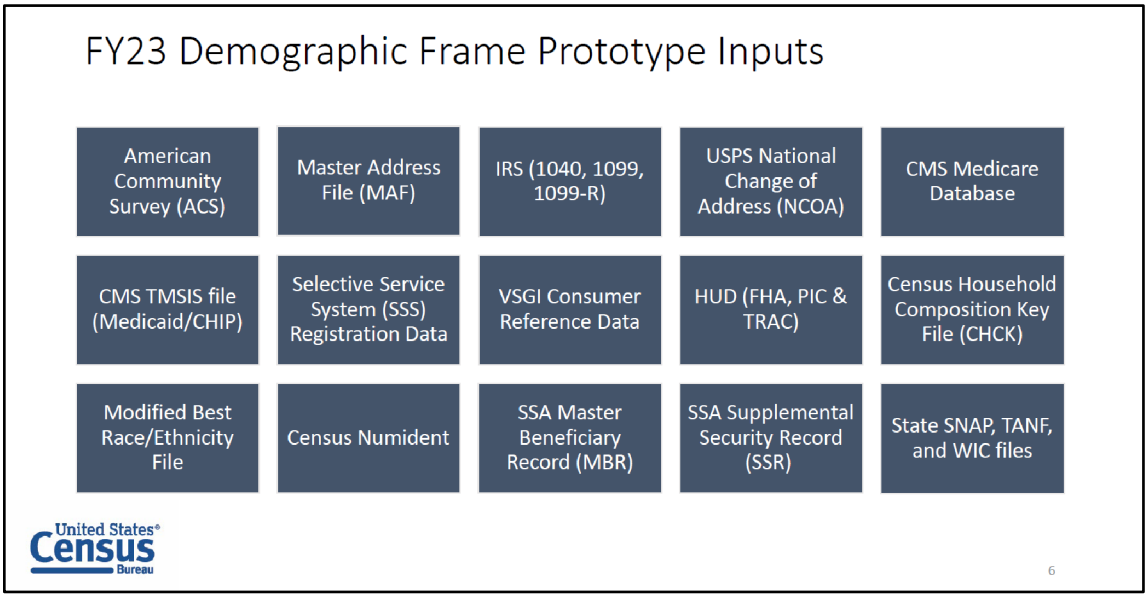
Idenqficaqon Keys (PIKs) at the Census Bureau being deployed in *Frames*, and not only provide confidenqality but are commonly used to establish an unduplicated count of persons. Similar data sources along these lines may be found with State Longitudinal Data Systems (SLDS) or P20W data iniqaqves, which expend resources "following" individuals from early childhood through school and into the workforce[19]. Entities such as the Data Quality Campaign[23] have made governance and the development of standards priorities as part of their advocacy for state and local longitudinal data development.

Readiness for Collaboration at the State Level

Ideally, data through a collaboration for young children could be provided by a state entity to the Census Bureau on a confidential basis through a Federal Statistical Research Data Center (FSRDC) for linkage to Census Bureau data from 2020. Such an exercise could be integrated into existing platforms being developed for persons as part of the demographic frame for the larger *Frames* program, where multiple inputs are being evaluated and probabilistic matching employed to create an unduplicated dataset of persons (U.S. Census Bureau, 2023). Such data may enhance counts from current census programs and, ultimately, the 2030 Census. The unique characteristics that local ECIDS-related information can bring to the table in the form of identifying young children missing from the censuses and surveys have the potential to enhance this process. This is especially the case, given that the same programs used to identify these children as part of ECIDS are programs frequented by the very children missed in surveys. Moreover, ECIDS can likely provide additional data that are especially pertinent to young children and go beyond the state files already available or proposed for use by the Census Bureau (Exhibit 7).

[19] https://dataqualitycampaign.org/wp-content/uploads/2022/05/DQC-P-20W-fact-sheet.pdf [23] https://dataqualitycampaign.org/

Exhibit 7



While ECIDS-related data exist in some form in 31 states, this review of the content, methods, and accessibility of data put some states in a good position for potential collaboration. The 11 states, highlighted in Exhibit 8, possess data that could be used to inform the estimation of young children as a component of the *Frames* program and show a willingness to share information. Still, it is important to bear in mind that this conclusion is preliminary and awaits a test of feasibility on multiple dimensions: content, methods, administrative, and legal/contractual. The point is that this problem needs to be researched further – beyond discovery – using the Curated Data Enterprise Framework to assess if local data and methods will enhance the Census Bureau's ability to enumerate young children. Again, no single survey or administrative dataset is perfect, but together, value is added based on local input.

To assess the feasibility of using ECIDS to address the undercount of young children, we completed an inventory of ECIDS data by state, using publicly available information posted on their websites. While this exercise was useful in identifying those states where there was some indication that key data were present, it was just an initial limited look of what may be available. We used several criteria to make judgments and "leaps of faith" regarding how ready a state would be to collaborate on a Use Case involving better counts of young children. It is important at the outset to emphasize that this exercise was limited by what was posted on State websites and may neither reflect the data that are actually available nor the willingness of states to work with external collaborators. Nevertheless, we identified 31 states with ECIDSstyle data programs and searched, summarized, and evaluated the readiness of a state ECIDS program to collaborate using four main criteria:

- Availability of data for young children across different local agencies
- Evidence of data integration and quality control for analytical applications related to local issues, such as program planning and resource distribution
- A receptivity to external requests for data
- Documentation on sources and methods

The descriptions vary because each state is different in how they bring together, manage, and allow researcher access to the data. For example, in some states, the early childhood information is linked as a subset of the K-12 data, and follow-up is required in order to determine the degree to which early childhood data are available as a separate entity. In several cases (e.g., Mississippi and Utah), any such determination from their website presentations is almost impossible to determine; nonetheless, there is enough there to make follow-up worthwhile. This underscores an important point about the list in Exhibit 8. *This list below is meant to be the beginning of a data discovery exercise and is not intended to be a final statement on the availability of data on young children.* It is meant as a starting point for follow-up, to see what data may be available. It is entirely possible that some states on the list do not have the data available for collaboration, while others not on the list do indeed have such data.

Exhibit 8. Promising State ECIDS for Potential Collaboration

State	SLDS/ECIDS Name	Link	eScholar*
Pennsylvania	Pennsylvania Informaqon Management System	Link	Yes
Minnesota	Early Childhood Longitudinal Data System	Link	No
Georgia	Cross Agency Child Data System	Link	No
North Carolina	Early Childhood Integrated Data System	Link	Yes
Kentucky	KYStats	Link	No
Iowa	Iowa Integrated Data System for Decision Making	Link	No
Mississippi	LifeTracks	Link	No
Utah	Early Childhood Integrated Data System	Link	No
Connecqcut	P20 WIN	Link	No
New Mexico	New Mexico Longitudinal Data System (NMLDS)	Link	Yes
Louisiana	EdLink 360	Link	Yes

Note: **eScholar** is a software company that provides software and services used by educators for student record management and career planning.
*There may be some error in which states use eScholar, as not all states report their contracts.
Below is a list of states with systems we think are most ready for potential collaboration

Listening Session on Data Sharing to Improve the Counts of Young Children

As part of Data Discovery, the University of Virginia's Biocomplexity Institute conducted a listening session with key state representatives who develop and use ECIDS-related data. Four states on our high readiness list participated: North Carolina, Minnesota, Iowa, and Kentucky all provided information on the data that they curate and utilize. Early indications are that these states possess records from a variety of programs that may be helpful as part of the Census Bureau's demographic frame. Data from education, human services, birth and related health records, all of which may improve coverage. Having high-quality data on children who are most likely to be missed in the Census, such as those who are poor or in need of government intervention, is important in correcting the undercount of young children.

It is impossible, however, to know how helpful such data may be until a sample of their records can be obtained for a proper evaluation. There are a variety of administrative and legal hurdles identified by participants that need to be overcome in order for a successful collaboration to occur. A first step is to develop data sharing agreements with the UVA Biocomplexity Institute, for aggregate statistics that can be compared with the 2020 decennial census. Participants from the states indicated that it would be easier to share their data with UVA initially, than sharing directly with the Census Bureau, given the Bureau's more formidable legal and administrative requirements. Most important, despite the difficulties with the data sharing agreement process, the state participants were excited to find ways to contribute their data as a potential pathway to improving the counts of young children.

Conclusion

Looking forward, there is little doubt that new, creative ways of addressing the undercount of young children will be required. The Curated Data Enterprise (CDE) proposes to combine multiple data sets to create new statistical products on young children. Each data source has its limitations, but taken together, stands to provide a better product. This paper has highlighted one pathway to better counts of young children – local administrative records as part of the Early Childhood Integrated Data Systems (ECIDS) movement among the states. As data science is increasingly seen as a path to address local problems, more local governments are joining the ranks of the data savvy. While it is impossible to tell, at first glance, whether ECIDS will be of benefit to those who are charged with developing better data products for young children, it is in the spirit of data discovery within the CDE to find out. Early indications are encouraging enough to warrant a move to the next level – obtaining a sample of records, initially in aggregate form, to see how these data compare with data from the 2020 decennial census.

Appendix A: Descriptions of Promising State ECIDS

Eleven state ECIDS systems are briefly described below. Each state takes a different approach. There are similarities across states as well. Some integrated the ECIDS system into their State Longitudinal Data Systems that provide detailed K-12 student and teacher data by school (e.g., classes taken, grades obtained, teacher credentials, and much more). Other ECIDS systems are separate systems. Some allow researcher access through an online application process, others allow direct requests to the agency or university center that acts as a liaison, and some will do the calculations for the analyst requesting the data.

1. Pennsylvania:
   https://www.education.pa.gov/DataAndReporting/PIMS/Pages/default.aspx

   Pennsylvania's Information Management System (PIMS) is a longitudinal, nationally recognized program and was one of the first programs. It uses the *eScholar* backend, which provides a suite of software products many states use. All pre-K through grade 12 and Adult Affidavit Program (AAP) students enrolled in a public school and students enrolled in an Approved Private School (APS) or Private Residential Rehabilitation Institution (PRRI) in Pennsylvania, are in the PIMS[20], which includes hundreds of variables student DOB's, grade level, ethnic and racial classification, course enrollment, and academic performance. However, there is limited data tracked on student participation in welfare programs. Data dashboards[21] provide an overview of what is and may be available upon request.

2. Minnesota: https://ecIds.mn.gov/
   Using data from a whole array of state agencies and departments, the Early Childhood Longitudinal Data System (ECLDS) provides an interactive platform for database queries[22]. Decision-makers use this tool to make access programs efficacy and the return on investments, now and looking forward. In addition to using a single platform to integrate data, this effort emphasizes the *linkage* between data from different agencies[23] to see what outcomes were generated across multiple programs. Thus, the interactive nature of this effort and the links between the data provide valuable examples of what should be possible on the ECIDS front.

[20] Note: APS, PRRI, and schools for individuals with special needs or circumstances should not be confused with private schools for non-disabled students.
[21] https://www.dhs.pa.gov/about/Pages/Early-Learning-Dashboards.aspx
[22] https://ecIds.mn.gov/#data
[23] https://ecIds.mn.gov/#about

The Minnesota Statewide Longitudinal Education Data System (SLEDS) and ECLDS are complementary systems providing useful functionality. The P20W data warehouse serves as an umbrella structure for two data systems – ECLDS, birth to grade 3 data, and the SLEDS, kindergarten through postsecondary, and workforce. Program participation information is vast and includes a student identification process for children registered in Early Childhood Family Education (ECFE) and School Readiness to better understand children's early learning experiences before kindergarten. Also, there are programs for children age three to pre-K to prepare children to enter kindergarten. Most important are programs aimed at Early Childhood Screening (ECS), a free program required for **all** children before kindergarten. Required screening components include: developmental screening for risk factors affecting potential learning issues prior to school entry.

3. Georgia: https://www.gacacds.com/
   Georgia's Cross Agency Child Data System (CACDS) explicitly focuses on the demographic data relevant to early childhood. Many programs are almost universal for children of certain ages, especially programs sponsored by the Departments of Public Health and Education and the Georgia Head Start Association. The data are compiled, housed, and managed by the Department of Early Care and Learning. The system contains data on early start, head start, home visits, childcare/parent services, and Census Bureau demographic data[24]. Standard data are open access for children greater than 5 and for children 0-5, a request process is being set up for CACDS.

4. North Carolina: https://www.ncdhhs.gov/about/department-initiatives/early-childhood/earlychildhood-data/north-carolina-early-childhood-integrated-data-system

   North Carolina ECIDS provides high quality linked data across several government programs. Included in ECIDS are data from Food and Nutrition Services, Infant Toddler programs, and pre-K. The ECIDS is intended for public, state agency, and policymaker use, with approval processes in place for research purposes. NC ECIDS provides counts of children who receive multiple early childhood services from participating programs. Some of these data are publicly available,[29] other data require a special request.

5. Kentucky: https://kystats.ky.gov/
   The Kentucky Center for Statistics (KYSTATS) collects and links data to evaluate education and workforce efforts in the Commonwealth. This includes developing reports, responding to research requests, and providing statistical data about these efforts so policymakers, practitioners, and the general public can make better informed decisions. KYSTATS provides a webpage for researchers to request data:

[24] https://www.povertyactionlab.org/admindatacatalog/state-level-early-childhood-integrated-data-systems-ecids

https://kystats.ky.gov/Reports/DataRequest . This indicates a willingness to engage with researchers. Birth record data appear to come from the Kentucky State Data Center (KSDC). KSDC is the state's lead agency for the Census Bureau's State Data Center Program and is Kentucky's official clearinghouse for Census data. KSDC lists research on their website with the Bureau using administrative records to create population estimates.[25]

6. Iowa: https://i2d2.iastate.edu/
   Iowa's Integrated Data System for Decision-Making (I2D2) is a state-university partnership that is about the integration of administrative data from public and private agencies, aimed at better decisions for Iowa's children and their families. I2D2 is governed by strict data sharing and data access protocols that ensure data are securely used to inform Iowa priorities. I2D2 is a robust system of stakeholders who share a common goal to bring together comprehensive systematic data regarding Iowa's children and families[26]. The focus is on integrating data from a variety of sources[27] – successfully dealing with the technical, administrative, and legal challenges that come with linking records and making those data available to decisionmakers[28]. (Iowa I2D2 is part of the Actionable Intelligence for Social Policy (AISP) network.) Follow-up has indicated that researcher access can be challenging, as multiple sign-offs are likely required by agencies that contribute data to the I2D2 system.

7. Mississippi: https://lifetracks.ms.gov/PK12/Home.aspx
   The *LifeTracks* system in Mississippi receives data from over 20 different agencies. The public interface is lacking and how to access detailed information cannot be determined. There are some indications that higher quality data is collected and that

[29] https://www.ncdhhs.gov/about/department-initiatives/early-childhood/early-childhood-data/northcarolina-early-childhood-integrated-data-system

   the public interface is not representative of the capabilities of the system. It is worth a further look, especially since online requests for data access are possible.[29]

8. Utah: https://ecids.utah.gov/
   Utah's ECIDS integrates early care and education data from participating agencies and programs providing services to families with young children and is hosted by the Utah Department of Health and Human Services. The goal is to evaluate outcomes related to participation in early case and education (ECE) programs, and help agencies with funding opportunities to support these programs. Their mission is to facilitate the use of

[25] See http://ksdc.louisville.edu/research/kprc/
[26] Iowa's Integrated Data System for Decision Making | I2D2 (iastate.edu)
[27] https://i2d2.iastate.edu/portfolio_page/eci-longitudinal-study-ecils/
[28] https://earlychildhood.iowa.gov/document/indicators-and-data-dictionary
[29] https://lifetracks.ms.gov/Account/AccountRequest.aspx

integrated data to ensure young children receive the care and support they need to enter school healthy and ready to learn. They are developing data-sharing agreements and procedures that will provide researchers access to ECIDS data sources. These include de-identified, unit-level, matched student assessment data housed at Utah Data Research Center.

9. Connecticut: https://portal.ct.gov/OPM/P20Win/
   P20 WIN is a longitudinal data system, which brings together data across state agencies and other organizations to promote data-driven policy decision making. The Data and Policy Analytics (DAPA) unit oversees the collection, analysis, coordination, and sharing of data at the Office of Policy and Management. The Office of Early Childhood (OEC) is one of the agencies that work with DAPA to provide information on young children and their program participation. This system has extensive processes in place for sharing data among state agencies that puts it in a good place to discuss the sharing of data with the Census Bureau.

10. New Mexico: https://rise.nm.gov/about-us

    Known as RISE-NM, this is a Longitudinal Data System Project, that aggregates and links data from partner agencies into a single cloud data platform to evaluate education and workforce efforts in New Mexico. One of their partners is the Early Care and Education Department. Educaqonal program parqcipaqon and student enrollment data are among the data categories available through this site. The goal is to develop one data environment where all data are brought, analyzed, and matched by state-of-the-art technology as a cost-effecqve way to promote data-driven decisions and provide beuer access to services. Very recently, a data sharing process has been put in place for organizaqons within the state.

11. Louisiana: hups://www.louisianabelieves.com/resources/library/data-center/research
    Through its EDLINK 360 and a process for special requests, this state provides access to a variety of data and staqsqcs related to early childhood, as part of the subject mauer

---

seen here. In addiqon, they have a process in place for collaboraqon with other agencies on research projects, both within the state and with external researchers.

References

Becker-Medina, Erika and Nate Ramsey. (2023) "FY23 Priority 3: Engagement, Addressing New and Complex Data User Needs." U.S. Census Bureau, February 14 Presentation.

Fernandez, Leticia, R. Shattuck, J. Noon, (2018). Using Administrative Records and the American Community Survey to Study the Characteristics of Undercounted Young Children in the 2010 Census. Paper Presented at the Annual Forum of the Federal Committee on Statistical Methodology, November. https://nces.ed.gov/FCSM/pdf/H_4FernandezetalARChildUndercountFCSM2018.pdf

First Focus on Children. (2021). Children's Budget 2021. Washington DC, November. https://firstfocus.org/resources/report/childrens-budget-2021

Griffin, D., and O'Hare, W.P. (2022). "Counting Young children in the U.S. Census – Important Differences Between 2010 and 2020," Posted on the Leadership Conference Education Fund website. March 29, 2022. https://civilrightsdocs.info/pdf/reports/Counting-Young-Children-inthe-Census-Report.pdf

Griffin, D., and O'Hare, W.P. (2020). "Are Census Omissions of Young Children Due to Respondent Misconceptions about the Census? International Journal of Social Science Studies, Vol 8. No. 6 November pp 59-72. http://redfame.com/journal/index.php/ijsss/article/view/4994/5223

Griffin, D., and S. Konicki (2019) "Investigating the 2010 Undercount of Young Children – Further Analysis of Census Coverage Measurement Results." U.S Census Bureau, January 15. https://www.census.gov/programs-surveys/decennial-census/decade/2020/planningmanagement/plan/final-analysis/2020-report-2010-undercount-children-further-analysiscensus-coverage.html

Hogan, Howard R. (2019). "Investigating the 2010 Undercount of Young Children – Net Census Coverage of Very Young Children." U.S. Census Bureau, January. https://www2.census.gov/programs-surveys/decennial/2020/program-management/final-analysis-reports/2020-report-2010-undercount-children-net_census_coverage.pdf

Jensen, Eric (2022). "Census Bureau Expands Focus on Improving Data for Young Children. U.S. Census Bureau, March. https://www.census.gov/library/stories/2022/03/despite-effortscensus-undercount-of-young-children-persists.html

Jensen, E. and Johnson, S. (2021). "Using Demographic Benchmarks to Help Evaluate the 2020 Census results." U.S. Census Bureau, Washington, DC. Nov. https://www.census.gov/newsroom/blogs/random-samplings/2021/11/demographicbenchmarks-2020-census.html

Jensen, Eric. (2019). Investigating the 2010 Undercount of Young Children – Examining Coverage in Demographic Surveys. U.S. Census Bureau.

https://www2.census.gov/programssurveys/decennial/2020/program-management/final-analysis-reports/2020-report-2010undercount-children-summary-recent-research.pdf

Keller S, Prewitt K, Thompson J, Jost S, Barrett C, Nusser S, Salvo J, & Shipp S. (2022) A 21st Century Census Curated Data Enterprise. A Bold New Approach to Create Official Statistics, Technical Report BI-2022-1115. Proceedings of the Biocomplexity Institute, University of Virginia; 2022. DOI: https://doi.org/10.18130/r174-yk24 .

Keller, Sallie Ann (2023). "Evolving a Data Enterprise to Support Relevant, Timely, and Equitable Statistical Products – We Need Your Help." Presentation to the Association of Public Data Users, U.S. Census Bureau May 11. https://www.apdu.org/events/census-bureau-seeks-input-onplans-to-re-envision-its-data-enterprise-05-2023/

Khubba, Shadie, Krista Heim, and Jinhee Hong. (2022) "2020 Post-Enumeration Survey Estimation Report: National Census Coverage Estimates for People in the United States by Demographic Characteristics," U.S. Census Bureau, PES20-G-01, Washington, DC, U.S. Government Publishing Office, March. https://www2.census.gov/programssurveys/decennial/coverage-measurement/pes/national-census-coverage-estimates-bydemographic-characteristics.pdf

Mulry, Mary H. (2011). Encyclopedia of Survey Research Methods. Edited by Paul J. Lavrakas, Sage Publications. https://methods.sagepub.com/reference/encyclopedia-of-survey-researchmethods/n115.xml

Nusser, Sarah, Sallie Keller, Zhengyuan Zhu, Ed Wu, Stephanie Shipp. Proceedings of the Biocomplexity Institute, University of Virginia (Forthcoming).

O'Hare, William (2023). Counties with High Undercounts of Children in 2020 U.S. Census. https://2hj858.a2cdn1.secureserver.net/wp-content/uploads/2023/03/Counties-with-HighUndercounts-of-Children-in-2020-U.S.-Census.pdf

O'Hare, W. P. (2022). Race and Hispanic Origin Composition of Population in Counties with High Net Undercount Rates for Children in the 2020 Census, presentation at the Southern Demographic Association annual conference, October17-19, Knoxville, TN. October.

O'Hare, William. (2022a). March). New Census Bureau Data Show Young Children Have a High Net Undercount in the 2020 Census. Count All Kids. March. https://countallkids.org/resources/newcensus-bureau-data-show-young-children-have-a-highnet-under

O'Hare, W.P. (2021). "New Census Bureau Data Show Young Children Have a High Net Undercount in the 2020 Census," March, Count All Kids website. https://countallkids.org/resources/new-census-bureau-data-show-young-childrenhave-a-high-net-undercount-in-the-2020-census/

O'Hare, William P., D. Griffin, and S. Konicki (2019) *Investigating the 2010 Undercount of Young Children –Summary of Recent Research.* February 14, 2019. U.S. Census Bureau. Washington, DC. Retrieved from https://www.census.gov/programs-surveys/decennialcensus/decade/2020/planning-management/plan/final-analysis/2020-report-2010-undercountchildren-summary-recent-research.html

O'Hare W.P. (2017). "Geographic Variation in 2010 U.S. Census Coverage Rates for Young Children: A Look at Counties," *International Journal of Social Science Studies,* Vol. 5, No. 9 Sept. Redframe Publishing, 295180286.pdf (core.ac.uk)

O'Hare, W.P. (2015). *The Undercount of Young Children in the U.S. Decennial Census. Springer Publishers.* https://www.springer.com/gp/book/9783319189161#

Reamer, A. (2020). Counting for Dollars 2020: The Role of the Decennial Census in the Geographic Distribution of Federal Funds. George Washington Institute of Public Policy, George Washington University. https://gwipp.gwu.edu/counting-dollars-2020-role-decennial-censusgeographic-distribution-federal-funds

U.S. Census Bureau (2023). "Demographic Frame: Leveraging Person-Level Data to Enhance Census and Survey Taking," Presentation to the UVA Census Working Group Meeting, May.

U.S. Census Bureau (2022). " Adapting Population Estimates to Address COVID-19 Impacts on Data Availability," Luke Rogers and Christine Hartley, U.S. Census Bureau April 14. https://www.census.gov/newsroom/blogs/random-samplings/2022/04/population-estimatescovid-19-impacts.html

U.S. Census Bureau (2017). "Investigating the 2010 Undercount of Young Children – Analysis of Census Coverage Measurement Results." January. https://www2.census.gov/programssurveys/decennial/2020/program-management/final-analysis-reports/2020-2017_04undercount-children-analysis-coverage.pdf

Villa Ross, Ceci A. (2023) "Uses of Decennial Census Programs Data in Federal Funds Distribution: Fiscal Year 2021." U.S. Census Bureau, June. https://www.census.gov/newsroom/press-releases/2023/decennial-census-federal-fundsdistribution.html