sciendo

Journal of Official Statistics, Vol. 39, No. 3, 2023, pp. 411–420, http://dx.doi.org/10.2478/JOS-2023-0019

A Rejoinder to Garfinkel (2023) – Legacy Statistical Disclosure Limitation Techniques for Protecting 2020 Decennial US Census: Still a Viable Option

Krishnamurty Muralidhar[1] *and Josep Domingo-Ferrer*[2]

In our article "Database Reconstruction Is Not So Easy and Is Different from Reidentification", we show that reconstruction can be averted by properly using traditional statistical disclosure control (SDC) techniques, also sometimes called legacy statistical disclosure limitation (SDL) techniques. Furthermore, we also point out that, even if reconstruction can be performed, it does not imply reidentification. Hence, the risk of reconstruction does not seem to warrant replacing traditional SDC techniques with differential privacy (DP) based protection. In "Legacy Statistical Disclosure Limitation Techniques Were Not an Option for the 2020 US Census of Population and Housing", by Simson Garfinkel, the author insists that the 2020 Census move to DP was justified. In our view, this latter article contains some misconceptions that we identify and discuss in some detail below. Consequently, we stand by the arguments given in "Database Reconstruction Is Not So Easy. . .".

1. Introduction

In this article, we address some of the issues that Garfinkel (2023) raises regarding our earlier article Muralidhar and Domingo-Ferrer (2023). Before we do however, we feel compelled to make the following preliminary remarks:

(1) Garfinkel makes the following statements:

> "Even if more than half of the matches are incorrect, and an external attacker has no means of confirming them" (Jarmin 2019), a group of hoodlums out to terrorize mixed race families can easily confirm the putative matches by visiting the addresses with the intent to commit acts of violence.
>
> More accuracy means that demographers will likely publish better papers with improved p values using the published data. It also means that the hoodlums will be more likely to find mixed race households when they pound on the door.

There is no evidence to support the above claims. Using real data from the 2010 Census, Muralidhar (2022, Table 11, 321) has shown that the reconstruction procedure assigns race and ethnicity randomly. Specifically, due to the uncertainty in the reconstruction of age, he showed that, for a given (sex, age-group) every individual in the block could potentially be assigned every (race, ethnicity) combination present for that (sex, age-group). Hence, the putative identification is essentially random and is useless in assigning the correct (race,

[1] University of Oklahoma, Price College of Business, Dept. of Marketing and Supply Chain Management, 307 West Brooks, Adams Hall Room 10, Norman, OK 73019, U.S.A. Email: krishm@ou.edu
[2] Universitat Rovira i Virgili, Dept. of Computer Engineering and Mathematics, CYBERCAT-Center for Cybersecurity Research of Catalonia, Av. Països Catalans 26, 43007 Tarragona, Catalonia, Spain. Email: josep.domingo@urv.cat

© Statistics Sweden

ethnicity) to a particular individual. The adversary will have to visit every household. Hence, this claim is baseless fearmongering intended to inflame rather than inform.

(2) The author makes another similar statement relating to trans children in Texas, extracted from a non-refereed Opinion in the Scientific American (Keyes and Flaxman 2022), where the authors claim that they can identify children who changed gender between 2010 and 2020. It is important to note that the results are based on simulated data and not on real data. Perhaps more importantly, this claim of reidentification is not based on identifying individuals; it is based exclusively on comparing unique records within a given geography whose attributes matched on (age-adjusted-by-ten-years, race, ethnicity), but differed on gender (Flaxman and Keyes 2022). But there are many reasons that may explain such a difference, the most common being that households had moved between the two censuses. In fact, even for their own simulated data, the authors find that, of the 70,184 individuals identified, only 657 (0.94%) were trans children, while the remaining 69,527 (99.06%) individuals were incorrectly identified as trans children (Flaxman and Keyes 2022, 11). Note that this verification was only possible only because the data were simulated, and the authors knew the correct identity of the trans children. Without this information, there is no way to establish the identity of the trans children. Hence, there is no basis to the claim that trans children can be identified. Again, baseless fearmongering intended to inflame rather than inform.

We now proceed to address the main content of Garfinkel, much of which is only tangentially relevant to Muralidhar and Domingo-Ferrer. Sections 1, 2, and 3 in Garfinkel are a primer on differential privacy; Section 6 offers a defense of differential privacy. Garfinkel engages in an extensive discussion of suppression which we never addressed in our article. We ignore this discussion as well. We focus our attention primarily on Sections 4 and 5 in Garfinkel and, where appropriate, on specific comments in the other sections.

To facilitate the readability of this article, we first summarize the specific issues raised by Muralidhar and Domingo-Ferrer:

(1) The relevance of Dinur and Nissim (2003) in Census database reconstruction.
(2) The Census reconstruction procedure.
(3) The Census reidentification procedure.
(4) The impact of geography and policy choices on the performance of swapping and the DP-based method.
(5) A discussion of the privacy offered by the DP-based method; and
(6) A discussion of the accuracy of the data generated using the DP-based method.

In the following sections, we address each issue above in greater detail in view of Garfinkel.

2. The Dinur-Nissim Reconstruction

One of the primary motivations of our study was to highlight the fact that, while the Dinur and Nissim (2003) study is an important contribution, it has limited relevance in the context of the Census. Specifically, we wanted to bring attention to the "CD Model" described in Dinur and Nissim (2003), which states that a data curator may be able to modify the original data set to produce a "modified" data set that could possibly be

securely shared. It is easy to see that this is exactly the procedure that the U.S. Census Bureau (Census Bureau) employs. Statistical disclosure limitation is applied to the Census Edited File to create the Hundred-percent Detail File. All tabular data are then generated from the Hundred-percent Detail File (and no access is provided to the Census Edited File) (Zayatz et al. 2009). Furthermore, unlike the Dinur and Nissim (2003) reconstruction procedure where the adversary can issue any query, in the Census context only the results of a pre-determined set of queries (the tabular data results) are allowed. The result of this is that the adversary can only reconstruct the Hundred-percent Detail File and never the Census Edited File. Another possible way to state this is as follows: you can only reconstruct what the Census allows you to reconstruct and nothing else.

Garfinkel contends that there are two types of protection models, local and trusted curator. According to him, the local model involves adding noise to microdata while the trusted curator model involves adding noise to the output. But this is a misunderstanding of the methods. Indeed, the local and trusted curator models represent *where* the protection is applied, not what protection is applied. In a local model, the respondent modifies the response value prior to sharing with the curator. In a trusted curator model, the respondent shares the true value with the curator. In fact, it is possible to implement noise addition to microdata in the trusted curator model (which is the case with the American Community Survey (McKenna 2019). Note that, by definition, a local model precludes adding noise to the output.

Garfinkel also questions our statements regarding the level of noise required to prevent reconstruction. Our statements are based directly on the Dinur and Nissim (2003) conclusions (our emphasis): "a database algorithm that is within $O(\sqrt{n})$ perturbation and is private against polynomial adversaries in the strongest possible sense." Thus, our original statement is accurate. That there are other bounds is irrelevant.

Garfinkel also makes another dubious comment that adding noise at the microdata level is less efficient than adding noise at the query level: "With counting queries, the amount of noise added to each record in local model is the same as the total amount of noise added to the final result in the trusted curator model." This claim is made without any substantiation or support. We already addressed this exact question in our article (Muralidhar and Domingo-Ferrer 2023, subsec. 3.5) and showed that when many queries are issued, the noise added by a DP-based method will be higher than with an alternate procedure such as randomized response. The statement of Bach (2022) is consistent with our analysis.

Finally, the 2010 Decennial Census tabular data release did not involve any noise addition. Instead, data swapping was used as the disclosure limitation tool (Zayatz et al. 2009). One of the biggest benefits of data swapping is that it does not modify the original values of each attribute. In addition, data swapping was implemented with the explicit requirement that voting age population must be preserved exactly at the block level. As a result, the question how many children are in this block will be answered exactly, without any error.

3. The Census Reconstruction Procedure

The entire justification for adopting differential privacy for the 2020 Census was based on the results of the reconstruction and reidentification experiments of the 2010 tabular data release. The Census Bureau has also made the argument explicit in Abowd (2021):

> "... only over the last few years have computing power and the sophisticated numerical optimization software necessary to perform these types of reconstructions advanced enough to permit reconstruction attacks at any significant scale."

Ruggles (2018) has argued from the very beginning that the Census Bureau reconstruction amounts to little more than creating a list of individuals from the tabular data: "any tabular data can be expressed as microdata." This was proved by Muralidhar (2022), who was able to perform the reconstruction using only the data released to the public and an Excel spreadsheet on a laptop, without any numerical optimization software. He observed that, at the *tract* level, the Census Bureau provides the count of (Sex × Age × Race × Ethnicity). Similar data are provided at the *block* level, but with predefined Age Bins in place of individual year of age. Muralidhar (2022) showed that the data at the block level can be reconstructed by randomly assigning the counts of the individual year of age at the tract level to the corresponding Age Bins at the block level. This essentially proves the claim by Ruggles (2018) that the Census tabular data can be expressed as microdata, albeit with a little arithmetic. More importantly, it disproves the claim by Abowd (2021) that extensive computing power and numerical optimization is necessary to perform the reconstruction.

In 2022, the Census Bureau even stopped reconstructing individual years of age. Instead, reconstruction was performed using only Age Bins. With this change, reconstruction is little more than creating a list based on the cell counts from the tabular data, which goes to further support Ruggles (2018) and contradict Abowd (2021).

Note that Garfinkel does not acknowledge Muralidhar (2022) and the detailed explanation the latter provides as to how the reconstruction is simple and straightforward and that the technology needed to perform reconstruction has been available for 50+ years.

An even more important contribution of Muralidhar (2022) is that he showed that, even for small geography (at the Census tract level), there are billions of feasible reconstructions that are equally likely. This has significant implications for the claims of reidentification as we now discuss.

4. The Impact of Policy and Geography Choices

Of all the claims made by the Census Bureau, the most dramatic is the following (Abowd 2021, app B):

> "If the external data on name, address, sex, and age are comparable to the 2010 Census, then the attacker will putatively re-identify 238 million persons (77% of the 2010 Census resident U.S. population). Confirmed re-identifications will be 179 million (58% of the same population). This means that with the best quality external data, relative to the 2010 Census, as many as 179 million persons could be correctly re-identified using the attack strategy outlined here."

Abowd (2021) leaves no doubt as to what he means by reidentification:

> ". . . link the reconstructed microdata to a real name and address associated with the block, sex, age, race, and ethnicity on the reconstructed microdata."

These are dramatic statements because they directly contradict the legal requirements in Title 13 of the United States Code, described in Abowd et al. (2020) as follows:

> "Additionally, the Department of Commerce (2017), in which the Census Bureau is housed, has issued directives regarding the protection of personally identifiable information (PII) and business identifiable information (BII). These directives largely mirror those issued by other government agencies and prohibit release of information that can be used "to distinguish or trace an individual's identity, such as their name, social security number, biometric records, etc., alone or when combined with other personal or identifying information which is linked or linkable to a specific individual, such as date and place of birth, mother's maiden name, etc."

Abowd (2021) states explicitly that the primary reason for adopting differential privacy for the 2020 Decennial Census was that the tabular data release from the 2010 Decennial Census violated Title 13 requirements:

> "To defend against this known vulnerability, the Census Bureau explored different confidentiality methods that explicitly defend against database reconstruction attacks and concluded that the best tool to protect against this modern attack while also preserving the accuracy and usability of data products comes from the body of scientific work called "differential privacy"."

Muralidhar and Domingo-Ferrer pointed out that the claims of reidentification were vastly overstated. These results are supported by the empirical analysis of 2010 Census tabular data by Muralidhar (2022). The reconstruction procedure in Abowd (2021) was shown to directly contradict the reidentification procedure suggested by the Census Bureau's own Research and Methodology Directorate (McKenna 2019). The claims of reidentification also contradict Dr. Ron Jarmin, Acting Director of the Census Bureau, who said: "more than half the matches are incorrect." (Jarmin 2019) Based on all this information, we were pointing out the fact that what Abowd (2021) claims as reidentification was actually the accuracy of reconstruction, as illustrated in Subsection 5.2. of Muralidhar and Domingo-Ferrer. What is lacking is the ability to uniquely link the respondents in the reconstructed data to their identity (see Subsection 5.2 in Muralidhar and Domingo-Ferrer for more details).

Garfinkel claims that "there are many ways that even unidentified census microdata with block-level geography can cause harm." This may be true, but it is a far cry from claiming that the reconstructed data can be linked to individual names and addresses. Garfinkel now claims that even if the reidentification was incorrect, it can be verified by the adversary by visiting the address. If we take this to its logical conclusion, for the adversary to make the same claim as Abowd (2021) regarding reidentification, the adversary will have to conduct a new Census.

In summary, Garfinkel offers nothing to contradict our conclusion that the reidentification claims regarding the 2010 Census tabular data release have been vastly overstated by the Census Bureau reconstruction attacks.

5. The Impact of Policy and Geography Choices

Muralidhar and Domingo-Ferrer pointed out that the performance of data swapping and the DP-based method were evaluated using different benchmarks. Perhaps the most important of this is reflected in the *invariants*, that is, counts that are preserved and the

geographic level at which these counts are preserved. The 2010 Decennial Census tabular data release preserved exact counts at the block level for the total population, population aged 18 and above, and population below age 18. Regarding the 2020 Decennial Census, a 2017 Census publication (Dajani et al. 2017) states that:

> By agreement with the Department of Justice (2000), the Census Bureau will provide exact counts at the Census block level for the following variables:
> - Number of people: total, age 18+ (voting age), and less than age 18,

The updated version of the same document in 2020 (Abowd et al. 2020) states:

> The Census Bureau currently plans the following invariants for the 2020 Census data publications:
> - Total number of people by state;

This is a remarkable change in policy (and one that has had a significant detrimental impact on data accuracy as we will discuss later). While the 2017 invariants would have been consistent with the 2010 invariants, the 2020 invariants have remarkably relaxed the requirement, from the block level to the state level. From the Census demonstration data products released to the public, we know that the DP-based method for the demonstration products only preserved totals at the state level, while the original 2010 tabular data release using swapping preserved totals at the block level. Considering that the number of blocks varies from 6,012 (Washington D.C.) to 668,757 (Texas), this represents a massive change in policy. Muralidhar and Domingo-Ferrer point out that such changes in policy have a direct impact on the protection afforded by the statistical disclosure limitation procedure.

It is true that the Census Bureau has recently released comparisons of applying data swapping with different parameters (the proportion of swapped records) (Hawes 2022). However, it is still not clear whether these new parameters were applied with the old policy (invariants at the block level) or the new policy (invariants at the state level). Just changing the proportion of records swapped without changing the invariant requirement is unlikely to change the performance of the swapping algorithm. It would have been relatively easy for the Census Bureau to release data where both data swapping and the DP-based method were held to the same standard. By not doing so, the Census Bureau has made any meaningful comparison impossible. Unfortunately, Garfinkel does not address this issue at all.

Instead, Garfinkel chooses to make new policy by stating "the primary risks that DP was designed to protect: the release of confidential information about race." This is quite a remarkable statement and, in some ways, a curious one. Garfinkel claims that information regarding race and ethnicity data are "in general, not for sale." While this claim is arguable, let us assume it is true. If this is the case, then race and ethnicity are the confidential variables that the adversary does not know. Remember that reconstruction was performed by matching age and sex from the reconstructed data to the externally obtained data. This is a common phenomenon encountered in statistical disclosure limitation where the attributes used to perform the matching are referred to as quasi-identifiers. To prevent the adversary from using the quasi-identifiers (age and sex) to learn the confidential attributes (race and ethnicity), privacy models such as k-anonymity

(Samarati 2001) would modify the quasi-identifiers and not the confidential attributes. Thus, if the objective is to prevent reconstruction, it would make more sense to protect the age and sex attributes rather than the race and ethnicity attributes. The rationale is simple – if you cannot reconstruct the data accurately, you cannot infer the confidential attributes.

6. The Privacy Provided by DP

We must admit that we were astounded by the statement in Garfinkel that the privacy loss parameter ϵ and the corresponding randomized response probability is "not relevant to the use of differential privacy for the Census." Garfinkel contends that "there is no single query over the 2020 Census confidential data for which this privacy loss value is used. Instead, the privacy loss is spread over tens of millions of individual queries to produce the so-called noisy measurements."

According to Dwork et al. (2019) Differential privacy allows us to quantify cumulative privacy loss as data are analyzed and re-analyzed, shared, and linked. These differentially private systems will, for the first time, allow us to measure and compare the total privacy loss due to these personal data-intensive activities.

The privacy parameter ϵ represents this total (cumulative) privacy loss. It would be a profound misinterpretation to claim that, since the privacy loss is distributed over tens of millions of individual queries, the cumulative privacy loss does not matter.

Muralidhar and Domingo-Ferrer does not say that "there are objectively correct values for data utility and privacy loss," as Garfinkel asserts. In our opinion, the statements "Taking a very small ϵ entails unaffordable utility loss, but taking very large ϵ entails very little noise addition and offers little to no protection against reconstruction" (Muralidhar and Domingo-Ferrer 2023) and "choosing a small ϵ means that the data owner has chosen to highly value privacy at a cost of data utility, while large ϵ values mean the data owner has chosen to highly value data utility at the cost of privacy" (Garfinkel 2023), say the same thing albeit in different words.

7. Data Accuracy of the DP-based Method

The lack of accuracy of the DP-based method is well documented, as shown in Muralidhar and Domingo-Ferrer. Unfortunately, Garfinkel does not address any of the multiple research studies cited in Muralidhar and Domingo-Ferrer. He only affirms that "the Census Bureau was largely successful, as evidenced by numerous data analyses published since that make use of the data." According to Menger (2021), as many as 11% of the blocks have "impossibilities" (more households than people, zero housing units with non-zero population, zero population but non-zero occupied housing units, children but no adults, and very large households). Whether this constitutes a large successful implementation is up to question. As in many other instances, Garfinkel makes the statement of numerous data analyses, but provides no citation.

Garfinkel has chosen to address one humorous example highlighted in the NY Times of people living under water. He categorically states that this error was not due to DP. But if this were the case, one would have expected the Census Bureau to issue some type of formal explanation for this. Garfinkel further contends that this error could not have happened since, according to Abowd et al. (2022), there is a constraint that "the number of

householders (person one on the questionnaire) cannot be greater than the number of housing units. But there have been documented instances where 4.83% [of blocks] have zero occupied housing units which have people living in them (that is, there are no occupied dwellings, but there are people living there)" (Menger 2021). Clearly, the constraint described by Abowd et al. (2022) is being violated.

Finally, we address the rather serious claim by Garfinkel: "MDF fails to address a very real harm to the statistics of vulnerable populations done by suppression and swapping. That harm is erasure. By design, these techniques hide the statistical contribution of minorities." As we have emphasized before, we absolutely deny that we recommend suppression. Removing suppression from the discussion and focusing only on swapping, we can categorically state that data swapping does not result in erasure. Data swapping simply exchanges the attribute values between records. As such, there is no addition or subtraction to the counts. Hence, we reiterate that data swapping does not result in erasure.

The real irony is that the DP-based method does result in erasure. With the DP-based method, the frequency counts are modified by adding or subtracting a random noise term. As a result, every time negative noise is applied to a count of minorities, there is erasure. And there is no way to prevent this from happening. Thus, contrary to the claim made by Garfinkel, it is the DP-based method that is likely to erase the contribution of the minorities, not data swapping.

8. Conclusions

We have carefully gone through the article of Garfinkel. The article offers no direct substantive, justifiable criticism of Muralidhar and Domingo-Ferrer. We have gone to great lengths to address every issue that we found relevant.

Garfinkel calls for finding a constructive way forward. The first step in moving forward is to perform a comprehensive comparison of data swapping and the DP-based method using the same criteria as in the 2010 tabular data release. The second step is to re-evaluate the reidentification claims. Moving forward also requires acknowledging the work of researchers on the other side and not ignoring all their arguments.

9. References

Abowd, J.M. 2021. *Declaration of John M. Abowd. Case no. 3:21-CV-211-RAH-ECM-KCN*, U.S. District Court for the Middle District of Alabama. Available at: https://censusproject.files.wordpress.com/2021/04/2021.04.13-abowd-declaration-alabama-v.-commerce-ii-final-signed.pdf.

Abowd, J., R. Ashmead, R. Cumings-Menon, S. Garfinkel, M. Heineck, C. Heiss, and R. Johns. 2022. "The 2020 Census Disclosure Avoidance System TopDown Algorithm." *Harvard Data Science Review 2*. Available at: https://hdsr.mitpress.mit.edu/pub/7evz361i (accessed May 2023).

Abowd, J.M., G.L. Benedetto, S.L. Garfinkel, S.A. Dahl, A.N. Dajani, M. Graham, and M.B. Hawes. 2020. *The Modernization of Statistical Disclosure Limitation at the U.S. Census Bureau*. Available at: https://www.census.gov/library/working-papers/2020/adrm/CED-WP-2020-009.html (accessed May 2023).

Bach, F. 2022. "Differential Privacy and Noisy Confidentiality Concepts for European Population Statistics." *Journal of Survey Statistics and Methodology*, 10: 642–687. DOI: https://doi.org/10.1093/jssam/smab044.

Dajani, A.N., A.D. Lauger, P.E. Singer, D. Kifer, J.P. Reiter, A., Machanavajjhala, S.L. Garfinkel, S.A. Dahl, M. Graham, V. Karwa, H. Kim, P. Leclerc, I.M. Schmutte, W.N. Sexton, L. Vilhuber, and J.M. Abowd. 2017. "The modernization of statistical disclosure limitation at the U.S. Census Bureau." In: Census Scientific Advisory Committee Meeting, September 14–15, Suitland MD, USA. Available at: https://www.census.gov/library/video/2017/2017-09-sac.html (accessed May 2023).

Dinur, I., and Nissim. 2003. "Revealing Information While Preserving Privacy." In: Proceedings of the 22nd ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems: 202-210, June 9–12, 2003, San Diego CA, USA. DOI: https://doi.org/10.1145/773153.773173.

Dwork, C., N. Kohli, and D. Mulligan. 2019. "Differential Privacy in Practice: Expose Your Epsilons!" *Journal of Privacy and Confidentiality* 9(2). DOI: https://doi.org/10.29012/jpc.689.

Flaxman, A.D., and O. Keyes. 2022. "The Risk of Linked Census Data to Transgender Children: A Simulation Study," In: 2020 Census Data Products: Workshop on the Demographic and Housing Characteristics Files, June 21–22, Washington D.C. Available at: https://www.nationalacademies.org/event/06-21-2022/docs/DF3CD148E8CAEC0B93FB88C4A7820AC2D54503ACDD18?noSaveAs = 1 (accessed June 2022).

Garfinkel, S. 2023. "Legacy Statistical Disclosure Limitation Techniques Were Not An Option for the 2020 US Census of Population and Housing." *Journal of Official Statistics*, this issue.

Hawes, M. 2022. "Reconstruction and Reidentification of the Demographic and Housing Characteristics File (DHC)", Presentation to the Census Scientific Advisory Committee, September 29–30, 2022, Washington D.C. Available at: https://www2.census.gov/about/partners/cac/sac/meetings/2022-09/presentation-reconstruction-and-re-dentification-of-dhc-file.pdf (accessed March 2023).

Jarmin, R. 2019. *Census Bureau Adopts Cutting Edge Privacy Protection for 2020 Census.* Available at: https://www.census.gov/newsroom/blogs/random-samplings/2019/02/-census_bureau_adopts.html (accessed May 2023).

Keyes, O., and A.D. Flaxman. 2022. "How Census Data Put Trans Children at Risk." *Scientific American* 21. Available at: https://www.scientificamerican.com/article/how-census-data-put-trans-children-at-risk/ (accessed May 2023).

McKenna, L. 2019. *Disclosure Avoidance Techniques Used for the 1960 Through 2010 Decennial Censuses of Population and Housing Public Use Microdata Samples Research and Methodology Directorate*, U.S. Census Bureau. Available at: https://www.census.gov/library/working-papers/2019/adrm/CED-WP-2019-007.html (accessed May 2023).

Menger, G. 2021. "Using 2020 Census Data." *Applied Geographic Solutions*. Available at: https://appliedgeographic.com/2021/09/using-2020-census-data/ (accessed May 2023).

Muralidhar, K., J. Domingo-Ferrer. 2023. "Database Reconstruction Is Not So Easy and Is Different from Reidentification." *Journal of Official Statistics*, this issue.

McKenna, L. 2019. *U.S. Census Bureau Reidentification Techniques.* Research and Methodology Directorate, U.S. Census Bureau. Available at: https://www2.census.gov/adrm/CED/Papers/CY19/2019-04-Reidentification%20studies-20210331FinRed.pdf (accessed May 2023).

Muralidhar, K. 2022. "A Re-Examination of the Census Bureau Reconstruction and Reidentification Attack." In *Privacy in Statistical Databases – PSD 2022*, Paris, France. Springer: 312–323. DOI: https://doi.org/10.1007/978-3-031-13945-1.

Ruggles, S. 2018. *Implications of Differential Privacy for Census Bureau Data Dissemination*, Institute for Social Research and Data Innovation, University of Minnesota. Available at: https://apps.bea.gov/fesac/meetings/Ruggles%20Presentation%20Revised.pdf (accessed May 2023).

Samarati, P. 2001. "Protecting Respondents Identities in Microdata Release." *IEEE Transactions on Knowledge and Data Engineering* 13(6): 1010–1027.

Zayatz, L., J. Lucero, P. Massell, A. Ramanayake. 2009. *Disclosure Avoidance for Census 2010 and American Community Survey Five-year Tabular Data Products*, Census Bureau Research Report RRS2009-10. Available at: https://www.census.gov/content/dam/Census/library/working-papers/2009/adrm/rrs2009-10.pdf (accessed May 2023).

Received May 2023
Revised June 2023
Accepted June 2023