IPUMS Working Papers

Unraveling Geographic Complexities in the Current Population Survey

David Van Riper†
University of Minnesota

Sarah Flood
University of Minnesota

Finn Roberts
University of Minnesota

September 2021

IPUMS Working Paper No. 2021-02
DOI: https://doi.org/10.18128/IPUMS2021-02

Cross-posted as MPC Working Paper No. 2021-04
and LCC Working Paper No. 2021-01

†Address correspondence to David Van Riper: vanriper@umn.edu. Support for this research was funded by the Minnesota Population Center (P2C HD041023), the Life Course Center at the University of Minnesota (P30AG066613), and IPUMS CPS (R01HD067258).

Abstract

The Current Population Survey (CPS) is the nation's primary source of information on labor force participation and is typically used to generate national- and state-level estimates for a wide variety of indicators. Scientists are increasingly interested in linking contextual data measured over finer-grained geographies to the CPS. Geographic units such as metropolitan areas, cities, and counties are identified in the CPS, but a close examination reveals counties with large populations absent from the data while counties with smaller populations in the same metropolitan area are identified. We describe the process used to assign geographic identifiers to respondents and how that process leads to counties with large population being left off. We also describe strategies to recover or assign county identifiers for respondents who are missing them. These strategies consider the complex spatial relationships among metropolitan areas, cities, and counties.

Introduction

The Current Population Survey (CPS) is the nation's primary source of information on labor force participation. With data collected monthly and large sample sizes, it is routinely used to generate national- and state-level estimates. Finer geographic units are identified in the CPS, including metropolitan areas, cities, and counties. Counties, one of the most detailed geographies available in the public-use files, are identified for about 45% of households participating in the CPS. The identification of sub-state geographic units is particularly useful for researchers who want to link contextual data to the CPS. A careful look at the geographic data in the CPS, however, reveals counties with large populations that are curiously absent from the data even when counties with smaller populations are identified in the same state. In this paper, we describe the ways that CPS geographies are assigned and provide an example to illustrate the nuances, describe our effort to assign county identifiers to individuals living in sampled dwellings based on the geographic information available in the data, and discuss the implications for researchers.

IPUMS CPS delivers harmonized CPS data to the research community free of charge via cps.ipums.org. Through harmonization, we navigate changes to variable names, codes, and labels in the underlying data to produce variables with consistent names and identical coding schemes across datasets. Harmonization simplifies analyses that pool multiple months and/or years of data, eliminates redundant effort by researchers, and provides researchers with a common starting point for their analyses, thus increasing the reproducibility of research results. IPUMS also adds value to the data by documenting complex variables (e.g., geography) and constructing enhanced variables that simplify their use.[1] This paper on CPS geography complements existing variable-level documentation for geographic variables.

[1] The improvements described in this paper also extend to IPUMS ATUS (American Time Use Survey), since all ATUS respondents were previously in the CPS and any improvements made to IPUMS CPS data may also be used in

Identification of additional counties in the public-use CPS data will serve researchers investigating many important research questions. For example, sub-state geographic variation has been employed in analyses of smoking (Chahine, Subramanian, and Levy 2011), occupational licensing and earnings (Ingram 2019), and income inequality and health (Mellor and Milyo 2002). The CPS has been one of the primary datasets for understanding the consequences of the COVID-19 pandemic for the US population, and researchers will almost certainly be combining county-level COVID vaccination, infection, hospitalization, and death rates as well as local public health mandates with the CPS. For these purposes, the ability to locate individuals more precisely in space is a real benefit of these data. It is also important, however, for researchers to understand the caveats associated with locating CPS respondents spatially, which we elaborate on in the discussion.

Assignment of geographic identifiers in the CPS

Geographic identifiers are assigned to CPS individuals in a hierarchical manner based on the place of residence when the survey is completed (Figure 1). Assignment of geographic identifiers is based on the size of the population in specific geographic areas as opposed to the number of people in the sample that are located in a particular geographic area. Each of these geographic units corresponds to variables that are available in the public-use microdata. Table 1 provides variable names in the CPS data available from the Census Bureau and in the IPUMS CPS version of the data.

Individuals first receive a code based on their state of residence. All US states and the District of Columbia have large enough populations to be identified; therefore, every CPS respondent receives a state code.[2] Next, they receive a code indicating whether they are located in a metropolitan statistical

---

conjunction with the ATUS. Linkages between IPUMS CPS and IPUMS ATUS are easy using the variable CPSIDP, which is created by IPUMS and available via both the IPUMS CPS and IPUMS ATUS websites.

[2] States, metropolitan areas, and counties in the CPS are assigned codes based on the Federal Information Processing Series (FIPS). The FIPS classifications impose standard coding systems for geographic entities such as states, counties, metropolitan areas, or American Indian Areas. Codes are assigned to geographic entities in alphabetical order based on the entities' names.

area (MSA). It is possible but uncommon for an individual residing in a non-metropolitan area to be assigned additional geographic identifiers.

Individuals in metropolitan areas receive one of three “In metro area” codes based on their location within an MSA:

- Central city - individual resides in a central city in an MSA
- Outside central city - individual resides outside of an MSA’s central city(ies) but still within the MSA
- Central city status unknown - individual’s central city status cannot be identified because that would violate the CPS suppression rules[3]

Next, an individual living in a metropolitan area is assigned the code for their specific MSA if the MSA’s population is large enough to be identified. For example, there are eight official metropolitan areas in the state of Minnesota in 2019 (Office of Management and Budget 2018). Respondents in the Minneapolis-St. Paul-Bloomington, MN-WI metropolitan area receive its specific code (33640). Individuals living in other metropolitan areas (e.g., Duluth-Superior, MN-WI; St Cloud, MN; Mankato, MN) receive one of the three “in metro area” codes but not a specific MSA FIPS code.

Then, individuals residing in a central city[4] may be assigned a code for their specific central city if that city meets the population thresholds for identification. For example, the Minneapolis-St. Paul-Bloomington, MN-WI metropolitan area contains six central cities. Two of these cities, Minneapolis and St. Paul, meet the thresholds; therefore, respondents in Minneapolis and St. Paul are assigned codes for

[3] The CPS identifies geographic units if they have a minimum population of 100,000 and their identification does not indirectly reveal geographic units with populations less than 100,000 (Weyland 2020). Indirect identification is possible because of the assignment order for geographic identifiers (Figure 1) and the complex spatial relationships among metropolitan areas, central cities, and counties.

[4] Central cities, also known as principal cities, form the cores of metropolitan areas. The largest city within a metropolitan area serves as its “central city”. Additional cities may qualify as central cities if they meet certain population and employment thresholds as defined by the Office of Management and Budget (Office of Management and Budget 2018).

their central city of residence.

Finally, county codes are assigned if doing so does not violate suppression rules for previously assigned geographies. Respondents residing in metropolitan areas who are located outside a central city or with an unknown central city status are considered for county codes. If identifying the individual's county of residence does not violate suppression rules, then the appropriate county code is assigned. If identifying the county violates suppression rules, then the metro resident will not be assigned a county code. Individuals located in central cities may be assigned a county code if the population in the county that is outside of the central city is large enough to ensure confidentiality. Table 2 shows how common it is in the 2019 Annual Social and Economic (ASEC), both for the nation and for Minnesota specifically, to have each combination of sub-state geography.

Assigning geographic identifiers to ASEC respondents in Minnesota

To illustrate the nuances of the geographic identifier assignment process to CPS respondents, we outline the procedure by which geographic identifiers are assigned to individuals residing in the state of Minnesota. Data for this example come from the 2019 ASEC of the CPS. This concrete example will help researchers understand why certain geographic units are identified and others are not.

Metro versus Non-metro

Of the 2,300 respondents residing in Minnesota in the 2019 ASEC (see Table 2), 1,773 lived in metropolitan areas and 527 lived outside of metropolitan areas (Figure 2). Thus, 77 percent of the 2019 ASEC respondents in Minnesota live in one of the gray areas on the map, and the remaining 23 percent live in the white area. In addition to indicating metro versus non-metro status, this variable also denotes the central city status of the respondents residing in metropolitan areas. In Minnesota, of the metropolitan area respondents, 20 percent resided in a central city, 60 percent resided outside of a central city, and 20 percent had a “central city status unknown”[5].

Specific metropolitan statistical area

While the state of Minnesota contains multiple MSAs (eight as of 2019), only one is specifically identified in the CPS. CPS respondents in the Minneapolis-St Paul, MN-WI MSA receive its Federal Information Processing Standard (FIPS) code of 33460. Of the 1,773 Minnesota respondents living in a metropolitan area, 1,447 live in the Minneapolis-St Paul MSA and 326 live in one of the other MSAs in the state.

Central cities

The Minneapolis-St Paul, MN-WI metropolitan area contains six central cities (Figure 3). Respondents who reside in one of these central cities receive the “Central city” code, and respondents who live in the metro area but outside these central cities receive the “Outside central city” code. Of the 1,447 Minneapolis-St Paul respondents, 377 reside in one of the six central cities and 1,070 reside outside of the central cities.

Two of the six central cities (Minneapolis and St. Paul) meet the threshold for identification; therefore, respondents who live in these central cities receive specific codes. Of the 377 metropolitan area respondents residing in a central city, 156 reside in Minneapolis and 122 reside in St Paul. The remaining 99 respondents reside in one of the four other central cities and do not receive a specific central city code.

[5] The phrase “central city status unknown” indicates that the CPS respondent lives in a metropolitan area, but whether they live in a central city is not disclosed. Respondents with a “central city status unknown” code generally live in metropolitan areas with relatively small populations. If their central city status was disclosed, suppression rules may be broken.

Counties

The final geographic area that is identified is the county. Respondents receive a county code if the county meets the threshold for identification. This threshold is not solely related to the population of the county; instead, it is related to the other geographic identifiers (e.g., metro vs. non-metro codes, specific metropolitan statistical area codes, specific central city codes) already assigned to respondents and to the counties included in the CPS sampling frame.

The Minneapolis-St Paul, MN-WI metropolitan area consists of 14 Minnesota counties (Figure 4). In the 2019 ASEC data, the five counties with bolded names in Figure 4 are identified. These are five of the largest counties by population in the metropolitan area. Interestingly, the first (Hennepin) and third (Dakota) largest counties by population are not identified in the ASEC. Their lack of identification is due to the complex relationship between central cities and counties.

The six central cities in Minneapolis-St Paul are spread across three counties, with Hennepin county containing four and Ramsey and Dakota counties containing one each. ASEC respondents residing in St Paul receive a specific central city code and a county code (Ramsey). ASEC respondents residing in the part of Ramsey county outside St Paul receive an "outside central city" code and a county code. Both central city and county are identified in this case because the part of Ramsey county outside St Paul meets the threshold for identification - it contained 223,571 persons in 2010.

Neither Hennepin nor Dakota county are identified despite being the first and third largest in the Minneapolis-St Paul, MN-WI metropolitan area. This is the result of the identification of central cities before the assignment of counties in the CPS (Figure 1). As such, the omission of Hennepin and Dakota counties from the identified CPS counties is intentional for the following reasons. First, three counties contain central cities in the metropolitan area. Ramsey county, which includes St Paul, is identified, leaving only Hennepin and Dakota counties as the two other counties containing central cities. Second, Dakota county is not identified because the central city of Eagan (the only central city in Dakota county) does not meet the threshold for identification. Even though Dakota county is large enough to be identified, doing so would make it possible for researchers to identify the respondents who lived in the central city of Eagan and violate confidentiality rules. Third, and as a result, Hennepin county is not identified because doing so in addition to identifying Ramsey county would, by process of elimination, identify individuals living in Dakota county and Eagan specifically.

Identifying additional counties in the 2019 ASEC

Thus far we have described the spatial relationships among metropolitan areas, central cities, and counties in the CPS using Minnesota as a concrete example. In this section, we propose strategies for recovering and imputing county information where metropolitan area and/or central city geographies exist but county does not.

Table 2 shows the number of 2019 ASEC respondents (and those living in Minnesota) with various combinations of specific metropolitan area, central city, and county codes. As the data are delivered by the Census Bureau, 137,134 have some specific sub-state geography identified and 75,991 (42.2%) are assigned county codes. For 61,143 individuals, county assignment may be possible.

The geographic assignment rules for CPS respondents privilege metropolitan areas and central cities over counties; thus, we observe thousands of CPS respondents residing in metropolitan areas and central cities with no county codes. As we examined the geographic relationships between metropolitan areas, central cities and counties, we realized that we could impute county identifiers for some CPS respondents. Of the six combinations in Table 2, we are particularly interested in the 6,760 respondents who have a metro area and central city code but no county code and the 54,383 respondents with a specific metro code but no central city or county code. We developed two imputation strategies based on the geographic identifiers assigned to respondents. We describe each strategy below.

Strategy 1: Imputing county identifiers for residents in specific central cities

If we know what central city a resident lives in, we can also know the county where they reside. For CPS respondents who receive a specific central city code but no county code, we use the spatial relationship between central cities and counties to guide our imputation. We illustrated this above: Minneapolis central city residents all live in Hennepin county. Our strategies for this case are:

A. One to one correspondence between a specific central city and county

   1. If a specific central city nests within a single county, all respondents with the specific central city code are assigned the code for the county that contains the central city.

      i. *Example*: Minneapolis nests within Hennepin County; therefore, all CPS respondents in Minneapolis are assigned Hennepin's county code. Thus, we can assign the Hennepin county code to the 156 cases in the 2019 Minnesota ASEC who have a Minneapolis-St Paul, MN-WI metropolitan area code and a Minneapolis central city code (Table 2).

   2. If a specific central city falls in two or more counties, but 100% of the central city population resides in one county, all respondents in the specific central city code are assigned the code for the county that contains the central city.[6]

      i. *Example*: The city of Chicago falls into Cook and DuPage counties, but 100 percent of Chicago's population resides in Cook county. Therefore, there is no uncertainty when we assign Cook county's code to respondents residing in the city of Chicago.

B. If a specific central city falls in two or more counties, we compute the proportion of the central city population that resides in each county. We then randomly assign county codes to the central city respondents using the proportions.

[6] In the 2019 ASEC, other examples of this situation include Raleigh, NC; Milwaukee, WI; and Kansas City, KS.

1. *Example*: Naperville, IL, is a central city in the Chicago-Naperville-Elgin, IL-IN-WI metropolitan area. It falls into DuPage and Will counties. Sixty-seven percent of Naperville's population is in DuPage county and 33 percent is in Will county. Thus, 67 percent of the cases in Naperville receive DuPage's county code and 33 percent receive Will's county code.

We have high confidence in strategy 1A for central city county assignment since the central city nests within the county (1A1) or the population of the central city nests within the county (1A2). Our confidence in the second strategy for central city county assignment (strategy 1B) varies based on the magnitude of the central city proportions in each county. In the Naperville, IL, example, our county assignment could be wrong for a large fraction of cases since it's split fairly evenly across two counties.

Table 3 shows a breakdown of the number of 2019 ASEC respondents assigned a county code by central city county assignment strategy. Of the 6,760 respondents with a specific central city code, 3,841 reside either in a central city that nests within a single county or has 100% of its population residing in one county (if the central city falls into two or more counties); therefore, we can assign them a county code with 100 percent confidence. The remaining 2,919 respondents reside in a central city that falls into two or more counties and has population in more than one county. Those respondents are assigned a county code using the proportional method. While this strategy does assign a county, we have less than 100 percent confidence that the assignment is correct.

Strategy 2: Imputing county identifiers for residents who reside in specific metropolitan areas but not in specific central cities

Recall that metropolitan areas often encompass more than one county. For 54,383 cases in the 2019 ASEC, specific metropolitan areas are identified (and central cities are not) but county of residence is not. We use a proportional population strategy similar to the one used in (2) above to assign individuals to counties. Our multi-step strategy for this case is:

A. For each specific metropolitan area in the CPS, generate a list of the metropolitan area's counties that are not already identified in the CPS.

   1. *Example*: For the Albuquerque, NM metropolitan area, the non-identified counties are Sandoval, Torrance, and Valencia.[7]

B. Sum the total population for the non-identified counties and then, for each non-identified county, compute its proportion of the total.

   1. *Example*: The total population for the three non-identified counties in Albuquerque, NM is 224,513 (Census 2010 population). The proportions for the non-identified counties are 0.586 (Sandoval), 0.073 (Torrance), and 0.341 (Valencia).

C. Randomly assign county codes to CPS respondents in the specific metropolitan area using the proportions.

   1. *Example*: Sandoval county contains 58.6 percent of the total population of the non-identified counties; therefore, 58.6 percent of the Albuquerque metropolitan area respondents without a county code receive the code for Sandoval county.

We use this strategy to assign county codes to the 54,383 ASEC cases that have a specific metropolitan area code but no central city or county code. Our confidence in these county assignments varies based on the magnitude of the non-identified counties' proportions. If the proportions for the non-identified counties are relatively similar, then our confidence in our assignment will be lower. If one

[7] The Albuquerque, NM metropolitan area consists of four counties: Bernalillo, Sandoval, Torrance, and Valencia (Office of Management and Budget 2018). Bernalillo County is identified in the CPS; therefore, CPS respondents who reside in Bernalillo receive its county code. CPS respondents who reside in the other three counties receive the Albuquerque, NM metropolitan area code but no county code.

non-identified county makes up a large proportion, then our confidence will be higher for cases assigned the code for that specific county.

Discussion

We set out to unravel the complexities of geographic assignment in the Current Population Survey (CPS). The measures taken to assign geographic identifiers in CPS are to protect respondent confidentiality, but the nuances of this assignment are not clearly documented. We developed a diagram illustrating CPS geographic assignment. It shows that CPS geography is not assigned in top-down fashion from largest to smallest geographies, but rather that the identification of cities is prioritized over the identification of counties. The result is that some of the most populous counties in a given state are not identified while less populous counties are. This is because large cities in those high population counties are identified first.

Our motivation for understanding and documenting geography in the CPS was driven by an interest in linking contextual data for counties to the CPS. As a result of our work, we have shown that recovering and imputing county codes for CPS respondents has the potential to increase geographic coverage dramatically. In some instances, we have 100% certainty that we have located the individual in the correct county, but in most cases, there is less certainty about the accuracy of assignment.

We are most certain about our assignment using Strategy 1A when a respondent's central city is identified in the data and the central city is entirely located in a single county or the entire population in the central city lives in a single county. In the 2019 ASEC, we increase the number of individuals with an identified county of residence by 5% (3,841/75,991=.0505) by assigning county identifiers to these cases based on the known relationship between cities and counties. At a minimum, this is useful because individuals residing in counties like Hennepin in Minnesota, which is the most populous in the state, can now be included in analyses of the relationship between where an individual is located spatially and their social and economic outcomes.

In the majority of cases, however, we have less certainty about the exact county in which individuals reside. Even when cities are identified, they sometimes cross county boundaries (Strategy 1B), meaning that with the public-use CPS data there is no way to know for sure which county an individual lives in. And the same is true when metropolitan areas are identified without additional clues about specific county locations (Strategy 2). These cases consist of 2,919 individuals (Strategy 1B) and 54,383 individuals (Strategy 2) for a total of 93.7% of individuals whose county is not identified, but the central city or metropolitan area is ((2,919+54,383)/61,143=.937). These uncertainties create issues that researchers should carefully consider if they want to use imputed county information.

First, researchers should avoid generating county-level estimates with the CPS. The data are representative at the state level, but not necessarily at lower levels of geography. We have not investigated, for example, the representativeness of the respondents who receive county codes or the suitability of weights for generating sub-state estimates. In some cases, like Hennepin county, where the only CPS respondents who receive the Hennepin County code with 100% confidence reside in the central city of Minneapolis, generating estimates of the people in Hennepin county from the CPS would be especially problematic. Neither the Minneapolis central city residents nor the respondents assigned a Hennepin county code via imputation are necessarily representative of all Hennepin county residents. The representativeness of respondents in a county is not as big of a concern for attaching contextual information to the CPS as it is for trying to make inferences about people who live in a specific sub-state location.

Second, researchers need to think about misclassification risk and the implications for their analyses. Misclassification risk is possible for most of the additional counties that we can identify, but the consequences of misclassification for analyses depend on the reasons for wanting to locate respondents in counties. In some instances, misclassification bias will have little impact on analyses. For example, environmental effects like temperature and air quality tend to be similar across large spaces and are not influenced by county boundaries. In other cases, misclassification bias is more consequential, as might be the case with policies that are enacted at the county-level such as social distancing guidelines during the COVID-19 pandemic or effects of county-level racial segregation, on some outcome of interest, which may vary dramatically across counties.

Limitations and Future plans

Our proportional assignment strategy is easy to implement and explain to data users, but it does have a high risk of misclassification error. Future work will investigate more sophisticated imputation methods, including the penalized maximum entropy dasymetric model (P-MEDM) developed by Leyk et al. (2013) and other re-weighting techniques developed for spatial microsimulation (Whitworth et al. 2017; Hermes and Poulsen 2012). These alternative methods use ancillary demographic data, typically from the American Community Survey, with finer-grained geographic resolution to improve the imputation results.

We have illustrated our approach using the 2019 ASEC, but incorporating additional months and years of CPS data will be especially challenging because of changes to metropolitan areas over time. The Office of Management and Budget updates metropolitan area boundaries following each decennial census. These updates may include the delineation of new metropolitan areas, the removal of existing metropolitan areas, or the addition/subtraction of counties or central cities from existing metropolitan areas. The CPS then adopts the updated metropolitan area boundaries. As a result of the changes in the metropolitan areas from decade to decade, we must implement our imputation approach for each set of metropolitan areas used in the CPS. In the last 40 years, for example, there have been four distinct versions of metropolitan areas used by the CPS. As part of this process, we will add a geography vintage variable to each CPS respondent. This variable will indicate which version of metropolitan area definition was used to assign geographic identifiers.

We have also omitted metropolitan statistical areas in New England from our current work. New England's MSAs present a unique challenge because their component parts have changed over time. Prior to 2015, the New England MSAs were built from cities and towns and not counties as they were in the rest of the United States. Starting in 2015, the New England MSAs used in the CPS are composed of counties. When MSAs were built from cities or towns, we observe parts of counties assigned to two or more MSAs, or counties where parts are in an MSA and parts are outside an MSA. Thus, we need to develop a methodology for the city/town-based MSAs that accounts for the complex relationship between cities/town, counties, and MSAs.

To help users identify the respondents who receive imputed county identifiers, we will develop a set of flags that indicate: (1) whether the county identifier was imputed; (2) the strategy used for the county imputation; (3) our level of confidence in the imputed identifier. These flags will help users decide whether to include or exclude these respondents from their analysis.

Conclusion

Linking contextual data to microdata is increasingly popular, but many datasets, including the CPS, are not produced for this purpose. The result is that the data are not as good as they could be for the purpose of incorporating contextual data, and some important places—like Hennepin county in Minnesota—are missed if the data are used as they are delivered by the Census Bureau. We demystify the nuances of CPS geographic assignment by describing how CPS geographies are assigned and identified in the public-use CPS data. Then we propose strategies to recover and impute counties to increase the utility of CPS data for analyses that incorporate contextual information. The work we have done to date would be challenging and inefficient for every researcher using the CPS to do themselves. Our efforts reduce mistakes, save time, and increase reproducibility of research results.

Table 1. Variables indicating geographic locations in the Current Population Survey. The Original CPS column contains the labels of the variables provided by the US Census Bureau, and the IPUMS CPS column contains the labels of equivalent variables provided by IPUMS.

Geographic location	**Original CPS**	**IPUMS CPS**	**Description**
Metro vs. Non-Metro	GTMETSTA	METRO	Indicates whether a respondent lived in a metropolitan or non-metropolitan area
Central city status	GTCBSAST	METRO	Indicates whether a metropolitan respondent lived in a central city, outside the central city or an unknown location within the metropolitan area
Metropolitan area	GTCBSA	METAREA	Provides the FIPS code for the metropolitan area of residence
County	GTCO	COUNTY	Provides the FIPS code for the county of residence

Table 2. Combinations of sub-state geographic detail available in the full 2019 ASEC and in Minnesota in the 2019 ASEC (Flood et al. 2020).

	Specific metro area code	**Specific central city code**	**County code**	**Full ASEC (N)**	**MN ASEC (N)**
(a)	Yes	Yes	Yes	17,266	122
(b)	Yes	No	Yes	57,981	465
(c)	No	No	Yes	744[8]	0
(d)	Yes	Yes	No	6,760	156
(e)	Yes	No	No	54,383	704
(f)	No	No	No	42,967	853[9]
(g) Subtotal	136,390	24,026	75,991		
(h) Total				180,101	2,300

[8] The 744 respondents with a county code but no specific metropolitan area or central city code reside in one of six counties with populations that exceed the suppression threshold but are not classified as metropolitan areas by the Office of Management and Budget (US Census Bureau 2019).

[9] The 853 respondents include the 527 individuals who lived in the non-metropolitan part of Minnesota plus the 326 individuals who resided in a metropolitan area but did not receive a specific metropolitan area FIPS code.

Table 3. Central city county assignment and confidence in assignment.

Assignment Strategy / Confidence	**N (%)**
Central city nests within county or 100% of central city population resides in one county (100% confidence)	3,841 (57%)
Central city's population resides in two or more counties (< 100% confidence)	2,919 (43%)
Total	6,760

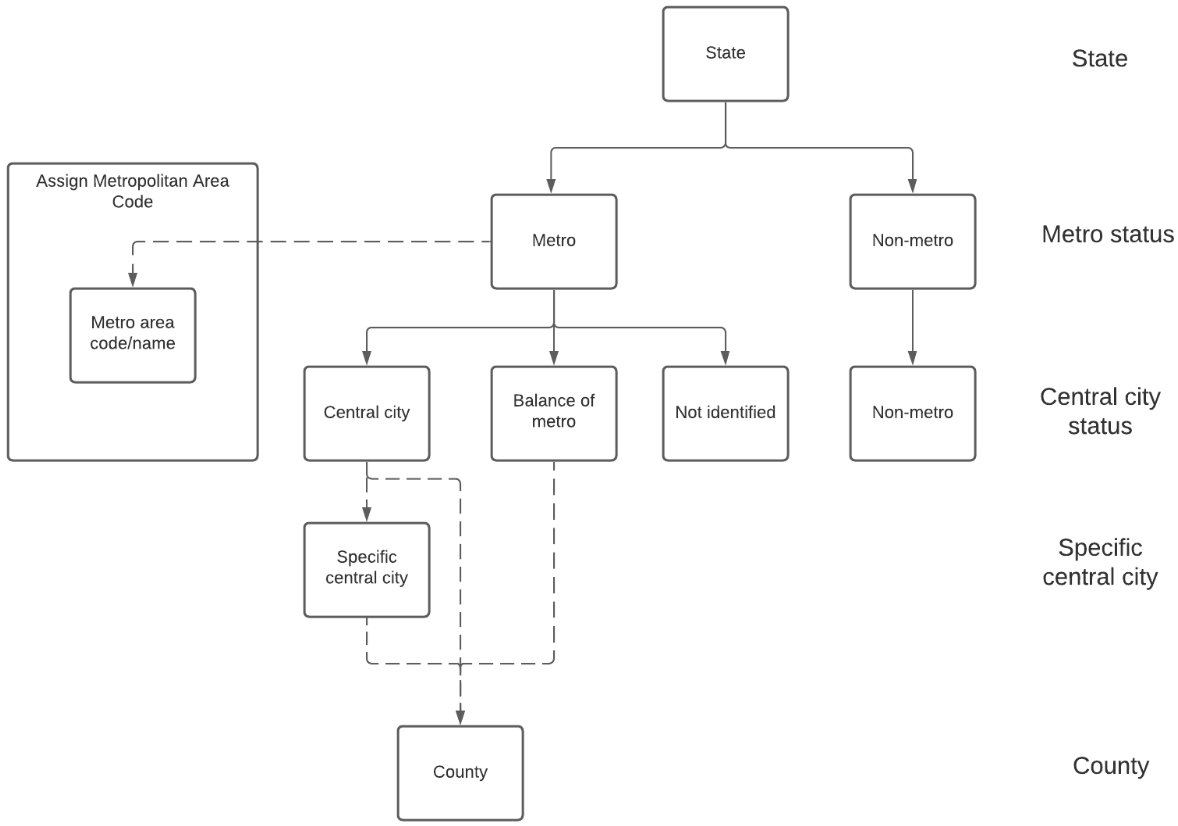

Figure 1. Hierarchy for the assignment of geographic identifiers in the CPS. The solid lines indicate codes assigned to all CPS respondents. The dashed lines indicate that metropolitan area, central city, or county codes are assigned to some CPS respondents.

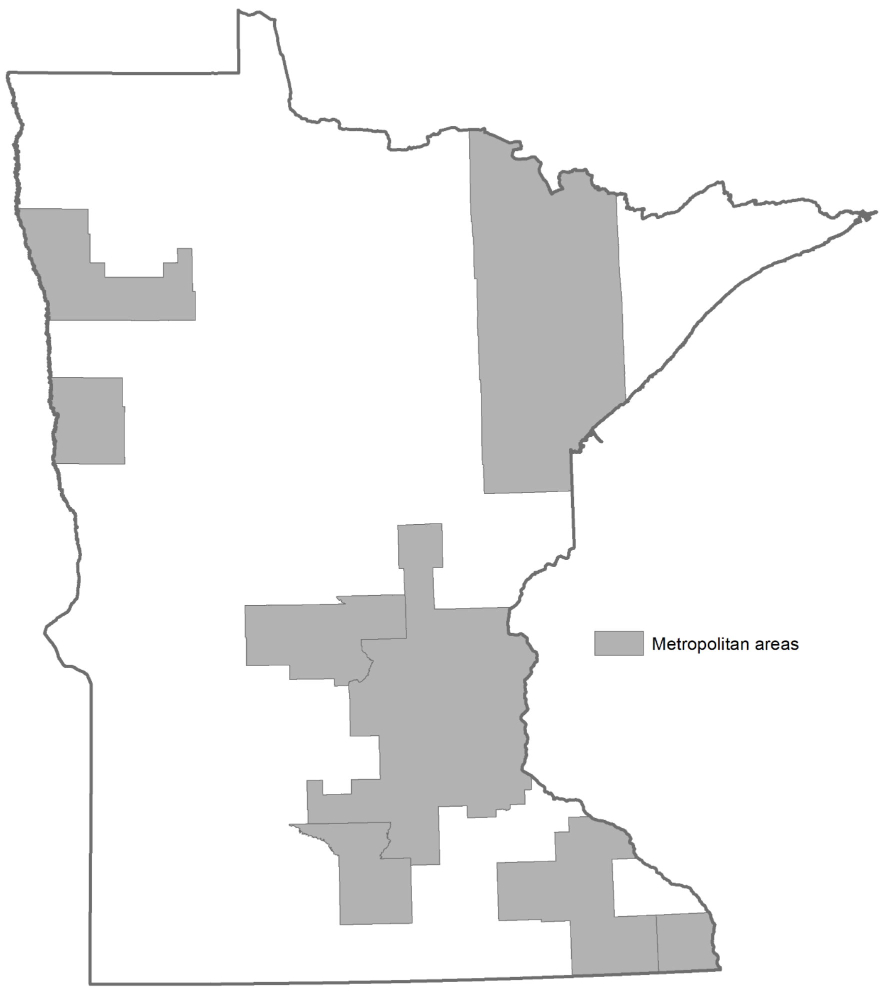

Figure 2. Metropolitan and non-metropolitan areas within the state of Minnesota. The gray areas represent metropolitan areas, and the white areas represent non-metropolitan areas.

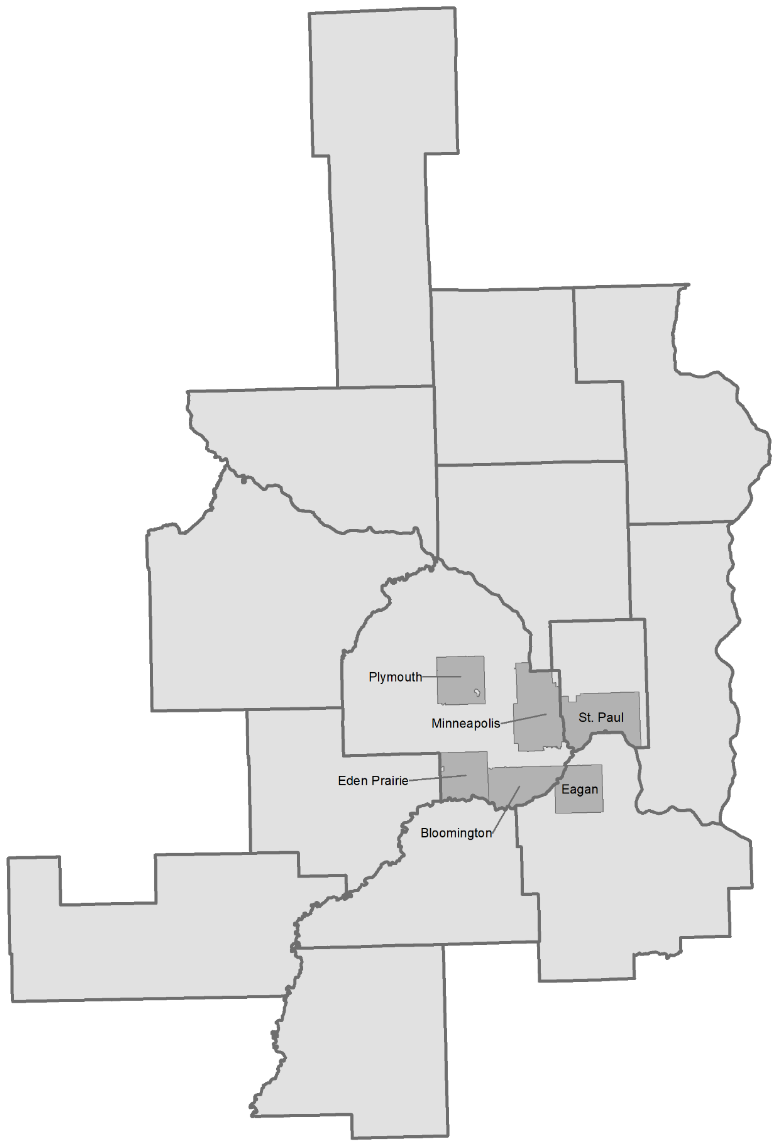

Figure 3. Central cities in the Minneapolis-St Paul, MN-WI metropolitan area. All CPS respondents who reside in these central cities are flagged as residing in a central city. Respondents residing in Minneapolis or St Paul receive codes specific to those cities.

Figure 4. Counties and central cities in the Minneapolis-St Paul, MN-WI metropolitan area. County names in bold are identified in the CPS. The number in parenthesis denotes the rank of the county by population within the metropolitan area. Additionally, central city names in bold are also identified in the CPS; thus, we know which respondents live in the cities of Minneapolis or St. Paul.

Chahine, T., S. V. Subramanian, and J. I. Levy. 2011. “Sociodemographic and Geographic Variability in Smoking in the U.S.: A Multilevel Analysis of the 2006–2007 Current Population Survey, Tobacco Use Supplement.” *Social Science & Medicine* 73 (5): 752–58. https://doi.org/10.1016/j.socscimed.2011.06.032.

Flood, Sarah, Miriam King, Renae Rodgers, Steven Ruggles, and J. Robert Warren. 2020. “Integrated Public Use Microdata Series, Current Population Survey: Version 8.0 [Dataset].” IPUMS. https://doi.org/10.18128/D030.V8.0.

Hermes, Kerstin, and Michael Poulsen. 2012. “A Review of Current Methods to Generate Synthetic Spatial Microdata Using Reweighting and Future Directions.” *Computers, Environment and Urban Systems* 36 (4): 281–90. https://doi.org/10.1016/j.compenvurbsys.2012.03.005.

Ingram, Samuel J. 2019. “Occupational Licensing and the Earnings Premium in the United States: Updated Evidence from the Current Population Survey.” *British Journal of Industrial Relations* 57 (4): 732–63. https://doi.org/10.1111/bjir.12469.

Mellor, Jennifer M., and Jeffrey Milyo. 2002. “Income Inequality and Health Status in the United States: Evidence from the Current Population Survey.” *The Journal of Human Resources* 37 (3): 510–39. https://doi.org/10.2307/3069680.

Office of Management and Budget. 2018. “OMB BULLETIN NO. 18-04: Revised Delineations of Metropolitan Statistical Areas, Micropolitan Statistical Areas, and Combined Statistical Areas, and Guidance on Uses Ofthe Delineations of These Areas.” https://www.whitehouse.gov/wp-content/uploads/2018/09/Bulletin-18-04.pdf.

US Census Bureau. 2019. “Current Population Survey: 2019 Annual Social and Economic (ASEC) Supplement Technical Documentation.” US Department of Commerce. https://www2.census.gov/programs-surveys/cps/techdocs/cpsmar19.pdf.

Weyland, Gregory. Personal communication. 2020. “CPS Geography,” November 23, 2020.

Whitworth, A., E. Carter, D. Ballas, and G. Moon. 2017. “Estimating Uncertainty in Spatial Microsimulation Approaches to Small Area Estimation: A New Approach to Solving an Old Problem.” *Computers, Environment and Urban Systems*, Spatial analysis with census data: emerging issues and innovative approaches, 63 (May): 50–57. https://doi.org/10.1016/j.compenvurbsys.2016.06.004.