POTENTIAL APPROACHES FOR UPDATING URBANIZED AREA POPULATION AND POPULATION DENSITY

Prepared for:
Transit Cooperative Research Board
Transportation Research Board
of The National Academies

Prepared by:
Ken Hodges
Nielsen

August 2011

The information contained in this report was prepared as part of TCRP Project J-11, Task 11, Transit Cooperative Research Program.

SPECIAL NOTE: This report IS NOT an official publication of the Transit Cooperative Research Program, Transportation Research Board, National Research Council, or The National Academies.
Acknowledgements

This study was conducted for the American Public Transportation Association, with funding provided through the Transit Cooperative Research Program (TCRP) Project J-11, *Quick-Response Research on Long-Term Strategic Issues*. The TCRP is sponsored by the Federal Transit Administration; directed by the Transit Development Corporation, the education and research arm of the American Public Transportation Association; and administered by The National Academies, through the Transportation Research Board. Project J-11 is intended to fund quick response studies on behalf of the TCRP Oversight and Project Selection (TOPS) Committee, the Federal Transit Administration, and the American Public Transportation Association and its committees.

The work was guided by a technical working group with representatives from transit agencies, metropolitan planning organizations, private consulting organizations, U.S. Bureau of the Census, the Federal transit Administration, and the American Public Transportation Association.

Disclaimer

The opinions and conclusions expressed or implied are those of the research agency that performed the research and are not necessarily those of the Transportation Research Board or its sponsoring agencies. This report has not been reviewed or accepted by the Transportation Research Board Executive Committee or the Governing Board of the National Research Council.
# Table of Contents

Table of Contents ................................................................. 2  
Executive Summary ................................................................. 3  
Note on Population Change and Urbanized Areas ................................. 4  
History of Small Area Population Estimates in the United States .............. 4  
Geography ............................................................................... 5  
Some Basic Population Estimation Methods ..................................... 6  
    Extrapolative Methods .......................................................... 6  
    Component Methods ............................................................ 7  
    Housing Unit Methods ......................................................... 7  
    Vital Rates Methods .......................................................... 8  
    Other Methods ..................................................................... 9  
Private Sector Methods and Evaluations ......................................... 10  
    Methods ............................................................................. 10  
    Evaluations ....................................................................... 11  
Broad Options for Urbanized Area Estimates ..................................... 11  
    Produce Estimates or Acquire a Completed Set ......................... 12  
    Estimate UAs Directly or Estimate Small Areas and Aggregate ......... 13  
    Build Estimates from Census Counts or Use An Independent Source “As Is” .... 14  
    Estimates with Top-Down Control Totals or Pure Bottom-Up Estimates ........ 16  
    Estimates Reflecting “Usual Residence” or Another Residence Rule .......... 18  
    Public Sources Only or Proprietary Resources ......................... 19  
    Estimates As a Time Series or Current Estimates Only .................. 20  
Examples of Input Data Sources ................................................ 20  
    USPS Address Counts ......................................................... 20  
    Private Sector Consumer Databases ......................................... 21  
    National Change of Address Data ........................................... 21  
Population Density ................................................................... 22  
Conclusions ............................................................................. 24  
Notes and References ............................................................... 25
Executive Summary

Urbanized areas are widely used in transportation planning, but they are updated only once per decade, following the decennial census. Census population totals can be up to 10 years out of date, so there is interest in estimates of population and population density for urbanized areas that could supplement the most recent census counts. This report describes approaches that could be used to produce such estimates for dates following the most recent census.

Note: throughout the report, the abbreviation UA is often used in place of urbanized area.

The delineation of urbanized areas (UAs) is tied to the population of census blocks, such that changes in population imply changes to UA boundaries. However, the objective of the report is to describe options for estimating the population and population density of UAs as fixed areas defined by previous census counts. The re-delineation, or updating of UA boundaries, is not a consideration.

Because some options involve the estimation of population for small areas that can be aggregated to urbanized areas, the report provides a brief history of small area population estimation, and background on geography – both urbanized areas and the small areas from which UA estimates could be built. The report describes some common estimation methods, including extrapolative, component, housing unit, and vital rates methods, as well as those based on surveys and statistical modeling. Rather than describing methods in “cook book” detail, the emphasis is on the basic approach, and suggested advantages and limitations with respect to UA estimation.

The report also identifies the major decisions to be made in designing an estimates program for UAs – decisions that have methodological implications, but which can cut across different methods. For example, one could either estimate UAs directly, or estimate small areas, and then sum to UA totals. Most methods start with census counts, and estimate change since the census date, but there is also the option of estimates derived from sources independent of the census. One also has to decide whether to use control total estimates for larger areas, and whether a periodic series of estimates is expected to serve as a time series. Another consideration is whether the estimates should be produced only with input data sources that are publicly available, as opposed to proprietary private sector data, or confidential government data. Several data resources – public and otherwise – that could be useful for UA estimates are described. The report finishes with a less extensive description of some of the options and issues related to the estimation of population density, which generally involve the specification of land area.

The report strives to provide a balanced presentation of options. Advantages and limitations are suggested, where relevant, but no recommendations are made concerning optimal methods or options for UA estimates.
Note on Population Change and Urbanized Areas

Urbanized areas are collections of census blocks classified as urban (as opposed to rural) based on their population density and proximity to large densely settled areas. They are delineated once per decade in a complex process applied to population counts from the decennial census. The designation of which blocks are part of UAs (and thus the boundaries of UAs) is a function of the population totals for blocks. Therefore, any estimation of block population totals implies a change in UA boundaries that would be realized if the delineation process were applied to them.

The Census Bureau has no mandate to update UAs based on unofficial population estimates. So it should be noted that this report describes options for estimating population and density for UAs as fixed areas as delineated following the previous census. UA boundary changes implied by the post-census population estimates are not reflected.

History of Small Area Population Estimates in the United States

The frequent production of population estimates for small areas nationwide is a relatively recent practice. The Census Bureau produced its first set of population estimates for all counties in 1966, and started producing population estimates for cities and towns in response to the State and Local Fiscal Assistance Act of 1972. The “Revenue Sharing” act called for the distribution of federal funds based on the most recent Census Bureau data, and the task of producing nationwide small area estimates was new and ambitious enough that a National Academy of Sciences panel was convened to review the Census Bureau’s methods.1 The Census Bureau does not produce population estimates for small statistical geographies such as census tracts and block groups – although population totals for such areas are a byproduct of the new American Community Survey.

Also in the 1970s, census data were made available to the public for the first time in computer readable form. Businesses became prolific and creative users of population data, often acquiring data in value-added form from private sector suppliers. Census data were valued for applications such as site selection, where small areas (such as census tracts) could be aggregated by the private suppliers to custom areas, such as a three mile radius or 20 minute drive time around a store.

The success of these applications, and the business sector’s impatience with aging census counts, led to demand for small area population estimates. Some local governments produced estimates for their census tracts, but most did not. And while the Census Bureau’s revenue sharing estimates were produced regularly for cities and towns nationwide, they did not provide small area detail within large cities or for unincorporated areas. The demand was for population estimates for small areas with “wall-to-wall” nationwide coverage.

Available methods and data resources were not well developed to meet this demand. But in business, demand does not yield to methodological adversity, so the private suppliers
started producing annual population estimates for all census tracts and minor civil
divisions nationwide. By 1980, several companies were producing annual census tract
level estimates with methods ranging from simple to ambitious. In the 1990s, computing
capacity, and competitive pressures, led the suppliers to produce estimates for even
smaller areas called block groups, and in recent years, suppliers have started estimating
population for census blocks – the smallest unit of census geography, and the building
blocks for UAs.

Many demographers have looked askance at the mass-produced, for-profit estimates, and
wondered how accurate they could be. Early efforts were rudimentary, but the forces that
pressed the suppliers to increasingly audacious feats of small area estimation, also
promoted the use of some impressive, if sometimes mysterious resources. And while the
industry is not known for transparency, there is some history of evaluations dating back
to the 1980 census, with results occasionally reported in papers presented at professional
meetings (see discussion of evaluations in the “Private Sector Methods and Evaluations”
section).

**Geography**

Because some options for UA estimates involve the estimation and aggregation of small
areas, it is important to review the relevant geographic levels, and how they relate to each
other.

Most estimation methods are executed within a hierarchy of geographic units that nest
neatly within each other. For example, the U.S. is subdivided into 50 states plus the
District of Columbia, and the states are subdivided into 3,141 counties (or equivalent
units). Counties are divided into about 65,000 census tracts, which are made up of about
208,000 block groups. Block groups are groups of the roughly 8.2 million census blocks
defined nationwide, and which are the smallest units for which the census reports
population totals.

The correspondence between geographic levels can have a bearing on the approach taken
for population estimates. Because blocks, block groups, tracts and counties nest within
each other, they facilitate the use of control total estimates – an option referenced
frequently below. For example, if the county level sum of census tract estimates is three
percent lower than an independent estimate for the county, the estimates for all tracts in
the county can be adjusted upward by three percent to make them sum to the county
“control” estimate.

Blocks, block groups, and tracts are redefined with each new census, but they have the
advantage of being stable for the 10 years between censuses. Thus, these small
“statistical geographies” are updated only as frequently, and on the same schedule as the
updating of Urbanized Areas. In contrast, ZIP Codes, cities, towns, and even counties are
subject to change at anytime – although county changes are relatively infrequent.
For historical perspective, it is useful to note that while blocks, block groups and tracts are now defined wall-to-wall nationwide, that was not always the case. Prior to 1990, blocks were not defined for many nonmetropolitan counties, so they had no block groups or census tracts. In those counties, estimates were typically provided for either minor civil divisions (MCDs) or census-defined alternatives called census county divisions (CCDs). MCDs and CCDs were subdivided into smaller units called enumeration districts for which estimates were not usually produced.

Urbanized Areas are defined by the Census Bureau to separate urban and rural territory and populations. They are densely settled areas with populations of 50,000 or more, and consist of clusters of blocks or block groups that meet a complex set of conditions related to population density and proximity to densely settled areas. Blocks are the only units of census geography that are always completely urban or completely rural, so blocks also are the only units that can be summed cleanly to urbanized areas.

**Some Basic Population Estimation Methods**

The methods described below are among those commonly used for population estimates.

**Extrapolative Methods**

Extrapolative methods typically project population based on trends observed in the recent past. For example, one could observe the population count for the 1990 and 2000 censuses, and extrapolate the rate of change to the current year target date. Extrapolations can be linear or nonlinear, or can even incorporate more than two points in time. But they are projections of previous trends, and do not incorporate measures of population change that extend to the target date.

Another simple approach is to make use of estimates already prepared for larger areas, such as states or counties, and simply apply the larger area’s rates of population growth or decline to the smaller component areas – such as census tracts or block groups. Again, the approach is simple and inexpensive to execute, but benefits from no current data at the level being estimated.

This approach is sometimes referred to as the share, or constant-share method, since the sub-areas maintain a constant share of the population of the larger area. In a variant shift-share method, one estimates the change in each sub-area’s share of the larger area’s population. Estimates of the change in share can be based on either extrapolation or data specific to the sub-areas.

Extrapolative methods have the advantage of being easy and inexpensive to execute, but in the absence of data reflecting post-census change, they tend to be inaccurate for areas where population trends have changed or even reversed. And for small areas, there are logistic challenges. Because census tracts, block groups and blocks are redefined with each census, one cannot determine the trend from one census count to the next without reconciling the changes in geography. For example, in order to project change from 2000
to 2010 out to later years, one would need to express the 2000 census counts for the blocks and block groups for which the 2010 census population counts are reported.

**Component Methods**

Component methods start with a population at Time 1, and estimate the components of population change to derive estimates of population at the estimation date – Time 2. Specifically, the population estimate reflects the original population, plus births, minus deaths, plus persons who move into the area, minus persons who move out of the area.

The basic component method formula is as follows.

\[ P_2 = P_1 + B - D + I - O \]

where:
- \( P_2 \) = Population at Time 2
- \( P_1 \) = Population at Time 1
- \( B \) = Births from Time 1 to Time 2
- \( D \) = Deaths from Time 1 to Time 2
- \( I \) = In-migrants from Time 1 to Time 2
- \( O \) = Out-migrants from Time 1 to Time 2.

The formula is a definition of population at the estimation date – the challenge is in acquiring data that accurately reflect the components of change for small areas nationwide. Data on births and deaths typically derive from birth and death records collected locally, and compiled nationwide by the National Center for Health Statistics. Like all administrative data, birth and death statistics are subject to error, but the biggest source of error in component population estimates traces to the estimates of in and out migration (or net migration). Where a county, tract, or block group experiences dramatic population growth or loss, it is usually the result dramatic migration, in or out of the area. The Census Bureau uses this component approach in producing its population estimates for U.S. counties.

Component methods have an intuitive appeal in that they reflect the basic process of population change – births, deaths and migration. But data tracking these components are difficult to come by for small areas, and one might have to estimate them based on the age/sex composition of small area populations (probably as of the most recent census year). For example, in the absence of recent birth and death counts for blocks or block groups, one might estimate deaths by applying survival probabilities to population by age/sex, and estimate births with a rough fertility measure, such as the child/woman ratio (population age 0-4 / female population age 15-44). In short, component methods are conceptually straightforward, but can be cumbersome to execute.

**Housing Unit Methods**

The housing unit method is another popular approach to population estimates, and another based on a formula that defines population for the estimation year. Reduced to
The method estimates housing units for the target date, then derives households (or occupied units), then population in households, and finally total population. 

The basic housing unit method formula is expressed below.

\[ P2 = (((HU1 + NU – DU) * OC) * PPH) + GQ \]

where:
- \( P2 \) = Population at time 2.
- \( HU1 \) = Housing units at time 1
- \( NU \) = New units constructed from time 1 to time 2
- \( DU \) = Units demolished from time 1 to time 2
- \( OC \) = Percent of housing units that are occupied
- \( PPH \) = Average number of persons per household
- \( GQ \) = Group quarters population estimated for time 2.

The method starts with total housing units for the estimation year – either from a direct count (such as from administrative data) or by adding new construction and subtracting demolitions from a census (or other) housing unit count. With the estimates of housing units established, estimated occupancy rates are used to derive estimates of households (or occupied housing units). Estimates of average household size (persons per household) are applied to estimate the number of persons living in households, and estimates of persons in group quarters (persons not living in households) are added to derive the estimates of total population.

Like component methods, housing unit methods are based on a formula that defines population for the estimation year – and again, the challenge is in acquiring data that accurately reflect the elements of the formula. The estimates of housing units are subject to error, but the estimates of occupancy rates and persons per household often are considered weak links in housing unit method estimates. And estimates of persons in group quarters also are notoriously subject to error. One could accurately estimate the growth or decline in housing units, but introduce substantial errors to the population estimates through error in estimating occupancy rates, average household size, and the size of the group quarters population. On the other hand, where occupancy, household size, and group quarters population are relatively stable, housing unit estimates can be very accurate.

Many state and local governments use the housing unit method to produce population estimates for areas in their jurisdiction, and the Census Bureau uses the method to produce its population estimates for cities and towns. In fact, the Census Bureau has been conducting research on the potential application of housing unit methods for its county population estimates.²

**Vital Rates Methods**

Vital rates methods estimate change in population based on change in vital events, such as births and deaths. While counts of total population are not regularly available (the
U.S. Census is once per decade), counts of events such as births and deaths are frequently updated, and the idea is that change in total population is reflected in observed changes in vital events.

The key assumption is that the ratio of births or deaths to total population is reasonably stable over time, and that as population grows or declines, that change will be reflected in increases or decreases in vital events. In addition to events such as births and deaths, the method can estimate population based on changes in other indicators, such as utility hookups, automobile registrations, and voter registrations. For estimation programs that are nationwide in scope, one needs to consider that some indicator data are more readily available in some locations than others.

Vital rates methods can be easy to execute, and require less data than component and housing unit methods. One needs only total population for an origin date (such as the most recent census count) and the vital events count for the origin date and the estimation date. If the vital events count (let’s say deaths) doubled from the origin date to the estimation date, the method estimates a doubling of population from the origin to the estimation date.

An example of a vital rates method formula (based on deaths) is presented below.

\[ P_2 = \left( \frac{P_1}{D_1} \right) \times D_2 \]

where:
P2 = Population at time 2
P1 = Population at time 1
D1 = Deaths at time 1
D2 = Deaths at time 2.

The primary risk with the vital events approach is that the ratio of vital events to population can change over time, resulting in erroneous estimates of population change. Also, assigning vital events to small areas (or even counties or UAs) can be problematic. Events such as births or deaths might not be assigned to the newborn’s or deceased’s home address, and the assignment of addresses to small census geographies is subject to error. Data availability can pose logistic obstacles, especially if vital statistics data are needed for small areas nationwide. Event data might be more readily available in some areas than others, and even where data available on a regular schedule for small areas, one might need to acquire, process, and assess the suitability of data from multiple sources.

**Other Methods**

Other estimation methods, with which the author has less experience and expertise, include statistical model based estimates and survey based estimates. A model based method might build a model expressing the relationship between total population and indicator variables, such as vital events, school enrollment, or utility hookups – or the relationship between population change and change in the indicator variables. The model
is then applied to the indicator data available for the estimation year to produce estimates of population. Statistical models can determine more refined relationships between population and indicator variables than simple vital rates ratios, but they typically require more data, and can pose even greater logistic challenges. And like vital rates methods, model-based estimates depend on the quality of indicator data, and rest on the assumption that the statistical relationships specified by the model are stable over time.

Survey methods start with a sampling frame, such as a list of addresses, and conduct field work to determine the number of persons living at those addresses. Even if based on just a sample of the addresses in an area, the survey results provide an estimate of the number of people living in the area. Survey estimates can be labor intensive, and thus expensive to complete for large areas, or where the scope is nationwide. And they require the existence of a complete and up to date list of addresses for use as a sampling frame. However, the approach has been useful for selected areas in specific circumstances. For example, in late 2005 and early 2006, the City of New Orleans Emergency Operations Center fielded surveys to develop population estimates in the wake of Hurricane Katrina.3

Private Sector Methods and Evaluations

Because annual small area population estimates on a nationwide scale have been produced primarily in the private sector, it is worth taking a closer look at the methods that have been used, and the results that have been achieved. The private sector experience might help inform the decisions to be made in developing an estimates program for urbanized areas.

Methods

The private suppliers’ descriptions of their methods sometimes are designed more to promote than to inform, so the history of their methods is not well documented. The account that follows draws on the author’s experience in the industry dating to 1982. Company identities are omitted, as they are not relevant to the objectives of this report.

The production of nationwide small area population estimates by private suppliers dates to the early to mid 1970s, when estimates were produced for census tracts in metropolitan areas and MCDs (or CCDs) in nonmetropolitan counties. Some methods were rudimentary – relying on extrapolation or the top-down distribution of trends for larger areas. Some were elaborate – applying regression models to tracts categorized by likelihood of population growth. One supplier acquired and incorporated tract estimates produced by local governments (where available), and another tracked change in a large consumer name/address list that the company maintained for its direct mail business.

During the 1990s, as estimation was extended from the tract to the block group level, some companies continued with extrapolative methods, but typology-based methods were discontinued, and the use of household counts from proprietary databases was expanded. And a number of companies started using address counts from the US Postal Service as a basis for their estimates. USPS provides counts of active residential addresses for carrier routes (the area covered by a mail carrier), so the suppliers have to estimate their
distribution to block groups in order to use them as input to estimates for conventional geographic units.

In the most recent decade, extrapolative methods have all but vanished, as most if not all suppliers use publicly available USPS address counts and/or proprietary databases. And with advances in computing capacity and address coding, combined with industry competition and elevated user expectations, estimates at the block level have become more common. Most suppliers still control to independent county estimates, such as those from the Census Bureau, but few control to estimates for cities and towns (given the difficulties described below), and some suppliers might be bypassing control totals altogether in favor of straight sums from proprietary databases.

Evaluations
The accuracy of population estimates can be measured once every decade by comparing estimates for the census year with subsequent counts from the census. Few companies publish evaluation results, but the author has been involved with such evaluations dating to an evaluation of 1980 estimates that was probably the first evaluation reported by a private supplier. The author also conducted and reported an evaluation of private supplier estimates produced for 1990 and for 2000. As of this writing, preliminary evaluations of 2010 estimates are underway, and while the results have not yet been formally reported, they are cited below where relevant.

The evaluation results conform to expected patterns. Errors have been larger for areas with small populations, and those that changed the most between census years. In 2000, for example, the mean absolute percent error (MAPE) for county population estimates was 3.4, meaning the 2000 population estimate was, on average, 3.4 percent different from the census count. For tracts, the MAPE was 11.6 and for block groups it was 15.3. And error rates have improved over the decades. For tracts, the MAPE for population estimates has dropped from 15.8 in 1980 to 13.1 in 1990 and 11.6 in 2000. Preliminary results indicate a further reduction to about 9.7 percent for 2010.

Perhaps more important, the evaluations have found a relatively even mix of overestimation and underestimation – a lack of bias likely enhanced by the use of control total estimates for larger areas. The payoff is that the sometimes large positive and negative errors observed for individual small areas tend to offset, such that error is much lower for aggregate areas.

The results from past evaluations, and preliminary results from 2010 evaluations are cited below where they are relevant to the consideration of options for producing population estimates for urbanized areas.

**Broad Options for Urbanized Area Estimates**

A variety of options is available for producing population estimates for urbanized areas (UAs). Some of the major options are listed below.
• Produce estimates – or – acquire a completed set
• Produce directly for UAs – or – produce for small areas and aggregate
• Produce from census counts – or – independent source “as is”
• Produce with top-down controls – or – pure bottom up.
• Produce for usual residence definition – or – for another residence rule.
• Public sources only – or – proprietary resources
• Estimates as a time series – or – current estimate only

These options define the major decisions to be made in establishing an approach to UA estimation. A discussion of strengths and limitations follows.

Produce Estimates or Acquire a Completed Set
At least on paper, there is the option to acquire population estimates for UAs from a third party rather than producing them. Most likely, this option would involve the annual purchase of population estimates from a private sector supplier. Specifically, one would acquire the supplier’s small area estimates aggregated to UAs (with aggregations performed either by the supplier or those acquiring the supplier estimates). A variant of this option would be to acquire estimates from several suppliers, and average them to a final estimate. Estimates are often averaged to minimize extreme errors, and in this case, averaging would preclude any perceived endorsement of a single supplier. The risks are that one supplier with notably inaccurate estimates could undermine the advantage of averaging, or that a changing list of suppliers producing estimates might yield a changing base for averaged estimates.

Acquiring from a supplier would be simple – eliminating the need to determine a method; acquire and understand input data, and actually produce estimates. Acquiring such estimates might even be less expensive than producing them. But the use of commercial estimates for some purposes might be a non-starter, and even if the option could be considered, one would have to work with the suppliers, and determine which one to use (unless all are acquired and averaged). One might also have to be willing to accept a less than complete understanding of the suppliers’ estimation methods. Again, continuity could be an issue, as a private supplier could change methods or even discontinue its estimation program in response to market conditions or corporate restructuring.

Another option, at least on paper, would be estimates of population (and density) from the American Community Survey (ACS). The ACS is the large survey that is now the source of data previously provided by the long form census. Every year, the ACS provides estimates of a wide array of demographic and socioeconomic variables nationwide for areas as small as block groups.

The ACS reports population totals, so one could aggregate the block group totals to UAs for use as population estimates. However, the ACS block group population totals are not official Census Bureau population estimates, and the Census Bureau stresses that the ACS is best viewed as a source of estimates of characteristics, not population totals. Among the complications are that 1) ACS data do not conform to the residence rules used by the census (some seasonal populations might be included), and 2) the small area data
are based on five years of data collection, and therefore reflect periods of time rather than a point in time. For example, the ACS block group data released in late 2010 reflect population totals not for a single year, but for the period 2005-2009. For larger areas, ACS also provides 1-year and 3-year estimates that could be considered for use as control totals – although for those levels, the Census Bureau would recommend its official population estimates. A potential obstacle to the use of block group ACS population totals is the fact that ACS block group data are provided only on the large and complex ACS Summary File product – and not through the more readily accessible American FactFinder products.

Despite the complications, there is reason to expect ACS totals to reflect household (and population) growth and decline, since its sample is based on the Census Bureau’s Master Address File, which benefits from updates from the U.S. Postal Service. In fact, preliminary evaluations of the 2005-2009 ACS household totals confirm their ability to reflect growth. For example, a block group that was part of Denver’s former Stapleton Airport, now Stapleton development, had just three households in 2000, but 4,092 in the 2010 census. The ACS household total of 2,901 is well short of the 2010 total, but reflects the growth, and is probably reasonable for the 2005-2009 period that it reflects. If used as estimates for 2010, the ACS 2005-2009 household totals would have a mean absolute error of 14.8 percent – not too different from some of the other options described below.\(^8\)

The remaining options assume the production of population estimates.

**Estimate UAs Directly or Estimate Small Areas and Aggregate**

The objective is to produce estimates for urbanized areas, but the methodology need not be applied directly to UAs. Instead, estimates could be produced for smaller “building block” geographic units, and aggregated to UAs. For example, one could produce estimates for all census blocks or block groups, and aggregate to UAs.

The direct estimation of UAs would require fewer steps, and might be quicker and less expensive to execute. However, few input data sources would come already coded to UAs, so that would be a required task. Geocoding to UAs (because they are large) would be subject to less error than geocoding to blocks or block groups – although the impact of this advantage would largely disappear once the small area estimates are aggregated to UAs. Direct estimation also eliminates the need to produce estimates for geographic units that are not part of urbanized areas. Building block approaches typically produce estimates for the full universe of building block geographies, and then aggregate those that are components of the areas in question.

The estimation of wall-to-wall building block geographies would enable a flexible response to any re-delineation of UAs (official or unofficial), and facilitates the use of control totals, in which small area estimates can be adjusted to independent estimates for larger areas (see discussion of control total estimates below).
If a building block approach is used, one must choose the unit of building block geography. Nine-digit ZIP Codes (ZIP+4s) have become popular with some data users, as they are smaller than census blocks (often corresponding to a block face), and thus offer powerful granularity for aggregations. However, ZIP+4s can be unstable over time, and are so small that census counts are not available for them – thus precluding estimates that build upon census counts. Also, postal codes reflect locations or points for the delivery of mail. They are not defined by boundaries, and therefore lack definitive land areas from which population density could be computed (although this limitation might be of little practical concern in establishing land areas for areas as large as UAs).

Census tracts, block groups, and blocks are larger than ZIP+4s, but they have stable and definitive boundaries, and can be summed (and therefore controlled) to counties. Tracts have the advantage of larger size and lower estimation error, but since they do not sum cleanly to UAs, proportioning would be required in aggregation. For example, a UA might be defined as 100 percent of 20 tracts plus 72 percent of another and 4 percent of another – and the inclusion percentages are subject to error. Block groups are smaller, and less dependent on splitting for UA aggregation, but splitting (and its associated error) still would be required. Only blocks could be aggregated straight to UAs without splitting, but block estimates are a greater challenge, in terms of execution and available input data, and are subject to greater error than those for block groups or tracts.

Build Estimates from Census Counts or Use An Independent Source “As Is”

Another decision concerns whether to build estimates using previous census counts as a starting point, or adopt counts from another source as finished estimates – with no reference to census counts.

The use of census counts as a starting point is common enough to seem a default option. Component methods typically start with census population counts, then add births, subtract deaths, and add/subtract net migration to estimate population for the post-census year. The housing unit method usually starts with census housing unit counts, adds newly constructed units, subtracts demolitions, then applies estimated occupancy, persons per household, and group quarters population to estimate population for the post-census year. Even methods based on address/household databases often use database trends to advance census counts to post-census years.

Estimates building from census counts have the advantage of starting with numbers which, while not perfect, are the standard for accuracy. Many small areas experience little change from one census to the next. For example, from 1990 to 2000, 54 percent of block groups experienced household growth or decline of less than 10 percent, and only 9 percent experienced growth or loss of 50 percent or more. So the task is to detect which areas experienced significant change – and how much – relative to the previous census count. Of course, census counts can be in error, and to the extent they are, estimates built from them are likely to perpetuate those errors. For example, the 2000 census placed numerous prison and college dormitory populations in the wrong location, resulting in large errors in the population counts for the tracts and block groups involved – errors that census-based methods cannot feasibly correct until they are based on 2010 census counts.
An alternative to census-based estimates is to use small area counts from an alternative source, such as an address or consumer database (or public sector administrative data), as finished estimates. For example, one could start with current address counts from USPS as estimates of total housing units, or a consumer name/address list as estimates of total households. Because they are independent of previous census counts, these estimates would be immune to the impact of error in the census, and could reflect changes that might be dismissed as unrealistic, or too extreme relative to census counts.

A drawback to using database counts “as is” is that most reflect either addresses (approximating total housing units) or name/address records (approximating occupied housing units, or households). Few administrative sources reflect total persons or population. To produce population estimates, one might start with total housing units, estimate occupancy to derive total households (occupied units), estimate persons per household to derive population in households, and add estimated persons in group quarters to complete the estimate of total population. In other words, one would apply the latter steps of the housing unit method (described above), and each additional step is an opportunity to introduce additional error.

Another drawback is that many database sources have incomplete coverage. For example, telephone book listings tend to miss households with unlisted telephone numbers, and those with no landline telephone. Similarly, records based on tax returns exclude those who do not file. Geocoding, or the assignment of addresses to geographic areas, is another limitation. Although improved over the years, address-coding is still subject to error. A database could have all 3,000 addresses in Area A, but if 500 are mistakenly assigned to Area B – the result is an erroneous count for Areas A and B.

To compensate for coverage and geocoding issues, some methods use database rates of change to project census counts to current dates. For example, a tract with a 2000 census count of 1,500 households might have had 1,100 database households in 2000, and 1,300 database households in 2009. Rather than estimating 1,300 households for 2009 (an implicit decline), the method notes the 18.2 percent increase in “list” households, and applies that growth rate to the census household count to produce an estimate of 1,773 households.

The rate-of-change approach combines the strength of continuously updated database counts with the solid foundation of census counts, but is subject to its own error. Like all census-based methods, the rate-of-change approach will perpetuate errors in the census counts, and the database sources can introduce error if coverage is not stable. In the example above, the 2000 database count of 1,100 households relative to the census count of 1,500 represents a coverage of 73.3 percent. The rate-of-change based estimate of 1,773 households assumes coverage stable at 73.3 percent (if coverage is stable, database change reflects actual change). But if the list increase from 1,100 to 1,300 households reflects only an increase in database coverage, the rate-of-change based estimate gives a false indication of growth. The rate-of-change approach also requires that one have database counts back to the previous census year. And compilation and address coding
need to be consistent over the years to increase the likelihood that database changes reflect actual changes.

An evaluation in 2000 found lower error rates for estimates building from census counts than those based on 2000 database counts used as 2000 household estimates. The census-based household estimates for block groups had a mean absolute percent error (MAPE) of 13.3, while USPS address counts had a MAPE of 28.6, and counts from four private databases recorded MAPEs ranging from 25.7 to 31.4. However, the situation might be changing. Preliminary 2010 findings indicate that mean error for USPS address counts dropped in 2010, but was still high at 20.9 percent (some of that error tracing to the need to convert USPS counts from carrier routes to census geography). However, a set of 2010 address counts developed by a private supplier from private sector address lists had a block group MAPE of only 12.5 relative to the 2010 census household counts. While less accurate than the supplier’s census-based estimates (with a MAPE of 10.4), the address counts were consistently more accurate in rapid growth areas. For block groups with household growth of 50 percent or more (from 2000 to 2010), the census-based estimates had a MAPE of 23.9, while the database estimates had a MAPE of only 12.6. In fact, the database estimates were more accurate than the census-based estimates in about half of all block groups.

The choice between census-based and database estimates is not simple. Address counts reflect housing units, or maybe occupied units, and still require error-prone estimates of persons per household and persons in group quarters to build estimates of total population. Perhaps more important, census-based estimates have very low error immediately following a census, when change has been minimal. The 10.4 MAPE for census-based household estimates in 2010 likely reflects a high, or maximum, as error increased from very low levels in 2001 and 2002. Because database estimates are independent of the census, their error would be as high immediately following a census as it is 10 years after a census. So even where database estimates might have an advantage 10 years after a census, they would not necessarily have produced more accurate estimates for all years between census counts.

Given these results, one could contemplate a hybrid method that uses census-based methods where change has been modest, and database estimates where change has been substantial. Such a method presumes that one could accurately determine which areas have experienced substantial change, and determine a threshold of change at which one would switch to database-based estimates. The hybrid approach also could be resource intensive, as one would need to produce a full set of census-based estimates, and maintain a full set of database counts, and combine the two to the final set of estimates.

Estimates with Top-Down Control Totals or Pure Bottom-Up Estimates
As described above, the use of top-down control total estimates involves the ratio-adjustment of estimates at one geographic level to independent estimates for larger areas. For example, if estimates for census tracts sum to a population 2.5 percent higher than an independent estimate for the county, the tract estimates can be reduced by 2.5 percent to force them to sum to the county “control total” estimate.
The use of control totals ensures the consistency of estimates between geographic levels – block groups sum to tracts, tracts sum to counties, and so on. But more important, it ensures that small area estimates sum to totals believed to be accurate, or are consistent with estimates from a trusted source. For example, one might want tract population estimates to sum to county totals that are consistent with population estimates from the Census Bureau or a state demographer.

Evaluations following the 2000 census found that the use of control totals added little to the average accuracy of census tract estimates. In fact, tract level pre-control household estimates had a slightly lower MAPE (9.7) than the final post-control estimates (9.9). However, when summed to county level, the pre-control estimates were less accurate (MAPE = 6.8) than the county control estimates (MAPE = 4.1). The payoff to the control process was not in the accuracy of the tract estimates, but in enhanced accuracy at the control total level. Enhanced accuracy at the control level reduces bias, and promotes improved accuracy as small area estimates are aggregated to larger areas – such as UAs.

Preliminary 2010 evaluations do not fully replicate the 2000 finding. Results run this time for block groups (the level where controls were applied), indicate pre-control estimates that were once again about as accurate as post-control or final estimates (pre-control MAPE = 6.2 compared with final “controlled” estimates MAPE = 6.7). But in contrast to 2000, the 2010 pre-control estimates summed to county level (MAPE = 3.5) were about as accurate as the control total estimates (MAPE = 3.7). Again, these results are preliminary, but if county sums of 2010 pre-control estimates were as accurate as the county control totals, the finding holds only for census-based estimates. County sums of the USPS and address database totals (described above) had markedly higher MAPEs (14.6 and 19.7) than the final county level estimates. Thus, the use of control totals might be a more important component of database methods than census-based methods.

There are potential drawbacks to the use of control totals. Even where the control estimates are accurate, they will reduce the accuracy of some small area estimates. And error in the control estimates can introduce additional error to otherwise accurate small area estimates. For example, if a control estimate for a county is 12 percent too high, the estimates for the block groups in that county will be biased to overestimation, and would contribute to overestimation of UAs that they are part of (unless offset by underestimation in other block groups). The tradeoff is that control totals give one greater control over estimates at the control total level, but at the expense of somewhat less control over the estimates for individual small areas. And not knowing exactly where control totals increase or decrease accuracy is one of the realities of their use.

The use of control totals also adds time and effort to the production of estimates, and even if the use of controls is a given, one has to determine the geographic levels for which controls will be applied. For example, if one chooses to estimate blocks as the building block for UA estimates, one could control directly to the county level, and ensure consistency with the Census Bureau’s county population estimates. But one could also control the block estimates to independent estimates for block groups, which in turn
could be controlled to independent estimates for tracts, which are controlled to county estimates. This sequential set of control totals might compensate for the greater risk of error in block estimates, and might enhance accuracy for estimates aggregated to UAs. But it adds significantly to the time and effort required to produce annual estimates.

The Census Bureau also produces population estimates for cities and towns that can be projected to current dates to provide intermediate control estimates between the county and tract levels. But control estimates at this level can be difficult to work with. Tracts and block groups do not nest cleanly within cities and towns, so one would need files indicating the complex correspondence between these levels (including percents for partial inclusions). The correspondence is especially elusive where city and town boundaries change between censuses. Many of these changes are reflected in the Census Bureau’s population estimates, so one would have to detect such changes, and then somehow account for them when controlling tract or block group estimates to city/town estimates.

Also complicated is the controlling of city/town estimates to the county level. Some cities cross county lines, so the control process must limit to city/county parts (the parts within a county). And in many states, unincorporated areas need to be added to the city/town estimates in order to sum to county. On paper, this is not a major obstacle, since one can treat the unincorporated “balance of county” as just another city. But the Census Bureau’s city/town population estimates are produced with a different method than that used to produce the county population estimates, and experience suggests that population growth in some “balance of county” areas may be dampened, as the county estimate caps total growth for the sub-county areas. For example, estimated population growth in incorporated areas (governments that often challenge estimates thought to be low) can leave little room for estimated growth in the unincorporated county balance. The bottom line is that city/town can be a valuable control total level, but it is a complicated moving target, and the results are sometimes questionable.

Because urbanized areas do not sum cleanly to larger areas (such as counties or states), the use of control totals most likely requires a building block approach, as opposed to the direct estimation of UAs. There is reason to expect that controlling small area estimates to independent control totals can reduce bias, and promote the enhanced accuracy often achieved through aggregation. And while UAs cannot be controlled to Census Bureau estimates for cities, counties or states, the adjustment of small area building block estimates to control totals can be expected to give UA estimates some degree of consistency with Census Bureau population estimates.

**Estimates Reflecting “Usual Residence” or Another Residence Rule**

The U.S. census counts people according to a concept of “usual residence,” or where they live “most of the time.” Thus, all persons whose usual residence is in the U.S. (regardless of citizenship or legal status) are to be counted. The census also is charged with counting each person only once and “in the right place” – again according to usual residence. For example, a person who lives in New York from spring through fall, but at a second home in Florida during winter, should be counted in New York, since it is where they live most
of the time. Where people vote, pay taxes, or have legal residency status is not the criterion – it is where they live most of the time.

Population estimates can conform to the census usual residence concept or some alternative residence rule. For example, some areas experience significant population swings as seasonal residents arrive and then leave, and for some purposes, it might be useful for population estimates to include the seasonal population. However, such estimates often “double count” to numbers in excess of total U.S. population, as seasonal populations are more easily added to seasonal areas than they are subtracted from areas of usual residence. For this reason, one would not want to adjust small area estimates based on alternative residence rules to control total estimates based on census residence rules (although this is what the American Community Survey population totals do).

Estimates using census counts as a starting point are generally locked into the census residence definition. This definition has the advantage of being clearly defined, and guards against counting individuals in more than one place. On the other hand, consumer and administrative databases often do not conform to census residence (which is one reason coverage can be an issue). For example, a list of electric utility customers might legitimately include people who live in the area only for a few months each year. Estimates based on such sources, used “as is,” often will reflect something other than the census residence definition. The challenge is in determining exactly what that alternative definition is.

Public Sources Only or Proprietary Resources
Another decision concerns the use of proprietary resources from for-profit companies. Do the UA estimates need to be produced only with resources that are publicly available or is the use of proprietary resources acceptable? If proprietary resources can be considered, a wider range of options would be available, but costs likely would be higher, and one would have to be comfortable with the sometimes less than transparent nature of private sector databases. Proprietary databases also introduce the risk of discontinuities and disrupted availability, as companies can dissolve or change or even discontinue their database products. Of course, the continuity of publicly available public sector databases also can be at risk, especially in times of tight budgets.

Limiting to publicly available resources likely would reduce costs, and promote transparency, but it would preclude the use of resources that are powerful, and perhaps easier to use – as ease of use is one of the primary ways private data companies add value to public data resources.

If the UA estimates are produced for government purposes, one might also consider the possibility of government data resources that are neither proprietary to private companies, nor available to the public. For example, the Census Bureau produces county population estimates using migration data from the Internal Revenue Service, and used summaries of USPS change-of-address data to track population change in the aftermath of Hurricane Katrina. If access could be arranged (and that might be a big “if”), UA estimates might
benefit from otherwise unavailable data such as USPS change-of-address summaries, or counts of addresses maintained on the Census Bureau’s Master Address File.

**Estimates As a Time Series or Current Estimates Only**

When producing estimates on a regular basis, one needs to decide if the estimates constitute a time series, or independent estimates for sequential dates. For example, if one produces population estimates once per year from 2011 through 2014, one accumulates a set of estimates for 2011, 2012, 2013, and 2014. But those estimates are only a time series if they reflect change from year to year during that period.

If the method starts with 2010 census population, and estimates change based growth or decline in a data source such as address counts, the resulting estimates are not necessarily a time series, as change is always relative to the 2010 census count, and there might be revisions or corrections in the input data source. For example, when the 2012 estimates are produced, the address counts might include a downward correction relative to the address counts used for the 2011 estimates. The result is 2012 population estimates that are lower than the 2011 estimates – but the difference is not an estimate of population decline from 2011 to 2012. A downward revision to the 2011 estimates is implied, but not produced. To serve as a time series, the 2011 estimates would have to be revised to reflect the downward correction in the 2012 estimates. To maintain a time series, either the annual input data would have to always be correct the first time, or each time new estimates are produced, the estimates for all previous years in the series would have to be revised to reflect changes in input data, as well as changes in any control total estimates.

As an alternative, the estimates for a new year could always build from the estimates for the previous year – and estimate change since the previous year. This approach might produce better estimates of year to year population change (without revisions to previous years), but one would forfeit the ability to correct for known errors. For example, if the 2012 estimates for an area were found to be too high, the excess population would be carried forward in the estimates for 2013, 2014, and later years.

Time series estimates have the obvious benefit of tracking patterns of change over time, but they require a lot of additional work, and input data sources capable of supporting successive revisions to earlier estimates.

**Examples of Input Data Sources**

Several data sources have been cited as input sources for population estimates. They are examples only, and this section describes them in greater detail for illustrative purposes – not by way of recommendation.

**USPS Address Counts**

The US Postal Service provides “USPS Carrier Route Delivery Statistics” that several companies are using as input to nationwide small area estimates. Released monthly, the delivery statistics include counts of a variety of residential and business addresses, with definitions relevant mostly to mail delivery operations. For population estimation, the
focus is typically on active mailable residential addresses. The address counts are reported by local postal officials for carrier routes. A carrier route is defined by the list of addresses to which a carrier delivers mail, so the challenge is in converting the address counts from carrier routes to the areas being estimated—such as block groups, census tracts, or even UAs. Conversions from carrier routes to census geography are more precise in some areas than others, and typically involve the assignment of addresses to nine digit ZIP Codes (ZIP+4s), and translation from ZIP+4 to block and block group. So the compilation of USPS address counts can involve significant work that might depend on resources from third parties (such as ZIP+4-to-block correspondence), and conversions to non-postal geography are subject to error. But the effort yields address counts from a reliable source (I do not recall the USPS delivery statistics ever being late) that can be updated as frequently as once per month.

**Private Sector Consumer Databases**

For decades, companies in the direct mail marketing business have compiled lists of addresses that can be useful as input to small area population estimates. Some databases consist of name-and-address records compiled from sources such as telephone book listings, telephone customer lists, automobile registrations, and other sources. The databases are updated as the source data are updated, and the association of a name with an address (year after year) increases the odds that the housing unit is occupied. However, it is clear that some seasonal residents are included in these databases.

Other databases consist of “resident lists,” or lists of addresses without associated names. Such lists are updated with assistance from the US Postal Service. If a company has a list with most addresses, and submits it to USPS in carrier walk order, USPS will identify addresses to be deleted, and addresses to be added to make the file current—the apparent benefit to USPS being a reduction in undeliverable mail in large direct mail campaigns.

Databases of this type have the benefit of being nationwide in scope, updated frequently, and geocoded to small area census geography. Their limitations include imprecision in the coding of addresses to census geography, coverage issues (missing some households, and including some seasonal households), and changes in compilation procedures (or address sources) that can disrupt year-to-year trends in household counts. Such databases also are proprietary, and can be expensive to acquire.

**National Change of Address Data**

A wealth of information on population change is contained in the US Postal Service National Change of Address (NCOA) files, which are compilations of forwarding address information. For example, a person moving from Houston to Chicago might file a change-of-address card with the post office, requesting that mail sent to their Houston address be forwarded to their Chicago address. Compiled nationwide, with addresses coded to small areas, and specifying the month and year of the address change, NCOA records provide a detailed, but imperfect, picture of population movement throughout the US.
However, NCOA is strictly a tracking of requests for the forwarding of mail – not necessarily of moves. Seasonal and other temporary changes of address are included, and many address changes may relate only to some persons in a household. Also, access to NCOA data is strictly regulated by the Postal Service. Companies engaged in large direct mail campaigns are permitted access to forwarding addresses, as this practice reduces the volume of undeliverable mail that USPS has to deal with. But tabulations identifying migration streams are generally not permitted. Consequently, NCOA data have intriguing potential, but have not been used widely as regular inputs to small area population estimates.

**Population Density**

The second component of the project is the computation of population density for Urbanized Areas (UAs). Population density is generally expressed as a ratio of population per square unit of land area – such as square miles as indicated below.

\[
\text{Population density} = \frac{\text{Total population}}{\text{total area in square miles}}
\]

With total population estimated for UAs, the calculation of population density would seem a simple matter of establishing total land area for each UA, and completing the simple calculation.

In fact, the Census Bureau provides area measurements for all geographic entities in its geographic database – including UAs. These areas are calculated based on the boundaries maintained for UAs, and are expressed in whole square meters (with instructions to divide by 2.58999 to convert to square miles).

One option would be to simply use the UA land areas provided by the Census Bureau. This option might be the most convenient, but might require some care to ensure that the land area figures are based on UA boundaries consistent with those for which the population estimates were prepared.

A second option (if population estimates are produced from small area building blocks) is to determine land area for the blocks included in each UA, and aggregate to a total land area for the UA. If the UA population estimates were produced for census blocks (serving as building blocks to UA), this approach would ensure that the land area and population estimates are for the same area. However, if the building blocks consist of block groups, some block groups will be partial inclusions in the UA, and cross reference files are needed to determine the percent of the block group’s data to be allocated to the UA. The percent inclusion factors developed for this purpose often are based on block population, and can be inappropriate for use in allocating land area.

Consider a hypothetical block group with 20 blocks, one of which is in a UA. What percent of the block group’s population and land area should be allocated to the UA? The block population of 200 is 50 percent of the block group’s total population of 400. But the block’s land area is small – just one percent of the block group’s total area.
population-based allocation would correctly assign 50 percent of the block group’s population to the UA, but it would erroneously assign 50 percent of the block group’s land area to the UA. If building UAs from block groups, one would want separate cross-reference files – one with inclusion factors based on population (for use in allocating population), and another with inclusion factors based on land area (for use in allocating land area).

If aggregating land area from the census block level, one would want to consider the distinction between land area and water area. Again, the Census Bureau provides area measures for all entities for which it tabulates data – based on the boundaries recorded for each entity on the TIGER database. The measures include both land area and water area, so one would need to decide if it is important to either include or exclude water area from the square miles (or square meters) measure of area used to calculate population density.

Water area includes inland, coastal, Great Lakes, and territorial water. Inland waters include lakes, reservoirs, ponds, or other bodies of water that are recorded in the Census Bureau’s geographic database as a two dimensional feature rather than a line. Coastal waters include the portions of oceans and bodies such as the Gulf of Mexico that belong to the United States. In some cases, water areas are large, as block boundaries can extend a considerable distance from a shoreline. Census Bureau documents caution that features identified as “intermittent water” and “glacier” are reported as land area, and thus it is not always possible to derive the land area of a larger entity by aggregating the land area if its component blocks.

For purposes of computing population density, one might wonder why one would consider including water area. An answer might be that some water areas are populated by persons living on boats (crews of vessels or houseboats). Eliminating water areas implicitly adds these populations to adjacent land area, thus increasing their population density. The increase might be inconsequential when aggregating to UAs level, and it might even make sense, as such populations are likely proximate to land. The alternative is the distortion of adding sometimes large amounts of water area to the calculation of population density – effectively treating unpopulated water area as land. A relevant consideration is that, when calculating the population density of blocks for the delineation of urbanized areas, the Census Bureau does not include water area. Thus, the use of land area only is more consistent with the original delineation of UAs.

Finally, it is important to keep in mind that the proposed population density measures relate only to the land area of UAs as originally delineated by the Census Bureau. Post-census population estimates for blocks (if that is the approach taken) imply updated, UA boundaries and therefore updated population densities, but such updates are beyond the scope of this report.
Conclusions

The widespread use of urbanized areas in transportation planning, combined with the infrequent decennial updating of UA data, has promoted interest in post-census estimates for these areas. True updates would require the estimation of population and population density for census blocks, and a re-application of the complex process that the Census Bureau uses to delineate UAs. Such an effort is beyond the scope of this report. Instead, the report considers the options for estimating population and population density for UAs as delineated following the previous census.

Methodological options range from simple extrapolation, which requires only counts from recent censuses, to labor-intensive survey based methods, to component, housing unit, and model based estimates that require data reflecting post-census change (with such data often being scarce or imperfect). In addition to selecting a method and input data sources, decisions on broader options are required. Are UAs to be estimated directly, or as the aggregation of estimates for smaller “building block” geographic units? Will the estimates be built from census counts, or taken directly from an independent administrative source? Will top-down control totals be used, will the census residence rules be reflected, and if estimates are produced at regular intervals, should they reflect a time series? And if UA estimates are to be used for official purposes, is it appropriate for them to incorporate proprietary input data sources, or should they be based only on publicly available sources?

The field of population estimation is established enough that in developing a new program, one can choose from a range of methods and resources that have been used by others. But the field remains sufficiently new and wide open that the existing body of work does not define right or wrong approaches to the development of an estimates program.

Instead, the applications for which estimates are intended, and the expectations of those who will use them, are major factors in making one approach preferable to otherwise reasonable alternatives. So in the absence of a set formula for designing a population estimates program, and without knowing the specific needs or expectations related to UA estimates, this report has sought to describe a range of options and the types of decisions to be made in designing a UA estimates program. Rather than outlining a completed UA estimates program, or recommending the best path toward such a program, the report describes some paths traveled by others, and some of what one can expect to encounter on those paths.
Notes and References


2 Since 2007, the Census Bureau’s Housing Unit-Based Estimates Research Team (HUBERT) has been conducting research on the potential contributions of housing unit-based methods to the Census Bureau’s county population estimates.


12 TIGER (Topologically Integrated Geographic Encoding and Referencing) is the Census Bureau’s computerized mapping system, with digitized boundaries and features covering the entire United States.
If current trends in population density continue and all areas with high probabilities of urban expansion undergo change, then by 2030, urban land cover will increase by 1.2 million km², nearly tripling the global urban land area circa 2000. Less than 1% of all hotspot areas were urbanized circa 2000. By 2030, new urban expansion will take up an additional 1.8% of all hotspot areas (Table 2). Five biodiversity hotspots are forecasted to have the largest percentages of their areas to become urban: the Guinean forests of West Africa (7%), Japan (6%), the Caribbean Islands (4%), the Philippines (4%), and the Western Ghats and. Counties in large metropolitan areas (1 million population or more) saw the largest population gains. As a group, their populations increased 8 percent between 2010 and 2018, and nearly half of them grew faster than the national average. In contrast, noncore counties—those located outside metropolitan and micropolitan areas—have been the biggest demographic losers since 2010. Noncore counties as a group had a net loss of about 2 percent of their population between 2010 and 2018. Census, the Census Bureau redefines urban and rural areas based on criteria related to population thresholds, density, distance and land use. Population density. Quite the same Wikipedia. Just better. Biological population densities. Population density is population divided by total land area or water volume, as appropriate.[1]. Low densities may cause an extinction vortex and lead to further reduced fertility. This is called the Allee effect after the scientist who identified it. The potential to maintain the agricultural aspects of deserts is extremely limited as there is not enough precipitation to support a sustainable land. The population in these areas are extremely low. Therefore, cities in the Middle East, such as Dubai, have been increasing in population and infrastructure growth at a fast pace.[7]. Population density (in agriculture: standing stock or plant density) is a measurement of population per unit area, or exceptionally unit volume; it is a quantity of type number density. It is frequently applied to living organisms, most of the time to humans. It is a key geographical term. In simple terms, population density refers to the number of people living in an area per square kilometre. Population density is population divided by total land area or water volume, as appropriate. Estimates on urban populations vary mainly as a result of disagreements on the exact definition of an urban area and what this includes. Just under 1-in-3 people in urban areas globally live in a slum household. For most of human history, populations lived in very low-density rural settings. Urbanization is a trend unique to the past few centuries. By 2050 it’s projected that more than two-thirds of the world population will live in urban areas.