Population Estimates and the Needs of Local Governments

Joseph J. Salvo
Population Division
New York City Department of City Planning

Warren A. Brown
Program on Applied Demographics
Cornell Institute for Social and Economic Research
Cornell University

Population Estimates and the Needs of Local Governments

OVERVIEW

Local governments need population estimates in order to draw a portrait of population change, an integral part of assessing needs, and developing policies and programs that serve local residents. The U.S. Census Bureau prepares estimates of total population for all counties in the United States on an annual basis, using a demographic procedure known as the "administrative records component of population change" method. This method assumes that post-census population change can be closely approximated by administrative data on births and deaths, along with other data that are symptomatic of migration. The administrative data fall short of what is needed to calculate the components of change in many counties, given recent changes in the factors affecting population change, especially immigration. Moreover, with changes in the quality of administrative statistics and the precarious nature of population change in the post-2000 era, reliance on this single method leaves the Census Bureau vulnerable.

Therefore, the Census Bureau needs to consider creation of a composite estimate of total population, utilizing both a components-of-change method and a housing-unit based method. The Bureau currently develops housing unit estimates for a variety of applications but has not directly incorporated them into the county population estimates program. Before that can occur, the Census Bureau needs to improve its housing unit estimates and test a strategy for the creation of intercensal estimates of vacancy levels and household size. We point to several aspects of the current method that are lacking, make suggestions for improvement, and discuss strategies for estimates of vacancy levels and average household size.

MAKING THE CASE FOR ALTERNATE METHODS

New Trends That Affect Population Estimates

1. **The legal status of immigrants.** A body of research done over the past five years has consistently shown that the fastest growing segment of the foreign-born population consists of undocumented immigrants (Massey and Bartley, 2005; Passel and Suro, 2005; Passel, 2005). The numbers have risen to such a large degree, that the number of foreign-born who are undocumented rivals those who are here as legal permanent residents.
2. **The residential distribution of immigrants.** Immigrants are showing a pattern of increasing dispersion of settlement. There was a time when immigrants were concentrated in just a few places in the nation. Six states accounted for 57 percent of all immigrants in 2004, down substantially from 73 percent in 1990 (Passel and Suro, 1995). While part of this pattern includes seed migration as a result of targeted job recruitment efforts, this phenomenon has now grown way beyond the logging mills of the Northwest or the food processing plants of the South.

Immigrants are showing a propensity to settle in new areas, in many parts of the nation where they were not seen before in any substantial way (Singer, 2004).

3. **Changes in new housing construction.** When history looks back at this decade, the creation of large numbers of new housing units will emerge as an important factor that defines the lives of Americans. In New York City alone, the number of building permits for new residential construction has soared from just 4,500 in 1994 to more than 30,000 in 2005 (Figure 1).

Figure 1
Number of New Privately-Owned Residential Building Permits Issued
New York City, 1994–2005
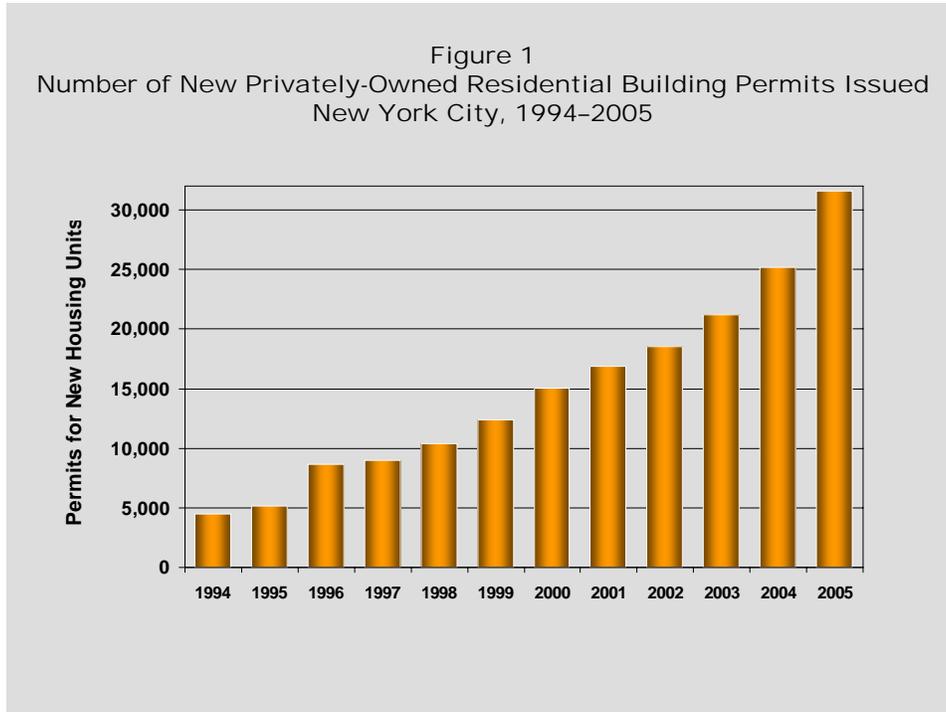
4. **The need for current information.** Pro-active strategies in reaction to natural and man made catastrophes have become the order of the day, in the wake of lessons learned from 9/11 and from Katrina. We live in a new era, one where we need a barometer capable of reacting to sudden changes in circumstances across the nation, making the once-in-a-decade census no longer plausible as the sole source of data for local areas. This underscores the importance of the American Community Survey (ACS) and those items that are the foundation of the program: the Master Address File (MAF), which is used to select the sample and the population and housing unit estimates which are used as control totals for weighting and post-stratification purposes. Therefore, it is not only important for the ACS to be implemented each year, but it is equally important that the MAF be updated on a regular basis and that the population estimates be capable of gauging change in an accurate fashion.

Limits of Current Methods: Components of Population Change

The Census Bureau constructs intercensal population estimates for all counties in the United States by separately estimating the components of change. Births and deaths are compiled using data from the national vital statistics system. Net migration is calculated by estimating the rate of net migration for persons coming in from and leaving for other counties in the 50 states (*net internal migration*) and the balance of people who immigrate from and emigrate to other nations and Puerto Rico (*net international migration*). The net internal migration rate is derived using income tax returns from the Internal Revenue Service and Medicare enrollment data from the Social Security Administration.

The limitations of the Census Bureau's components of change method for counties are well known. ***Net internal migration*** represents the net exchange between a county and other counties in the 50 states. Rates are calculated by comparing the addresses of income tax filers from year to year to determine residence at two points in time. For the July 1, 2005 estimates, the addresses of tax filers were compared for 2004 and 2005. They are subdivided into persons who were deemed in-migrants to a county (address in county in 2005 but outside the county in 2004), those who moved out of a county (address in county in 2004, but outside the county in 2005), and those who filed tax returns with an address in the same county at both points in time (non-migrants). The number of taxpayers moving out of an area is then subtracted from those who move in to determine net internal migration and a net migration rate is then derived based on tax filers and their dependents. This rate is then applied to the general population—assuming there is no difference in the rate of migration between tax filers and their dependents and the population at large. The calculations are limited to tax filers and their exemptions that are under age 65. For persons 65 years and over, addresses from Medicare enrollment data are used.

Net International Migration is the result of net flows to and from foreign countries and Puerto Rico. Data from the 2000 Census are used to allocate each county's share of the national non-citizen, foreign-born population that arrived in the U.S. between 1995 and 2000 for persons under age 65. Net movement from Puerto Rico is also allocated based on the county distribution of Puerto Ricans who entered the 50 states between 1995 and 2000. In addition, the Census Bureau creates a national estimate of emigration for the native-born. Each county receives a share of total emigration, based on its share of the national native-born population from the 2000 Census.

While the data on births and deaths are generally considered to be reliable, the data on migration can be very problematic because the method assumes that tax filers represent the migration experience of the total population. The Census Bureau uses Medicare enrollment data for persons 65 years and over to create migration rates because many retired persons do not file tax returns. Yet, there are other groups that have a low propensity to file returns where no procedure is available to compensate for the shortfall: persons who are marginal to the formal economy, those who fear government, groups

with serious language problems, and those who are otherwise alienated from the mainstream. For example, recent estimates of unauthorized migrants for New York State put the number in the range of 650,000 persons (Passel, 2005). Most of these immigrants probably lived in New York City, with few having a reason or sufficient incentive to file income tax returns. In addition, in places with large numbers of students, such as Manhattan, some enter as students, having never filed an income tax return. After living and working in Manhattan, many do then file returns and migrate to other parts of the nation. These individuals are detected as they exit Manhattan, but were not factored into the rate when they first arrived.

Since the Census Bureau is using a relatively new method to calculate net international migration, it will be some time before the efficacy of this method can be demonstrated. Suffice it to say that use of the 2000 Census to determine international migration for the post-2000 period is based on an assumption that the immigration picture has remained relatively unchanged, which is not likely the case, as discussed above. In addition, emigration remains enigmatic, given the absence of reliable data sources on persons who leave the U.S. We do know from past experience that immigration is malleable, with shifts in the country composition and in the patterns of settlement among immigrants commonplace. As it currently stands, the Bureau's methods are unable to incorporate any of the more dynamic aspects of international migration flows and are probably biased based on the current research, cited above.

Alternative Method for Estimating Population

Overall

An alternative method that is used frequently in jurisdictions where population growth is heavily driven by new housing construction is the housing unit method.   The housing unit method calculates the population in households as the product of housing units, occupancy rates and average household size. When persons in group quarters (i.e., prisons, nursing homes, dormitories and other facilities) are added to persons in households, an estimate can be created for the total population.

The Census Bureau does have a program that estimates the number of housing units for states and counties for public consumption, and for sub-county areas for internal use. County housing unit estimates are used as controls for the ACS and sub-county estimates are currently used to allocate the county population estimates, which are derived using the component method, to sub-county areas. What the previous discussion points to is a need for an estimates program that does not rely on a single method of population estimation for all areas of the nation, but one that is flexible enough to employ alternate methods in a formula that creates a composite picture. The strengths and weaknesses of various methods vary depending upon local circumstances, so no one formula can be expected to work everywhere. Currently, the Bureau compensates for this shortfall by offering localities the opportunity to employ alternate projection methods as "challenges" to the component method of producing population estimates. While many localities

genuinely appreciate the opportunity to take issue with the Census Bureau's estimates, the fact is that many localities do not have the resources to conduct a challenge.

The Census Bureau needs to explore the feasibility of preparing alternate sets of estimates that can be combined into a single composite number. The 2010 Census will offer the Census Bureau an opportunity to evaluate alternative sets of estimates -- component and housing unit-based -- so that they can determine what combination works best for different counties and places throughout the nation. Using standard measures, the evaluation will give the Bureau a tool that can be used to determine weights that can be applied in a formula for producing a single set of estimates. Areas with large increases in population as a result of new or rehabilitated housing would be treated differently than places where population change is the result of household size shifts that are a consequence of population aging, increases in immigration, or changes in patterns of domestic migration. This composite method provides some degree of customization, thereby making it more tailored to the characteristics of an area; all of this, without making the process too unwieldy from the standpoint of execution. In the interim, as the critical mass of data on household size and vacancy levels is acquired, the state members of the Federal State Cooperative for Population Estimates should be tapped, to compare estimates from each method against those produced by individual states and localities that prepare their own estimates.

Housing Estimates: The Specifics

The Census Bureau uses a simplified components-of-change approach to estimate the number of housing units in a sub-county area. There are three components: new housing units that are not mobile homes; new mobile homes; and loss of housing units. In order to estimate additions to the housing stock of new units that are not mobile homes, the Census Bureau uses building permit data from the internal files of the Manufacturing and Construction Division (MCD). A hefty dose of imputation is required for jurisdictions that do not report permits for the entire year and estimates are employed in places where new residential construction occurs without a permit being issued. The Census Bureau estimates that two percent of these permits never result in the construction of a housing unit and that a six-month lag time is required for a permit to result in construction of a unit ready for occupancy. In order to estimate additions of new mobile homes, the Bureau allocates statewide administrative data on mobile home shipments to sub-county areas based on their share of mobile homes as reported in Census 2000.

Estimating housing loss has been a serious stumbling block for years, since the Bureau does not collect data on permits for demolition of existing buildings. In order to estimate the loss of housing units—due to demolition, disaster, structural problems or the structure being moved—the Census Bureau uses data from the American Housing Survey's national sample to create rates based on age and type of housing unit. For mobile homes one rate is created and applied to stocks of mobile homes regardless of their age. For all other housing types—including single and multi-unit structures—rates of loss are derived based solely on vintage of the housing unit without regard for type of unit. These rates are then applied to the housing stock of a sub-county area based on characteristics

reported in Census 2000. The approach is "bottom-up," in that housing units for sub-county areas are summed to counties, which are in turn summed-up to states.

The construction of housing unit estimates is not new for the Census Bureau. The Bureau has, in fact, evaluated their housing unit estimates based on building permits and found them to be superior to those produced as a by-product of the component method (Devine and Coleman, 2003). There are, however, at least six issues regarding housing data inputs that need to be addressed before housing unit estimates can be incorporated into a set of population estimates:

1. **The Census Bureau needs to improve its model regarding the relationship between the age of housing and housing loss**. Assumptions about the state of aging housing units, based on national experience, may be generally inappropriate for many older cities. Results from the national sample of the American Housing Survey for the nation have led the Census Bureau to assume that the rate of loss for housing units rises in an S-curve fashion (see Figure 2) with age, reaching an annual loss rate of 0.364 percent for units built prior to 1940. At 36 percent of all units in 2000, New York City's pre-1940 buildings contained some 1.15 million units; a large and permanent part of the city's housing stock. About 43 percent of the housing stock in Manhattan and Brooklyn was built prior to 1940. Nationally, just 15 percent of all housing units in the nation were built prior to 1940. Many of New York City's pre-1940 buildings are among the most elegant and well-maintained in its housing stock. Moreover, local government has succeeded in rehabilitating many pre-1940 buildings that left the stock or laid vacant in the 1960s and 1970s. Research on housing depreciation and loss has shown that age alone is not a good predictor, and that tenure, type of structure and housing market conditions must be included (Follain and Malpezzi, 1980).

Figure 2
Loss Rates Used in the Census Bureau Estimates of Housing Units
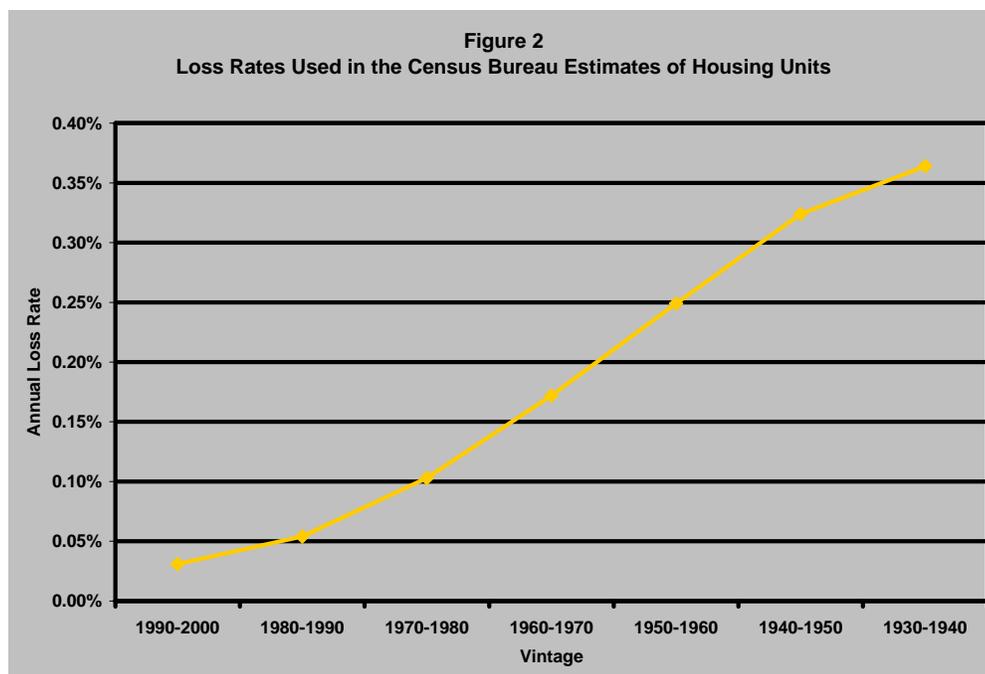
2. **The nature of housing unit change in many older metropolitan areas may cause the ratio of demolitions to new construction to be low because of the availability of land to build and because of increases in the density of new construction.** In some cities, a considerable amount of vacant land is available in neighborhoods where houses were razed in the 1970s and 1980s. Moreover, changes have been enacted in local zoning ordinances to permit formerly nonresidential land to be used for residential purposes; for example, land in formerly manufacturing and/or commercial districts has been rehabilitated for the purposes of residential occupancy. Finally, given the changes in immigration and resulting pressure on housing, it has become commonplace in cities of all sizes to demolish a small home and replace it with buildings that "push" the zoning envelope, frequently from a one-family house to a larger building containing several more units.

3. **The lag time between the issuance of a building permit and the creation of units ready for occupancy can vary from jurisdiction to jurisdiction.** The issuance of the certificate of occupancy requires a statement of compliance by the architect, engineer, or other person who supervised the construction work, stating that the building has been erected in compliance with the approved plans and complies with the provisions of administrative code and/or other applicable laws and regulations. Thus, the compilation of units with certificates of occupancy needs to be clearly defined, with a recognition that administrative issues can cause the time between issuance of a building permit and certificate of occupancy to vary from one jurisdiction to the next.

4. **Not all legal additions to the housing stock are covered by building permits for new residential construction.** Some examples of such additions include the following. Group quarters may be converted to individual housing units, such as the conversion of a fraternity house to a private home or apartment building. Existing homes may be converted from a one unit to a multi-unit structure. An uninhabitable structure may be rehabilitated and made ready for occupancy. A commercial or industrial structure may be converted to a multi-unit residential structure (Baer, 1990).

5. **Non-permitted additions to the housing stock will always be a "wild card" in the estimates process.** The degree to which an area has a sizable number of housing units created outside of the permitting process, will always affect the utility of the housing unit method. In some parts of the nation, especially those with a large influx of recent immigrants, illegal subdivision may occur frequently. These new addresses may not be recorded by local authorities nor can they be added to the Master Address File during intercensal years. While complete coverage of such units is unlikely, opportunities to capture such units can occur during the Local Update of Census Addresses (LUCA) program or during block canvassing, just prior to the 2010 Census enumeration. Further, regional census offices can work with local jurisdictions to identify areas with problems of this type. Clearly, for any estimation methodology, the presence of units or people

that are undocumented is a limitation that adds a healthy dose of modesty to our activities.

6. **Utilize information from the Master Address File (MAF) to inform the housing unit estimates.** The ACS provides the Census Bureau with constant input from the field on the status of housing units. A process needs to be designed to capture this experience and apply it to the estimates program on a continuous basis. The best method would be to ensure that MAF updates, which are triggered by ACS experience, actually occur.

Turning Housing Change into Population

A big problem for the Census Bureau has been the development of adequate estimates of average household size and occupancy rates, which allow for the translation of housing unit changes into population changes. With full implementation of the ACS sample of households for 2005 will come more statistically reliable data on households than has previously been available from sample surveys. With the one-year samples, the Census Bureau will produce estimates of household size and occupancy rates for geographic areas with populations of 65,000 and more. The accumulation of additional years of sample data will eventually push the estimates down to the smallest levels of local government and statistical areas including census tracts and block groups.

There are at least three approaches that utilize the ACS as a resource to move estimates of average household size forward. The first is to directly use the estimates of average household size as produced in the ACS for the geographic areas of various population thresholds. The values for counties and municipalities with populations greater than 65,000 can be implemented with release of the 2005 ACS. Estimates of average household size for areas with populations of 20,000 to 65,000 can be implemented with release of the 2007 ACS; and finally for areas with population less than 20,000 with release of the 2009 ACS. A decision rule may be that average household size from the ACS sample survey is not used unless an estimate is significantly different from the prior census full count. That is, if Census 2000 reported an average household size of 2.78 and the ACS reports an estimate of 2.82 with a confidence interval of +/- 0.08 then the factor remains at its level as reported in Census 2000.

A second approach to estimating average household size would be to utilize a regression model that uses a combination of administrative records and data from the ACS. Smith, Nogle and Cody (2002) have developed a regression approach to estimate average household size for counties. Their tests of this approach in four states have shown that the estimates were more precise and less biased than other traditional approaches. The administrative records they used were for births, school enrollment, and Medicare enrollment. Data from the ACS on fertility, marital status, and household composition might be combined with the administrative records to improve the quality of estimates.

9

A third approach would be to develop models of household change for small geographic areas when households tend to change in a predictable manner. Brown (2006) has proposed to model changes in household composition at the level of Block Groups. Brown's proposal is to classify Block Groups using matrices of household composition based on the age of householders and other household members. Transitions in these matrices would be used to track the interactions between aging, migration and household composition that affect average household size. Rather than waiting for the accumulation of five years of sample data—as in the first approach above—this approach would create large samples by grouping responses of households from thousands of Block Groups that have similar household composition.

The Census Bureau needs to continue its quest for new ways to improve data quality and estimates. The ACS makes possible a number of different approaches to estimating factors used in the estimation of population and housing. Programs such as the Center for Economic Studies' Census Research Data Centers that foster collaboration of the Census Bureau with research demographers and the partnership programs of the State Data Center and Federal State Cooperative for Population Estimates with applied demographers are excellent examples of the potential for continued and expanded collaboration and experimentation.

CONCLUSION

The time has come for the Census Bureau to utilize a multi-pronged approach to estimates of total population for counties and sub-county areas. Based on their own evaluation work, the Census Bureau is well aware that the complexities of population change frequently overwhelm the utility of a single approach. Their current answer – permitting jurisdictions to challenge the census results with other methods – does serve the cause of permitting alternative estimation techniques into the program. The major shortfall of this approach, however, is that it is uneven, favoring those jurisdictions with resources to calculate population using alternative methods. By adopting a composite method using two approaches, the Census Bureau provides for a more inclusive program, one that will provide all areas with an opportunity to have high quality population estimates.

BIBLIOGRAPHY

Baer, William C. "Aging of the Housing Stock and Components of Inventory Change." In Dowell Myers (Ed.), *Housing Demography: Linking Demographic Structure and Housing Markets.* (pp. 249-273). Madison, WI: The University of Wisconsin Press, 1990.

Brown, W. "An Alternative to the Census Bureau's Distributive Housing Unit Method for Estimating Population Size for Sub-County Areas: Updating Persons Per Household." Presented at the 2006 Population Association of America meetings, March 2006.

Devine, J and C. Coleman. "People Might Move but Housing Units Don't: An Evaluation of the State and County Housing Unit Estimates." Population Division Working Paper, No. 71, April 2003.

Follain, J. R., Jr., and S. Malpezzi. *Dissecting Housing Value and Rent*. Washington, DC: The Urban Institute, 1980.

Massey, D. and K. Bartley. "The Changing Legal Status Distribution of Immigrants." International Migration Review, Summer 2005.

Passel, J. "Unauthorized Immigrants: Numbers and Characteristics," Pew Hispanic Center, 2005.

Passel, J and R. Suro, "Rise, Peak and Decline: Trends in U.S. Immigration: 1992-2004," Pew Hispanic Center, 2005.

Singer, A. "The Rise of New Immigrant Gateways," Brookings Paper, 2004.

Smith, S., J. Nogle and S. Cody. "A Regression Approach to Estimating The Average Number of Persons Per Household." *Demography*, Vol. 35-4, (pp.697-712), November 2002.