

Making the Housing Unit Method Work: An Evaluation of 2010 Population Estimates in Florida

Stanley K. Smith · Scott Cody

Received: 6 June 2012 / Accepted: 30 December 2012 / Published online: 11 January 2013
© Springer Science+Business Media Dordrecht 2013

Abstract The housing unit (HU) method—in which population estimates are derived from estimates of occupied HUs—is the most commonly used method for making small-area population estimates in the United States. It is widely used because it is conceptually simple, can utilize a wide variety of data sources, can be applied at virtually any level of geography, and often produces reliable estimates. Yet the HU method is more nearly a general approach to population estimation than it is a specific methodology. In this paper, we describe and evaluate several data sources and estimation techniques that can be used in applying the HU method. Using a set of county and subcounty estimates produced by the Bureau of Economic and Business Research (BEER) at the University of Florida for April 1, 2010, we analyze errors by population size and growth rate, calculate errors for each of the three components of the HU method, and investigate the accuracy of estimates based on several different data sources and estimation techniques. We compare the accuracy of the 2010 estimates with previous BEER estimates and estimates produced by the U.S. Census Bureau. We conclude that although some data sources and estimation techniques work better than others, the HU method can be tailored to produce reliable population estimates for a wide variety of geographic areas.

Keywords Population estimates · Housing unit method · Estimation accuracy · Households · Household size · Group quarters population

Population estimates are used for a wide variety of planning, budgeting, and analytical purposes. Examples include choosing sites for schools, shopping centers, and housing developments; planning transportation routes; tracking growing or

S. K. Smith (✉) · S. Cody
Bureau of Economic and Business Research, University of Florida, 221 Matherly Hall, Gainesville,
FL 32611-7145, USA
e-mail: sksmith@ufl.edu

declining markets; analyzing demographic trends; setting geographic boundaries for political districts; and determining eligibility for government programs. In addition, they determine—in whole or in part—the distribution of hundreds of billions of dollars in government funding each year (see Blumberman and Vidal (2009) for a description of federal programs). Given their importance, it is not surprising that population estimates are of so much interest to small businesses, giant corporations, scholars, the media, and federal, state, and local government agencies.

The housing unit (HU) method—in which population estimates are derived from estimates of occupied HUs—is the most commonly used method for making small-area population estimates in the United States (Siegel 2002). It is widely used because it is conceptually simple, can utilize a wide variety of data sources, can be applied at any level of geography, and often produces reliable estimates. Yet the HU method is more nearly a general approach to producing population estimates than it is a specific methodology. Many different data sources and estimation techniques can be used in its application, each with its own strengths and weaknesses. Some are clearly better than others. What are the characteristics of these data sources and techniques? How accurate are the estimates they produce? How can the HU method be applied in ways that maximize its strengths and minimize its weaknesses?

In this study, we provide some answers to these questions. We start by describing the data sources and estimation techniques used by the Bureau of Economic and Business Research (BEBR) at the University of Florida to produce state, county, and subcounty population estimates in Florida. Then, we evaluate a set of estimates for April 1, 2010 by comparing them with 2010 census counts. We analyze errors by population size and growth rate, calculate errors for each of the three components of the HU method, and investigate the accuracy of estimates based on several different data sources and estimation techniques. We consider the role played by professional judgment and compare the accuracy of the 2010 estimates with previous BEBR estimates and estimates produced by the U.S. Census Bureau (hereinafter referred to simply as the Census Bureau). We conclude that although some data sources and estimation techniques work better than others, the HU method can be tailored to produce reliable population estimates in a wide variety of circumstances. We believe the results presented here will be useful to practitioners who use the HU method to produce small-area population estimates and to data users who rely on those estimates for decision-making purposes.

Methodology

The foundation of the HU method is the fact that nearly everyone lives in some type of housing structure, whether a traditional single family unit, an apartment, a mobile home, a college dormitory, or a state prison. The population of any geographic area can be calculated as the number of occupied HUs (i.e., households) times the average number of persons per household (PPH), plus the number of persons living in group quarters such as college dormitories, military barracks, nursing homes, and prisons:

$$P_t = (H_t \times PPH_t) + GQ_t$$

where P_t is the population at time t , H_t is the number of occupied HUs at time t , PPH_t is the average number of persons per household at time t , and GQ_t is the group quarters population at time t . Estimates of the group quarters population typically include persons without permanent living quarters (e.g., the homeless population).

This is an identity, not an estimate. If these three components were known exactly, the total population would also be known. The problem, of course, is that these components are almost never known exactly. Rather, they must be estimated from various data sources, using one or more of several possible techniques. In this section, we describe the data and techniques BEBR uses to estimate these components for counties and subcounty areas in Florida (subcounty areas are defined as incorporated cities and the unincorporated balance of each county). Other descriptions of the HU method can be found in Smith (1986), Siegel (2002), and Swanson and Tayman (2012).

Households

Census Bureau guidelines require a person to be counted as an inhabitant of his/her usual place of residence, which is generally construed to mean the place where he/she lives and sleeps most of the time. This place is not necessarily the same as one's legal or voting residence. A household is the person or group of people occupying a HU; by definition, the number of occupied HUs is the same as the number of households. Households refer solely to permanent residents; consequently, HUs are classified as vacant even when they are occupied by temporary residents staying for a few days, weeks, or months. BEBR follows Census Bureau guidelines in defining households and permanent residents.

BEBR uses three different data sources to estimate the number of households. The first is residential building permits, as collected and distributed by the U.S. Department of Commerce. The housing inventory for a city or county in 2010 can be estimated by adding permits issued since 2000 to the units counted in the 2000 census and subtracting units lost to destruction, demolition, or conversion to other uses. The time lag between the issuance of a permit and the completion of a unit is estimated to be 3 months for single-family units and 15 months for multifamily units; these time lags are based on surveys of developers in Florida. Estimates of units lost to destruction, demolition, and conversion are based on data from the Census Bureau (U.S. Census Bureau 2010a). Building permits are no longer issued for mobile homes, but proxies can be derived from state licensing and property appraisal records or data on shipments to mobile home dealers.

Using building permit data to construct a housing inventory requires complete data from every permitting agency. Although such data are not always available, complete (or nearly complete) data for 2000–2010 were available for all counties and for 84 % of subcounty areas in Florida.¹

¹ Some subcounty areas in Florida do not issue building permits; in these places, permits are issued by the county permitting agency. As a result, counties can have complete data even when some subcounty areas do not. For places in which data are available for most but not all years, estimates for the missing year(s) can be imputed from the available data.

Estimates of households can be derived from the housing inventory by applying occupancy rates. Until recently, there were no readily available data sources providing comprehensive post-censal information on occupancy rates. Accurate information could be obtained only through special censuses or sample surveys, which were generally too expensive to be feasible. This problem was typically dealt with by using the occupancy rates reported in the most recent census. This is the procedure BEBR uses, occasionally making adjustments to account for factors such as changes in the seasonal population.²

Data from the American Community Survey (ACS) provide an alternative to holding previous census values constant. Begun on a trial basis in the mid-1990s, the ACS was fully implemented in 2005 and replaced the long form of the decennial census in 2010. This survey collects demographic, socioeconomic, and housing data from approximately 3 million households each year and can be used to develop estimates of occupancy rates for states, counties, and subcounty areas. However, the ACS uses different residence rules and weighting procedures than those used in the decennial census and the proportion of households it reaches is substantially lower than that reached by the long form of the decennial census, even when ACS data are aggregated over a 5-year period. Also, ACS data are collected through a series of monthly surveys conducted over 1-, 3-, and 5-year time periods, rather than at a single point in time. All these differences raise questions regarding the usefulness of ACS data for the construction of population estimates. Although ACS data have been used in an exploratory way for estimating occupancy rates (Yowell and Devine 2012), further research is needed before we can draw any firm conclusions regarding the accuracy of those estimates.

The product of the inventory figure and the occupancy rate (performed separately for single family, multi-family, and mobile home units) provides an estimate of the number of households. There are several potential problems with this estimate. Time lags between the issuance of permits and the completion of units vary from place to place and from year to year. The proportion of permits resulting in completed units is usually unknown. Data on demolitions, conversions, and mobile homes must be estimated indirectly. Reliable estimates of changes in occupancy rates are generally unavailable. Certificate-of-occupancy data can eliminate problems related to completion rates and time lags but not those related to occupancy rates, demolitions, and conversions. These problems affect the accuracy of household estimates based on building permit data.³

The second data source is active residential electric customers. These data are collected annually from each of Florida's 53 electric utility companies and can be used in several different ways to construct household estimates (Smith and Cody 1994, 2004; Starsinic and Zitter 1968). BEBR uses a technique in which households are estimated by constructing a ratio of households to active residential electric

² Seasonality adjustments are somewhat subjective, but can be based on extrapolations of previous trends in the proportion of total housing units designated as seasonal, changes in the composition of units added since the most recent census (e.g., single-family, multi-family, and mobile home units), information provided by local planners, and consumption data provided by electric power companies.

³ Further discussion of these problems—and how they are handled by the Census Bureau—may be found in U.S. Census Bureau (2010a) and Yowell and Devine (2012).

customers using data from the most recent census year (e.g., 2000) and multiplying that ratio times the number of active residential customers in some later year (e.g., 2010).⁴ In some places, the household/customer ratio is adjusted to account for expected changes in the proportion of HUs occupied by permanent residents.⁵ The accuracy of estimates based on electric customer data is affected by the quality of the data and by changes over time in the proportion of customers who are permanent residents. Previous research has shown that household estimates based on electric customer data are—on average—more accurate than those based on building permit data (Rives and Serow 1984; Smith and Cody 1994, 2004; Starsinic and Zitter 1968).

The third data source is the number of homestead exemptions reported by the Florida Department of Revenue.⁶ Households can be estimated by constructing a ratio of households to exemptions using data from the most recent census year (e.g., 2000) and multiplying that ratio times the number of exemptions in some later year (e.g., 2010). An important advantage of these data is that they cover only HUs occupied by permanent residents, thereby excluding the impact of seasonal and other non-permanent residents. The primary disadvantage is that they do not include households occupied by renters or other non-homeowners. BEBR only recently began using homestead exemption data and currently uses them only for estimates at the county level. To our knowledge, this is the first study to evaluate the accuracy of household estimates based on homestead exemption data.

These data sources provide three household estimates for counties and two for subcounty areas. Based on the apparent quality of each data series (e.g., completeness and consistency over time), BEBR applies professional judgment in choosing which data source(s) to use for its final estimate of each county and subcounty area. In many instances, averages of estimates from two or even three data sources are used. Averages of several estimates draw on more information than a single estimate and reduce the risk of making large errors; the potential benefits of averaging are well-known, both for population estimates (Siegel 2002; Smith and Mandell 1984; Smith and Cody 2004) and population projections (Isserman 1977; Rayer and Smith 2010; Sanderson 1999).

Persons Per Household

The second component of the HU method is the average number of persons per household. Florida's PPH dropped from 3.22 in 1950 to 2.46 in 1990 but then

⁴ BEBR uses a ratio because there is seldom an exact one-to-one relationship between households and active residential electric customers. For example, places with master-metered units or illegal electrical connections may have more households than customers, whereas places with large numbers of second homes and other seasonal units typically have more customers than households. The record-keeping practices of electric power companies may affect the relationship as well. Using a ratio accounts for these differences.

⁵ These adjustments are based on the same factors noted in Footnote 2.

⁶ By law, Florida homeowners can apply for a property tax exemption on their primary residence. Most homeowners do so because the exemption often leads to a substantial reduction in property taxes. Homestead exemption data provide information on the number of homeowners who are permanent residents of a particular geographic area.

leveled off, remaining constant between 1990 and 2000 before rising to 2.48 in 2010. PPH values in 2010 ranged from 2.1 to 3.1 for counties and from less than 1.5 to greater than 4.0 for subcounty areas. PPH values have risen over time in some local areas and declined in others.

For each county and subcounty area, BEBR bases its PPH estimates on the PPH value reported in the most recent census, the state-level change in PPH since that census (as measured by the Current Population Survey), and the local change in the mix of single-family, multifamily, and mobile home units since that census. For counties, BEBR also uses a regression model in which changes in PPH are determined by changes in births, school enrollment, and Medicare enrollees (Smith et al. 2002). In some instances, indirect indicators of changes in PPH are used to adjust the estimates (e.g., changes in the racial composition of public school enrollees). Again, professional judgment based on previous trends, evaluations of data quality, and input from local planners is used to determine which approach to use for each county and subcounty area.

Group Quarters Population

The household population is calculated as the product of households and PPH. To obtain an estimate of total population, the population residing in group quarters facilities (GQ) must be estimated. In most places, BEBR estimates the GQ population by assuming that it accounts for the same proportion of total population in 2010 as it did in 2000. In places where the GQ population represents a substantial proportion of the total population, data are collected directly from the administrators of each GQ facility. Inmates in state and federal institutions are accounted for separately in all local areas; these data are available from the federal government, the Florida Department of Corrections, and the Florida Department of Children and Families. The total population estimate is made by adding the estimate of the GQ population to the estimate of the household population.

Evaluating Precision and Bias

BEBR constructed population estimates for April 1, 2010 for each incorporated city, each county, and the unincorporated balance of each county in Florida.⁷ Subcounty estimates were adjusted to be consistent with county estimates and the state estimate was calculated as the sum of the county estimates. We evaluate these estimates by comparing them with census counts for the same date. We define *percent error (PE)* as the percent difference between an estimate (E) and a census count (C) for an area at a given point in time:

$$PE = [(E - C)/C] * 100$$

It should be noted that percent errors as defined here may be caused by census enumeration errors—and changes in those errors from one decennial census to the

⁷ These estimates (and estimates for previous census years) are available from BEBR on request.

next—as well as by errors in the estimates themselves. We do not attempt to adjust for the impact of enumeration errors in this study because those errors are generally quite small and changed very little between 2000 and 2010, at least at the national level (U.S. Census Bureau 2012). Furthermore, estimates of enumeration errors themselves are subject to error, especially at the state and local levels. As noted by Swanson and Tayman (2012), most empirical studies do not adjust for census enumeration errors when evaluating the accuracy of population estimates.

We use five accuracy measures. Mean absolute percent error (MAPE) is the average error when the direction of the error is ignored. The proportion of errors less than 5 % and greater than 10 % indicate the frequency of relatively small and large errors, respectively. These are measures of precision, or how close the estimates were to census counts, regardless of whether they were too high or too low. Mean algebraic percent error (MALPE) is the average error when the direction of error is included. It is a measure of bias: a positive error indicates a tendency for estimates to be too high and a negative error indicates a tendency for estimates to be too low. Since a few extreme errors in one direction can strongly influence the MALPE, we use the proportion of estimates above the census count (%POS) as another measure of bias. These measures have been widely used for evaluating the accuracy of population estimates (Harper and Dixon 2011; Hoque 2008, 2010; Smith and Cody 1994, 2004; Yowell and Devine 2012).

Errors in State Estimates

BEBR's state population estimate for April 1, 2010 was 18,772,352, less than 0.2 % below the census count of 18,801,310. This error was remarkably small for a state that grew by almost 18 % during the decade; had large numbers of interstate migrants, seasonal residents, and foreign immigrants; was battered by several devastating and disruptive hurricanes; experienced a housing boom and bust; and was struck by a severe economic recession. To put this number into perspective, the MAPE for state estimates produced by the Census Bureau was 1.0 % in 2010 (Yowell and Devine 2012).

Table 1 shows errors for the state population estimates for each census year since 1980. Errors were below census counts in 3 years and above the count in 1 year. Although there was not a perfectly monotonic relationship, errors have generally become smaller over time.

Errors by Size and Growth

County Estimates

Table 2 summarizes the errors for the 2010 county population estimates. The MAPE for all counties was 2.7 %. Most counties had errors of less than 5 % and only one had an error greater than 10 %. The estimates displayed very little bias, as the MALPE was 0.5 % and errors were about evenly split between those that were too high and those that were too low.

Table 1 Errors in state population estimates for Florida, 1980–2010

Years	Estimate	Count	% Error
1980	9,486,455	9,746,324	−2.7
1990	13,150,027	12,938,071	1.6
2000	15,693,075	15,982,824	−1.8
2010	18,772,352	18,801,310	−0.2

Source: BEBR, unpublished data

Table 2 Population estimation errors by population size and growth rate: counties in Florida, 2010

Size (2000) and growth rate (2000–2010)	N	MAPE	MALPE	%POS	Percent of absolute errors	
					<5 %	>10 %
<25,000	17	3.6	−0.1	41.2	77.5	5.9
25,000–99,999	17	2.3	1.6	58.8	88.2	0.0
100,000–249,999	15	2.6	0.5	53.3	93.3	0.0
250,000+	18	2.2	0.0	44.4	94.4	0.0
<10 %	15	3.3	2.9	80.0	73.3	0.0
10–19 %	28	1.9	0.8	50.0	96.4	0.0
20–39 %	17	3.8	−2.1	23.5	82.4	5.9
40+ %	7	1.8	0.2	42.9	100.0	0.0
Total	67	2.7	0.5	49.3	88.1	1.5

Source: BEBR, unpublished data

MAPEs were substantially larger for small counties than large counties, but there was no clear relationship between errors and population size for counties with more than 25,000 residents. This is not unusual; errors are often found to decline as population size increases up to a certain point but to level off thereafter (Davis 2001; Harper and Dixon 2011; Hoque 2010; Yowell and Devine 2012). There was no clear relationship between population size and the tendency for estimates to be too high or too low. This too is a common empirical finding (Davis 1994; Hoque 2010; Smith and Cody 1994, 2004).

There was little indication of any relationship between errors and population growth rates. MAPEs were largest in the first and third growth-rate categories, smallest in the second and fourth. MALPEs followed a similar pattern, but the proportion of positive errors showed a slight upward bias for the most slowly growing counties and a slight downward bias for the most rapidly growing counties.

The lack of a consistent relationship between errors and growth rates is unusual. We believe this may have been caused by the relatively small sample size and by the lack of diversity in growth rates among Florida counties (e.g., only two counties lost population between 2000 and 2010 and most grew by less than 20 %). The interaction of size and growth-rate effects may have played a role as well. As we show in the next section, the typical relationship between errors and growth rates is clearly evident in the analysis of subcounty areas.

Subcounty Estimates

Table 3 shows errors for subcounty areas (i.e., incorporated cities and the unincorporated balance of each county). The MAPE for all subcounty areas was 9.2 %, more than three times larger than the MAPE for counties. This is not surprising, given that many subcounty areas have very small populations. Almost half of the errors were less than 5 %, but more than one-quarter were greater than 10 %. There was a small upward bias in the subcounty estimates, as indicated by a MALPE of 2.0 % and 55 % positive errors.

Differences in population size and growth rate had a much greater impact on estimation errors for subcounty areas than for counties. This occurred because the number of observations was substantially greater for subcounty areas and there was more variation in population size and growth-rate characteristics. The MAPE was 37 % for places with fewer than 250 residents and declined as population size increased, reaching approximately 3 % for places with 50,000 residents or more.

Table 3 Population estimation errors by population size and growth rate: subcounty areas in Florida, 2010

Size (2000) and growth rate (2000–2010)	N	MAPE	MALPE	%POS	Percent of absolute errors	
					<5 %	>10 %
<250	23	37.4	1.5	47.8	8.7	69.6
250–499	23	13.0	3.2	47.8	26.1	56.5
500–999	46	11.3	1.8	47.8	37.0	39.1
1,000–2,499	62	13.5	2.9	61.3	24.2	53.2
2,500–4,999	49	9.8	4.1	59.2	38.8	30.6
5,000–9,999	60	7.6	2.3	43.3	45.0	25.0
10,000–14,999	49	5.5	2.2	69.4	55.1	14.3
15,000–24,999	34	4.6	1.4	55.9	70.6	2.9
25,000–49,999	50	4.2	0.6	56.0	70.0	6.0
50,000–99,999	31	2.6	0.5	54.8	83.9	0.0
100,000–199,999	24	3.6	1.2	54.2	75.0	4.2
200,000+	17	3.3	–0.3	58.8	82.4	0.0
<–10 %	40	29.0	28.8	97.5	2.5	92.5
–10–0 %	97	8.5	8.0	90.7	41.2	27.8
0–4.9 %	53	5.9	1.1	56.6	64.2	11.3
5–9.9 %	48	5.3	1.2	52.1	64.6	8.3
10–14.9 %	49	4.1	–0.5	44.9	65.3	6.1
15–24.9 %	57	4.2	–2.2	33.3	73.7	10.5
25–49.9 %	81	7.3	–4.2	35.8	45.7	23.5
50–99.9 %	27	8.3	–5.2	18.5	40.7	29.6
100+ %	16	30.9	–30.0	6.2	12.5	75.0
Total	468	9.2	2.0	55.1	49.1	26.1

Source: BEBR, unpublished data

Above 50,000, there was no clear relationship between MAPEs and population size. There was a slight tendency for estimates to be too low in small places and too high in large places, but this relationship was fairly weak.

In contrast to the results for counties, differences in population growth rates had a strong, consistent impact on errors for subcounty areas. There was a distinct U-shaped relationship between MAPEs and growth rates: MAPEs were smallest in places with moderate positive growth rates but increased as growth rates deviated in either direction from those levels. MAPEs were less than 5 % for places growing between 10 % and 25 % during the decade, but were 29 % for places losing more than 10 % of their residents and 31 % for places whose populations at least doubled. Similar results have been found in many other studies (Harper and Dixon 2011; Hoque 2010; Smith and Cody 1994, 2004; Yowell and Devine 2012).

There was a strong tendency for estimates to be too high for places losing population and too low for rapidly growing places. This result has also been noted previously (Harper and Dixon 2011; Hoque 2010; Smith and Cody 1994, 2004). On average, places losing more than 10 % of their residents were *overestimated* by 29 %, whereas places whose populations more than doubled were *underestimated* by 30 %. Estimates were too high in 98 % of places losing more than 10 % of their residents and too low in 94 % of places that more than doubled. Places with low but positive growth rates displayed little bias, with small MALPEs and roughly half of the places overestimated and half underestimated.

Conclusions

Based on previous studies and the results presented here, it appears that the precision of population estimates generally increases with population size up to a certain point, but then levels off; that precision generally declines as growth rates deviate in either direction from moderate but positive levels; and that bias is largely unaffected by differences in population size but strongly affected by differences in population growth rates (places losing population are generally overestimated and rapidly growing places are generally underestimated).⁸ Although there may be occasional exceptions (e.g., the results for counties shown in Table 2), we believe these patterns occur frequently enough to be accepted as general characteristics of population estimates. The consistency of these results helps explain why errors for some places are generally larger (or smaller) than errors for other places.

Errors by Component

Which component of the HU method can be estimated most accurately? Table 4 shows that errors were smallest for PPH and largest for the group quarters population (GQ). For counties, MAPEs were 2.0 % for PPH, 2.6 % for households, and 18.4 % for GQ. There was a slight tendency for PPH estimates to be too high

⁸ We don't go into interactions between population size and growth rate in this study, but other studies have found that the relationship between size and errors does not change appreciably when differences in growth rates are accounted for, and the relationship between growth rates and errors does not change appreciably when differences in size are accounted for (e.g., Smith and Cody 1994, 2004).

Table 4 Estimation errors by component: Florida counties and subcounty areas, 2010

	Component	MAPE	MALPE	%POS	Percent of absolute errors	
					<5 %	>10 %
Counties	Households	2.6	-1.0	40.3	86.6	1.5
	PPH	2.0	1.1	68.7	97.0	0.0
	GQ	18.4	13.6	80.6	20.9	52.2
Subcounty areas	Households	7.8	1.6	50.2	57.9	20.3
	PPH	4.0	0.4	57.1	75.4	6.2
	GQ	110.1	86.4	42.3	42.0	52.4

Source: BEBR, unpublished data

and household estimates to be too low and a strong tendency for GQ estimates to be too high. For subcounty areas, MAPEs were 4.0, 7.8, and 110.1 % for the three components, respectively. There was a slight tendency for both PPH and household estimates to be too high. GQ had a large positive MALPE but more than half the errors were negative, implying that positive errors were often substantially larger than the negative errors.

A number of studies have found errors for households to be greater than errors for PPH (Lowe et al. 1984; Smith and Cody 1994, 2004; Starsinic and Zitter 1968). This most likely reflects the fact that growth rates are generally higher and more variable for households than for PPH. Whereas PPH changed by less than 5 % between 2000 and 2010 for most counties and subcounty areas in Florida, households often changed by 20, 30, 40 %, or more. There is simply more potential for error in estimates of households than in estimates of PPH.

Errors for GQ estimates have often been found to be particularly large (Smith and Cody 1994, 2004). There are several possible explanations for this finding. One is related to population size: small numerical errors can produce large percent errors when based on small populations, which is often the case for GQ. Other factors contributing to large GQ errors include uncertainties regarding definitions of the GQ population (e.g., a group home may be incorrectly classified as a household rather than as a group quarters facility), differences in counting procedures (e.g., average daily head count vs. the count on a specific day), and difficulties in maintaining up-to-date lists of GQ facilities. Voss and Marston (2012) provide a discussion of these issues, particularly as they relate to the collection of ACS data.

For both counties and subcounty areas, errors for the GQ population were much larger than errors for households and PPH. Does this mean that GQ errors contributed the most to overall estimation error? One way to answer this question is to construct synthetic population estimates using a combination of estimated values and census counts. We made estimates for counties and subcounty areas under three scenarios. The first used estimates of households and census counts for PPH and GQ; the second used estimates of PPH and census counts for households and GQ; and the third used estimates of GQ and census counts for households and PPH. For each scenario, then, errors in the population estimates were due solely to errors in the single estimated component. The results are shown in Table 5.

Table 5 2010 Population estimation errors under alternative scenarios

	Scenario	MAPE	MALPE	%POS	Percent of absolute errors	
					<5 %	>10 %
Counties	1	2.4	-0.9	40.3	91.0	1.5
	2	1.9	1.0	68.7	98.5	0.0
	3	0.7	0.4	80.6	100.0	0.0
Subcounty areas	1	7.6	1.6	50.2	59.4	19.2
	2	3.9	0.3	57.1	76.1	5.8
	3	1.4	0.3	72.9	95.9	1.9

Source: BEBR, unpublished data

Scenario 1: Estimates of households and census counts for PPH and GQ, *Scenario 2:* Estimates of PPH and census counts for households and GQ, *Scenario 3:* Estimates of GQ and census counts for households and PPH

It is clear that GQ errors did *not* contribute the most to overall estimation errors; in fact, they contributed the least. For both counties and subcounty areas, Scenario 1 had the largest MAPE, the most large errors, and the fewest small errors. Even with perfect estimates of PPH and GQ, errors in household estimates would have led to MAPEs of 2.4 % for counties and 7.6 % for subcounty areas. With perfect estimates of households and GQ, errors in PPH estimates would have created MAPEs of 1.9 % for counties and 3.9 % for subcounty areas. With perfect estimates of households and PPH, errors in GQ estimates would have created MAPEs of only 0.7 % for counties and 1.4 % for subcounty areas. Although errors were much larger for GQ estimates than for either household or PPH estimates, those errors contributed relatively little to overall estimation error because the GQ population generally accounts for a small proportion of total population. Similar results were found in previous evaluations of population estimates in Florida (Smith and Cody 1994, 2004).

Errors by Technique

Households

A variety of data sources and estimation techniques can be used to construct household estimates. Which ones provide the most accurate estimates? To answer this question, we evaluated the precision and bias of household estimates derived from several data sources and techniques:

BEER—The judgmental estimate based on all data sources and techniques, as described above.

EC—The household/electric customer ratio in 2000 is multiplied by the number of active residential electric customers in 2010.

EXTRAP—The linear change in the household/electric customer ratio between 1980 and 2000 is extrapolated to 2010 and multiplied by the number of active residential electric customers in 2010.

- BP—Household estimates based on building permits, as described above.
- HOMSTD—The household/homestead exemption ratio in 2000 is multiplied by the number of homestead exemptions in 2010.
- CONST—The number of households in 2000 is held constant.
- TREND—The linear trend in the number of households between 1990 and 2000 is extrapolated to 2010.
- AVE-2—Average of EC and BP estimates.
- AVE-3—Average of EC, BP, and HOMSTD estimates.

Results for counties are shown in the top panel of Table 6. The BEBR (judgmental) estimates were the most precise, with the smallest MAPE, the highest proportion of small errors, and the lowest proportion of large errors. The estimates had a small downward bias, with a MALPE of -1.0 and 40.3 % positive errors.

HOMSTD also performed well, with the second-smallest MAPE, the second-highest proportion of small errors, and the second-lowest proportion of large errors. It displayed very little bias, with a MALPE of 0.5 % and a fairly even split between positive and negative errors. Although BEBR only recently began using homestead exemption data, it appears that those data (and perhaps other types of property tax data as well) hold a great deal of promise for the production of household estimates.

Both electric customer techniques performed well. EC had a MAPE of 3.3 %, with 79.1 % small errors and 6.0 % large errors. EXTRAP had a MAPE of 4.6 % with 62.7 % small errors and 11.9 % large errors. However, EC had a small upward bias, with a MALPE of 1.5 and 65.7 % positive errors, whereas EXTRAP had

Table 6 Errors by technique for household estimates: counties and subcounty areas in Florida, 2010

	Technique	MAPE	MALPE	%POS	Percent of absolute errors	
					<5 %	>10 %
Counties	BEBR	2.6	-1.0	40.3	86.6	1.5
	EC	3.3	1.5	65.7	79.1	6.0
	EXTRAP	4.6	-0.9	50.7	62.7	11.9
	BP	5.5	4.7	86.6	59.7	19.4
	HOMSTD	3.1	0.5	52.2	79.1	3.0
	CONST	15.7	-15.5	1.5	9.0	74.6
	TREND	6.6	2.2	67.2	44.7	17.9
	AVE-2	4.2	3.1	74.6	68.7	9.0
	AVE-3	3.3	2.2	76.1	79.1	4.5
Subcounty areas	BEBR	7.8	1.6	50.2	57.9	20.3
	EC	8.4	3.1	60.9	54.0	25.6
	EXTRAP	11.0	-1.1	47.4	45.0	32.7
	BP	12.1	5.5	69.2	41.0	35.7
	CONST	15.1	-9.2	27.8	24.8	51.9
	TREND	15.1	1.9	57.5	27.8	50.0
	AVE-2	9.3	4.3	65.2	50.9	27.6

Source: BEBR, unpublished data

virtually no bias, with a MALPE of -0.9% and a nearly even split between positive and negative errors.

The BP technique was considerably less precise and more biased than the EC, EXTRAP, and HOMSTD techniques. It had a MAPE of 5.5% , with 59.7% small errors and 19.4% large errors. It also displayed a strong upward bias, with a MALPE of 4.7% and almost 87% positive errors.

The averages performed well, but not quite as well as several of the individual techniques. Their performance was affected by the inclusion of the BP estimates, which reduced the precision of both averages and caused them to have a strong upward bias. The average based on all three data sources (AVE-3) was more precise and less biased than the average based on electric customers and building permits (AVE-2). These results generally support the use of averages, but also highlight the importance of being careful about which individual techniques to include.

Whereas the first five individual techniques incorporated post-censal data as well as data from previous censuses, the last two did not. Holding household numbers constant at their 2000 levels (CONST) produced by far the worst estimates, with large absolute errors and a strong downward bias. Extrapolating 1990–2000 trends (TREND) performed better than holding numbers constant (especially for counties), but not as well as techniques incorporating post-censal data (although it displayed less upward bias than the BP technique).⁹

Results for subcounty areas are shown in the bottom panel of Table 6. Results for BEBR, EC, and CONST were based on all 468 places in the sample, but results for the other techniques were based on smaller numbers of places because some lacked complete data (there were 454 observations for TREND, 447 for EXTRAP, and 395 for BP).

Again, the BEBR estimates were the most precise, with the smallest MAPE, the highest proportion of small errors, and the lowest proportion of large errors. They exhibited very little bias, with a MALPE of 1.6% and a nearly even split between positive and negative errors. The EC technique performed almost as well in terms of precision, but had a stronger upward bias than BEBR. EXTRAP had a small downward bias and was less precise than either BEBR or EC. The BP estimates were less precise than those produced by the first three techniques and displayed a substantial upward bias. AVE-2 had a relatively small MAPE, but had a substantial upward bias. CONST and TREND performed very poorly, with larger MAPEs, lower proportions of small errors, and larger proportions of large errors than the other techniques; the former had a strong negative bias and latter had a slight positive bias.

As Table 6 shows, electric customer data produced more precise and less biased household estimates than did building permit data; as noted above, similar results have been reported in several other studies (Rives and Serow 1984; Smith and Cody 1994, 2004; Starsinic and Zitter 1968). Both data sources produced better estimates than could be obtained by holding previous census values constant or by

⁹ Although estimates incorporating post-censal data performed substantially better than estimates based on historical trends or previous census values in this study, there may be circumstances in which this would not be true (e.g., places with little population change). This is an issue deserving further study.

extrapolating historical trends. Averages generally performed well, but not quite as well as several of the individual techniques. Homestead exemption data performed very well at the county level. Finally, the application of professional judgment led to better estimates than any of the other techniques.

These results are quite powerful but many questions remain to be answered. What other data sources might be useful, such as certificates of occupancy, address lists, and postal delivery data? What business records or other types of proprietary data might be used? Can the ACS provide useful information on occupancy rates? Do some states or local areas have unique data series that might be used in those places even though similar data are not available elsewhere? Are there circumstances in which holding previous census values constant would lead to more accurate estimates than incorporating post-censal data? Further research will undoubtedly lead to additional improvements in household estimates.

Persons Per Household

We evaluated three different approaches to estimating PPH for both counties and subcounty areas: the judgmental BEBR estimates described previously; an estimate based on holding 2000 PPH values constant (CONST); and an estimate based on the linear extrapolation of 1990–2000 trends in PPH values (TREND). For counties, we also evaluated a regression model (REGR) in which changes in PPH were determined by changes in births, school enrollment, and Medicare enrollees (Smith et al. 2002). The results are shown in Table 7.

For counties, results for BEBR, REGR, and CONST were almost identical in terms of precision. MAPEs were about 2 % for all three techniques and there were very few errors greater than 5 %. BEBR and CONST had a small upward bias and REGR a small downward bias. Estimates for TREND were less precise and more biased than estimates for the other three techniques; the downward bias for TREND reflects the reversal of the downward PPH trend that occurred in many counties during the 1990s.

For subcounty areas, results for BEBR and CONST were almost identical. MAPEs and the proportions of large and small errors were exactly the same, but

Table 7 Errors by technique for persons per household: counties and subcounty areas in Florida, 2010

	Technique	MAPE	MALPE	%POS	Percent of absolute errors	
					<5 %	>10 %
Counties	BEBR	2.0	1.1	68.7	97.0	0.0
	REGR	2.0	-1.5	20.9	98.5	0.0
	CONST	1.9	0.4	58.2	97.0	1.5
	TREND	2.8	-2.3	16.4	86.6	0.0
Subcounty areas	BEBR	4.0	0.4	57.1	75.4	6.2
	CONST	4.0	0.0	53.4	75.4	6.2
	TREND	6.2	-1.6	30.4	64.5	14.5

Source: BEBR, unpublished data

BEBR had a slightly greater upward bias than CONST. Again, TREND estimates were substantially less precise and displayed a distinct downward bias.

More research is needed before we can draw firm conclusions regarding the performance of PPH estimation techniques. CONST performed well in the present analysis because PPH values changed very little between 2000 and 2010 for most places in Florida; if they had changed substantially, CONST would not have performed nearly as well. Similarly, TREND may perform reasonably well when PPH values are following a consistent path over time but not when that path changes. To illustrate these points: BEBR's judgmental estimates of PPH performed substantially better than CONST and TREND in an evaluation of 1990 estimates because PPH values changed substantially during the 1970s and 1980s (Smith and Cody 1994).

REGR performed well in our analysis of county PPH estimates. We believe regression techniques will be particularly useful when PPH values are changing rapidly because they are based on variables that pick up the impact of changes over time.

The ACS provides another potential source of information on changes in PPH. It was used in an exploratory way in an analysis of county-level errors in 2010 (Yowell and Devine 2012), but preliminary evidence suggests that PPH estimates based on ACS data exhibit too much year-to-year variability to be useful for the production of population estimates (Swanson and Hough 2012). Again, further research is needed before we can draw firm conclusions on this point.

Discussion

Comparison with Previous BEBR Estimates

As shown in Table 1, BEBR's estimate for the state of Florida was 2.7 % below the census count in 1980, 1.6 % above the count in 1990, 1.8 % below the count in 2000, and less than 0.2 % below the count in 2010. The change in the direction of error from negative in 1980 to positive in 1990 and back to negative in 2000 was most likely caused—at least in part—by changes in census coverage. Nationally, the net undercount declined between 1970 and 1980, rose between 1980 and 1990, and declined again between 1990 and 2000 (Brown et al. 2010). Because each set of estimates is based on previous census counts, a decline in the undercount can lead to estimates that are too low and an increase in the undercount can lead to estimates that are too high. Nationally, net census coverage changed very little between 2000 and 2010 (U.S. Census Bureau 2012). We believe this stability contributed to the high degree of accuracy of Florida's 2010 population estimate.

Table 8 compares errors for 2010 with errors for 1980, 1990, and 2000 for counties and subcounty areas. For both levels of geography, MAPEs and the proportion of large errors in 2010 were smaller than in any previous year and the proportion of small errors was larger. Improvements in precision were particularly notable for the county estimates.

MALPEs in 2010 were closer to zero than in any previous year and the proportion of positive errors was fairly close to a 50/50 split, especially for counties. Whereas the 1980 estimates had a downward bias and the 1990 estimates had an upward bias, the 2000 and 2010 estimates displayed relatively little bias. Viewed as a whole, these results suggest that the HU method as applied by BEBR had no systematic bias toward either overestimation or underestimation.

Why were the 2010 estimates more accurate than previous estimates? There are several possible explanations. Changes in PPH and household/customer ratios were smaller between 2000 and 2010 than in previous decades, making it easier to develop accurate estimates of PPH and households. Population sizes were generally larger and growth rates slower than in previous decades; both of these factors are associated with greater accuracy. The insights gained through an additional 10 years of studying estimation methods, sources of data, and the dynamics of population growth in Florida may have contributed to greater accuracy. Luck may have played a role as well. Whatever the causes, the 2010 estimates were the most accurate ever produced by BEBR.

Comparison with Other Estimates

How do the BEBR estimates stack up against those produced by other agencies? The only other agency making independent population estimates for all cities and counties in Florida is the Census Bureau. Some local governments make estimates for their own jurisdictions, but not for other places in the state. Several private data companies produce small-area estimates, but typically base them on estimates produced by the Census Bureau or by state demographic agencies.

The Census Bureau provides a good standard for comparison because it is the nation’s premier demographic agency. It has been producing state and local population estimates for many years and has pioneered in the development of several estimation techniques and data sources. In 2010, the Census Bureau produced county estimates using an administrative records (AR) method in which population estimates were based on births, deaths, Medicare enrollees, residents in group quarters facilities,

Table 8 Errors in county and subcounty estimates, 1980–2010

	Years	MAPE	MALPE	%POS	Percent of absolute errors	
					<5 %	>10 %
Counties	1980	5.4	−2.9	34.3	53.7	10.4
	1990	5.5	3.3	74.6	58.2	16.4
	2000	4.2	0.8	50.7	73.1	10.4
	2010	2.7	0.5	49.3	88.1	1.5
Subcounty areas	1980	14.4	3.5	46.7	33.6	42.4
	1990	11.9	6.0	68.4	36.5	40.5
	2000	10.4	2.3	51.2	46.6	32.3
	2010	9.2	2.0	55.1	49.1	26.1

Source: BEBR, unpublished data

foreign immigration, and estimates of internal migration derived from federal income tax returns (U.S. Census Bureau 2010b). County estimates were controlled to add to the Census Bureau's national population estimate and state estimates were calculated as the sum of each state's county estimates. Subcounty estimates were developed using a HU method based on building permit data and were controlled to add to the county estimates (U.S. Census Bureau 2010c).

Table 9 provides a summary of BEBR and Census Bureau estimation errors from 1980 to 2010. At the state level, BEBR estimates were more accurate in three of the four years (especially 1980 and 2000). At the county level, BEBR estimates were more precise and less biased than Census Bureau estimates in every year except 1990. Complete data for the Census Bureau's subcounty estimates are not available, but BEBR estimates had smaller MAPEs and MALPEs in all the years for which data were available.¹⁰

Why were the BEBR estimates more accurate than those produced by the Census Bureau? Again, there are several possible explanations. First, the Census Bureau applies a uniform methodology throughout the United States, whereas BEBR uses whatever data and techniques appear to produce the best estimate for any given place. This greater flexibility allows BEBR to draw on a greater variety of data sources and estimation techniques than the Census Bureau. Second, BEBR is knowledgeable regarding data idiosyncrasies and local population dynamics in Florida and uses this knowledge to improve the quality of its estimates. The Census Bureau lacks this degree of knowledge and eschews the application of location-specific professional judgment in the estimation process. Third, the Census Bureau's application of the HU method uses building permit data for estimates of households, whereas BEBR relies more heavily on electric customer and homestead exemption data. Both of these data sources have been found to provide more accurate household estimates than building permit data. Any (or all) of these factors may have played a role in the greater accuracy of the BEBR estimates.

Role of Professional Judgment

Professional judgment comes into play in several ways in the application of the HU method. One is in the collection of data. Collecting accurate electric customer, building permit, and other types of data requires a careful eye to note inconsistencies and an ability to correct errors before using the data to construct the estimates. This requires an intimate knowledge of the idiosyncrasies of each data source and how each source reflects population change. Even seemingly mechanical techniques require the application of professional judgment in the collection and revision of input data.

Professional judgment also comes into play in the choice of techniques and the formulation of assumptions. A variety of techniques can be used to estimate each component of the HU method and each requires that choices be made. Which

¹⁰ The Census Bureau's 2010 estimates for cities and counties in Florida are available on the Census Bureau website. Estimates for previous census years were provided in the early years of each decade but—to our knowledge—are no longer available.

Table 9 Population estimation errors for Florida, BEBR and U.S. Census Bureau (USCB), 1980–2010

STATE		Percent error		
Year	BEBR		USCB	
1980	-2.7		-5.6	
1990	1.6		0.3	
2000	-1.8		-4.4	
2010	-0.2		-0.9	

COUNTY	MAPE		MALPE	
	BEBR	USCB	BEBR	USCB
1980	5.4	5.7	-2.9	-5.1
1990	5.5	4.9	3.3	2.7
2000	4.2	5.5	0.8	-5.1
2010	2.7	3.2	0.5	-1.8

SUBCOUNTY	MAPE		MALPE	
	BEBR	USCB	BEBR	USCB
1980	14.4	15.7	3.5	-
1990	11.9	-	6.0	-
2000	10.4	16.1	2.3	4.2
2010	9.2	11.2	2.0	2.8

Source: BEBR and U.S. Census Bureau, unpublished data

technique or techniques should be used? If building permit data are used to estimate households for a particular place, what assumptions should be made regarding changes in occupancy rates? If electric customer data are used, should household/customer ratios be held constant or altered in some way? What other data sources might be available? What assumptions should be made regarding PPH values and the GQ population? Choosing techniques and developing assumptions requires not only an understanding of data sources and estimation techniques, but also of the population characteristics and trends for the places to be estimated.

In our evaluation of 2010 estimates, we found that applying professional judgment generally led to more accurate estimates of households than could be achieved through the mechanical application of any individual technique (or an average of those techniques). The differences were large for some techniques, small for others. Similar results were found in evaluations of the 1990 and 2000 estimates in Florida (Smith and Cody 1994, 2004). For PPH, the application of professional judgment did not lead to more accurate estimates than could be achieved by holding previous values constant in 2010, but led to substantially more accurate estimates than either holding previous values constant or extrapolating historical trends in both 1990 and 2000 (Ibid).

It is difficult to draw general conclusions regarding the application of professional judgment. Whose judgment should be used? How should it be

applied? Under what circumstances is it likely to lead to more accurate estimates? Can the role of judgment itself be formally incorporated into the estimation process? Many questions remain to be answered, but we believe the application of professional judgment—in the collection and revision of data, the formulation of assumptions, and the choice of techniques—will improve the quality of population estimates in most applications of the HU method.

Conclusions

Florida is a challenging state in which to produce population estimates. Many places are very small, are growing rapidly, have large numbers of seasonal residents, or are undergoing rapid changes in demographic composition. All these factors make it difficult to produce accurate estimates. Yet BEBR's estimates have performed well over the years, showing that the HU method can be made to work even under trying circumstances. Although errors for some places have been large, the overall performance has been quite good, especially in 2010.

Would similar results be found elsewhere? Given that the HU method is more nearly a general approach to population estimation than a specific methodology, its performance will depend on the choice of data, techniques, and assumptions used in its application. In contrast to the results reported here, several studies have found county-level estimates based on the HU method to be less accurate than estimates based on other methods (Hoque 2008, 2010; Yowell and Devine 2012). It must be noted, however, that those studies used building permit data in their application of the method; as shown here and elsewhere, estimates based on building permit data tend to be less accurate than estimates based on electric customer data. Smith (1986) and Smith and Mandell (1984) found that when electric customer data were used, the HU method produced county-level estimates that were *at least as accurate* as those produced by any other method (and were sometimes substantially more accurate).

In addition, several studies have found the HU method to produce population estimates at the subcounty level that were *more accurate* than those produced by other methods (Hoque 2008, 2010; Smith 1986). We are not aware of any studies finding the opposite to be true. Why does the HU method perform particularly well—compared to other methods—at the subcounty level? We believe the answer lies with data availability and reliability. Whereas the data required by the HU method are readily available for most subcounty areas (or can be obtained with sufficient effort), the data required by other methods are seldom available at the subcounty level or—when they are available—tend to be unreliable.

The HU method does not automatically produce accurate estimates, however. It is essential to choose appropriate techniques, make realistic assumptions, and take advantage of all available data sources (including those that may be unique to a particular area). It is also essential to pay close attention to data quality, resolving anomalies and cleaning up any errors that might be found. This is especially important for small areas, where data errors are likely to have their greatest impact.

Although the HU method can produce accurate estimates in a wide variety of circumstances, diligence and creativity will be required to obtain the best results.

The flexibility of the HU method is one of its outstanding attributes, not only in terms of the variety of data sources and estimation techniques it can utilize but also in terms of the variety of geographic areas to which it can be applied. Whereas the lack of relevant data severely limits (or completely precludes) the use of many estimation methods below the county level, the HU method can be applied at virtually any level of geography, from states down to counties, cities, census tracts, block groups, and even individual blocks. Perhaps the greatest strength of the HU method is its ability to produce reasonably accurate estimates for a wide variety of geographic areas.

Further improvements can be made, of course. Research opportunities include investigating additional uses of property tax data (Jarosz 2008), developing new regression models (Smith et al. 2002), evaluating the use of data from the Master Address File (Swanson and McKibben 2010), and exploring creative ways to incorporate data from the American Community Survey (Hogan 2008). This research promises to improve the accuracy of the HU method and create an ever-wider set of circumstances in which it can be applied. We believe research on the HU method (and other estimation methods as well) will have a large social and economic payoff, as better population estimates lead to better decision-making in both the public and private sectors.

References

- Blumerman, L. M. & Vidal, P. M. (2009). Uses of population and income statistics in federal funds distribution—With a focus on Census Bureau data. Governments Division Report Series, Research Report 2009-1.
- Brown, L. D., Cohen, M. L., Cork, D. L., & Citro, C. F. (Eds.). (2010). *Envisioning the 2020 census*. Washington, DC: National Academies Press.
- Davis, S. T. (1994). *Evaluation of the postcensal county population estimates for the 1980s*. Technical Working Paper No. 5. Washington, DC: U.S. Census Bureau.
- Davis, S. T. (2001). *Evaluating county population estimates: The big picture, hard to enumerate counties, and counties with special treatment*. Paper presented at the annual meeting of the Southern Demographic Association, Miami Beach.
- Harper, G. & Dixon, D. (2011). *Evaluation of 2010 subcounty population estimates*. Paper presented at the annual meeting of the Southern Demographic Association, Tallahassee.
- Hogan, H. (2008). Measuring population change using the American Community Survey. In S. H. Murdock & D. A. Swanson (Eds.), *Applied demography in the 21st century* (pp. 13–30). New York: Springer.
- Hoque, N. (2008). An evaluation of population estimates for counties and places in Texas for 2000. In S. H. Murdock & D. A. Swanson (Eds.), *Applied demography in the 21st century* (pp. 125–148). New York: Springer.
- Hoque, N. (2010). An evaluation of small area population estimates produced by component method II, ratio-correlation and housing unit methods for 1990. *The Open Demography Journal*, 3, 18–30.
- Isserman, A. (1977). The accuracy of population projections for subcounty areas. *Journal of the American Institute of Planners*, 43, 247–259.
- Jarosz, B. (2008). Using assessor parcel data to maintain housing unit counts for small area population estimates. In S. H. Murdock & D. A. Swanson (Eds.), *Applied demography in the 21st century* (pp. 89–101). New York: Springer.

- Lowe, T. J., Myers, W. R. & Weisser, L. M. (1984). *A special consideration in improving housing unit estimates: The interaction effect*. Paper presented at the annual meeting of the Population Association of America, Minneapolis, MN.
- Rayer, S., & Smith, S. K. (2010). Factors affecting the accuracy of subcounty population forecasts. *Journal of Planning Education and Research*, 30, 147–161.
- Rives, N. W., & Serow, W. J. (1984). *Introduction to applied demography*. Beverly Hills: Sage Publications.
- Sanderson, W. (1999). Knowledge can improve forecasts: A review of selected socioeconomic population projection models. In W. Lutz, J. Vaupel, & D. Ahlburg (Eds.), *Frontiers of Population Forecasting* (pp. 88–117). New York: The Population Council (A supplement to *Population and Development Review*, 24).
- Siegel, J. S. (2002). *Applied demography*. San Diego: Academic Press.
- Smith, S. K. (1986). A review and evaluation of the housing unit method of population estimation. *Journal of the American Statistical Association*, 81, 287–296.
- Smith, S. K., & Cody, S. (1994). Evaluating the housing unit method: A case study of 1990 population estimates in Florida. *Journal of the American Planning Association*, 60, 209–221.
- Smith, S. K., & Cody, S. (2004). An evaluation of population estimates in Florida: April 1, 2000. *Population Research and Policy Review*, 23, 1–24.
- Smith, S. K., & Mandell, M. (1984). A comparison of population estimation methods: Housing unit versus component II, ratio correlation, and administrative records. *Journal of the American Statistical Association*, 79, 282–289.
- Smith, S. K., Nogle, J., & Cody, S. (2002). A regression approach to estimating the average number of persons per household. *Demography*, 39, 697–712.
- Starsinic, D. E., & Zitter, M. (1968). Accuracy of the housing unit method in preparing population estimates for cities. *Demography*, 5, 475–484.
- Swanson, D. A., & Hough, G. C, Jr. (2012). An evaluation of person per household (PPH) estimates generated by the American Community Survey: A demographic perspective. *Population Research and Policy Review*, 31, 23–266.
- Swanson, D. A., & McKibben, J. N. (2010). New directions in the development of population estimates in the United States? *Population Research and Policy Review*, 29, 797–818.
- Swanson, D. A., & Tayman, J. (2012). *Subnational population estimates*. New York: Springer.
- U.S. Census Bureau. (2010a). Methodology for state and county total housing unit estimates (Vintage 2009): April 1, 2000 to July 1, 2009. (<http://www.census.gov/popest/topics/methodology/2009-hu-meth.pdf>).
- U.S. Census Bureau. (2010b). Methodology for the state and county total resident population estimates (Vintage 2009): April 1, 2000 to July 1, 2009. (<http://www.census.gov/popest/topics/methodology/2009-st-co-meth.pdf>).
- U.S. Census Bureau. (2010c). Methodology for the subcounty total resident population estimates (Vintage 2009): April 1, 2000 to July 1, 2009. (<http://www.census.gov/popest/topics/methodology/2009-su-meth.pdf>).
- U.S. Census Bureau. (2012). Census bureau releases estimates of undercount and overcount in the 2010 census. Press release 12-95. (http://www.census.gov/newsroom/releases/archives/2010_census/cb12-95.html).
- Voss, P. R., & Marston, K. (Eds.). (2012). *Small populations, large effects: Improving the measurement of the group quarters population in the American Community Survey*. Washington, DC: National Academies Press.
- Yowell, T. & Devine, J. (2012). *Evaluating current and alternative methods to produce 2010 county population estimates*. Paper presented at the U.S. Census Bureau Seminar on 2010 Estimates Evaluation, Suitland MD.