THE USE OF CADASTRAL AND CAMA DATA TO ESTIMATE UNINCORPORATED SUB-COUNTY POPULATIONS

Jim Sylvester
Bureau of Business and Economic Research
The University of Montana-Missoula

The number of housing units is the basis for any housing unit method for estimating current population. Building permits have served as the primary data source for estimating additions to the housing stock from new construction; however building permits are usually limited to a particular political entity such as a town or county. Some states, including Montana, do not even require building permits outside certain incorporated areas. Other states such as Florida require building permits statewide, but permit data at the address level are not readily available for most areas.

Montana has adopted a Geographical Information System (GIS) data base associated with appraisal data from the Montana Department of Revenue. This effort, called the Montana Cadastral Mapping Project, shows boundaries and ownership of land. The Montana project also combines the GIS component with its Computer Assisted Mass Appraisal (CAMA) data base. Florida and many other states have similar databases.

A cadastral map shows political and surveyed boundaries as well as ownership of land. Cadastral maps are often linked to CAMA files so that information about parcels can be located using GIS techniques. Cadastral maps are used to keep track of which properties are included in political and tax boundaries. Cadastral maps are updated when deeds are registered showing an ownership or boundary change. CAMA files are updated when improvements are made to a property.

Geographic Information is integral to cadastral maps so changes in political boundaries and the housing and populations associated with these changes can be immediately updated by re-estimating the housing counts for the new boundaries. US Census Bureau housing unit estimation procedures are used to update the populations. This paper demonstrates how population estimates can be continuously updated as sub-county political boundaries change.

Methodology

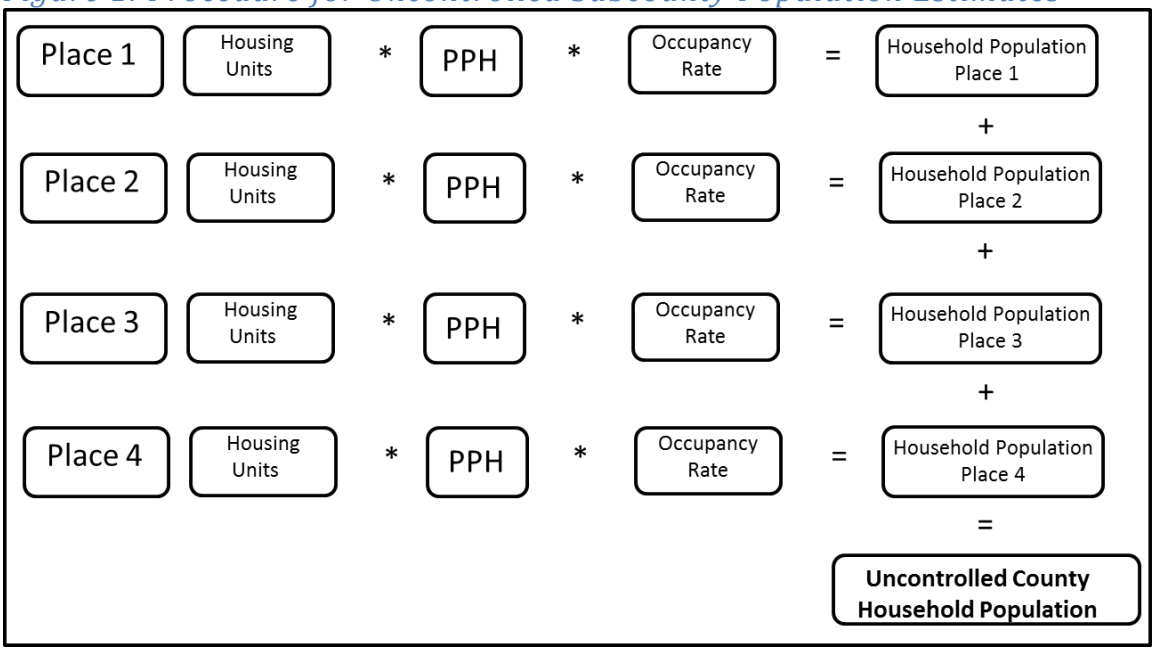
Sub-county estimates are produced by the US Census Bureau using a distributed housing unit method controlled to the official administrative county total. The following equation summarizes the basic housing unit method. A population is equal to the number of housing units multiplied by the persons per household times the percentage of housing units that are occupied plus the population in group quarters.

$$POP_t = HU_t * PPH_t * OCC\_RATE_t + GQ_t$$

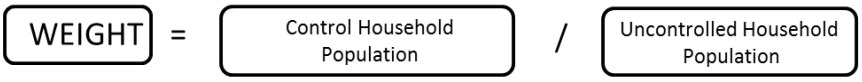
Estimating subcounty population estimate uses this basic equation for each subcounty area in a county where population estimates are desired. The subcounty estimates are summed to reach

a housing unit based county household population. This county estimate is independent of the administrative records county estimate used as a control. All areas of the county must be covered by an estimate so that the entire county population is estimated. Figure 1 illustrates the procedure for one year. The process is continued for each year in the time series.

Figure 1: Procedure for Uncontrolled Subcounty Population Estimates



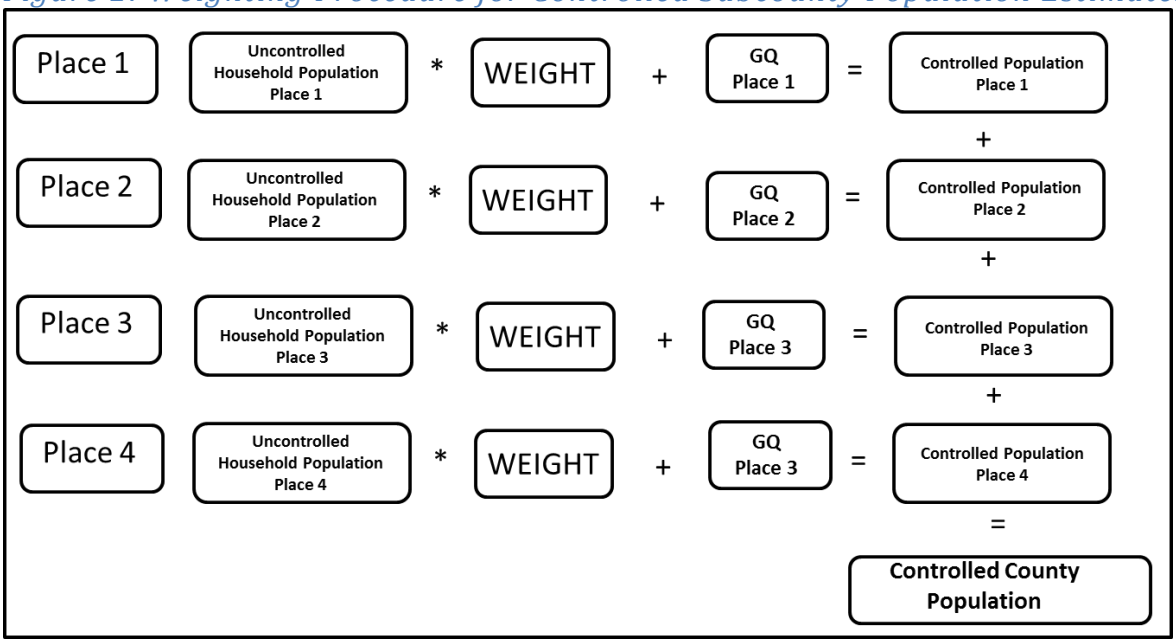
After housing unit estimates are obtained for each year a weighting variable for each year is computed using the following equation:



The weight for each year will be greater than one if the control is larger than the uncontrolled estimate for that year and less than one if the control is smaller.

Figure 2 shows the procedure for estimating each subcounty controlled household population. The total population for each area is derived by adding the group quarters population living in the subcounty area to the uncontrolled household population estimate multiplied by the weight and adding the group quarters population living in the subcounty area.

Figure 2: Weighting Procedure for Controlled Subcounty Population Estimates



Required data

Data needs for the estimation procedure described in this paper are minimal and readily available. Block level population and housing unit counts from the 2000 and 2010 Decennial Censuses are necessary. The 2000 numbers are used for starting the time series and the 2010 Census gives comparison data for checking accuracy. Cadastral data merged with CAMA data for the county with the subcounty area or areas that population estimates are desired. The year a structure is built and the number of units on a parcel are mandatory data items from the CAMA data. Building permit data geocoded to the block level can be substituted for the cadastral and CAMA data. Two GIS layers are also required: a layer with the blocks contained in the subcounty areas and a layer of the subcounty areas where estimates are desired.

Table 1: American FactFinder Tables

Dataset		Required data element	Data file	Variable name
2000 STF1	Block level data	Total population	P001	VD001
		Household population	QTH3	HC01_VC02
		Housing units	QTH1	HC01_VC02
		Occupied Housing units	QTH1	HC01_VC03
	County level data	All fields	DP1	
2010 STF1				
	Place level data	All fields	DP1	
	County level data	All fields	DP1	HD01_S001 HD01_S131 HD01_S169 HD01_S170

Study Areas

Seven areas were chosen from the list of places in Florida and Montana; three from Florida and four from Montana. Florida and Montana were chosen because all the necessary files were readily available with Internet downloads. Montana does not require building permits outside of a few selected incorporated places. Florida requires building permits statewide.

Flathead, Gallatin, Missoula and Lewis & Clark counties were chosen from Montana. Missoula county requires permits county-wide, but has only one incorporated place, Missoula city. About half of Missoula County's population lives outside Missoula city in scattered urbanized areas. These urbanized areas are classified as Census Designated Places (CDP) for census purposes. The CDPs are delineated by local officials. Two of the CDPs are rural areas with substantial seasonal housing.

Flathead County was chosen as a fast growing county with very few building permits. There is a substantial seasonal housing component in the housing stock.

Gallatin County is another fast growing county with scattered building permit requirements. It also has areas with seasonal housing.

Lewis and Clark County has scattered growth. Helena is the county seat and state capitals surrounded by five unincorporated suburbs classified as CDPs. One of two outlying areas has a large number of seasonal housing units. One other outlying area is suburb of Great Falls in neighboring Cascade County.

Two counties in Florida were selected randomly. Okaloosa County is a tourist area along the Gulf coast on the Florida panhandle. It is anchored by Destin and Fort Walton Beach. A large number of housing units are seasonal.

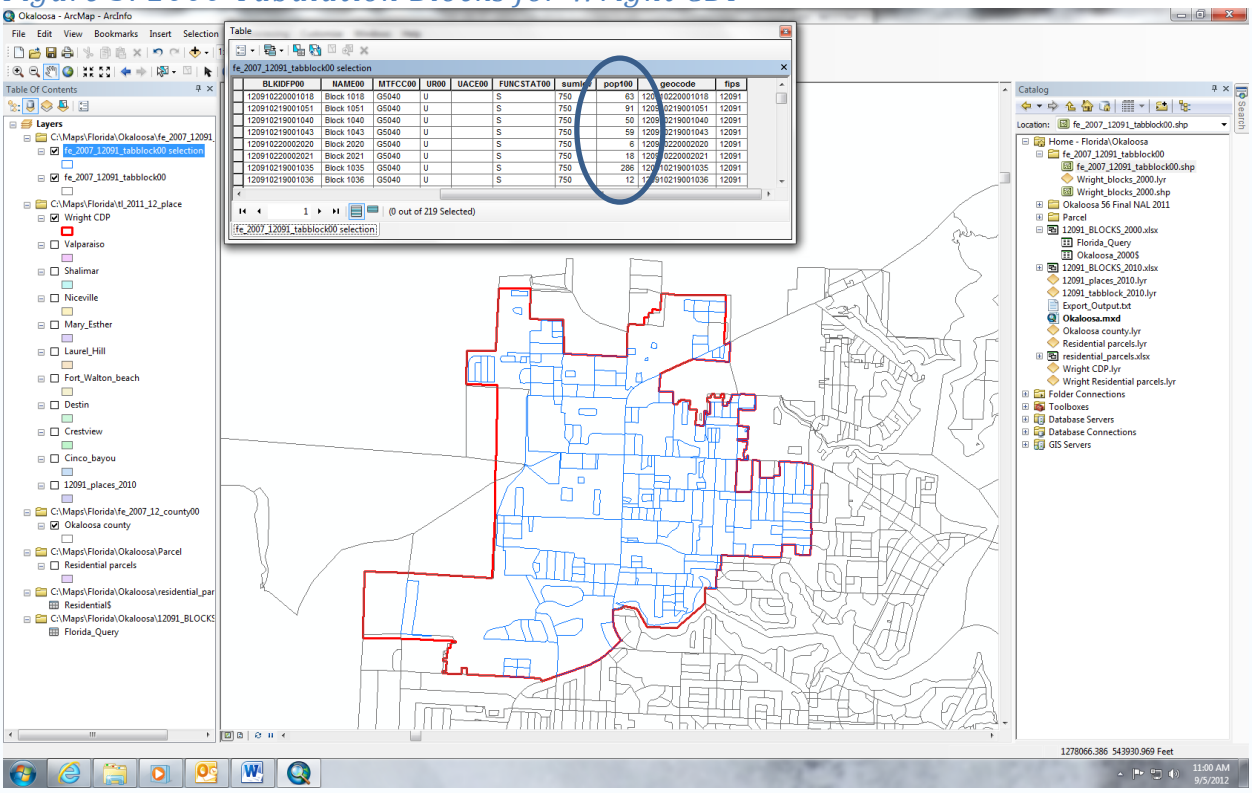
Osceola County, Florida is just outside Orlando. Kissimmee city is the county seat. One CDP is Celebration a community next to Disney World. Occupancy rates and household size are similar across the county.

The third Florida area, Marion-Sumter Counties, was chosen as the focus of a Wall Street Journal article about a growing CDP called the Villages that crosses county lines.

Procedure

The first step in making population estimates of unincorporated areas is deriving a starting population for each area. Figure 3 show a screen shot of the ARCGIS result for extracting the 2000 Census population for Wright CDP, Okaloosa County, Florida. The GIS layers used are the tabulation blocks for Okaloosa County joined with the block level population and housing unit counts. The 2000 level population for Wright can be calculated by summarizing POP100 from the attribute table. Similar selections and tabulations are done for other areas where estimates are desired. A similar selection and tabulation is done to adjust for geographic boundary changes where the 2000 and 2010 boundaries are different. The Census Bureau makes these changes annually for incorporated places through adjustments from the Boundary and Annexation Survey (BAS).

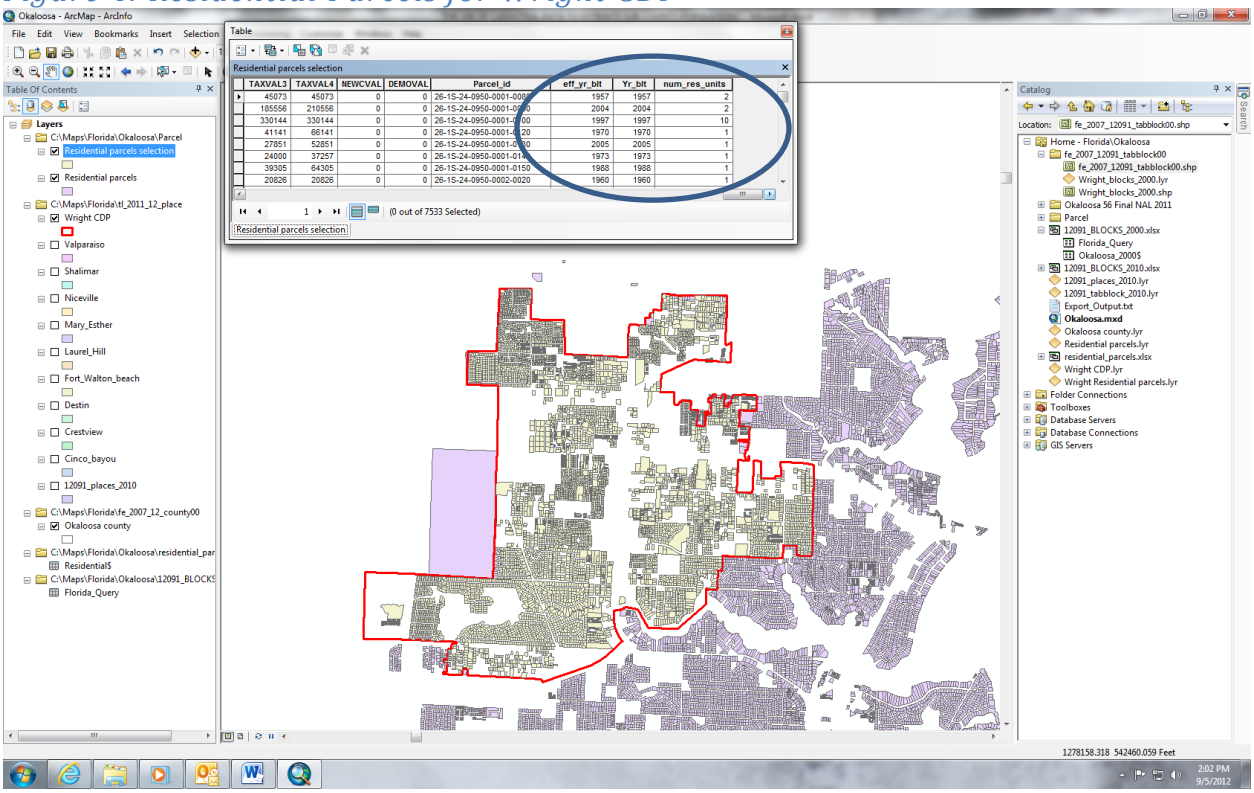
Figure 3: 2000 Tabulation Blocks for Wright CDP



The next step is estimating the number of new housing units added to the housing stock each year for each unincorporated areas. This is done by selecting the residential parcels for each unincorporated place in a process similar to the starting population (Figure 4). The relevant data is number of units by year built. The easiest way of tabulation by year is exporting the attribute table as an Excel file and using statistical software. Similar procedures are done for the remaining areas and the county total.

Using the numbers gathered from the GIS and 2000 Census, housing unit estimates are calculated for each subcounty area using the methods described previously. Weight the household populations and add back the group quarters population to arrive at total population for each subcounty area.

Figure 4: Residential Parcels for Wright CDP



Results

The subcounty estimates are evaluated by comparing the 2010 estimate based on 2000 Census data with the 2010 Census count for each area. The percent difference between the CAMA based subcounty population estimates and Census Bureau permit based subcounty population estimates are displayed in Figures 5 through Figure 11. A negative difference suggests that the 2010 Census is lower than the estimate; a positive difference the 2010 Census number is larger than the estimate. Some CAMA based estimates are closer and some building permit estimates are closer to the actual population as measured by the 2010 Census. The following discussion examines some potential reasons for the differences.

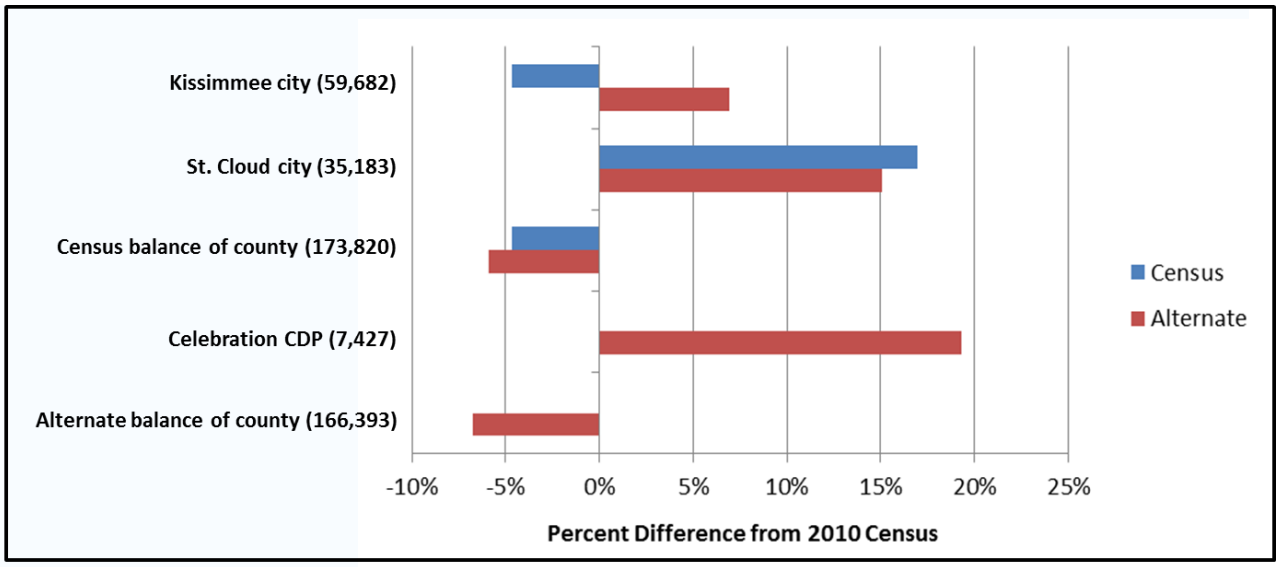
Each county used for demonstration purposes has at least one large subcounty area with a disproportionate error. There are five general possibilities for the large variation in errors:

- Small numbers
- Quality of the housing unit count
- Changing PPH and occupancy rates
- Boundary files and parcel data problems
- Increases or decreases in group quarter populations

Osceola County Florida (Figure 5) has two areas with large errors, St. Cloud City and Celebration CDP. The error associated with St Cloud City is about 17 percent for the Census Bureau estimate and 15 percent for alternative estimate. Both the Census Bureau estimate and the CAMA estimate of housing units are low when compared to the 2010 Census count of

housing units. The large difference for the Celebration CDP can be attributed to the CAMA estimate of 2010 Housing units being 1,200 below the 2010 Census count. The balance of county estimates, both Census and Alternate, are high because occupancy rates declined from 84 percent in 2000 to 65 percent in 2010. Census 2000 occupancy rates are used for estimate calculations.

Figure 5: Difference of Osceola County Florida Subcounty Estimates from 2010 Census



The Census Bureau estimate for Valparaiso City Florida in Okaloosa County Florida (Figure 6) was about 14 percent larger than the 2010 Census. The Census Bureau overestimated the number of housing units in Okaloosa County. The CAMA housing unit estimate was higher than the 2010 Census for Valparaiso City. The number of housing units for Wright CDP was underestimated by the CAMA methodology. Both PPH and occupancy rates remained relatively constant between 2000 and 2010 for both areas so that has little influence on the error.

Figure 7 shows the results from Marion-Sumter counties. Marion and Sumter counties were combined because the Villages CDP crosses county lines. The Census estimate for Dunnellon City was 11 percent above the 2010 Census count primarily due to a decline in occupancy rate and to many new housing units. The alternative method was much closer. The large error for McIntosh by both Census and the CAMA methods are mainly due to small numbers, the numerical difference was less than 100 persons. The Census estimate for Reddick town was nearly 30 percent higher primarily due to small number effect. The alternative estimate for Center Hill City was way off due to 268 units being built in 2009 according to the CAMA data. These units may have had a different occupancy rate than older units. The large error for Wildwood City was primarily due to a declining occupancy rate from 88 percent in 2000 to 65 percent in 2010. On the whole the unincorporated place estimates were not unreasonable, ranging from -6 percent for the Villages CDP to 9 percent for the alternative balance of county. Balance of county was affected by an increase in group quarters population between 2000 and 2010. Changing GQ data was not available for the alternative method.

Figure 6: Difference of Okaloosa County Florida Subcounty Estimates from 2010 Census

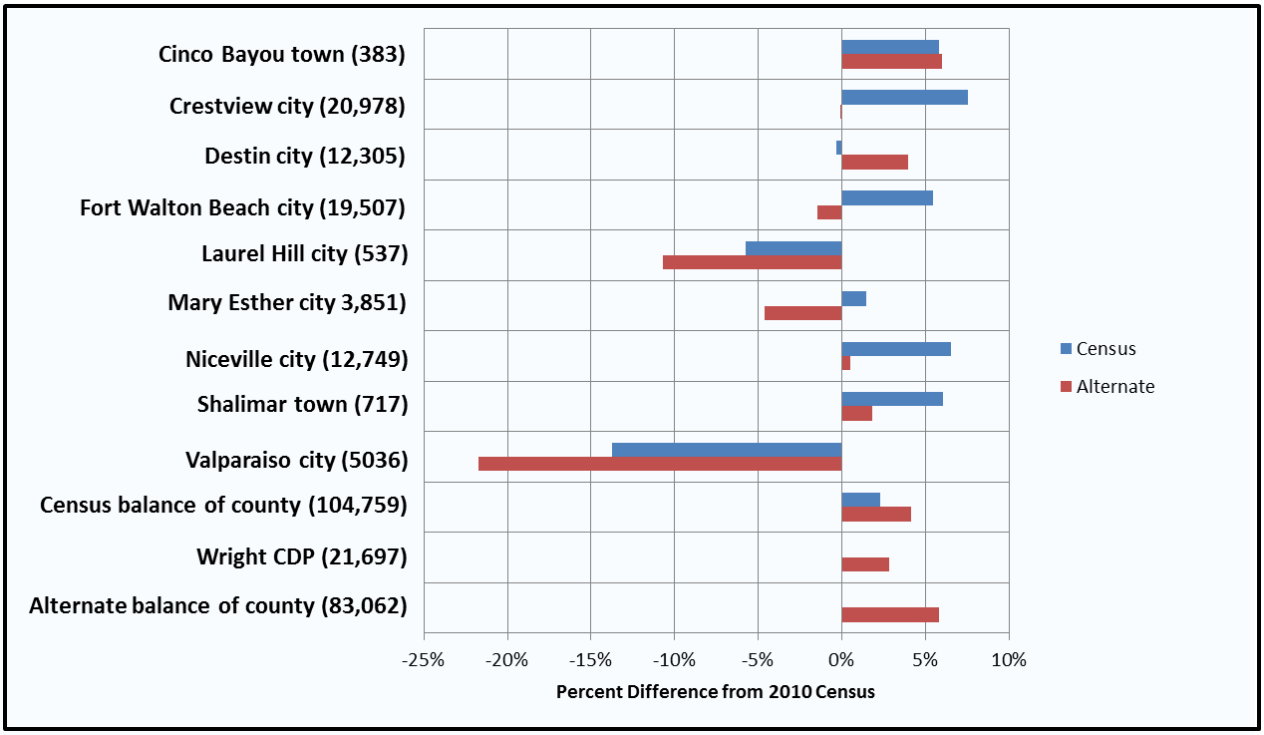
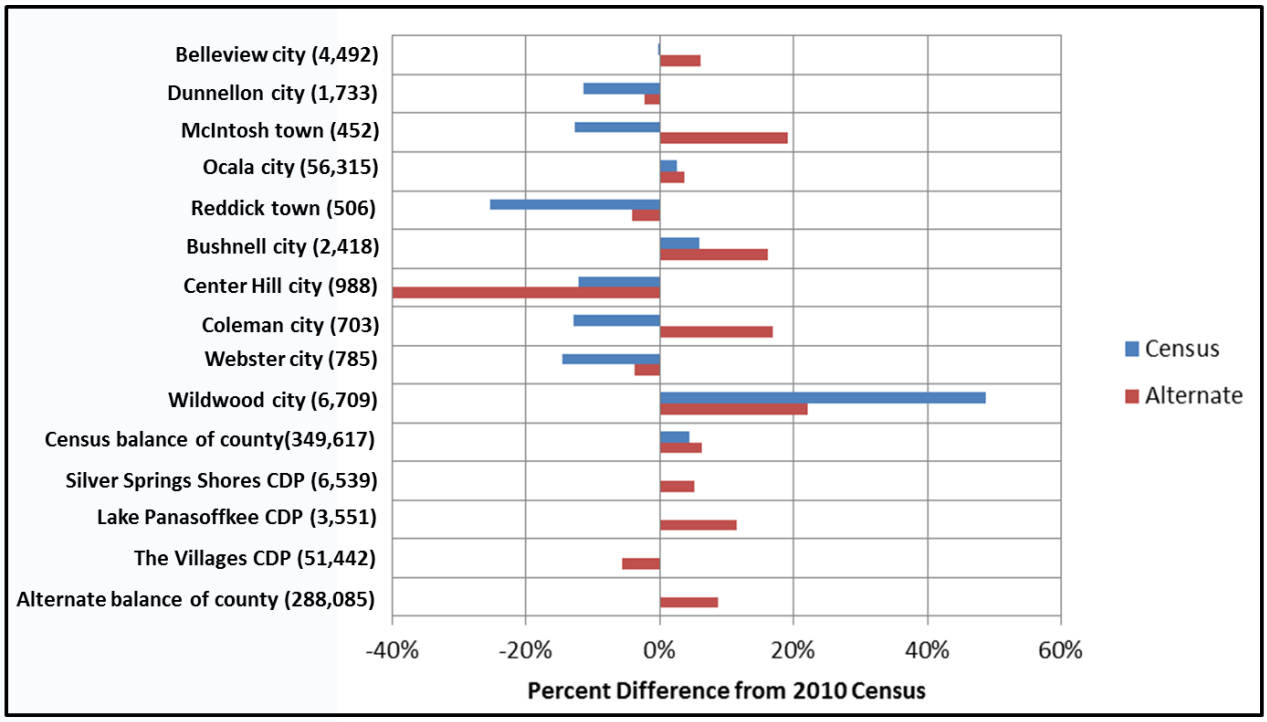
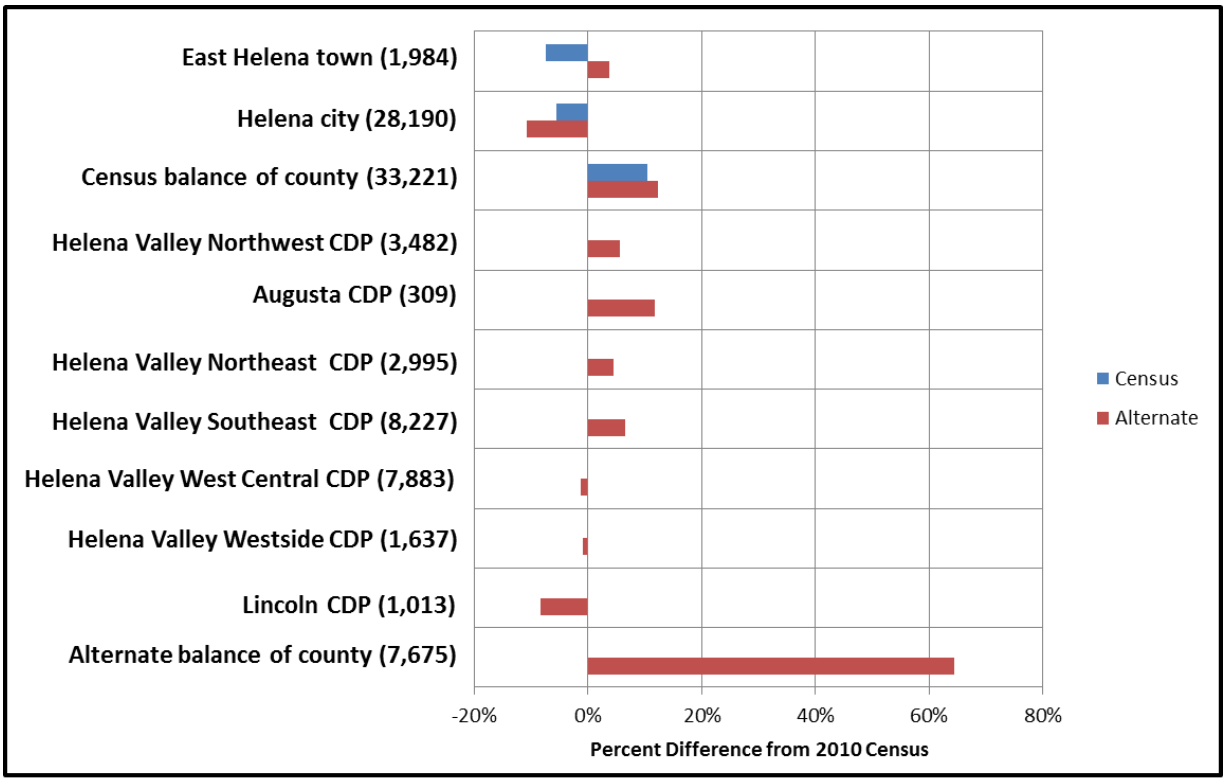


Figure 7: Difference of Marion-Sumter Counties Florida Subcounty Estimates from 2010 Census



Population estimates for subcounty areas in Lewis & Clark County Montana (Figure 8) were very close to their 2010 population count with the exception of the balance of county with a large percentage discrepancy (64 percent) but the numeric difference was a small percentage of the overall population (5 percent). The large percentage error was a function of small numbers. The population estimates for CDPs surrounding Helena City were very close.

Figure 8: Difference of Lewis & Clark County Montana Subcounty Estimates from 2010 Census



The CAMA based subcounty estimates worked extremely well for Missoula County, Montana (Figure 9) with one exception, Frenchtown CDP. The housing count number was low for Frenchtown CDP by about half. One explanation for this difference is how mobile homes are counted. There are several mobile home parks in the CDP. Mobile homes in parks are classified as personal property in Montana and are thus outside the CAMA system. Population estimates for other unincorporated areas were overestimated. The PPH for the three CDPs declined. The occupancy rate for Seeley Lake CDP also declined. The estimate for the alternate balance of county was only 3 percent; the main reason was a decline in occupancy rate as the number of housing units was remarkably close, a difference of only 268 units.

Figure 9: Difference of Missoula County Montana Subcounty Estimates from 2010 Census

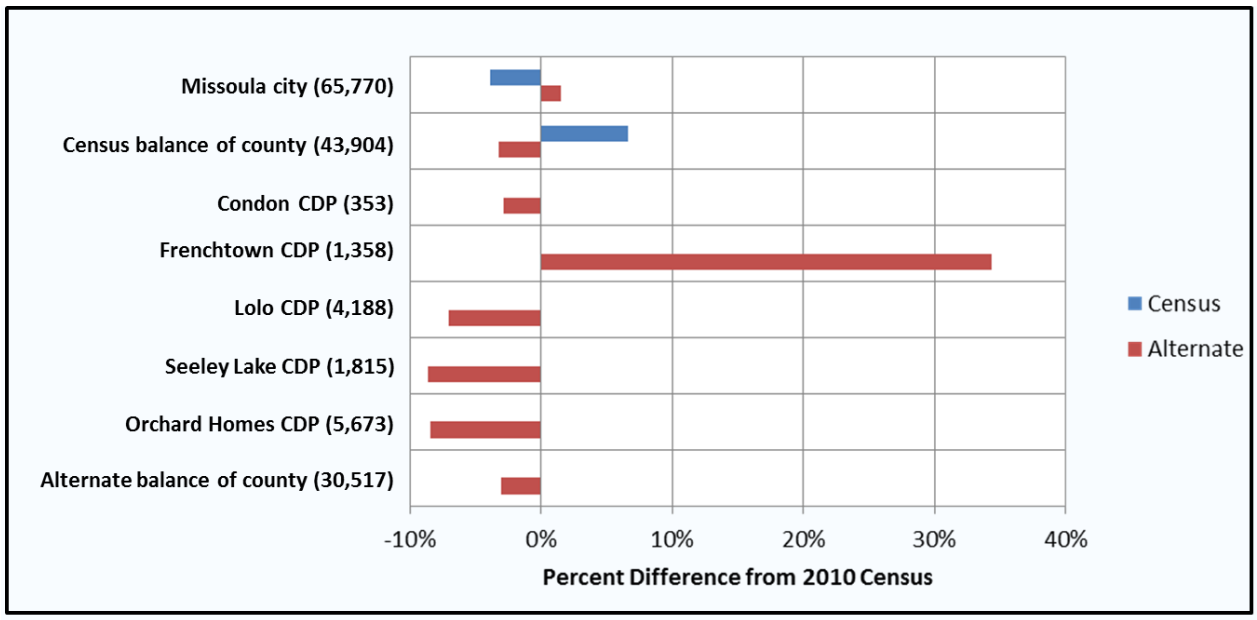


Figure 10, shows the GIS/CAMA method outperformed the Census Bureau estimates for all areas estimated by the Census Bureau in Flathead County. It performed well for estimates of unincorporated areas with the exception of Little Bitterroot Lake CDP where the error was more than 50 percent. The reason for the large error was twofold: a large downward change in occupancy rates and a small number effect. Other larger errors are primary due to small number effects.

The GIS/CAMA method performs better than the Census Bureau for all incorporated places in Gallatin County, Montana (Figure 11) estimated by the Census Bureau except Manhattan town. Both PPH and occupancy declined in Manhattan town; PPH from 2.53 to 2.40 and occupancy from 0.95 to 0.84. The CAMA estimate of housing units underestimated the in Big Sky CDP population where occupancy rates increased along with group quarters population. The GIS/CAMA method over- estimated Gallatin Gateway and King Arthur Park CDPs population. The reason was an overestimation of housing units in those areas.

Conclusion

Use of GIS and CAMA has potential for making small area population estimates, especially for areas with poor building permit coverage. The main advantage of the GIS approach is that political boundaries are not a hindrance. Boundaries can be manipulated for any area desired for planning or analysis purposes. The main disadvantage of the GIS/CAMA approach is the availability of quality data for many areas. If good building permit data is unavailable use of alternate estimates of housing units may be warranted.

References

http://www.census.gov/popest/methodology/index.html

Figure 10: Difference of Flathead County Montana Subcounty Estimates from 2010 Census

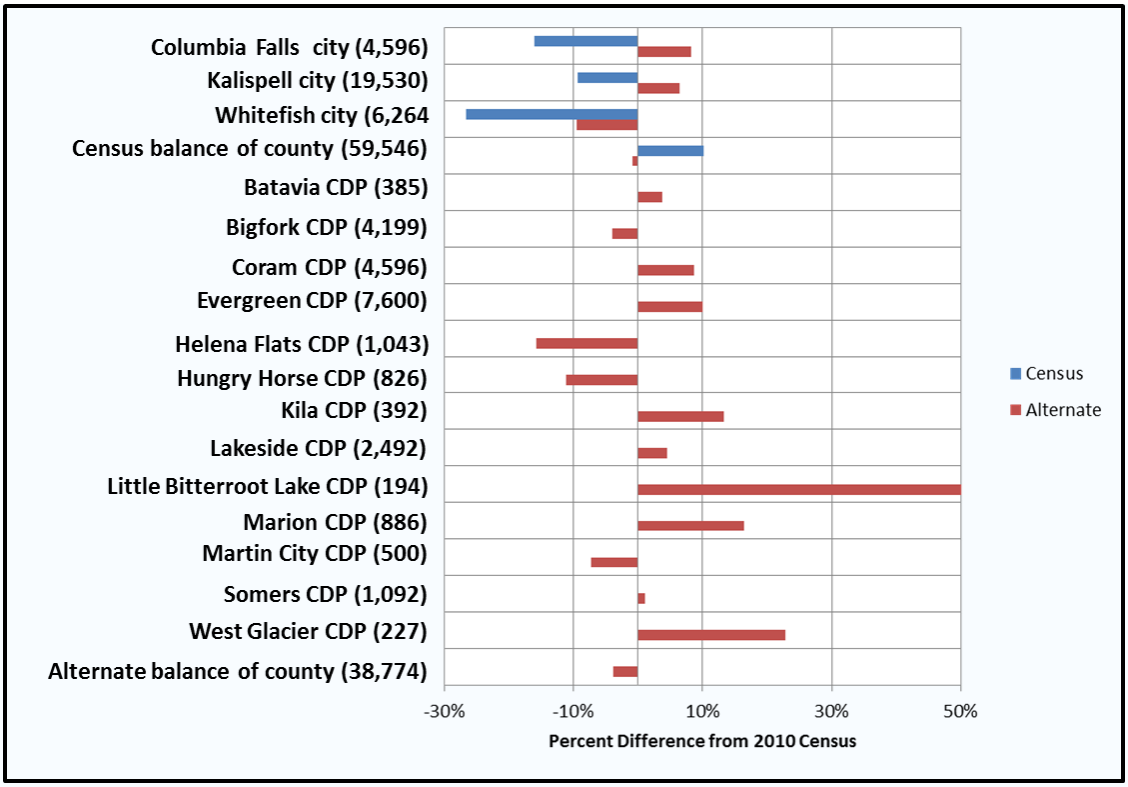


Figure 11: Difference of Gallatin County Montana Subcounty Estimates from 2010 Census