Evaluation of the U.S. Census Bureau's 2008 and 2009 National Population Projections

By

Jennifer M. Ortman

Population Division U.S. Census Bureau

Presented at the Annual Meetings of the Population Association of America, San Francisco, CA, May 3-5, 2012.

This paper is released to inform interested parties of ongoing research and to encourage discussion of work in progress. Any views expressed on statistical, methodological, technical, or operational issues are those of the authors and not necessarily those of the U.S. Census Bureau.

ABSTRACT

This paper evaluates the accuracy of the U.S. Census Bureau's Census 2000-based national population projections released in 2008 and 2009. The population projections are evaluated through comparisons to the 2010 Census counts and the Population Estimates Program's national population estimates. Projections of the components of population change (births, deaths, and net international migration) for the period from 2000 through 2010 are compared to the Population Estimates Program's estimates of these components. The results provide a tool to inform data users about the intrinsic uncertainty of projections data. The evaluations also yield valuable insights to assist in the development of the Census Bureau's 2010 Census-based population projections.

INTRODUCTION

The U.S. Census Bureau's national population projections provide a valuable resource for researchers, media, policy makers, and others with interest in the future direction of population change. This paper evaluates the accuracy of the Census 2000-based national population projections released by the Census Bureau in 2008 and 2009. Both the projections of the population and the components of population change – births, deaths, and net international migration – are assessed. The evaluation of the accuracy of these data offers insight to assist in the development of the Census Bureau's 2010 Census-based population projections, which are planned for release in 2012. The results will also inform data users about the intrinsic uncertainty of projections data.

DATA AND METHODS

This paper evaluates the 2008 National Population Projections (U.S. Census Bureau, 2008) and the series of supplemental projections produced in 2009 (U.S. Census Bureau, 2009).¹ These projections are of the resident population by age, sex, race, and Hispanic origin. The projections are based on Census 2000 and were produced using a cohort-component method. The components of change were projected into the future based on past trends. These projections data were produced for each year from July 1, 2000 to July 1, 2050. This paper will only include data for the period from July 1, 2000 to July 1, 2010.

The series produced in 2009 consists of four scenarios based on alternative international migration assumptions relative to those used in 2008: High Net International Migration, Low Net International Migration, Constant Net International Migration, and Zero Net International Migration. The 2009 Zero Net International Migration series is not included in the analysis. The

¹ A list of the data sources used in the evaluations presented in this paper is provided in Appendix 1.

observed divergence between the Census Bureau's annual estimates and projected levels of net international migration for the period from 2003 and 2008 provided the basis for developing the 2009 High and Low Net International Migration series. Estimates of net international migration produced by the Census Bureau's Population Estimates Program (U.S. Census Bureau, 2008) between 2003 and 2008 were lower than the projected values in the 2008 National Projections for the same period. In the series of projections released in 2009, the level of projected net international migration from the 2008 series was increased and reduced based on the ratio of the estimates of net international migration to the projections of net international migration for the years 2001 through 2008. The result is a ratio of 0.8586. Using this approach, the overall number of migrants projected to enter or leave the population is modified while maintaining the assumptions about the distributions of demographic characteristics (age, race, sex and Hispanic origin) and the projected trends in international migration used in the 2008 series. The level for the 2009 Constant series was developed by first reducing the projected level of net international migration from the 2008 series by the same ratio used to create the Low series. The reduced level of net international migration, 974,885, is held constant over the projection period. The distribution of net migration by age, sex, race, and Hispanic origin for the constant series is the same as the 2008 National Projections distribution from 2000-2001.

Projections of the population are evaluated through comparisons to the 2010 Census counts and the Vintage 2010 estimates of the national population produced by the Population Estimates Program (U.S. Census Bureau, 2010a; U.S. Census Bureau, 2010b). The 2010 Census represents the count of the population on April 1, 2010. The Vintage 2010 estimates of the national population are based on Census 2000 and were produced using the cohort-component method. The estimates provide an estimate of the resident population by age, sex, race, and

Hispanic origin for each year from 2000 to 2010. The evaluations include an assessment of the projected total size of the U.S. resident population from each series as well as the population distributed by age, sex, race, and Hispanic origin.

The projected number of births, deaths, and net international migrants for the period from July 1, 2000 to July 1, 2010 are also evaluated. The components of population change form the foundation of the population projections. The analysis of these components is based on comparisons to the estimates of these components produced by the Population Estimates Program for the same period. Differences between the projected and enumerated population can be understood in part through differences in the projected and estimated components of population change.

RESULTS

Projections and Counts of the Resident Population

The comparison of the 2008 and 2009 National Projections with the 2010 Census counts shows that the 2009 Low Net International Migration series was the closest to the 2010 Census count for the total resident population.² The projected total population in the low series was 308.2 million, which was 464,000 lower than the 2010 Census count of 308.7 million (see Table 1). The 2009 Constant Net International Migration series was also lower than the census count, but by 839,000, while both the 2008 projections and the 2009 High Net International Migration series were higher than the total census count by 1.5 and 3.8 million, respectively.

The projections of the Hispanic population were lower in all of the projection series with the exception of the 2009 High Net International Migration series. The Hispanic population

² The 2010 Census counts represent the U.S. resident population on April 1, 2010. The projections data represent the projected U.S. resident population on July 1, 2010. Since the data measure points in time that are three months apart, small differences are expected when comparing Census counts with projections data.

projected in the 2008 series was 752,000 fewer than in the 2010 Census and the 2009 low series had 1.8 million fewer Hispanics. The 2009 high series had just 440,000 more than the census count. In contrast, the projections of the non-Hispanic population were higher than the 2010 Census counts in all four projection series. The differences for the non-Hispanic population ranged from 3.3 million in the 2009 High Net International Migration series to 1.1 million in the 2009 Constant Net International Migration series.

With respect to race, the 2009 High Net International Migration series was the closest in total for each race group with the exception of the White population (see Table 2).³ The percent differences between the projection series and census counts are quite small for the White, Black, and Asian populations. For the White population, the differences range from 1.39 percent to 2.56 percent. The percent differences range from -0.36 percent to -1.28 percent for the Black population. The percent differences for the Asian population are slightly larger, ranging from - 1.57 to -9.91 percent. The differences are even larger for the American Indian or Alaska Native, Native Hawaiian and Other Pacific Islander, and Two or More Races populations.

Figure 1 presents the comparison of the 2010 Census counts and population projections by age. For most ages, the census counts and projections track very closely together. It is notable that the census counts are slightly lower than the projected population for ages 0 to 10. The percent differences between the population projections and the census counts are shown in

³ Race and Hispanic origin are collected according to the Office of Management and Budget (OMB) 1997 guidelines. For further information, see *Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity*, see http://www.whitehouse.gov/omb/fedreg_1997standards/. Race and Hispanic origin are treated as two separate and distinct concepts in the federal statistical system. People in each race group may be either Hispanic or not Hispanic, and people of Hispanic origin may be of any race. This document contains projections data for each of five racial categories (White, Black, American Indian and Alaska Native, Asian, and Native Hawaiian and Other Pacific Islander). This document refers to each of the races alone and uses the Two or More Races category to represent the population reporting more than one race.

Figure 2. In this figure, values less than zero indicate that the projections were lower than the census and values greater than zero indicate that the projections were higher than the census.

The largest differences are for the population under 5 years and projections of the population aged 0 to 9 were higher than the census counts. Evaluations of the coverage of the 1990 and 2000 censuses concluded that there was an undercount of the population under the age of 10 (Edmonston, 2002; Robinson, 2001; West and Robinson, 1999). Census 2010 counts of the population are also lower than estimates based on administrative records, suggesting that the undercount of children persists in the current census (O'Hare, 2012; Velkoff 2011). In light of these findings, the projections may better represent the population at the youngest ages than do the census counts.

The projections of the middle adult ages (25 to 44 years) were higher than the census counts. Differences for the populations aged 45 to 60 years and 70 and over are for the most part slightly lower than the census counts. For most ages between 60 and 70 years, the projected population was lower than the census count. A notable exception to this is for age 64, which spikes to a nearly six percent difference from the census count. Since the census counts represent the population on April 1 and the projections are for the population on July 1, some differences are expected. The magnitude of this difference is also in part an artifact of the method used to age the Census 2000 base population forward three months to create a projection of the population on July 1, 2000, which is the date from which the annual projections were developed. This three-month projection was generated by aging 25 percent of each birth cohort forward one year. The cohort in question was projected forwarded by subtracting one-quarter of the 54-year-old population (to added to the population aged 55 years) and then adding one-quarter of the 53-year-old population. An assumption implicit to this method is that births for

each cohort were distributed evenly by month within the year of birth. However, the onset of the baby boom in 1946 is characterized by an increase in births in the later months in 1946.⁴ The decision to age one-quarter of each birth cohort forward had the effect of inflating the 54-year-old cohort, thus accounting for some of the difference observed for that cohort 10 years later. We plan to address this issue in the next set of projections.

As was the case for the projections of the total population, the 2009 Low Net International Migration series was the closest to the census counts of the male and female populations. The projected male population in the low series was 40 thousand higher than the census count while the female population was lower than the census county by 503 thousand. In all series, the male population was projected to be just over 50 percent of the population while the female population was projected to be just less than 50 percent of the population. These distributions are consistent with the census distribution of the population by sex.

Projections and Estimates of the Components of Population Change

In this section, I evaluate the projected births, deaths, and net international migration through comparisons to the Population Estimates Program's estimates of these components for the period from 2001 and 2010.⁵ Figure 4 presents the percent difference in births between the projections and estimates. For the period from 2001 through 2008, the projected number of births was lower in all projection series compared to the estimates. The projected fertility rates were based on a time series of birth data ending in 2003. These data are the same inputs used by the Population Estimates Program, thus for the first few years the projected number of births

⁴ Hogan, Perez, and Bell, 2008, Who (Really) Are the First Baby Boomers?, In Joint Statistical Meetings

Proceedings, Social Statistics Section, Alexandria, VA: American Statistical Association, pp. 1009–1016.

⁵ For the components of change, the year indicates events occurring in the twelve-month period leading up to that time point. For example, the year 2001 represents events occurring from July 1, 2000 to June 30, 2001.

tracks closely with the estimates. The projected number of births was higher in all series compared to the estimated births for 2009 and 2010. The projected births from the 2009 High Net International Migration series are the closest to the estimates for 2001 through 2008, while the births in the 2009 Low and Constant Net International Migration series are the closest for 2009 and 2010. Higher levels of net international migration increased the size of the female population to which fertility rates were applied, thus raising the overall number of births. While the fertility rates may have been too low for most years, applying those rates to a larger population resulted in projections of births that were closer to what was estimated for 2001 through 2008.

To evaluate how well the assumptions about fertility performed for each race and ethnic group, I assessed the percent differences between the projected and estimated births by race and Hispanic origin (see Figure 5). The patterns of differences for White alone births follow a similar pattern to the differences for total births. The projected number of non-Hispanic White births was consistently higher than the estimates and the differences grew over the course of the decade. By 2010, the projected non-Hispanic White births exceeded the estimates by a range of 7.5 to 8.1 percent. The projected Black births were lower in all years and the same trend is found for Asian births, with the exception of births in 2010, where births in the 2009 High Net International Migration series were slightly higher than the estimates. Differences in the American Indian and Alaska Native births are similar to the patterns for Black births, indicating that the projected births for this group were also much lower than what was estimated over the course of the decade. The projected Native Hawaiian and Other Pacific Islander births fluctuated up and down, for the most part hovering around a difference of zero percent. In 2009 and 2010,

projected births for this group were higher than the estimates. Hispanic births were consistently under-projected in all four projection series.

The percent differences between the estimates and projections of deaths are shown in Figure 6. Projected deaths for 2001 through 2004 were very close to the estimates for this period. As was the case for fertility, mortality projections were based on a time series of death data that extended through 2003, so it is expected that the deaths would be similar for the first few years of the projections. The projected deaths were higher than the estimates in all other years, rising to a peak difference in the total number of deaths of about five percent in 2009.

Percent differences in deaths by race and ethnic group are shown in Figure 7. Deaths to the White and non-Hispanic White groups follow a trend similar to the total population. Differences were small in the early years while in later years the projected deaths exceeded the estimates. Projected deaths for the Black, Asian, Native Hawaiian and Other Pacific Islander, and Hispanic groups were much higher than what was estimated. Differences for the Asian and Native Hawaiian and Other Pacific Islander populations were quite dramatic, though these are relatively small populations. The projected American Indian and Alaska Native deaths were lower than the estimates from 2001 through 2006. In the later years, projected deaths were higher than the estimates for this group. Projected deaths for the Two or More Races population were consistently lower than the estimates, though the differences decreased in size over time. Projected deaths for this group were very similar to the estimates for the last years of the time series.

Projections and estimates of net international migration for the period from 2001 to 2010 are compared in Figures 8 and 9. Figure 8 presents the comparison of the levels of migration while Figure 9 shows the percent difference in the projections versus the estimates. For most

years, the projected level of net international migration in the 2009 constant series was closest to what was estimated. This is to be expected given that the overall level of net international migration declined over the course of the decade while the projections, based on time series trends in the preceding decades, projected an increase in level of migrants in all series with the exception of the constant series.

Figure 10 presents the percent differences by year between the projections and estimates of net international migration for Hispanics, non-Hispanic Asians, and non-Hispanic all other races. The patterns of differences are similar for all three groups. The assumptions of constant and lower levels of net international migration performed well for all groups through 2006. After 2006, the level of net international migration declined further, and the constant series provides the projected level of net international migration that is closest to what was estimated.

DISCUSSION

In this paper, I evaluated the U.S. Census Bureau's 2008 and 2009 population projections through comparisons to the 2010 Census counts and Population Estimates Program's estimates of the components of population change. The 2009 Low Net International Migration series was the closest to the total population as enumerated in the 2010 Census while the 2009 High Net International Migration series yielded projections of the population closest to the 2010 Census counts for the Hispanic population and all race groups, with the exception of the White population. There is little difference when comparing the projected distribution of the population by sex to the census counts. With respect to age, the population projections for 2010 have an age distribution that is overall very similar to the 2010 Census.

To better understand the pattern of differences between the projected and enumerated population, I also evaluated the projected components of change – births, deaths, and net international migration – for the period from 2001 to 2010. Projected births for the non-Hispanic White alone population were higher than the estimates while the births for all other groups were lower than what was estimated. Projections of deaths were higher than in the estimates, with the exception of the Two or More Races population where the projected deaths were lower than what was estimated. Net international migration was for the most part higher in the projections than the estimates for the decade. Projected levels of net international migration were based on trends in immigration through the 1990s, when immigration was growing rapidly. After 2001, the level of net international migration declined, falling below one million per year for much of the decade. As a result, the assumption of a constant level of net international migration resulted in projected levels of net international migration closest to what was estimated by the Population Estimates Program for this period. Changes in migration trends have the largest impact on the Hispanic population, since a large proportion of migrants to the United States are Hispanic. While we might have expected projections of the Hispanic population to be too high in all of the projection series, which was not the case, the overstatement of net international migration was offset by the tendency to under project the level of natural increase (births minus deaths) in each projection series.

The results of this evaluation will be used to inform research and development of the next series of national projections, which will be based on the 2010 Census and are planned for release in 2012. This evaluation has shown that while the 2008 and 2009 projection series were reasonably close to the official population counts for 2010, there are noteworthy differences between what was projected and what was enumerated within subgroups of the population.

Looking beyond 2010, these differences could have an impact on projections of the working age and older age populations as well as the year in which the total minority population is projected to become a numerical majority.

One avenue of research that we are undertaking is to reconsider the race-origin groupings used to project rates of fertility and mortality. Perhaps combining so many groups into the category of non-Hispanic non-Black resulted in rates that were too low (or high) for some groups. Non-Hispanic White births were 'over-projected' while American Indian and Alaska Native and Asian births were 'under-projected'. Deaths for the Asian and Native Hawaiian and Other Pacific Islander populations were 'over-projected' while the number of deaths for the White and American Indian and Alaska Native populations were close to the estimates. These patterns of difference suggest that the mortality rates for these populations are not as similar as it was assumed when these projections were developed. For the 2012 National Projections, we intend to expand the race-origin groups beyond the three groups used in the 2008 and 2009 series in an effort to develop rates that are more suitable to the groups that we project.

Differences in the estimated and projected levels of net international migration serve to validate the long held understanding the international migration is arguably the most difficult component to project. The adoption of a constant assumption about the level of international migration dominated the Census Bureau's population projections for most of the 20th century. Assuming constant levels of net international migration proved to miss the mark entirely in light of the substantial increases in the number of immigrants through the last decades of the 20th century and led to the development of migration assumptions that projected migration as a dynamic flow of individuals with levels and composition that changed over the projection horizon. This change resulted in an overestimation of net international migration for the period

from 2000 to 2010. Interestingly, the constant series released in 2009 actually came closest to the total estimates of net international migration for the decade. In the next series of projections, we plan to continue to develop and disseminate alternative assumptions for the level of net international migration to communicate uncertainty in the component as well as what alternative results would be if different levels, distributions, and flows of migration were considered.

REFERENCES

- Edmonston, Barry. 2002. *The Undercount in the 2000 Census*. Baltimore, MD and Washington, DC : The Annie E. Casey Foundation and The Population Reference Bureau.
- O'Hare, William P. 2012. "What Demographic Analysis Tells Us About Count of Children in the 2010 Census." Presented at the Applied Demography Conference, San Antonio, TX.
- Robinson, J. Gregory. 2001. "ESCAP II: Demographic Analysis Results." *Executive Steering Committee for A.C.E. Policy II*, Report No. 1. Washington, DC: U.S. Census Bureau.
- U.S. Census Bureau. 2008. 2008 National Population Projections. Washington, DC: U.S. Census Bureau. (http://www.census.gov/population/www/projections/2008projections.html).
- _____. 2009. 2009 National Population Projections (Supplemental). Washington, DC: U.S. Census Bureau. (http://www.census.gov/population/www/projections/2009projections.html).
- . 2008. Vintage 2008 Population Estimates. Washington, DC: U.S. Census Bureau. (http://www.census.gov/popest/data/historical/2000s/vintage_2008/index.html).
- _____. 2010a. 2010 Census. Washington, DC: U.S. Census Bureau. (http://2010.census.gov/2010census/)
- . 2010b. Vintage 2010 Population Estimates. Washington, DC: U.S. Census Bureau. (<u>http://www.census.gov/popest/research/eval-estimates/eval-est.html</u>).
- Velkoff, Victoria A. 2011. "Evaluating the United States Census." Panel presentation at the Annual Meetings of the Population Association of America, Washington, DC.
- West, Kirsten K. and J. Gregory Robinson. 1999. "What Do Know About the Undercount of Children?" *Population Division Working Paper*, No. 39. Washington, DC: U.S. Census Bureau.

	Total population		Hispanic		Non-Hispanic	
2010 Census Counts and		Difference from		Difference from		Difference from
Population Projections	Number	2010 Census	Number	2010 Census	Number	2010 Census
2010 Census	308,746	(X)	50,478	(X)	258,268	(X)
2008 National Projections	310,233	1,487	49,726	-752	260,507	2,239
2009 National Projections						
High Net International Migration Series	312,504	3,758	50,918	440	261,586	3,318
Low Net International Migration Series	308,282	-464	48,702	-1,776	259,580	1,312
Constant Net International Migration Series	307,907	-839	48,531	-1,947	259,375	1,107
(Not applicable						

differences are expected when comparing Census counts with projections data.

Source: U.S. Census Bureau, 2008-2010

Table 2. Population Projections and 2010 Ce Numbers in thousands		y nace and mop								
			2009 Net International Migration Serie							
		2008 National								
Race and Hispanic Origin ^{1, 2}	2010 Census		High	Low	Constant					
White	241,937	246,630		245,336	245,300					
Black	40,251	39,909	40,105	39,741	39,735					
American Indian or Alaska Native	3,740	3,188	3,206	3,172	3,169					
Asian	15,160	14,415	14,922	13,979	13,658					
Native Hawaiian and Other Pacific Islander	675	592	601	584	581					
Two or More Races	6,984	5,499	5,534	5,470	5,465					
Hispanic	50,478	49,726	50,918	48,702	48,531					
Non-Hispanic White	197,319	200,853	201,235	200,524	200,639					
Difference from 2010 Census (Number in thousands)										
White	(X)	4,693	6,200	3,399	3,363					
Black	(X)	-342	-146	-510	-516					
American Indian or Alaska Native	(X)	-552	-534	-568	-571					
Asian	(X)	-745	-238	-1,181	-1,502					
Native Hawaiian and Other Pacific Islander	(X)	-83	-74	-91	-94					
Two or More Races	(X)	-1,485	-1,450	-1,514	-1,519					
Hispanic	(X)	-752	440	-1,776	-1,947					
Non-Hispanic White	(X)	3,534	3,916	3,205	3,320					
Differe	nce from 2010	O Census (Perce	nt)							
White	(X)	1.94	2.56	1.40	1.39					
Black	(X)	-0.85	-0.36	-1.27	-1.28					
American Indian or Alaska Native	(X)	-14.76	-14.28	-15.19	-15.27					
Asian	(X)	-4.91	-1.57	-7.79	-9.91					
Native Hawaiian and Other Pacific Islander	(X)	-12.30	-10.96	-13.48	-13.93					
Two or More Races	(X)	-21.26	-20.76	-21.68	-21.75					
Hispanic	(X)	-1.49	0.87	-3.52	-3.86					
Non-Hispanic White	(X)	1.79	1.98	1.62	1.68					

Table 2. Population Projections and 2010 Census Counts by Race and Hispanic Origin for the United States: 2010

Race refers to each of the race groups alone and the Two or More Races category represents the population ² Hispanics may be of any race.

Notes:

The original race data from Census 2010 are modified to eliminate the "some other race" category. This modification is used for all Census Bureau projections products and is explained in the document entitled "Modified Race Data Summary File Technical Documentation and ASCII Layout" that can be found on the Census Bureau website at http://www.census.gov/popest/archives/files/MRSF-01-US1.html.

The 2010 Census counts represent the U.S. resident population on April 1, 2010. The projections data represent the projected U.S. resident population on July 1, 2010. Since the data measure points in time that are three months apart, small differences are expected when comparing Census counts with projections data.

Source: U.S. Census Bureau, 2008-2010









Note: Percent difference is calculated as: (Projections - Estimates) / Estimates Source: U.S. Census Bureau, 2008-2010





Note: Percent difference is calculated as: (Projections - Estimates) / Estimates * 100. Source: U.S. Census Bureau, 2008-2010







Appendix 1. Data Sources

National Projections for years 2000-2050

- 2008 National Population Projections
- 2009 National Population Projections (supplemental)
 - High Net International Migration Series
 - o Low Net International Migration Series
 - Constant Net International Migration Series

2010 Census

Vintage 2010 National Population Estimates for years 2000-2010