400voices

Analytics for Hispanic Population Data

by Dr. Donald Allen
Professor of Mathematics
Texas A&M University

2016

Analytics. To understand the estimates of how populations grow, we must accept that projections into the future involve making some assumptions. And the **assumptions** are often one of the two principle sources of the error. Assumptions comprise the greatest source of error in making projections. Here is a simple example.

> Example. When you read a report (Nations 2004) that the world's population in 2050 will be 9.4 billion, you might wonder as to the origin of this number. After a little inspection, the UN analysts seemed to assume the rate of increase of the world population is declining by about 0.02% per year, beginning with a rate of increase of about 1% in the year 2016. (The rate of increase year-by-year is decreasing, only slightly.) The assumptions about the rates of change came from current estimates.

The other principle source of error is with the **model**. Given the data and your assumptions, you need some mechanism to project ahead. The simplest way to do this is graph the data (you have to have a starting point) and draw a straight line at some angle with your assumptions in mind, and read from the graph its value at future times. When it comes to populations, straight lines often don't work well at least in the long term.

In the next few paragraphs, we will consider known data, examine assumptions, and create certain models to see how they predict what the future will bring. The first model will be somewhat phenomenological, looking at the data and looking for the data trends. The second model uses the classical growth function, similar to the way money accrues in an interest bearing savings account. The third model will detail how the United States Census Bureau makes its projections, and where the errors are. Yes, when one predicts the future for virtually every social or economic outcome, there will almost certainly be errors.

Our goal is to project current information to determine how the Hispanic population will grow in the decades ahead. We can answer questions such as (a) when will it reach 100 million citizens, and (b) will it overtake the White (non-Hispanic) population and when?

We need some notation: P_n - the population of the designated group in the year *n.* Normally, this is taken to be on July 1 of year *n*. For example, P_{2010} designates the population on July 1, 2010. In fact, *n* need not be an integer, if one wishes to discuss some point during the year. For convenience, in this note, we will use *P* for the total United States population or a general population, *H* for Hispanic populations, *W* for white (non-Hispanic) populations, and *B* for Black populations. Other designations will be defined as needed.

Method 1 - Spreadsheet modeling. Using past population data. In this method, a table of population data from past years is tabulated, and a trend line is constructed. This trend line is extended to future dates, and populations into the future are projected. We begin with historical data from 1940 to 2010. It is found online (Wikipedia 2015) and given in Tables 1a and 1b.

	Population (in millions)							
	1940	**1950**	**1960**	**1970**	**1980**	**1990**	**2000**	**2010**
Total Population	131.67	150.70	179.32	203.21	226.55	248.71	281.42	308.75
Non-Hispanic White	116.26	131.81	153.22	169.62	180.26	188.13	194.55	196.82
Hispanic (of any race)	2.02	3.23	5.81	8.92	14.61	22.35	35.31	50.48
Gap (White - Hispanic)	114.24	128.57	147.40	160.70	165.65	165.77	159.25	146.34
Black	12.87	15.04	18.87	22.54	26.50	29.99	34.66	38.93

Table 1a – Total Populations (in millions)

	1940	**1950**	**1960**	**1970**	**1980**	**1990**	**2000**	**2010**
Non-Hispanic White	88.30%	87.46%	85.44%	83.47%	79.57%	75.64%	69.13%	63.75%
Hispanic (of any race)	1.54%	2.14%	3.24%	4.39%	6.45%	8.99%	12.55%	16.35%
Gap (White - Hispanic)	86.76%	85.32%	82.20%	79.08%	73.12%	66.65%	56.59%	47.40%
Black	9.77%	9.98%	10.52%	11.09%	11.70%	12.06%	12.32%	12.61%

Table 1b – Percentages of the total population

As is evident, both populations of non-Hispanic White and Hispanic have changed considerably over these seven decades. With this data, we make a graphic to determine if anything can be discerned quantitatively. Shown in Figure 1, it shows both White (non-Hispanic) and Hispanic seem to curve and moreover at some time in the future will seem to intersect. Data for African Americans has been included for comparative purposes. The data for Asian populations and other races has been excluded in the interest of our focus. As is evident, there is a quadratic-like curve that seems to fit this data. With that in mind, we find the best fit of a quadratic to this data, and this can be accomplished using spreadsheet tools.

We have computed the quadratic trend lines to the first three populations. (That for the Black population is omitted here.) The Pearson R^2 correlation coefficients are exceptionally close to 1.000, which would be exact agreement. This makes a convincing argument the populations are changing quadratically. That this is not so is demonstrated later, but in the meantime, they can be used to estimate when the populations hypothetically converge. We solve the for what value the quadratic equations have the same value. Computing these equations to more places than shown we have

$$W = -15671.828 + 16.299x - 0.00421x^2$$
$$H = 10277.0697 - 10.607x + 0.00274x^2$$

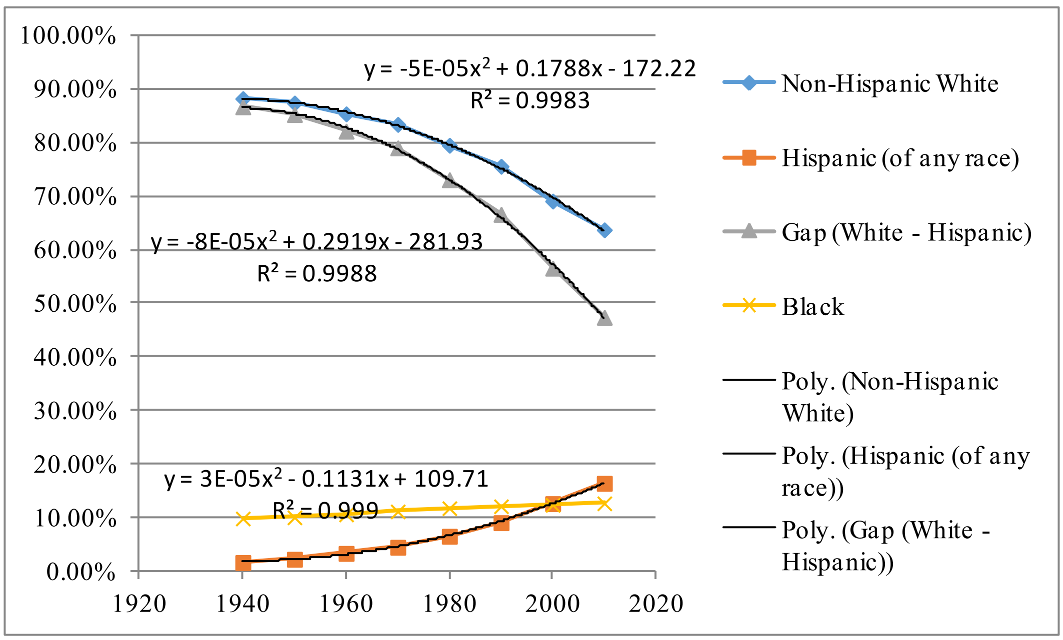

Figure 1- Populations Data 1940-2010

Solve $W = H$ for x to get $x = 2047.04$ and at this time each will share about 34.5% of the population. This value is the theoretical time when the populations should converge. Between the two populations, this gives only 69% of the total population, implying something may be incorrect, or that other racial populations are changing far more than the data suggests. In Figure 2, we show the same data extrapolated using the quadratic model through the year 2060. Doing so is notoriously risky for most data. It assumes the same conditions that prevailed during the seven decades of the data sets remains in play for the next 3.5 decades. Incidentally, the Hispanic population in this model would reach 100,000,000 sometime in the year 2037-2038.

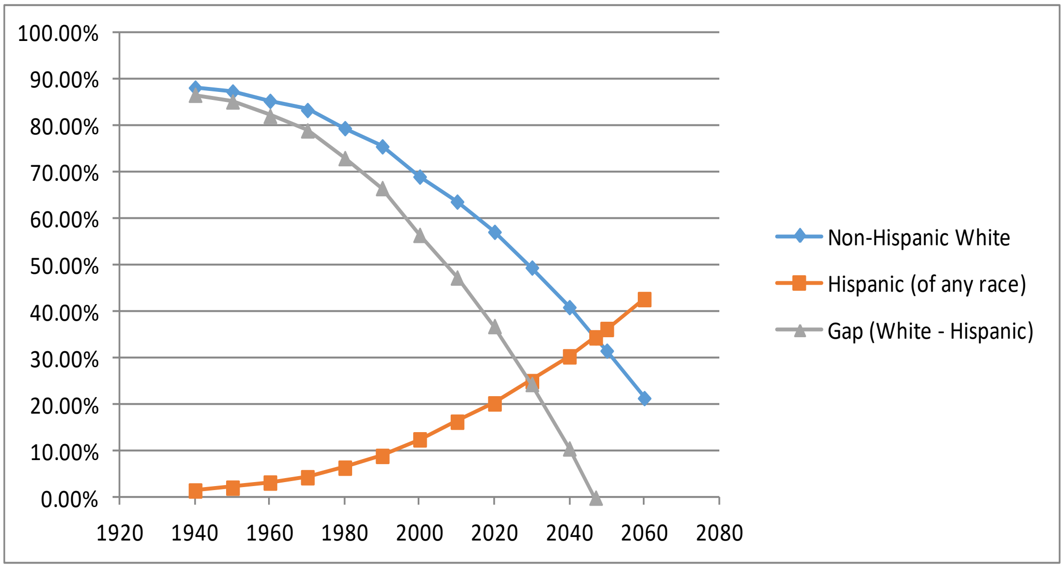

Figure 2 – White and Hispanic Populations – Projected to 2060

Recall, the quadratic fit is so good, we are tempted to make a causality assumption that the Hispanic and White populations do follow a quadratic curve. However, if we include more data, say for the years 2011-2014, we discover a different story. The enhanced data set is shown in Table 2, the graphic in Figure 3.

	2010	**2011**	**2012**	**2013**	**2014**
Non-Hispanic White	63.75%	63.25%	62.83%	62.44%	61.91%
Hispanic (of any race)	16.35%	16.67%	16.87%	17.08%	17.34%
Gap (White - Hispanic)	47.40%	46.58%	45.96%	45.36%	44.57%
Black	12.61%	12.24%	12.25%	12.28%	12.31%

Table 2. Population Data 2010-2014

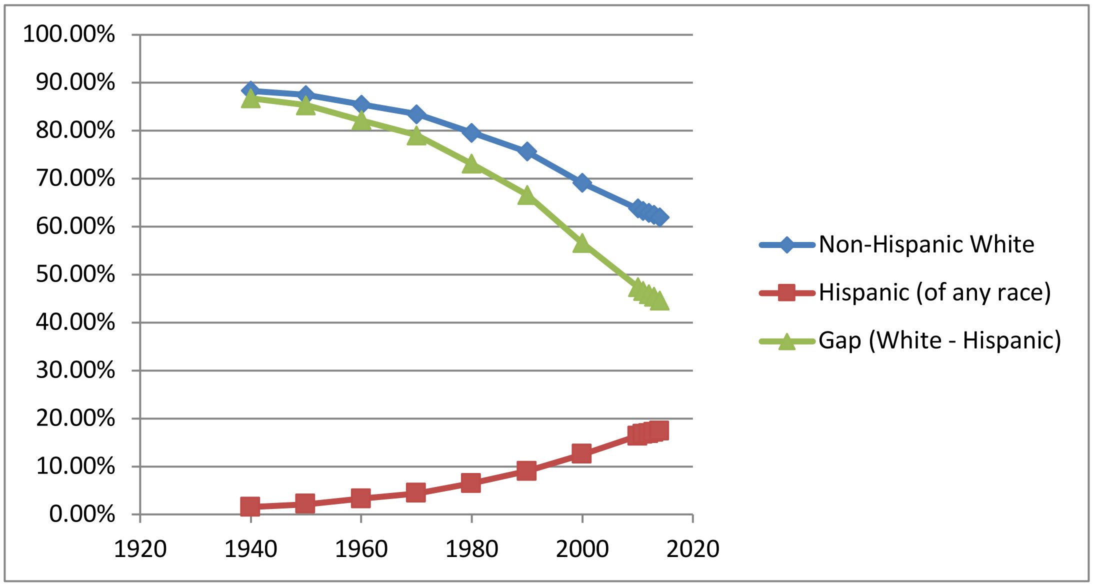

Figure 3. Population Percentages though 2014

Notice the ever so slight change in both the Hispanic and White data, with the Hispanic numbers still rising but at a decreased rate, and the White numbers falling at a lower rate. Not surprisingly, in the population projections game, the most recent numbers are usually given a far higher weight on calculations. It is of value in the next section to show the percentage change of these populations.

	Populations (in millions)							
Year	**Non-Hispanic White**	**Estimated Percentage Change (annual)**	**Hispanic (of any race)**	**Estimated Percentage Change (annual)**	**Gap (White - Hispanic)**	**Estimated Percentage Change (annual)**	**Black**	**Estimated Percentage Change (annual)**
1940	116.3		2.0		114.2		12.9	
1950	131.8	1.34%	3.2	5.98%	128.6	1.25%	15.0	1.69%
1960	153.2	1.62%	5.8	7.99%	147.4	1.46%	18.9	2.55%
1970	169.6	1.07%	8.9	5.34%	160.7	0.90%	22.5	1.94%
1980	180.3	0.63%	14.6	6.38%	165.6	0.31%	26.5	1.76%
1990	188.1	0.44%	22.4	5.30%	165.8	0.01%	30.0	1.32%
2000	194.6	0.34%	35.3	5.79%	159.2	-0.39%	34.7	1.56%
2010	196.8	0.12%	50.5	4.30%	146.3	-0.81%	38.9	1.23%
2011	197.1	0.14%	51.9	2.90%	145.1	-0.82%	38.1	-2.00%
2012	197.2	0.08%	53.0	1.97%	144.3	-0.59%	38.5	0.83%
2013	197.4	0.08%	54.0	1.94%	143.4	-0.61%	38.8	0.89%
2014	197.4	0.01%	55.3	2.40%	142.1	-0.89%	39.3	1.18%

Table 3 – Percentage Change of Population

Important also is the population change is Table 3, where we show the estimated percentage change of the respective populations. It is the last number that has some merit in the next section, Method 2.

As is apparent, the White population, while growing, is growing at a rate insufficient to sustain its numbers. For developed countries, the sub-replacement fertility is any rate below approximately 2.1 children born per woman. (Espenshade, Guzman, and Westoff 2003). The U.S. birth rate now is 1.9 births per woman over her lifetime.

Important also is the population change is Table 4, .where we show the estimated percentage change of the population percentages. Remember, the percentages are of the total population for the given year, and the total population is changing. Thus, this is different from the percentage increase in the population. There seems to be little by way of patterns in this data, except the Hispanic changes are more-or-less decreasing, while the White (non-Hispanic) charges are negative, even while the absolute numbers are increasing.

Year	Non-Hispanic White	Estimated Percentage Change (annual)	Hispanic (of any race)	Estimated Percentage Change (annual)	Gap (White - Hispanic)	Estimated Percentage Change (annual)	Black	Estimated Percentage Change (annual)
1940	88.30%		1.54%		86.76%		9.77%	#N/A
1950	87.46%	-0.09%	2.14%	3.96%	85.32%	-0.17%	9.98%	0.22%
1960	85.44%	-0.23%	3.24%	5.12%	82.20%	-0.37%	10.52%	0.54%
1970	83.47%	-0.23%	4.39%	3.54%	79.08%	-0.38%	11.09%	0.54%
1980	79.57%	-0.47%	6.45%	4.69%	73.12%	-0.75%	11.70%	0.54%
1990	75.64%	-0.49%	8.99%	3.94%	66.65%	-0.88%	12.06%	0.31%
2000	69.13%	-0.86%	12.55%	3.96%	56.59%	-1.51%	12.32%	0.21%
2010	63.75%	-0.78%	16.35%	3.03%	47.40%	-1.62%	12.61%	0.24%
2011	63.25%	-0.78%	16.67%	1.96%	46.58%	-1.72%	12.24%	-2.90%
2012	62.83%	-0.66%	16.87%	1.21%	45.96%	-1.33%	12.25%	0.08%
2013	62.44%	-0.63%	17.08%	1.22%	45.36%	-1.30%	12.28%	0.19%
2014	61.91%	-0.85%	17.34%	1.52%	44.57%	-1.74%	12.31%	0.32%

Table 4 – Percentage Change of Population Shares

Method 2 – Exponential Models. The exponential method based on growth rates. It is a well known and well-understood law that populations grow at a rate proportional to itself. This model is very effective when there are adequate resources for continued growth and conditions that created the rate are sustained. With population rate of growth, *r*, we can write

$$P_{n+1} = (1 + r)P_n$$

Note, this shows the next population to be the current population plus the growth based on the current population. If r is positive, the population increases; if r is negative, this indicates a decline in the population. Let's consider an example for the Hispanic population. The 2014 US Hispanic population was about 55.4 million, or 17.4% of the US population. This was an increase of 2.1% over the previous year. (Krogstad 2015). Thus $H_{2014} = 55.4$ and $r = 0.021$. Assuming the same for the next year, we expect the Hispanic population to be

$$H_{2015} = (1+0.021)H_{2014} = (1+0.021)55.4 = 55.6 \text{ million}$$

rounded of course. We can continue such calculations year after year. Doing so gives us a table of information. However, observe from the data above (Wikipedia 2015) we compute we see the rate of increase was 2.34% for the 2013-2014 increase in population. Both sources of information are reliable though both give slightly different numbers. For comparison, we will begin both sets of calculations beginning with the Pew number of 55.4 for $H_{2014} = 55.4$ million. See Table 5.

Year	Hispanic Population (in millions) r=2.10%	Hispanic Population (in millions) r=2.4%
2014	55.4	55.4
2015	56.6	56.7
2020	62.8	63.9
2030	77.3	81.0
2040	95.1	102.6
2050	117.1	130.1
2060	144.1	164.9

Table 5 – Projections of Hispanic Population

In this model, we can also answer "what if" or "when" type questions. For example, it is a simple calculation to establish that occurs in the year 2042 (resp. 2040 when r = 2.40%) the Hispanic population will reach 100 million when r = 2.10%. We can give the doubling time, that is, the time for the population to double. It takes 33.3 years (resp. 29.2 years for r = 2.40%) when r = 2.10%. Note that at r = 2.40% the population doubles about every generation and a quarter.

The actual population data is probably not exponential. However, we have used it with the growth rate taken as a composite of birth, death, and immigration factors. In any event, while we normally expect populations to change exponentially, this is a simplistic model that has built in many assumptions, not viable for recent population trends, and particularly for saturated populations. We will discuss some of them in the next section.

We also include a chart of such population growth with various models of amnesty for the 11 million undocumented Hispanics.

Method 3 – The cohort-component method. How population is projected. It is done using the cohort-component method wherein the factors of birth (fertility), death, and immigration are estimated for each period, and these are combined with the current population to compute the subsequent population. The period can be in years or even decades. The wider the time spread, the stronger considered and, therefore, less reliable the assumptions. This is graphically organized in triads with vertices marked at Birth, Death, and Immigration. See in Figure 4. Each point in a triangle corresponds to a unique triple of these vital numbers.

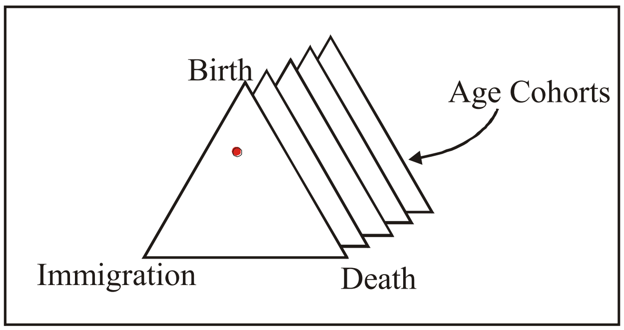

Figure 1 – Cohorts-Component Triads

The basic equation is Frequently the cohorts are divided by decades of population ages: 0-10 years, 10-20 years, and so on. These estimates are used to compute cohort populations for each subsequent year. Then the cohort populations are adjusted as there is necessarily some movement across cohorts as the cohort populations change. The basic population change is a simple balance model. Ignoring cohorts for the moment, the new (or next) population is simply the old population plus births and immigration minus deaths.

$$\text{Population}_{n+1} = \text{Population}_n + \text{Births}_n - \text{Deaths}_n + \text{Immigration}_n$$

However, when there are cohorts, the situation is one level more complicated. Here the new population of say cohort 3 will be the population from cohort 2 minus the deaths in cohort 2 plus the immigration into cohort 3. All births go into cohort 0, the youngest cohort. The old population of cohort 3 will advance to cohort 4, with the same type of provisos.

The Census Bureau uses cohorts of just one year, for ages from 0 (newborn) to 100. Those over 100 are lumped in this last cohort. For comparison purposes, we give the populations of groups up to 2060 in Table 6. Note the Hispanic population more than doubles during this period.

Year	Total	Hispanic	White (non Hispanic)	Black
2015	321,362,789	57,075,129	198,448,842	42,531,561
2020	333,895,553	63,784,157	199,312,742	44,809,572
2030	358,471,142	78,654,856	198,817,220	49,245,947
2040	380,015,683	94,875,732	193,887,051	53,411,745
2050	399,803,369	111,731,705	186,334,175	57,553,051
2060	420,267,733	128,780,232	178,950,774	61,821,604

Table 6

We can look at the gender split in the Hispanic population, as in Table 7.

Year	Hispanic	Hispanic male	Hispanic female
2015	57,075,129	29,063,619	28,011,510
2020	63,784,157	32,570,444	31,213,713
2030	78,654,856	40,374,878	38,279,978
2040	94,875,732	48,861,920	46,013,812
2050	111,731,705	57,618,964	54,112,741
2060	128,780,232	66,426,934	62,353,298

Table 7

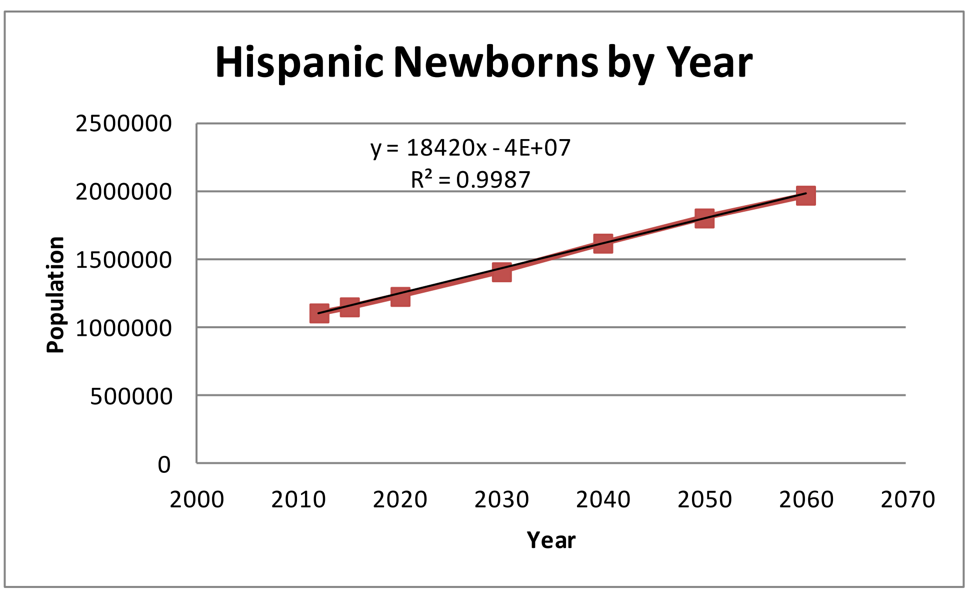

Table 8

We can look at some details. For example, in Table 8 we see the (Hispanic) newborn population projections. They grow linearly at a rate of 18,240 babies per years over the entire range. In real terms, this growth translates to a declining relative newborn population from about 0.4% currently to about 0.1% in 2060. However, it also shows a newborn population of about 200,000 in 2060.

The full 101 cohort Hispanic population projections are shown in Table 9.

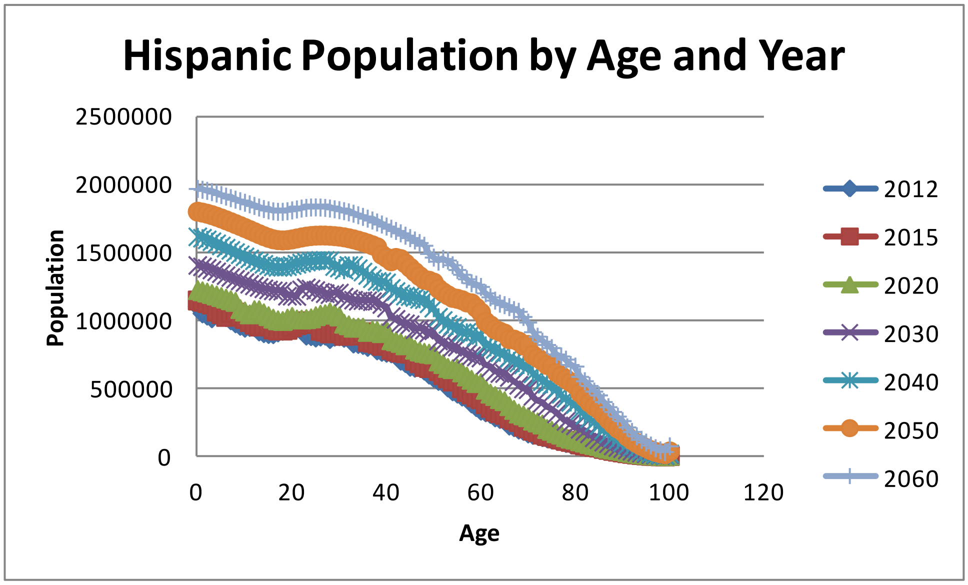

Table 9

Thousands of charts could be generated from the main table. One important thing to note is that a finely detailed look at the population shows significant perturbations in the populations by age and year of extant populations, but almost none for the projected populations – not year born. This is a consequence of modeling. Modeling cannot predict the unpredictable – or black Swan (Taleb, 2007).

All information used here is available among the many tables at http://www.census.gov/population/projections/data/national/2012/downloadablefiles.html.

Risk Factors. Or, what can go wrong with these predictions/projections? While the cohort-component model takes specific and detailed account of the three factors of birth, death, and immigration, the equations of projection must estimate these factors, year-by-year, into the unknown future. However, they do not and cannot account for any of the following, some affecting the population growth positively, and some negatively.

- Decline in fertility rates. The total fertility rate in the United States is estimated at 1.86 children per woman, which is below the replacement fertility rate of about 2.1 children per woman. The fertility rate for Hispanic women in 2013 is estimated at 2.149 children per woman, down from 2.189 in the previous year. (Martin JA 2015)

- The consequences of increased Hispanic affluence and education are factors, both of which lead to decreased birth rates, the age of marriage, and family size effects (Mathews 1997). As more and more Hispanic women achieve higher education, could this be a driver toward unexpected changes in the projections?
- The models seem not to consider the carrying capacity of the environment and a resultant logistic-type growth. When an environment becomes highly populated, there is invoked a natural carrying capacity, whereby the population is constrained in growth by that very environment. Another type of growth obtains – logistic growth is one model.
- Medical advances. A cure for any serious disease that affects women of childbearing age can affect fertility rates.
- Economic downturns such as depressions and recessions and consequent unemployment During economic downturns, from the great recession to the current recession, fertility rates have declined. A similar decline of fertility rates has been reported in Ireland, Italy, Spain, Sweden, and several other European countries (Mather 2012).
- Unforeseen calamities such as disease. Diseases spread more quickly among people who live near each other, such as urban areas. According to a recent report by the Census Bureau about 80.7% of Americans live in urban areas, up from 79% in 2000. In urban setting disease can be more rampant, having a profound effect on health (Alirol 2009).

What is completely unknown is how these factors may affect birth rates in the decades ahead. Nonetheless, population experts make their best estimates and then predictions. We should not be overly concerned, but we should accept their derived estimates are worthy of our attention.

Conclusions. In this article we have reviewed three models for the estimation of population trends, focusing on outcomes for the Hispanic segment of the United States population.

What we can rely on is the Hispanic population should reach 100 million before 2050. Populations numbers this great engender a formidable force, depending on how they are marshaled. With such numbers, it becomes a major player, if not the major player in American politics. We cannot say, nor can others, what kind of economic force such a large population will create, beyond of course consumerism. We can ask questions such as (a) Will a population of this size affect US relations in this hemisphere and how? (b) Will a population of this size affect American attitudes toward religion? (c) Will such a population demand a forced multiculturalism independent of the traditional American way? (d) Will the Hispanic population take the full measure of the educational process and opportunities to force its way into a role of leadership in American commerce? This measure goes beyond measures of higher purchasing power. (e) Will the Hispanic population continue to concentrate in California (now 15 million), Texas (10.4 million), and Florida (4.8 million), (Krogstad 2015), now totaling more than half the total US Hispanic population? (f) Excluding the Asian population, will this new majority-minority embrace the technological revolution or delimit its efforts toward more political and social concerns?

The population projections portend a profound future for the Hispanic population. What will happen? There is no data, no model, and no expectations beyond hope. Will true leaders emerge in sufficient numbers?

References

Alirol, Emilie et al. 2009. "Urbanisation and infectious diseases in a globalised world. *The Lancet Infectious Diseases* no. 11 (2):131-141.

Espenshade, Thomas J., Juan Carlos Guzman, and Charles F. Westoff. 2003. "The Surprising Global Variation in Replacement Fertility." *Population Research and Policy Review* no. 22 (5/6):575-583. doi: 10.1023/B:POPU.0000020882.29684.8e.

Krogstad, Jens Manuel and Lopez, Mark Hugo. *Hispanic population reaches record 55 million, but growth has cooled*. Pew Research Center 2015. Available from http://www.pewresearch.org/fact-tank/2015/06/25/u-s-hispanic-population-growth-surge-cools/.

Martin JA, Hamilton BE, Osterman MJK, et al. . 2015. Births: Final data for 2013. In *National vital statistics reports*. Hyattsville, MD: National Center for Health Statistics.

Mather, Mark. 2012. *Fact Sheet: The Decline in U.S. Fertility*. Population Reference Bureau 20122012]. Available from http://www.prb.org/publications/datasheets/2012/world-population-data-sheet/fact-sheet-us-population.aspx.

Mathews, T. J. , Ventura, Stephanie J. 1997. "Birth and Fertility Rates by Educational Attainment:United States, 1994." *Monthly Vital Statistics Report* no. 45 (10):1-20.

Taleb, Nassim Nicholas. *The Black Swan: The Impact of the Highly Improbable*. New York: Random House, 2007.

Nations, United. 2004. World Population in 2003. edited by Department of Economic and Social Affairs. New York: United Nations.

Wikipedia. 2015. *Historical racial and ethnic demographics of the United States*. Wikipedia 20152015]. Available from https://en.wikipedia.org/wiki/Historical_racial_and_ethnic_demographics_of_the_United_States.