New York population projection by age and sex

County Projections 2005-2035
Model description

Jan K. Vink
Program on Applied Demographics
Cornell University
May, 2009

Program on Applied Demographics

Web: http://pad.human.cornell.edu Email: PADinfo@cornell.edu

2008 PAD projections

Model description

Introduction

This document describes a projection model developed by Cornell Program on Applied Demographics.

The projection model consists out of two closely related parts: a set of algorithms and the numeric coefficients feeding the algorithms. The values of the coefficients differ between counties, the way the algorithms work with the coefficients is not dependent on the location.

Much effort is put in calculating the coefficients in a way that is in line with the way they are used in the algorithms. This document describes both the algorithms and the way the coefficients were derived.

The algorithms formula's are in Appendix A. Appendix B contains a list of institutions that are recognized for possible role in the county population counts in the past and the future. Appendix C describes changes made to the model that led to different versions of the projected counts.

The projected County population counts are published on the internet. The URL for this web page is: http://pad.human.cornell.edu/data/projections.cfm

General model overview

One of the most widely projection methods is called the Cohort-Component method (see [1]). It centers on the notion that the population can be split in age-sex cohorts and that changes in the size of each of the cohorts can be split into 3 components: births (only the youngest cohort), deaths and migration. The component method takes that notion and projects the size for each of those components for a period in time and than calculates the population size at the end of the period, giving the size of the population at the beginning and the component sizes.

In a formula this looks like:

$$Pop_{t2} = Pop_{t1} + B_{t1,t2} - D_{t1,t2} + NM_{t1,t2}$$

Where Pop_{t1} and Pop_{t2} are the cohort population count at respectively moment t1 and t2, $B_{t1,t2}$ the number of births between t1 and t2, $D_{t1,t2}$ the number of deaths between t1 and t2 and $NM_{t1,t2}$ the net migration between t1 and t2. The net migration in it self is the size of in-migration minus the size of out-migration.

The projection model we developed has the following characteristics:

- The start population is derived from the 2005 population estimates by the U.S. Census Bureau [2],
- The population is divided by sex and 5 year age groups, with 85

and older being the oldest age group,

- The population is projected in 5 year intervals, ending in 2035,
- **County specific coefficients are estimated using historic data and are held constant.**

Changes to these basic model characteristics can be made if necessary and will lead to publication of a new version of the results. Appendix C describes a history of the versions and the changes made in each version.

This model description first describes the start population, next it briefly explains the sub models for each of the components and the way County specific coefficients for these sub models were estimated.

The formulae for the sub models can be found in Appendix A.

Start population

Each year the U.S. Census Bureau produces post-Census estimates [3]. These estimates play an important role in distribution of regional funds, controls for surveys, denominator in several rates and ratios (e.g. birth rates, mortality rates employment rates), etc.

Besides total population, the Census Bureau also estimates population by sex and several age and race/ethnicity categories.

With the release of each new set of estimates, the old set is revised.

We wanted to use the newest release of these sex/age estimates for 2005 as our start population, but found problems with the age distributions as estimated by the Census Bureau.

Instead we use the latest population by sex estimate to estimate historic crude migration rates, but arrive at the start population by projecting a population based on the Census 2000 five year forward to 2005 and then adjust these such that the totals match the newest 2005 total population estimates by county.

Fertility

The number of births born to women in each cohort is calculated by multiplying the number of females in that cohort with an age specific birth rate. These numbers are than added over all cohorts and multiplied with a survival rate to get a total number of live births.

As the number of women in that age group we calculate a mid-point estimate, that is the average between the number of women in the beginning of the time interval and the number of women in the same cohorts at the end of the interval.

The age specific fertility rates are estimated by dividing an annual number of births to women from a certain age group by the estimated number of women in that age group. The annual number of births by age of the mother is derived from 1998-2001 annual detailed data we get through the New York State Department of Health.

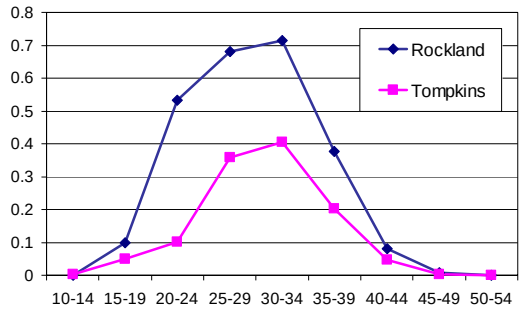

Figure 1: Examples of age specific fertility rates for 5 year age intervals

The number of women in the denominator is derived from Census 2000.

A sex-ratio determines the ratio between baby boys and baby girls.

Mortality

One of the first steps our projection model takes is projecting the portion of the current population that survives to the next moment in time. Survival probabilities depend on age, sex and location.

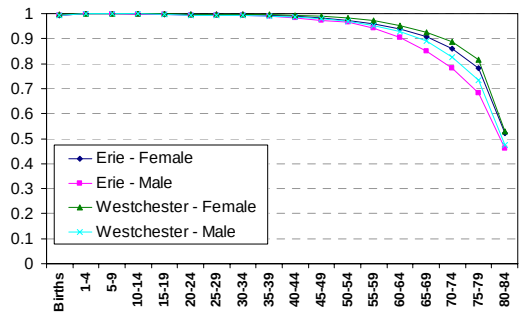

Figure 2: Examples of age specific survival rates

The survival rates are estimated using life tables that are created using detailed vital statistics from the New York State Department of Health (deaths by place of residence and single year of age in 1998-2001) and the Census 2000 estimates for population size in each of the age groups. The survival rate at Birth is derived completely from the NYSDOH vital statistics.

In the calculations of the life table (see e.g. [4] for more details on calculations of life tables) one has to make assumptions on the life expectancy in the oldest open age group. In our calculations we used the statewide average life expectancy of the 85+ age group as published by the NYSDOH [5].

We took that statewide average life expectancy for the 85+ age group and applied that to each of the county specific life tables.

The life tables were created by single year of age and as a last step abridged to accommodate 5 year age groups.

Migration

The projected net migration is the result of a number of projected flows:

- People moving into the area from elsewhere in the United States
- People moving in from outside the US
- People moving out of the county
- In and out flows of Special Populations
- The total net migration

Inflow from elsewhere in the US
The size of this flow is calculated by taking a fraction of the assumed US population in that sex/age group (outside of the county we are projecting) and increase that number if the projected Special Population in that sex/age group is growing.

The fractions are estimated from special tabulations derived from the 2000 Census question: "Where did you live 5 years ago?" The numbers in these tabulations were adjusted for fluctuations in the Special Populations between 1995 and 2000. The assumed US populations are based on U.S. Census Bureau projections [6].

People moving in from abroad
The projected flows of recent immigrants to each of the counties are calculated as shares of all projected recent immigrants to New York State.

The shares and the total State wide size of the recent immigrant flows are calculated using special tabulations derived from the 2000 Census question: "Where did you live 5 years ago?"

People moving out
The flow of people in each sex/age group that leave the County in a period is calculated as a fraction of the survivors that resided in that County at the beginning of the period. If there is a projected decrease of the Special Population in that sex/age group then the outflow is increased.

The fractions were estimated from the same special tabulations as the other migration flows.

Special Populations
In our Projection model we defined Special Populations to accommodate flows in and out of institutions that attract people from mostly outside the county and mostly for a limited time. Examples are people going in, staying at and leaving Colleges & Universities, Federal & State Prisons, Military posts and State long term care hospitals.

We estimated all migration rates and such over the whole population, thus including the special populations. But fluctuations in the size of these special populations in the past have an impact on these rates. We corrected the estimated rates so that they reflect a constant level of Special Populations. The projection model handles fluctuations in the level of Special Populations separately.

We limited the Special Populations to Special Populations in counties where fluctuations would noticeably impact the overall population. Appendix B contains the institutions we recognized for our projections. Sex and age characteristics for these Special Populations were derived from the age/sex characteristics of the areas where these populations resided at Census 2000. Historic population counts (1990, 1995, 2000 and 2005) were derived from the annual Group Quarter Reports the Program on Applied Demographics prepares for the Census Bureau.

Total Net Migration
The previous sections described how each of the flows is determined. They are all by sex and age group. Net flows of migration are calculated and controlled for assumed levels of net migration. This controlling is done with

a method called plus-minus method and is described in [1].

Historic levels of Net Migration are derived through residual methods. Starting in 1990 we worked with population estimates for 1995, 2000 and 2005 and by correcting for natural increase and one time changes in Special Populations we estimated historic Net Migration flows.

These historic flows are expressed as percentages of the population at the beginning. The three 5-year rates were averaged to get the rate we used for the projections.

In our current model we estimate the historic net migration rate for the total population but we control the female and the male migration separately using the same control rates.

Deviations from estimated historic rates

Migration into Cortland County
The special migration tables produced by the Census Bureau contain erratic number for the inflows of males into Cortland County.

We looked for a county that might show a similar migration pattern and found that the neighboring Otsego County has a very similar female in migration pattern. We then assumed the male in migration pattern is similar to that the Otsego male in migration pattern.

Rockland County
A special study was undertaken by Rockland County to come up with better estimates and assumptions using local data and knowledge. The results of this study are used in our projections.

References

[1] Smith, S.K., Tayman, J., & Swanson, D.A. (2001). State and local population projections: Methodology and analysis. New York: Kluwer Academic/Plenum Publishers.

[2] U.S. Census Bureau. County population estimates by age and sex http://www.census.gov/popest/counties/asrh/CC-EST2006-alldata.html

[3] U.S. Census Bureau. Post census estimates http://www.census.gov/popest/counties/

[4] NYSDOH life expectancy http://www.health.state.ny.us/nysdoh/vital_statistics/2000/table03.htm

[5] Hinde, A (1998). Demographic Methods. London: Arnold.

[6] U.S. Interim Projections by Age, Sex, Race, and Hispanic Origin: 2000-2050
web site: http://www.census.gov/ipc/www/usinterimproj/
data table: http://www.census.gov/ipc/www/usinterimproj/usproj2000-2050.xls

Appendix A

Formula

General

$Pop_t = Pop_{t-1} + Births – Deaths + Net\ Migration$

Fertility

Births module:

$$BIRTHS_{total} = \sum birthrate_{age} * FEMALES_{age}$$

$$BIRTHS_{sex} = \frac{sexratio_{sex}}{\left(sexratio_{male} + sexratio_{female}\right)} * BIRTHS_{total}, \text{ where:}$$

$$FEMALES_{age} = \frac{\left(FEMALES^{t}_{age} + FEMALES^{t+1}_{age+1}\right)}{2}$$

Coefficients estimation:

$$sexratio_{male} = 1.05, sexratio_{female} = 1$$

$$birthrate_{age} = \frac{BIRTHS^{C2000}}{FEMALES^{C2000}_{age}}, \text{ where}$$

$$BIRTHS^{C2000} = \frac{BIRTHS^{1998} + 4 * BIRTHS^{1999} + 4 * BIRTHS^{2000} + 3 * BIRTHS^{2001}}{12}$$

Mortality

Deaths module:

$$DEATHS = \left(1 - surv_{sex,0}\right) * BIRTHS_{sex} + \sum (1 - surv_{sex,age}) * POP^{t}_{sex,age}$$

Coefficients estimation:

$$surv_{sex,0} = 1 - \frac{DEATHS^{C2000}_{sex,0}}{BIRTHS^{C2000}_{sex}}$$

$surv_{sex,age}$ coefficients are derived from life tables constructed starting with q-type mortality rates:

$$q_{sex,age} = \frac{DEATHS_{sex,age}^{C2000}}{POP_{sex,age}^{C2000}}, \text{ where}$$

$$DEATHS^{C2000} = \frac{DEATHS^{1998} + 4 * DEATHS^{1999} + 4 * DEATHS^{2000} + 3 * DEATHS^{2001}}{12}$$

The life tables were calculated unabridged and then abridged to the appropriate age-groups. Another assumption that had to go in the life tables and only influences the survival rates for the oldest age-group is an assumption about the life-expectancy at the open-ended age group. The unabridged life table was calculated in single years upto age 84 and had open interval 85+. The assumption is that the life expectancy at that age is equal for all Counties and can be copied from a state wide table on life-expectanc.

Migration

$$BASESP = \sum_{SP} size_{SP}^{2000}, BASESP_{sex,age} = \sum_{SP} \left(size_{SP\,sex,age}^{2000} \right)$$

$$STARTSP = \sum_{SP} size_{SP}^{t}, STARTSP_{sex,age} = \sum_{SP} \left(size_{SP}^{t} * spchar_{SP\,sex,age}^{t} \right)$$

$$ENDSP = \sum_{SP} size_{SP}^{2000}, ENDSP_{sex,age} = \sum_{SP} \left(size_{SP}^{t+1} * spchar_{SP\,sex,age}^{t+1} \right)$$

$$INMIGR = \sum \begin{bmatrix} inmrate_{sex,age} * \left(uspop_{sex,age} - POP_{sex,age}^{t} + BASESP_{sex,age} - STARTSP_{sex,age} \right) + \\ inmshare_{sex,age} * nyimmigration_{sex,age} + \\ \max\left(0, ENDSP_{sex,age} - BASESP_{sex,age}\right) - \max\left(0, STARTSP_{sex,age-1} - BASESP_{sex,age-1}\right) \end{bmatrix}$$

$$OUTMIGR = \sum \begin{bmatrix} outmrate * \left(POP_{sex,age}^{t} - DEATHS_{sex,age} + BASESP_{sex,age} - STARTSP_{sex,age} \right) + \\ \max\left(0, STARTSP_{sex,age-1} - BASESP_{sex,age-1}\right) - \max\left(0, ENDSP_{sex,age} - BASESP_{sex,age}\right) \end{bmatrix}$$

$$NETMIGRATION = ADJIN * INMIGR - ADJOUT * OUTMIGR,$$

with *ADJIN* and *ADJOUT* such that :

$$NETMIGRATION = netrate * \left(POP^{t} + BASESP - STARTSP \right) + \left(ENDSP - BASESP \right)$$

Example to explain Special Population calculations:

Assume that *inmrate* and *outmrate* were based on an institution with 1500 male age 15-19 and 300 male age 20-24 (total 1800). Suppose further that we project for a certain projection period that this institution has 2100 residents at the beginning of the period and 1950 at the end. We keep the characteristics constant. In this example we calculate the Special Population corrections for male 20-24 year old at the end of the period. At the beginning of the projection period this cohort was 15-19 and is projected to be 1750 in

size (1500/1800 * 2100) and in the end 325 (300/1800 * 1950). The 1750 is 250 more than the 1500 in the base and we add those as extra's to the out-migration, the 325 is 25 more then the base and those 25 are added to the immigration, so the corrected net migration is 225 (25-250) lower then the uncorrected net migration. Another way to arrive at this total correction is: with constant base counts the net migration would be -1200 (300 – 1500), with the projected counts the net migration is 325 – 1750 = -1425, which is indeed 225 lower.

Coefficients estimation:

The coefficients are estimated solving he same formula for the coefficients using data from a special Census Bureau tabulation to fill in values for the migration flows.

The size of the Special Populations comes from the annual Group Quarter report the Program on Applied Demographics prepares for the U.S. Census Bureau. We gather that information through interviews and overseeing state agencies. The size of the military installations is estimated from the Census counts, where we took the corresponding CDP as the size of that installation. Characteristics are estimated using characteristics of the blocks where the institutions are located and where most of the residents are in that Group Quarters. There is one institution that was build and occupied since 2000. We derived its characteristics from the yearly "Profile of Inmate Population Under Custody" prepared by the New York state Department of Correctional Services.

Appendix B

Special Populations

The table underneath lists all the institutions that were used to adjust net migration rates in case there were fluctuations at that institution. A 'x' in the Sex/age column indicates that that we estimated sex and age characteristics for the residents of that institution and that detailed immigration and outmigration rates are adjusted for count variations in the past and in the projected future.

County	Institution	Sex/age
Albany	College Of Saint Rose	x
	Siena College	x
	State University Of New York-Albany	x
	Capitol District Psychiatric Center	
Allegany	Alfred University	x
	Houghton College - Houghton Campus	x
	Suny Ag Tech College-Alfred	x
Cayuga	Auburn Correctional Facility	x
	Cayuga Correctional Facility	x
Chemung	Elmira College	x
	Elmira Correctional Facility	x
	Southport Correctional Facility	x
Clinton	Altona Correctional Facility	x
	Clinton Correctional Facility	x
	State University Of New York-Plattsburgh	x
	Plattsburgh Air Force Base	
Cortland	State University College-Cortland	x
Dutchess	Bard College	x
	Downstate Correctional Facility	x
	Fishkill Correctional Facility	x
	Green Haven Correctional Facility	x
	Marist College	x
	Vassar College	x
	Harlem Valley State Hospital	
	Taconic//Wassaic Developmental Center	
	Hudson River State Hospital	
Essex	Adirondack Correctional Facility	x
	Moriah Shock Incarceration Correctional Facility	x
	Ray Brook Federal Correctional Institution (Ray Brook Fci)	x

Franklin	Camp Gabriels	x
	Franklin Correctional Facility	x
	Upstate Correctional Facility + Bare Hill Correctional Facility	x
	Sunmount Developmental Center	
Greene	Coxsackie Correctional Facility	x
	Greene Correctional Facility	x
Jefferson	Cape Vincent Corr Facility	x
	Fort Drum	x
	Watertown Correctional Facility	x
Livingston	Groveland Correctional Facility For Men + Livingston Corr Facility	x
	State University College-Geneseo	x
Madison	Cazenovia College	x
	Colgate University	x
	Suny Ag Tech College-Morrisville	x
Oneida	Hamilton College	x
	Institute Of Technology At Utica/Rome	x
	Marcy Correctional Facility	x
	Mid-State Correctional Facility	x
	Oneida Correctional Facility + Mohawk Correctional Facility	x
	Utica College	x
	Mohawk Valley Psych Center	
	Griffiss Air Force Base	
Orleans	Albion Fem Corr Facility	x
	Orleans Corr Facility	x
Otsego	Hartwick College	x
	State University College-Oneonta	x
St. Lawrence	Clarkson College	x
	Gouverneur Corr Facility	x
	Ogdensburg Correctional Facility	x
	Riverview Corr Facility	x
	Saint Lawrence University	x
	State University Of New York-Potsdam	x
	Suny Ag Tech College-Canton	x
	Saint Lawrence Psychiatric Center	
Schoharie	Suny Ag Tech College-Cobleskill	x
Seneca	Willard Drug Treatment Campus	x
	Five Points Correctional Facility	x
	Willard State Hospital	

Suffolk	Long Island Developmental Center//Suffolk Developmental Center	
	Central Islip State Hospital	
	Pilgrim Psychiatric Center	
	Kings Park State Hospital	
Tompkins	Cornell University	x
	Ithaca College	x
Ulster	Eastern Correctional Facility	x
	Shawangunk Correctional Facility	x
	State University College-New Paltz	x
	Ulster Corr Facility	x
	Wallkill Correctional Facility	x
Washington	Great Meadow Correctional Facility	x
	Washington Correctional Facility	x
Wyoming	Attica Correctional Facility	x
	Wyoming Correctional Facility	x

Appendix C

Model version history

Version 20080529

This is the first published version and contains the algorithms and coefficients as described in this document.

Version 200806nn

Cortland County male out migration
The out migration rates for men age 20-40 have unexplained high values. We knew from earlier analysis that the special migration tabulations contain an error for Cortland county. The female out-migration rates from Cortland County are very similar to the female rates in neighboring Otsego County. We copied the male out-migration rates from Otsego County to Cortland County and used those for our projections.

Start population
The 2005 sex/age characteristics as estimated by the US Census Bureau are not consistent with the characteristics 5 year earlier in Census 2000. As an alternative start population we 'projected' the 2005 population by using this same projection model and the Census 2000 population as the starting population. A raking factor was used to adjust the totals to the Census Bureau's estimated totals.

Version 20081128

New US Projections
The U.S. Census Bureau published new projections for the U.S. These projections are used in the calculation of migration into each of the counties. This version includes the new projections.

Rockland County
Results from a study with Rockland County are included in the assumptions.

Version 20090402

Start Population
The Census Bureau published new county estimates (Vintage 2008). This version's crude migration rate estimates is based on the Vintage 2007 county populations by sex. The start population totals are adjusted to match the Vintage 2008 totals.