New York Metropolitan Transportation Council

The Metropolitan Planning Organization

Demographic and Socioeconomic Forecasting

Technical Memorandum
Task 1.4.5.2
Modeling Methodology at County/Subregional Level

Submitted by:
Urbanomics
October 19, 2007

NEW YORK METROPOLITAN TRANSPORTATION COUNCIL

DEMOGRAPHIC AND SOCIOECONOMIC FORECASTING

TECHNICAL MEMORANDUM NO. 1.4.5.2
MODELING METHODOLOGY AT COUNTY/ SUBREGIONAL LEVEL

This study is funded by a matching grant from the Federal Highway Administration, under NYSDOT PIN PT 1949911.

PRIME CONSULTANT: **URBANOMICS**
115 FIFTH AVENUE
2ND FLOOR
NEW YORK, NEW YORK 10003

The preparation of this report was financed in part through funds from the Federal Highway Administration (FHWA) and Federal Transit Administration (FTA). This document is disseminated under the sponsorship of the U.S. Department of Transportation in the interest of information exchange. The contents of this report reflect the views of the author who is responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the FHWA, FTA, nor of the New York Metropolitan Transportation Council. This report does not constitute a standard, specification or regulation.

TECHNICAL MEMORANDUM NO. 1.4.5.2
MODELING METHODOLOGY AT COUNTY/SUBREGIONAL LEVEL

TABLE OF CONTENTS

MAP

TABLES

FIGURES

APPENDIX

Technical Memorandum No. 1.4.5.2
Modeling Methodology at County/Subregional Level

1.1 INTRODUCTION

Background

In accordance with the approved approach, this memorandum describes the four existing forecasting models at the county and subregional levels, adapted and re-calibrated with new national drivers of Global Insight, Inc. (GI) and updated time series on population, employment and earnings, labor force and housing/household formation data. In addition, two new sub-models will be developed: an enrollment model that forecasts primary, secondary and university enrollment by subregion of residence and institutional location; and an occupational model that converts industry forecasts to occupational forecasts. The models will be capable of producing forecasts at five-year increments through the year 2050.

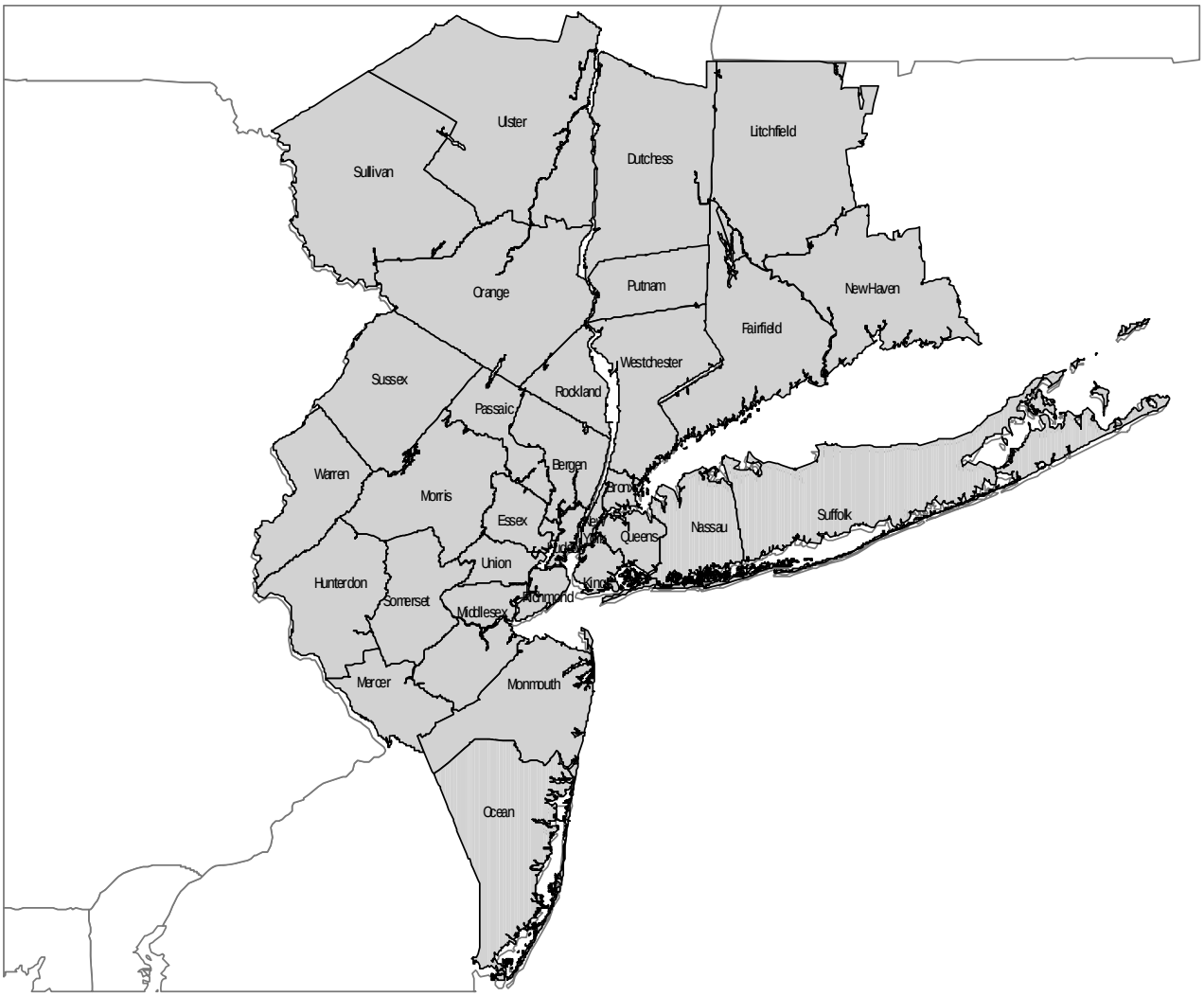
This memorandum describes the models in sequence, starting with the Population Model and concluding with the Housing/Household Formation Model. With the exception of new model enhancements, the methodologies described here represent a recalibration and extension of prior work performed under the Transportation Models and Data Initiative (TMDI) and the Demographic and Socioeconomic Forecasting (DSF) projects. The former program undertaken by the New York Metropolitan Transportation Council (NYMTC) forecasted transportation needs of the New York Metropolitan Region[1] through the year 2020. The latter program forecasted demographic and socioeconomic inputs of the Best Practices Model (BPM) to 2025 and 2030, with the most recent forecasts presented in TM 4.1.2*, County Level Demographic and Socioeconomic Forecast Extensions, 2025- 2030.*[2]. The current TM describes the enhanced methodology for forecasting demographic and socioeconomic trends to 2035 at the county and subregional level, on a five-year interval basis. The geographic area for which county and subregional forecasts will be prepared is shown in Map 1.

Population, employment, labor force, and household formation forecasts are necessary for transportation modeling because the size and distribution of such demographic trends directly affect the journey-to-work trip patterns. Journey-to-work trips account for a large proportion of all travel within the Region, especially peak hour commuting trips, which place the greatest demand on transportation infrastructure.

[1] The New York Metro Region includes the following counties, by subregion: New York City subregion: Bronx, Kings, New York, Queens, Richmond Counties; Long Island subregion: Nassau & Suffolk Counties; Mid-Hudson subregion: Dutchess, Orange, Putnam, Rockland, Sullivan, Ulster, Westchester Counties; New Jersey subregion: Bergen, Essex, Hudson, Hunterdon, Mercer, Middlesex, Monmouth, Morris, Ocean, Passaic, Somerset, Sussex, Union, Warren Counties; Connecticut subregion: Fairfield, Litchfield, New Haven Counties.

[2] Technical Memorandum 4.1.2, County Level Demographic and Socioeconomic Forecast Extensions, 2025-2030. March 18, 2004

Map 1. Thirty One County New York Metropolitan Region



Changes from NYMTC 2030 TM 4.1.2

The 2030 NYMTC long term demographic and socioeconomic forecasts were presented at the county and subregional level in *TM 4.1.2, County Level Demographic and Socioeconomic Forecast Extensions, 2025-2030.*[3] The methodology was based upon the approach that generated 2025 forecasts, as documented in TMs *1.2.1, Population Model, 1.2.2, Employment Model., 1.2.3, Labor Force Model, and 1.2.4, Household Model.*[4] However, several significant changes were made pertaining to the Employment Model. Notably, between the 2025 and 2030 forecast series, the federal

[3] Technical Memorandum 4.1.2, County Level Demographic and Socioeconomic Forecast Extensions, 2025-2030. March 18, 2004

[4] Technical Memorandum 1.2.1, Population Model. May 16, 2001; Technical Memorandum 1.2.2, Employment Model. December 5, 2001; Technical Memorandum 1.2.3, Labor Force Model. December 2, 2001; Technical Memorandum 1.2.4, Household Model. December 21, 2001;

system of industrial classification was substantially revised from the *SIC (Standard Industrial Classification)* to the *NAICS (North American Industrial Classification System)* basis. In addition to the change in industry nomenclature, consistent regional data were no longer available on an annual basis from 1970, requiring quarterly data from 1990 to model historical employment and earnings. While quarterly *NAICS* data were sufficient at the subregional level, far less industry detail was available at the county level, a situation that persists to the present day. County forecasts were thus prepared for private sector, government, proprietors and total employment.

National drivers for the 2030 forecasts were supplied by Global Insight, Inc., from the summer 2003 Long Term Trend forecast of the U.S. economy to 2027. TM 4.1.1*, Comparison of National Driver Assumptions, DRI 2000-2024 with GI 2003- 2027*, presented a comparison of national driver assumptions between GI's Summer 2003 series and the Summer 1999 series of Data Resources Inc. (DRI) which was utilized to drive NYMTC's 2025 Forecasts.[5] For the 2035 NYMTC forecasts, the national drivers of Global Insight's March 2007 series will be employed. It should be noted that, just as significant differences existed in national growth rates between GI 2003 and DRI 1999, changes can be expected in the robust character of the new GI 2007 long term assumptions.

Overview

Section 1.2, below, describes the Population Model, including all model inputs, the model's structure and outputs, and the current model enhancements. These enhancements include cohort-survival models for all NYMTC partner counties[6] and enrollment forecasts for primary, secondary and university enrollment. Section 1.3 offers a detailed discussion of the Employment Model that forecasts employment and earnings by sector, including a model enhancement that forecasts industry employment by occupational sector. Section 1.4 describes the Labor Force Model and Section 1.5 discusses the Housing/Household Formation Model. Enhancements to the Labor Force Model provide detailed demographics of the civilian labor force for all NYMTC partner counties. Lastly, Section 1.6 identifies Task 1.4.5.2 work products. All model forecasts and analyses will be presented in subsequent technical memorandum.

- Task 5.1.1 Summary and Analysis of Existing Trends and Key Factors for Forecasting
- Task 5.2.1 NYMTC County Level Population and Labor Force Models, and TAZ Modeling Methodology
- Task 5.2.2 Model Estimation, Calibration, Validation and Testing
- Task 5.2.3 Model Interoperability and Interface with GIS Databases
- Task 5.3.1 Regional, Subregional, County and TAZ Level Forecasts of Socioeconomic Variables from 2005 to 2035
- Task 5.3.2 Comparison with Other Forecasts

[5] Technical Memorandum 4.1.1, Comparison of National Driver Assumptions, DRI 2000-2024 with GI 2003-2027. May 18, 2004

[6] The NYMTC territory includes the following counties, by subregion: New York City subregion: Bronx, Kings, New York, Queens, & Richmond Counties; Long Island subregion: Nassau & Suffolk Counties; Mid-Hudson subregion: Putnam, Rockland & Westchester Counties.

- Task 5.3.3 Model Training and Users Manual
- Task 5.4.1 Meetings and In-Person Presentations for Review and Consensus on Forecasts
- Task 5.4.2 Project Website and Web Applications

1.2 POPULATION MODEL

The Population Model is based on the cohort-survival technique, a standard population forecasting methodology. In this method, the population is broken into separate age/sex-specific cohorts. A separate projection is made for every cohort for each time interval, based on the cohort's population at the beginning of the interval, its forecasted survival rate during the interval, and any net in- or out-migration over the interval. Population growth due to births for each interval is calculated based on forecasted age-specific fertility rates for the female population. The cohort-survival method provides a reliable measure of population change due to natural increase (i.e., births and deaths) and estimates the effects of net-migration. The Population Model accounts for the net-migration component as a function of historical rates and the future demand of employment for labor force. The model's structure is discussed in detail in Section 1.2.2 below.

The Population Model generates population forecasts by sex and five-year age cohort through 84 years-of-age, as well as for persons 85 years-of-age and over. A separate model will be run for each subregion and for each partner county of the NYMTC territory. Within each subregional model, submodels generate separate outputs for each mutually-exclusive racial/ethnic group (non-Hispanic White, non-Hispanic Black, non-Hispanic Asian/Other, and Hispanic). Each model includes an historical section, covering the years 1970 through 2005, and a forecast section, covering the years 2005 through 2035. All outputs are generated on a five-year interval basis. A separate forecasting routine disaggregates the subregional population forecasts to the county level for the non-partner counties of the NYMTC study area.[7]

The Population Model uses US Census Bureau historical population figures as a basis for all projections, including the interim population estimates by age/sex/race-ethnicity.[8] Forecasted survival and fertility rates will be developed in consultation with Census Bureau demographers, based upon national-level Census Bureau forecasts adjusted for subregional differences. The forecasts of births and deaths will be controlled by an adjustment factor that reflects the difference between actual historical births and deaths, on a racial/ethnic-specific basis, from state Department of Health (DOH) sources, and predicted births and deaths based on population and fertility figures from the Census Bureau historical series. Net migration forecasts will be based on historical levels forecasted in relation to natural increase, taking into consideration national policies and outlooks on immigration and the constraint imposed by regional employment and labor force forecasts. All Population Model inputs are discussed in detail in Section 1.2.1 below.

[7] For these 21 non-partner counties, extensive use was not made of county-level input data, other than to control for estimated births and deaths at the subregional level by aggregated county vital statistics. County level survival and fertility rates exist only for selected historical years and have not been forecasted, as have national rates of survival and fertility have been forecasted.

[8] United States Census Bureau, www.census.gov, Accessed Date March 2007.

1.2.1 DATA INPUTS

The Population Model incorporates a number of inputs at each five-year interval:

- Historical population figures by racial/ethnic group, sex and five-year age cohort.
- Estimated year 2005 population figures by racial/ethnic group, sex and five-year age cohort.
- Historical and forecasted survival rates by racial/ethnic group, sex and five-year age cohort.
- Historical and forecasted fertility rates by racial/ethnic group and five-year age cohort of women in reproductive ages.
- Historical births and deaths, by racial/ethnic group where available.
- Labor force net migration, by racial/ethnic group, sex and age.

This section discusses the various inputs to each model, gives their sources, and where relevant discusses any adjustments to source data.

As discussed above, each subregional and county model incorporates submodels for each racial/ethnic group. Where necessary, source data were disaggregated by racial/ethnic group within the model for incorporation into each racial/ethnic submodel. For the sake of convenience, this disaggregation process is described together with the discussion of each input, below.

Historical Population Data

The Population Model uses 2005 interim Census estimates as a base for all forecasts.[9] In addition, decennial Census population data for 1970 to 2000 are incorporated as inputs in the historical section where they are used in the calculation of control factors used in births and deaths forecasting. Data for 2000 come from the Census's Modified Race Data Summary File (MRS), while data for 1980 and 1990 come from the Modified Age/Race/Sex (MARS) data set; and data for 1970 from the Summary Tap File (STF) data set. All Census data are aggregated by sex and five-year age cohort through age 84 with an additional cohort including persons 85 years of age and older. MRS and MARS data are reported by the racial/ethnic groupings used in this study; data for 1970 were adjusted to fit these groupings. All figures are aggregated at the subregional level.

For intercensal years 1975 and 1985, the Census Bureau did not release age-, sex- and race-specific figures; for intercensal year 1995, such data were released but judged to be inaccurate upon enumeration of the 2000 Census. Therefore, for these years, figures were interpolated from the preceding and following decennial Census years.

[9] United States Census Bureau, www.census.gov, Accessed Date March 2007

Estimated 2005 Population

Age/sex/racial-ethnic population estimates are compiled for the year 2005 from the Census Bureau's Estimates program, released annually on the Bureau's website as: Estimates of the Population of Counties by Age, Sex, and Race/Hispanic Origin. The following Table 1 compares the regional summary of all races for 2005 with the population in 2000, on the July 1 basis used in cohort-survival forecasting models. For the NYMTC partner counties and subregions, Census estimates are shown in comparison with 2000 data on an age/sex/racial-ethnic basis in Appendix Table A.1.1 through A.1.1.16.

Table 1. Age/Sex Structure of All Races Population of the New York Metropolitan Region in 2000 & 2005

MALE ALL RACES	**2000**	**2005**	**Change**	**Percent Change**
POPULATION	**ESTIMATE**	**ESTIMATE**		
Under 5	743,303	766,837	23,534	3.17%
'5 - 9	787,431	731,559	-55,872	-7.10%
10 - 14	762,732	778,540	15,808	2.07%
15 - 19	702,722	748,709	45,987	6.54%
20 - 24	668,270	693,479	25,209	3.77%
25 - 29	734,281	700,752	-33,529	-4.57%
30 - 34	833,241	750,109	-83,132	-9.98%
35 - 39	898,212	835,584	-62,628	-6.97%
40 - 44	854,088	890,766	36,678	4.29%
45 - 49	737,544	837,565	100,021	13.56%
50 - 54	657,204	712,404	55,200	8.40%
55 - 59	500,016	619,810	119,794	23.96%
60 - 64	395,220	457,760	62,540	15.82%
65 - 69	326,859	332,263	5,404	1.65%
70 - 74	292,626	264,766	-27,860	-9.52%
75 - 79	228,316	236,021	7,705	3.37%
80 - 84	139,529	165,284	25,755	18.46%
85 & Over	99,156	130,845	31,689	31.96%
Total	10,360,750	10,653,053	292,303	2.82%
FEMALE ALL RACES	**2000**	**2005**	**Change**	**Percent Change**
POPULATION	**ESTIMATE**	**ESTIMATE**		
Under 5	710,198	734,740	24,542	3.46%
'5 - 9	751,424	699,641	-51,783	-6.89%
10 - 14	727,547	742,599	15,052	2.07%
15 - 19	662,064	715,951	53,887	8.14%
20 - 24	670,417	670,347	-70	-0.01%
25 - 29	762,199	692,704	-69,495	-9.12%
30 - 34	865,113	773,932	-91,181	-10.54%
35 - 39	931,890	865,241	-66,649	-7.15%
40 - 44	893,788	919,414	25,626	2.87%
45 - 49	792,316	879,469	87,153	11.00%
50 - 54	731,876	772,734	40,858	5.58%
55 - 59	570,499	700,942	130,443	22.86%
60 - 64	459,249	535,452	76,203	16.59%
65 - 69	403,909	410,578	6,669	1.65%
70 - 74	393,642	353,952	-39,690	-10.08%

75 - 79	347,311	339,511	-7,800	-2.25%
80 - 84	250,124	277,068	26,944	10.77%
85 & Over	250,371	295,789	45,418	18.14%
Total	11,173,937	11,380,064	206,127	1.84%
TOTAL POPULATION	21,534,687	22,0331,17	498,430	2.31%

Source: United States Census Bureau, Interim Population Estimates

Prior to 2000, the Census series provided aggregate county level population by four race categories (White, Black, American Indian and Alaska Native, Asian and Pacific Islander) cross-tabulated by two ethnicity categories (Hispanic, non-Hispanic), yielding eight racial-ethnic population estimates by county. With the 2000 Census, and subsequently with annual estimates, the Bureau provided race and Hispanic origin data by age and sex on a more detailed basis. In addition to five race categories now identified (White, Black, American Indian and Alaska Native, Asian, and Native Hawaiian and Other Pacific Islander), respondents were given the option to mark one or more races as their racial identity, or to indicate "Some other race" if none qualified.[10]

Cross-tabulated by two ethnicity categories, the multiplicity of racial choices resulted in 66 age/sex/race-ethnicity categories that required consolidation to four mutually-exclusive racial/ethnic groups used in forecasting. This was performed by segregating the difference between a non-Hispanic age/sex cohort of a racial category on a solitary basis and in combination with another race, that difference then being normalized by a control across all races for non-Hispanics of two or more racial groups in the age/sex cohort. Hispanics by race alone or two or more races were taken as reported. These procedures were followed for each subregion and NYMTC partner county in 2000 and 2005.

Survival Rates

Table 2 shows historical and forecasted survival rates used as inputs to the population model. For each five-year age group the model requires inputs of race/sex-specific rates. For any given population group, the survival rate is defined as the percentage of persons alive at the beginning of a time interval who survive to the end of that interval. Survival rates are typically presented in life tables, which show the attrition of a hypothetical population cohort (usually of 100,000 persons) on an annual basis given the age-specific survival rates for a given year. While complete life tables show survival rates by single year-of-age, abridged life tables present comparable figures for five-year age groups. The abridged data of Table 2 represent the latest available survival rates of the Census Bureau by age/sex/race-ethnicity. New rates will likely be available with the Bureau's release of revised long term population forecasts for the nation, in Summer 2007.

[10] Those that responded to the latter option were primarily of Hispanic origin.

Table 2. National Level Survival Rates by Age/Sex/Race-Ethnicity, for 1992-2032 Forecast Period

White Population:

Age	1992	1997	2002	2007	2012	2017	2022	2027	2032
White Males									
Births	0.99257	0.99381	0.99405	0.99428	0.99450	0.99473	0.99496	0.99519	0.99542
Under 5	0.99832	0.99858	0.99874	0.99879	0.99884	0.99889	0.99894	0.99899	0.99904
5 to 9	0.99891	0.99907	0.99910	0.99914	0.99918	0.99921	0.99925	0.99929	0.99932
10 to 14	0.99649	0.99675	0.99745	0.99754	0.99763	0.99772	0.99780	0.99789	0.99798
15 to 19	0.99348	0.99369	0.99454	0.99472	0.99490	0.99508	0.99526	0.99544	0.99562
20 to 24	0.99313	0.99305	0.99467	0.99484	0.99501	0.99518	0.99534	0.99551	0.99568
25 to 29	0.99262	0.99146	0.99419	0.99437	0.99455	0.99473	0.99490	0.99508	0.99526
30 to 34	0.99000	0.98769	0.99212	0.99237	0.99263	0.99288	0.99313	0.99339	0.99364
35 to 39	0.98669	0.98520	0.98934	0.98970	0.99006	0.99042	0.99077	0.99113	0.99149
40 to 44	0.98273	0.98282	0.98483	0.98536	0.98588	0.98641	0.98693	0.98746	0.98798
45 to 49	0.97441	0.97646	0.97694	0.97774	0.97854	0.97935	0.98015	0.98095	0.98176
50 to 54	0.95873	0.96272	0.96356	0.96483	0.96610	0.96736	0.96863	0.96990	0.97117
55 to 59	0.93329	0.93879	0.94194	0.94395	0.94597	0.94799	0.95000	0.95202	0.95404
60 to 64	0.89747	0.90639	0.90740	0.91042	0.91344	0.91645	0.91947	0.92249	0.92550
65 to 69	0.84726	0.86085	0.86011	0.86446	0.86881	0.87316	0.87751	0.88186	0.88621
70 to 74	0.77186	0.78733	0.78863	0.79513	0.80164	0.80815	0.81466	0.82116	0.82767
75 to 80	0.66673	0.67995	0.68806	0.69743	0.70680	0.71617	0.72555	0.73492	0.74429
80 to 84	0.53474	0.54848	0.56795	0.58048	0.59300	0.60553	0.61806	0.63058	0.64311
85+	0.42882	0.44975	0.47684	0.49607	0.51531	0.53454	0.55377	0.57301	0.59224
White Females									
Births	0.99389	0.99495	0.99531	0.99549	0.99566	0.99584	0.99601	0.99619	0.99636
Under 5	0.99872	0.99891	0.99903	0.99906	0.99910	0.99914	0.99918	0.99921	0.99925
5 to 9	0.99928	0.99937	0.99932	0.99934	0.99937	0.99940	0.99943	0.99946	0.99948
10 to 14	0.99839	0.99845	0.99858	0.99863	0.99868	0.99873	0.99879	0.99884	0.99889
15 to 19	0.99762	0.99776	0.99781	0.99789	0.99797	0.99804	0.99812	0.99820	0.99827
20 to 24	0.99762	0.99777	0.99776	0.99784	0.99791	0.99799	0.99807	0.99814	0.99822
25 to 29	0.99718	0.99724	0.99720	0.99730	0.99739	0.99749	0.99758	0.99768	0.99777
30 to 34	0.99614	0.99628	0.99605	0.99619	0.99633	0.99646	0.99660	0.99674	0.99687
35 to 39	0.99434	0.99482	0.99422	0.99442	0.99463	0.99483	0.99503	0.99524	0.99544
40 to 44	0.99110	0.99195	0.99124	0.99154	0.99185	0.99215	0.99246	0.99276	0.99307
45 to 49	0.98548	0.98649	0.98606	0.98654	0.98701	0.98749	0.98797	0.98845	0.98893
50 to 54	0.97555	0.97672	0.97696	0.97776	0.97856	0.97936	0.98016	0.98096	0.98176
55 to 59	0.95991	0.96121	0.96319	0.96448	0.96576	0.96705	0.96834	0.96963	0.97091
60 to 64	0.93893	0.94060	0.94186	0.94377	0.94568	0.94759	0.94950	0.95141	0.95332
65 to 69	0.90832	0.91102	0.91091	0.91372	0.91653	0.91934	0.92215	0.92496	0.92777
70 to 74	0.86009	0.86472	0.86241	0.86675	0.87109	0.87543	0.87977	0.88411	0.88845
75 to 80	0.78383	0.79192	0.78654	0.79317	0.79980	0.80644	0.81307	0.81970	0.82634
80 to 84	0.66701	0.68069	0.67318	0.68305	0.69292	0.70279	0.71266	0.72253	0.73240
85+	0.57178	0.59488	0.58607	0.60325	0.62044	0.63762	0.65481	0.67199	0.68918

Black Population:

Age	1992	1997	2002	2007	2012	2017	2022	2027	2032
Black Males									
Births	0.98158	0.98314	0.98522	0.98588	0.98654	0.98719	0.98785	0.98850	0.98916
Under 5	0.99699	0.99716	0.99757	0.99768	0.99778	0.99789	0.99800	0.99811	0.99821
5 to 9	0.99857	0.99865	0.99858	0.99864	0.99870	0.99876	0.99882	0.99889	0.99895
10 to 14	0.99345	0.99226	0.99563	0.99580	0.99598	0.99615	0.99633	0.99650	0.99668
15 to 19	0.98544	0.98227	0.98890	0.98934	0.98979	0.99023	0.99068	0.99112	0.99157
20 to 24	0.98378	0.98095	0.98825	0.98872	0.98919	0.98965	0.99012	0.99058	0.99105
25 to 29	0.98248	0.97813	0.98729	0.98778	0.98827	0.98877	0.98926	0.98976	0.99025
30 to 34	0.97601	0.96852	0.98280	0.98348	0.98417	0.98485	0.98553	0.98621	0.98689
35 to 39	0.96881	0.96195	0.97707	0.97799	0.97891	0.97984	0.98076	0.98168	0.98261
40 to 44	0.95997	0.95483	0.96680	0.96817	0.96954	0.97090	0.97227	0.97364	0.97500
45 to 49	0.94635	0.94537	0.95205	0.95402	0.95600	0.95797	0.95994	0.96192	0.96389
50 to 54	0.92225	0.92583	0.93185	0.93460	0.93735	0.94011	0.94286	0.94562	0.94837
55 to 59	0.88723	0.89224	0.90447	0.90825	0.91202	0.91579	0.91956	0.92333	0.92711
60 to 64	0.84081	0.84641	0.86183	0.86692	0.87200	0.87709	0.88218	0.88727	0.89235
65 to 69	0.78478	0.79158	0.81198	0.81846	0.82494	0.83141	0.83789	0.84437	0.85085
70 to 74	0.71862	0.72383	0.74077	0.74938	0.75799	0.76660	0.77521	0.78381	0.79242
75 to 80	0.63821	0.63519	0.66048	0.67104	0.68160	0.69215	0.70271	0.71327	0.72382
80 to 84	0.56164	0.56889	0.58836	0.60003	0.61170	0.62337	0.63503	0.64670	0.65837
85+	0.53172	0.56992	0.56270	0.57860	0.59450	0.61040	0.62630	0.64220	0.65810
Black Females									
Births	0.98439	0.98580	0.98774	0.98828	0.98882	0.98936	0.98990	0.99044	0.99097
Under 5	0.99754	0.99768	0.99800	0.99808	0.99817	0.99826	0.99835	0.99843	0.99852
5 to 9	0.99869	0.99871	0.99891	0.99896	0.99901	0.99906	0.99910	0.99915	0.99920
10 to 14	0.99783	0.99789	0.99822	0.99829	0.99836	0.99843	0.99850	0.99857	0.99864
15 to 19	0.99625	0.99607	0.99700	0.99711	0.99723	0.99734	0.99746	0.99757	0.99769
20 to 24	0.99485	0.99444	0.99569	0.99586	0.99603	0.99621	0.99638	0.99655	0.99672
25 to 29	0.99282	0.99156	0.99375	0.99400	0.99426	0.99451	0.99477	0.99502	0.99528
30 to 34	0.98910	0.98712	0.99043	0.99083	0.99123	0.99163	0.99203	0.99243	0.99283
35 to 39	0.98460	0.98370	0.98628	0.98686	0.98743	0.98801	0.98859	0.98917	0.98975
40 to 44	0.97910	0.97949	0.98054	0.98135	0.98216	0.98296	0.98377	0.98458	0.98539
45 to 49	0.96948	0.97107	0.97137	0.97254	0.97371	0.97488	0.97604	0.97721	0.97838
50 to 54	0.95427	0.95662	0.95776	0.95945	0.96115	0.96284	0.96454	0.96623	0.96793
55 to 59	0.93105	0.93238	0.93783	0.94029	0.94275	0.94521	0.94767	0.95013	0.95259
60 to 64	0.89805	0.89755	0.90422	0.90784	0.91147	0.91509	0.91871	0.92233	0.92596
65 to 69	0.86365	0.86585	0.86824	0.87297	0.87769	0.88241	0.88714	0.89186	0.89659
70 to 74	0.81886	0.82483	0.82556	0.83153	0.83750	0.84347	0.84944	0.85541	0.86138
75 to 80	0.75373	0.75470	0.75976	0.76755	0.77534	0.78313	0.79092	0.79871	0.80650
80 to 84	0.67785	0.68407	0.68244	0.69189	0.70135	0.71081	0.72027	0.72973	0.73919
85+	0.65436	0.68004	0.65646	0.67086	0.68526	0.69966	0.71406	0.72846	0.74286

Asian Population:

Age	1992	1997	2002	2007	2012	2017	2022	2027	2032
Asian Males									
Births	0.99483	0.99569	0.99602	0.99616	0.99629	0.99643	0.99657	0.99670	0.99684
Under 5	0.99866	0.99888	0.99893	0.99897	0.99901	0.99905	0.99909	0.99913	0.99917
5 to 9	0.99940	0.99950	0.99918	0.99921	0.99924	0.99928	0.99931	0.99934	0.99937
10 to 14	0.99789	0.99803	0.99787	0.99794	0.99801	0.99808	0.99815	0.99822	0.99829
15 to 19	0.99643	0.99674	0.99609	0.99620	0.99632	0.99643	0.99654	0.99665	0.99677
20 to 24	0.99601	0.99636	0.99623	0.99633	0.99643	0.99654	0.99664	0.99674	0.99684
25 to 29	0.99496	0.99492	0.99637	0.99645	0.99654	0.99663	0.99671	0.99680	0.99689
30 to 34	0.99311	0.99263	0.99578	0.99587	0.99596	0.99605	0.99614	0.99623	0.99633
35 to 39	0.99196	0.99180	0.99378	0.99394	0.99410	0.99426	0.99442	0.99457	0.99473
40 to 44	0.98981	0.99054	0.99049	0.99076	0.99103	0.99130	0.99157	0.99184	0.99211
45 to 49	0.98332	0.98513	0.98449	0.98495	0.98541	0.98588	0.98634	0.98681	0.98727
50 to 54	0.97649	0.97899	0.97620	0.97690	0.97760	0.97830	0.97900	0.97970	0.98039
55 to 59	0.96096	0.96439	0.96341	0.96447	0.96552	0.96658	0.96764	0.96870	0.96976
60 to 64	0.93829	0.94397	0.94166	0.94316	0.94466	0.94616	0.94766	0.94916	0.95066
65 to 69	0.90319	0.91223	0.91481	0.91674	0.91868	0.92062	0.92255	0.92449	0.92643
70 to 74	0.84694	0.85796	0.87341	0.87619	0.87898	0.88177	0.88456	0.88735	0.89014
75 to 80	0.78876	0.79808	0.80669	0.81091	0.81513	0.81935	0.82357	0.82779	0.83201
80 to 84	0.70632	0.71655	0.73274	0.73826	0.74377	0.74928	0.75479	0.76030	0.76582
85+	0.67852	0.69657	0.70372	0.71365	0.72359	0.73353	0.74347	0.75341	0.76335
Asian Females									
Births	0.99666	0.99724	0.99688	0.99698	0.99708	0.99718	0.99728	0.99737	0.99747
Under 5	0.99873	0.99893	0.99909	0.99912	0.99916	0.99919	0.99923	0.99926	0.99930
5 to 9	0.99925	0.99934	0.99928	0.99931	0.99934	0.99937	0.99940	0.99943	0.99946
10 to 14	0.99900	0.99905	0.99863	0.99868	0.99873	0.99879	0.99884	0.99889	0.99894
15 to 19	0.99891	0.99899	0.99824	0.99830	0.99835	0.99841	0.99847	0.99853	0.99859
20 to 24	0.99894	0.99903	0.99826	0.99832	0.99837	0.99842	0.99848	0.99853	0.99858
25 to 29	0.99822	0.99830	0.99841	0.99845	0.99849	0.99852	0.99856	0.99860	0.99864
30 to 34	0.99650	0.99671	0.99752	0.99759	0.99765	0.99772	0.99779	0.99786	0.99793
35 to 39	0.99457	0.99509	0.99608	0.99620	0.99631	0.99643	0.99655	0.99666	0.99678
40 to 44	0.99327	0.99394	0.99426	0.99443	0.99460	0.99476	0.99493	0.99509	0.99526
45 to 49	0.99149	0.99221	0.99155	0.99177	0.99200	0.99223	0.99245	0.99268	0.99290
50 to 54	0.98195	0.98304	0.98639	0.98675	0.98711	0.98747	0.98784	0.98820	0.98856
55 to 59	0.97269	0.97385	0.98003	0.98054	0.98105	0.98156	0.98207	0.98258	0.98309
60 to 64	0.96418	0.96536	0.96957	0.97025	0.97093	0.97160	0.97228	0.97296	0.97364
65 to 69	0.94749	0.94926	0.95333	0.95427	0.95521	0.95615	0.95710	0.95804	0.95898
70 to 74	0.91530	0.91838	0.92603	0.92753	0.92904	0.93054	0.93204	0.93355	0.93505
75 to 80	0.86913	0.87444	0.87997	0.88245	0.88493	0.88741	0.88989	0.89236	0.89484
80 to 84	0.80186	0.81095	0.81262	0.81637	0.82012	0.82388	0.82763	0.83139	0.83514
85+	0.80114	0.81879	0.80785	0.81556	0.82328	0.83100	0.83871	0.84643	0.85414

Hispanic Population:

Age	1992	1997	2002	2007	2012	2017	2022	2027	2032
Hispanic Males									
Births	0.9911	0.9926	0.9924	0.9927	0.9930	0.9933	0.9936	0.99396	0.99427
Under 5	0.9981	0.9982	0.9984	0.9985	0.9985	0.9986	0.9987	0.99873	0.99879
5 to 9	0.9991	0.9992	0.9991	0.9991	0.9992	0.9992	0.9992	0.99928	0.99931
10 to 14	0.9957	0.9961	0.9970	0.9971	0.9972	0.9974	0.9975	0.99757	0.99768
15 to 19	0.9917	0.9915	0.9934	0.9937	0.9939	0.9941	0.9944	0.99459	0.99483
20 to 24	0.9904	0.9895	0.9923	0.9926	0.9929	0.9931	0.9934	0.99370	0.99398
25 to 29	0.9888	0.9859	0.9915	0.9918	0.9921	0.9924	0.9927	0.99300	0.99331
30 to 34	0.9864	0.9814	0.9896	0.9900	0.9903	0.9907	0.9911	0.99146	0.99183
35 to 39	0.9828	0.9788	0.9878	0.9882	0.9887	0.9891	0.9895	0.98995	0.99038
40 to 44	0.9805	0.9780	0.9854	0.9859	0.9864	0.9869	0.9874	0.98789	0.98838
45 to 49	0.9737	0.9739	0.9792	0.9799	0.9806	0.9813	0.9820	0.98270	0.98339
50 to 54	0.9633	0.9658	0.9690	0.9701	0.9711	0.9721	0.9731	0.97413	0.97514
55 to 59	0.9461	0.9493	0.9529	0.9544	0.9560	0.9575	0.9590	0.96052	0.96204
60 to 64	0.9211	0.9269	0.9296	0.9316	0.9337	0.9357	0.9377	0.93975	0.94178
65 to 69	0.8864	0.8955	0.8973	0.9000	0.9027	0.9054	0.9081	0.91079	0.91350
70 to 74	0.8315	0.8423	0.8520	0.8557	0.8595	0.8632	0.8669	0.87063	0.87435
75 to 80	0.7528	0.7621	0.7838	0.7890	0.7942	0.7995	0.8047	0.80991	0.81513
80 to 84	0.6558	0.6664	0.6922	0.6994	0.7066	0.7139	0.7211	0.72836	0.73560
85+	0.6152	0.6333	0.6371	0.6498	0.6625	0.6752	0.6879	0.70057	0.71327
Hispanic Females									
Births	0.99281	0.99404	0.99377	0.99402	0.99427	0.99452	0.99477	0.99502	0.99527
Under 5	0.99852	0.99866	0.99882	0.99887	0.99892	0.99897	0.99901	0.99906	0.99911
5 to 9	0.99927	0.99932	0.99936	0.99939	0.99942	0.99944	0.99947	0.99949	0.99952
10 to 14	0.99850	0.99855	0.99886	0.99890	0.99893	0.99897	0.99901	0.99905	0.99909
15 to 19	0.99790	0.99822	0.99845	0.99850	0.99854	0.99859	0.99864	0.99868	0.99873
20 to 24	0.99763	0.99781	0.99811	0.99817	0.99823	0.99829	0.99835	0.99841	0.99847
25 to 29	0.99685	0.99633	0.99773	0.99780	0.99787	0.99794	0.99801	0.99808	0.99815
30 to 34	0.99604	0.99538	0.99696	0.99705	0.99715	0.99724	0.99733	0.99743	0.99752
35 to 39	0.99443	0.99422	0.99579	0.99592	0.99605	0.99617	0.99630	0.99643	0.99656
40 to 44	0.99189	0.99201	0.99333	0.99353	0.99374	0.99395	0.99415	0.99436	0.99456
45 to 49	0.98720	0.98785	0.98945	0.98978	0.99010	0.99042	0.99074	0.99106	0.99138
50 to 54	0.98044	0.98157	0.98330	0.98381	0.98431	0.98482	0.98533	0.98583	0.98634
55 to 59	0.96999	0.97108	0.97446	0.97523	0.97599	0.97676	0.97752	0.97829	0.97906
60 to 64	0.95277	0.95419	0.95795	0.95912	0.96029	0.96146	0.96263	0.96380	0.96497
65 to 69	0.92690	0.92920	0.93345	0.93524	0.93703	0.93881	0.94060	0.94239	0.94418
70 to 74	0.89383	0.89751	0.90380	0.90630	0.90880	0.91130	0.91380	0.91630	0.91880
75 to 80	0.84015	0.84642	0.84862	0.85251	0.85640	0.86030	0.86419	0.86808	0.87198
80 to 84	0.74919	0.76022	0.76765	0.77342	0.77918	0.78494	0.79071	0.79647	0.80223
85+	0.70746	0.72777	0.73986	0.75054	0.76122	0.77190	0.78258	0.79326	0.80394

Source: United States Census Bureau; Population Projections of the United States, by Age, Sex, Race, and Hispanic Origin: 1993 to 2050 (Current Population Reports series P25-1104, November 1993).

Census Bureau survival rates have been used as inputs to the population model for both the historical and forecast periods because they include the required level of racial/ethnic group detail. It should be

noted that all figures represent characteristics of the nation as a whole, since sub-national level survival rate data are not published by the required level of age/sex/race detail for the majority of the Region.

As part of its ongoing program of population projections, the Bureau regularly prepares updated base-year and forecasted life tables. However, life tables are not prepared for all years. As noted above, for any given forecast cycle, the Bureau prepares a base-year table and one or more forecast year tables; tables for intermediate years must be interpolated. Furthermore, not all vital rates used by the Bureau are published, nor does the Bureau publish compilations of their vital rates estimates for past years. Survival rates used as inputs to the population model are thus subject to the availability of published and unpublished figures.

For the model's historical section, Census survival rate estimates for the 1986 to 1991 period were adopted for all years and were obtained from report P-25, No. 1018, *Projections of the Population of the United States by Age, Sex, and Race: 1988 to 2080*.[11] Rates for the Hispanic population were obtained from report P-25, No. 995, *Projections of the Hispanic Population: 1983 to 2080*.[12]

For the period 1992 through 2000, survival rates are carried over from TMDI task 8.2. These are based on the complete life tables included in the Census Bureau's data file, "Population Projections of the United States, by Age, Sex, Race, and Hispanic Origin: 1993 to 2050 -- Middle Series Vital Rate Inputs." This file was released on floppy diskette and contains vital rate inputs used in preparation of the publication *Population Projections of the United States, by Age, Sex, Race, and Hispanic Origin: 1993 to 2050* (Current Population Reports series P25-1104, November 1993).

For the period 2000 through 2035, rates are based on the complete life tables released by the Census Bureau as *Population Projections of the United States by Age, Sex, Race, Hispanic Origin, and Nativity: 1999 to 2100* (also released as *Methodology and Assumptions for the Population Projections of the United States: 1999 to 2100,* Population Division Working Paper No. 38, January 13, 2000).[13] In both 1993 and 2000, the Census Bureau released only complete life tables (i.e. by single-year-of-age) for each racial/ethnic group. In 1993, tables were prepared for the years 1992, 2000 and 2050; in 2000, tables were prepared for the years 1999 and 2100. The complete tables were converted to abridged tables by five-year age group, as required by the Population Model. The methodology for calculation of five-year rates was supplied by Gregory Spencer of the Census Bureau's Population Division, and was verified against historical five-year survival rates data.

Deaths

Deaths records are maintained by state and local governments, and are assembled for the nation as a whole by the National Center for Health Statistics (NCHS) under the Vital Statistics Cooperative Program (VSCP). Death records are necessary inputs to the Population Model that regionalize the imputed mortality of national survival rates, applied at the county and subregional level, by controlling the Model's racial-ethnic output to the actual level of recorded deaths.

Locally maintained death records vary in quality of racial-ethnic detail, with not all areas providing mutually exclusive data with racial deaths reported on a non-Hispanic basis. Death statistics are also available from the U.S. Bureau of the Census, *Interim Population Estimates.* This source of state and county level post-censal population provides data on the demographic components of change,

11 US Government Printing Office, Washington, D.C., 1989, Appendix B.

12 US Government Printing Office, Washington, D.C., 1986, Appendix B.

13 Internet Release Date: May 2000. These rates may be superseded in mid-year 2007, when the Census Bureau releases the revised long term population forecasts for the U.S.

including births, deaths and net migration. Table 3 summarizes deaths reported by the Census *Interim Population Estimates* program for all counties in the New York Metropolitan Region, annually from 2001-2005.[14] Appendix Table 1.17. provides racial-ethnic and all races

Table 3. Total Deaths for All Races & Ethnicities by County & Subregion in the New York Metropolitan Region, 2001 -2005

All Races	***2001***	***2002***	***2003***	***2004***	***2005***	***2001-05***
New York City	***58,401***	***59,366***	***57,764***	***57,473***	***57,896***	***290,900***
Bronx	9,687	10,012	9,591	9,796	9,935	49,021
Kings	18,145	18,328	18,098	17,897	18,140	90,608
New York	11,036	11,311	10,934	10,884	10,807	54,972
Queens	16,042	16,022	15,593	15,369	15,526	78,552
Richmond	3,491	3,693	3,548	3,527	3,488	17,747
Long Island	***22,553***	***23,084***	***22,823***	***22,462***	***22,439***	***113,361***
Nassau	11,336	11,519	11,264	11,012	10,917	56,048
Suffolk	11,217	11,565	11,559	11,450	11,522	57,313
Mid Hudson	***17,250***	***17,295***	***17,035***	***16,834***	***17,036***	***85,450***
Dutchess	2,246	2,284	2,283	2,198	2,157	11,168
Orange	2,490	2,525	2,463	2,492	2,624	12,594
Putnam	563	598	627	591	590	2,969
Rockland	2,069	2,074	2,070	2,128	2,144	10,485
Sullivan	774	733	687	685	720	3,599
Ulster	1,631	1,665	1,566	1,517	1,513	7,892
Westchester	7,477	7,416	7,339	7,223	7,288	36,743
New Jersey	***56,331***	***57,707***	***56,205***	***55,882***	***55,824***	***281,949***
Bergen	7,095	7,681	7,535	7,352	7,363	37,026
Essex	7,212	7,229	6,832	6,897	6,995	35,165
Hudson	4,493	4,642	4,486	4,330	4,282	22,233
Hunterdon	764	807	836	797	854	4,058
Mercer	2,915	2,931	2,803	2,902	2,866	14,417
Middlesex	5,742	5,594	5,473	5,430	5,564	27,803
Monmouth	5,269	5,633	5,336	5,389	5,228	26,855
Morris	3,463	3,420	3,336	3,367	3,540	17,126
Ocean	6,845	7,079	6,900	6,896	6,946	34,666
Passaic	4,083	4,100	4,192	4,105	3,987	20,467
Somerset	2,028	2,040	2,145	2,120	2,027	10,360
Sussex	978	984	1,032	1,037	993	5,024
Union	4,686	4,677	4,407	4,335	4,237	22,342
Warren	758	890	892	925	942	4,407
Connecticut	***16,471***	***16,520***	***16,225***	***15,692***	***15,829***	***80,737***
Fairfield	6,991	6,993	7,091	6,694	6,558	34,327
Litchfield	1,609	1,726	1,648	1,610	1,658	8,251
New Haven	7,871	7,801	7,486	7,388	7,613	38,159
Region	***171,006***	***173,972***	***170,052***	***168,343***	***169,024***	***852,397***

Source: United States Census Bureau, Interim Population Estimates

[14] Year 2000 death data is reported for only one quarter.

mortality information on a county specific basis, from the New York, New Jersey and Connecticut Departments of Health for the period 2000 to 2004. As the Appendix Table notes, not all death data are mutually exclusive. The Health Departments' data will be utilized with the Census Bureau's *Interim Population Estimates* data to construct mutually exclusive racial-ethnic deaths by county and subregion for the period 2000-2005. Census Bureau deaths will be adopted as the control and Health Department deaths will be used to disaggregate the all races control to racial-ethnic components.

Fertility Rates

Table 4 presents the total fertility rate by race-ethnicity that underlies inputs to both the historical and forecast sections of the Population Model. Table 5 shows the age-specific fertility by racial/ethnic group in 2004 and future forecast years for the nation as a whole.[15] The model uses age-specific fertility rates, together with population figures for women of childbearing age, in the forecasting of births by racial/ethnic group. The age-specific fertility rate is the fertility rate for a specified age group, expressed in terms of live births per thousand women in the group. The Population Model uses fertility rate inputs specific to the racial/ethnic groups used in this study, by five-year age cohort for ages 10 through 49. Because of the lack of county and subregional age-specific fertility rates for the required level of racial/ethnic detail, fertility rates have been adopted from national level Census Bureau forecasts. Within the Population Model, births are estimated for each five-year time period as a whole, using rates for the midpoint of that period (e.g., 2003 for the interval 2000 to 2005). The Model output of births is then controlled, or regionalized, by the recorded births of State Departments of Health.

Table 4. National Level Total Fertility Rates by Race-Ethnicity, Middle Series for 1990-2100 Forecast Period, (in 000s)

	Middle Series							
Race & Hispanic Origin	***1998***	***2003***	***2008***	***2013***	***2018***	***2023***	***2028***	***2033***
White non-Hispanic	1824.5	1863.4	1901.2	1939.1	1977.0	2014.9	2031.6	2034.3
Black non-Hispanic	2390.5	2084.8	2092.8	2100.8	2108.8	2116.8	2119.2	2117.9
Asian non-Hispanic	1924.0	2220.2	2209.0	2197.9	2186.8	2175.7	2169.2	2165.9
Hispanic	2981.0	2883.0	2836.3	2789.5	2742.7	2696.0	2663.5	2640.6

Source: United States Census Bureau, *Population Division Working Paper No. 38, January 2000*

Table 5. National Level Age-Specific Fertility Rates by Race-Ethnicity, Middle Series for 1998 & the Forecast Period

Births per 1000 women:	***Age-Specific Fertility Rates by Race-Ethnicity***							
	1998	***2003***	***2008***	***2013***	***2018***	***2023***	***2028***	***2033***
White non-Hispanic								
10-14 years	0.50	0.44	0.45	0.47	0.46	0.45	0.48	0.48
15-19 years	41.80	41.15	42.39	43.37	44.01	44.50	44.76	44.98
20-24 years	96.60	94.45	96.42	98.41	100.35	102.21	103.01	103.15
25-29 years	113.30	108.50	110.74	112.98	115.26	117.45	118.40	118.51
30-34 years	77.70	85.66	88.49	90.26	91.80	93.26	94.00	94.22

[15] The age-specific fertility rate is the fertility rate for a specified age group, expressed in terms of live births per thousand women in the group. The total fertility rate provides a summary of the age-specific rates, and is equal to the number of live births that 1,000 women would have during their lifetime if they conformed to the prevailing age-specific birth levels throughout their reproductive years. Total fertility rates are also sometimes expressed as average births per woman, which is equal to the previously described figure divided by 1,000.

35-39 years	29.90	34.29	35.06	36.74	37.46	37.99	38.04	38.09
40-44 years	4.90	6.29	6.25	6.35	6.83	6.94	6.95	6.89
45-49 years	0.20	0.31	0.30	0.29	0.30	0.33	0.33	0.32
Black non-Hispanic								
10-14 years	5.00	3.65	3.83	3.62	3.60	3.65	3.70	3.69
15-19 years	109.80	92.96	94.95	96.06	95.00	95.03	95.54	95.87
20-24 years	156.00	133.83	134.67	134.60	134.82	135.82	136.14	135.91
25-29 years	109.80	91.77	91.96	92.51	92.52	92.70	93.13	93.12
30-34 years	65.10	60.66	61.09	61.17	61.64	61.51	61.38	61.68
35-39 years	27.00	27.83	28.24	28.59	28.56	28.87	28.56	28.39
40-44 years	5.20	5.79	5.70	5.80	5.93	5.89	5.97	5.83
45-49 years	0.20	0.27	0.26	0.25	0.26	0.27	0.26	0.27
Asian non-Hispanic								
10-14 years	0.70	0.55	0.59	0.59	0.59	0.59	0.59	0.59
15-19 years	24.50	24.78	23.95	24.29	24.19	24.11	24.03	23.96
20-24 years	75.50	77.47	77.59	76.57	76.62	76.25	76.07	75.94
25-29 years	121.70	134.86	133.66	133.23	132.14	131.66	131.32	131.17
30-34 years	102.40	130.97	130.11	129.91	128.97	128.77	128.10	127.91
35-39 years	48.60	61.02	60.54	60.07	60.26	59.63	59.93	59.60
40-44 years	10.40	14.15	14.07	13.89	13.80	13.89	13.74	13.88
45-49 years	1.00	0.90	0.89	0.87	0.86	0.86	0.87	0.86
Hispanic								
10-14 years	2.60	2.23	2.39	2.31	2.26	2.23	2.22	2.21
15-19 years	103.70	98.10	95.25	95.35	93.32	91.77	90.67	89.94
20-24 years	184.10	179.33	176.40	173.51	170.60	167.71	165.67	164.23
25-29 years	152.40	150.35	147.88	145.46	143.41	140.48	138.92	137.67
30-34 years	96.70	93.46	92.18	90.58	89.12	87.94	86.42	85.79
35-39 years	45.30	42.69	42.18	41.62	40.87	40.20	39.96	39.30
40-44 years	10.80	10.33	9.91	9.83	9.74	9.53	9.45	9.45

Source: United States Census Bureau, *Population Division Working Paper No. 38, January 2000*

The Census Bureau periodically revises fertility rate estimates as part of its ongoing population projections program. At the time of this writing, the latest projections have only been released in web format. Historical rates were obtained from the P-25 series, Nos. 917, 995, 1018, and 1092.[16] Rates for 1978 were interpolated from the preceding and following periods. Separate rates for Asians were not available and were therefore adjusted from rates for Whites for all periods from 1970 through 1990. For the period 1990 to 1994, fertility rates are carried over from *Transportation Models and Data Initiatives* Technical Memorandum 8.2, and are based on national-level figures published in the Census Bureau's report *Population Projections of the United States, by Age, Sex, Race, and Hispanic Origin: 1993 to 2050.*[17] For the period 1995 to 1999, figures were obtained from an updated version of the report, *Population Projections of the United States, by Age, Sex, Race, and Hispanic Origin: 1995 to 2050.*[18] Both reports include high-, middle- and low-series fertility rates for the racial/ethnic and age groups used in this study; model inputs are based on middle-series figures.

[16] United States Census Bureau; Population Projects. www.census.gov. Accessed date March 2007.

[17] US Bureau of the Census, Current Population Report, P25-1104, US Government Printing Office, Washington, DC, 1993, Appendix A.

[18] US Bureau of the Census, Current Population Report, P25-1130, US Government Printing Office, Washington, DC, 1996, Appendix A.

From 2000 onward, fertility rates were derived from *Population Projections of the United States by Age, Sex, Race, Hispanic Origin, and Nativity: 1999 to 2100*.[19] Rates for 2000 differ in format from previous releases. In previous years, rates were published in tables by mother's five-year age cohort for each forecast; model inputs were adopted from the middle year of each five-year forecast interval. In 2000, rates have been released in electronic format only by mother's single-year-of-age. These figures were converted to five-year figures using the single-year-of-age fertility rates in combination with comparable single-year-of-age projections of the female population, released concurrently by the Bureau. Births have been calculated as the product of female population and the fertility rate for each group. Five-year cohort fertility rates have then been calculated by dividing the total number of births for each cohort by the total number of women in the cohort.

Births

Inputs of births to the Population Model's historical section provide a subregional-level control to the national-level fertility behavior reflected in the age-specific fertility rate inputs. The Population Model requires inputs of reported resident live births, by racial/ethnic group, for each five-year interval to the year 2005. Historical data and recent estimates of total births were obtained from US Census Bureau and state Department of Health (DOH) sources for years through 2005.

Under the federal/state cooperative program, the Census Bureau works with states to produce an annual series of birth estimates that is consistent in methodology for the entire nation. Census Bureau estimates are released regularly on their web site, and are shown in Table 6 for the New York Metropolitan Region by county and subregion from 2001 through 2005.[20] Birth data compiled by state Departments of Health are published annually in registration reports and maintained on their web sites.[21] Appendix Table 1.18 provides racial-ethnic births for the Region from 2000 through 2004. Census and DOH birth totals are generally quite close. Census estimates have been adopted for years since 1990, while prior model years use DOH figures.

State Department of Health birth data are not uniformly available on a racial/ethnic basis, as required in the Population Model. Using Census Bureau births as a control, limited DOH shares

of racial-ethnic data were applied to disaggregate total births, supplemented where needed by the model-generated estimates of racial-ethnic births from national fertility rates. Total births by racial/ethnic group were divided between male and female on a 51-49 percent basis.

19 Internet Release Date: May 2000. These rates may be superseded in mid-year 2007, when the Census Bureau releases the revised long term population forecasts for the U.S.

20 United States Census Bureau, www.census.gov/population/www/estimates/popest.html. Accessed date March 2007

21 New York State: www.health.state.ny.us; New Jersey: www.state.nj.us/health/chs/; Connecticut: www.state.ct.us/dph/. Accessed date March 2007

Table 6. Total Births for All Races & Ethnicities by County & Subregion in the New York Metropolitan Region, 2001 -2005

All Races						
	2001	***2002***	***2003***	***2004***	***2005***	***2001-05***
New York City	***125,748***	***123,494***	***121,586***	***122,480***	***121,649***	***614,957***
Bronx	25,724	25,321	24,331	24,137	23,485	122,998
Kings	41,493	41,147	40,390	40,679	40,559	204,268
New York	21,142	20,540	20,597	20,999	20,871	104,149
Queens	31,675	30,750	30,428	30,732	30,798	154,383
Richmond	5,714	5,736	5,840	5,933	5,936	29,159
Long Island	***36,749***	***35,869***	***35,994***	***35,869***	***35,339***	***179,820***
Nassau	16,629	15,995	16,186	15,981	15,763	80,554
Suffolk	20,120	19,874	19,808	19,888	19,576	99,266
Mid Hudson	***29,461***	***29,077***	***29,510***	***29,566***	***29,725***	***147,339***
Dutchess	3,322	3,147	3,224	3,267	3,315	16,275
Orange	4,892	4,900	5,047	5,117	5,293	25,249
Putnam	1,139	1,201	1,168	1,153	1,178	5,839
Rockland	4,528	4,438	4,526	4,478	4,545	22,515
Sullivan	846	784	865	909	939	4,343
Ulster	1,732	1,794	1,773	1,811	1,893	9,003
Westchester	13,002	12,813	12,907	12,831	12,562	64,115
New Jersey	***91,317***	***91,196***	***94,099***	***93,588***	***92,185***	***462,385***
Bergen	9978	10104	10455	10555	10428	51,520
Essex	12372	12476	12600	12531	12196	62,175
Hudson	8599	8602	8967	9102	9081	44,351
Hunterdon	1418	1441	1424	1365	1387	7,035
Mercer	4782	4489	4723	4562	4385	22,941
Middlesex	10296	10437	10926	10884	10752	53,295
Monmouth	7737	7762	7925	7948	7950	39,322
Morris	6316	6156	6227	6235	6131	31,065
Ocean	6473	6745	7196	7272	7275	34,961
Passaic	8325	8179	8218	8066	8026	40,814
Somerset	4301	4343	4480	4354	4267	21,745
Sussex	1747	1648	1723	1698	1649	8,465
Union	7697	7504	7853	7714	7439	38,207
Warren	1276	1310	1382	1302	1219	6,489
Connecticut	***24438***	***23937***	***24515***	***24219***	***23641***	***120,750***
Fairfield	12206	11893	12212	12089	11896	60,296
Litchfield	1960	1905	1852	1883	1881	9,481
New Haven	10272	10139	10451	10247	9864	50,973
Region	***307,713***	***303,573***	***305,704***	***305,722***	***302,539***	***1,525,251***

Source: United States Census Bureau, Interim Population Estimates

Net Migration

For the historical period, net migration is calculated within the model as the residual of the population at the end of the interval and natural increase over the interval. Since 1990, the

Table 7. Net Domestic & Foreign Migration by County & Subregion in the New York Metropolitan Region, 2001 -2005

	Foreign		***Domestic***		Total
	2001-05	An'l Ave	2001-05	An'l Ave	2001-05
New York City	***484,241***	***96,848***	***-775,256***	***155,051***	**291,015**
Bronx County	65,286	13,057	-114,197	-22,839	-48,911
Kings County	151,530	30,306	-279,394	-55,879	127,864
New York County	85,758	17,152	-107,630	-21,526	-21,872
Queens County	171,708	34,342	-272,317	-54,463	100,609
Richmond County	9,959	1,992	-1,718	-344	8,241
Long Island	***48,025***	***9,605***	***-62,816***	***-12,563***	**-14,791**
Nassau County	25,008	5,002	-50,372	-10,074	-25,364
Suffolk County	23,017	4,603	-12,444	-2,489	10,573
Mid Hudson	***54,479***	***10,896***	***-37,687***	***-7,537***	**16,792**
Dutchess County	3,720	744	5,432	1,086	9,152
Orange County	4,136	827	13,343	2,669	17,479
Putnam County	1,222	244	347	69	1,569
Rockland County	7,825	1,565	-13,984	-2,797	-6,159
Sullivan County	650	130	1,190	238	1,840
Ulster County	1,169	234	2,709	542	3,878
Westchester County	35,757	7,151	-46,724	-9,345	-10,967
New Jersey	***255,424***	***51,085***	***-222,087***	***-44,417***	**33,337**
Bergen County	37,034	7,407	-33,159	-6,632	3,875
Essex County	31,613	6,323	-58,653	-11,731	-27,040
Hudson County	46,541	9,308	-73,341	-14,668	-26,800
Hunterdon County	1,168	234	3,880	776	5,048
Mercer County	11,412	2,282	-4,750	-950	6,662
Middlesex County	41,310	8,262	-29,200	-5,840	12,110
Monmouth County	10,242	2,048	-3,235	-647	7,007
Morris County	14,140	2,828	-8,433	-1,687	5,707
Ocean County	3,560	712	40,821	8,164	44,381
Passaic County	22,570	4,514	-34,173	-6,835	-11,603
Somerset County	11,207	2,241	-1,086	-217	10,121
Sussex County	788	158	4,389	878	5,177
Union County	22,773	4,555	-29,474	-5,895	-6,701
Warren County	1,066	213	4,327	865	5,393
Connecticut	***49,670***	***9,934***	***-39,875***	***-7,975***	**9,795**
Fairfield County	33,650	6,730	-40,181	-8,036	-6,531
Litchfield County	997	199	5,397	1,079	6,394
New Haven County	15,023	3,005	-5,091	-1,018	9,932
Region	***891,839***	***178,368***	***1,137,721***	***227,544***	**245,882**

Source: United States Census Bureau, Interim Population Estimates

Census Bureau's *Interim Population Estimates* have provided an annual estimate of net migration by county, with more recent years disaggregating this estimate by domestic versus foreign net migration. Table 7 presents the Bureau's estimate by county and subregion for the period 2001-2005. Components of change in the Population Model for the period 2000-2005 reflect the Bureau's migration

estimate, broken by mutually exclusive racial-ethnic groups based upon the 2000 and 2005 racial-ethnic population totals.

In the forecast period, net migration will be calculated in two components: an initial estimate based upon historical rates of net migration, and an adjustment reflecting labor force demand as computed in the Labor Force Model. The first component of net migration is calculated within the model based on rates of net migration for the previous period. The second component is input from the Labor Force Model and reflects the effect of employment demand on the migration of workers in or out of a given subregion. Foreign versus domestic net migration is not explicitly shown in the forecast period. However, racial-ethnic differences in future net migration suggest the relative importance of foreign versus domestic net migration.

1.2.2 METHODOLOGY

Population Models will be developed on a subregional basis and within subregions by racial/ethnic group, for purposes of forecasting population growth from 2005 to 2035. The models' conceptual organization, as illustrated before the Appendix in Figures 1 and 2, has not changed since the earlier forecast applications. Each subregional model will be functionally divided into an historical and a forecast section. In the forecast section, for the 2035 forecasts, base year 2005 population will be projected forward based on estimated future patterns of fertility and mortality, as well as recent subregional patterns of net migration. Because the fertility and mortality rates used in this section are based on national rather than subregional patterns, an adjustment for subregional conditions will be necessary. This is the primary purpose of the historical section. In the historical section, births and deaths adjustment factors will be generated based on the difference between actual reported subregional births and deaths in recent years, and the level of births and deaths that would have resulted if regional fertility and mortality rates had followed national patterns. The adjustment factors will be carried forward into future periods to modify the expected number of births and deaths that result from application of forecasted fertility and mortality rates.

The historical section will also generate estimates of past net migration rates, which are used in the forecast section in the projection of future levels of net migration. In addition, the forecast section of the model will depend, directly and indirectly, on the labor force and employment models, as shown in before the Appendix in Figure 3. These relationships can be stated briefly as follows: Employment will be forecasted independently of population and labor force. Labor force supply will be assumed to respond to employment demand, and net population migration will be assumed to be affected by subregional labor force requirements. Thus, forecasted net migration will be determined by both past patterns of migration and forecasted employment opportunities.

Overview of the Cohort-Survival Methodology

The Population Model is a modified version of the Cohort-Survival methodology. The latter is a standard population forecasting technique that is used to account for the effects of natural increase, i.e., the sum of births and deaths. In this technique a base year population is projected forward, subject to anticipated levels of fertility and mortality, to produce the estimated population for a future year. This future year population can then be used as the basis for a further round of projections. A projection interval of five years, as used in the Population Model, is typical; however one year, or any other time period, can be used.

In the Population model, the Cohort-Survival-Method will be extended to account for net migration, which as discussed above is in turn assumed to respond to forecasted employment levels.

Each component of change (births, deaths, and net migration) will be accounted for separately. While changes in natural increase (i.e., births and deaths) depend on socioeconomic factors that cannot always be foreseen (such as severe economic downturns or lifestyle changes) patterns display some degree of regularity over time and changes are usually gradual. In addition, accurate historical births and deaths are available from federal and state sources, and can be used to control modeling results. Net migration trends are more difficult to forecast reliably because they vary widely over time and are subject to influences that are often volatile in nature (e.g., government policies, political and economic events specific to foreign countries, etc.). In addition, historical net migration data are not as robust as births or deaths data.

In the cohort-survival method, population is divided into a number of separate age groups, or cohorts, and the behavior of each group is accounted for separately in terms of fertility and mortality. Although in principle any age grouping can be used, a five-year cohort is standard; in the Population Model, five-year cohorts will be used through age 84, with an additional cohort for ages 85 and over. Live births will be calculated by applying age-specific fertility rates to the female population cohorts of childbearing age, and then brought into the model at each interval as the youngest age cohort. Deaths will be accounted for by applying age/sex-specific survival rates to each cohort and then subtracted from the model during each interval. An initial ("closed") population projection for each age/sex-cohort will be calculated as the cohort's initial population (including births) minus deaths, based on natural increase alone without the effects of net migration.

In the Population Model net migration is accounted for based on two factors: historical rates of net migration by age/sex/race, and forecasted demand for labor on a racial/ethnic basis. Thus, an initial estimate of net migration by age/sex-cohort will be calculated as a function of the previous net migration levels expressed as age-specific rates of the "closed" population projection for the prior period. The predicted net in- and/or out-migration by age and sex will be supplemented by a labor induced migration component when the demand for labor in the Labor Force Model is not adequately matched by the labor supply from the "closed" population and expected migration.

A final ("open") forecast by age/sex-cohort will be made by adding the net migration component(s) to the "closed" projection, as expressed in the equation:

$$P = P' + M + L$$

where, for any given age/sex-cohort, *P* represents the "open" forecast, *P'* represents the "closed" projection, *M* represents net migration (from prior interval rates), and *L* represents labor induced in or out-migration. The open forecasts are functionally equivalent within the model to census-year population figures or intercensal estimates or interpolations. This method will be used to produce forecasts, at five-year intervals, for the years 2005 through 2035.

The structures of the historical and forecast sections differ and are therefore discussed in more detail separately, below. All discussions are at the level of the racial/ethnic submodels within each subregional model. These submodels are interdependent where necessary to aggregate historical births and deaths statistics by racial/ethnic group or to disaggregate employment and commutation shares in the Labor Force Model.

Historical Section

As noted previously, Figure 1 presents the conceptual organization of the Population Model's historical section. In the historical section of the model, the cohort survival method is applied to historical population, fertility and survival data. Generating estimates, for each interval, of the various components of population change (births, deaths, net migration). By comparing the model's results with actual reported births and deaths statistics, actual-to-estimated adjustment factors will be developed which are used to control birth and death estimates generated in the forecast section. The net migration figures generated by the model for the historical period 1995 to 2005 will be used as a base for forecasting net migration in subsequent years. In addition, the population model will supply historical labor force population figures to the labor force model. The outputs of the historical section can thus be summarized as follows:

- Births Adjustment Factor (to Population Model forecast section)
- Deaths Adjustment Factor (to Population Model forecast section)
- Net Migration Rate by Age and Sex (to Population Model forecast section)
- Labor Force Population (to Labor Force Model)

In the historical section, each five-year interval is discrete; the results of one period are not carried over to the next. New Census population data (or interim population estimates for non-Census years) will be introduced at the beginning of each interval as a base for estimates generated within that period. The sum of historical results will be used, however, in the computation of adjustment factors. A single birth and death adjustment factor will be generated for each racial/ethnic group; this factor will then be applied uniformly across age groups and for all time intervals in the model's section.

Each interval in the historical section will follow several steps. First, cohort-specific survival rates will be applied to each age group, generating an estimated number of deaths within that age group during the interval. Second, the estimated deaths will be controlled by the actual reported total deaths (for all ages) for the interval, resulting in an age-specific deaths estimate in agreement with historical deaths totals. Births over the five-year interval will also be entered and survived. Third, the adjusted deaths for each age cohort will be subtracted from the cohort's population at the beginning of the interval, resulting in a "closed" population projection for the end of the interval, which excludes the effect of net migration. Fourth, net migration will be estimated for each age cohort by comparing this projected "closed" population, based on natural increase alone, with the actual population at the end of the interval from Census data (or interim population estimates). The net migration estimate for each cohort will be equal to the difference between these two figures, a positive figure if the actual population for the cohort exceeds the "closed" projection, a negative figure if it is less. This process will then be repeated for the subsequent interval using the new Census figures.[22]

[22] Because end-of-period Census population figures make it possible to calculate net migration for a given period as a residual of actual final population and estimated natural increase, input from the Labor Force Model is not necessary in the Population Model's historical section. The historical section of the Labor Force Model incorporates Census-based estimates of actual labor force supply, and therefore produces no historical estimates of induced labor force immigration.

Births Adjustment Factors

Birth and death adjustment factors are generated within the historical section for application in the forecast section. The purpose of these factors is to act as a control to the model's births and deaths forecasting methodology by comparing outputs for historical periods with actual reported births and deaths data for the same years. The ratio of actual to estimated births/deaths in past years can then be used to adjust births/deaths estimates generated in the model's forecast section.

The Population Model includes separate sections within each racial/ethnic submodel for the estimation of live births. Births will be estimated for each five-year interval as a whole, based on population figures for women of childbearing age (ages 10 through 49) and fertility rates by age for the mid-point of each five-year interval. These population and fertility rate figures are specific to each racial/ethnic group and five-year age cohort. Fertility rates are expressed as the number of live births per thousand women. The number of live births in a given five-year interval will be calculated separately for each female age cohort, and these figures will then be summed to equal the total number of births for each racial/ethnic group.

Estimated births generated by the model for the period 1990 through 2005 will be totaled, as are actual reported births for the same period. The Births Adjustment Factor is calculated as the ratio of actual to estimated births. For the historical period, actual births will be used in the model. These totals will be divided between males and females on a 51-to-49 percent basis and incorporated within the main section of the model.

Deaths Adjustment Factors

Death estimates are generated within the main section of the model as a part of the cohort survival method. A separate deaths estimate will be made for each age/sex cohort at every time interval as a function of the cohort's population at the beginning of the interval and its five-year survival rate. These age/sex-specific death figures will then be summed within the model to equal the total number of deaths for each racial/ethnic group during each interval.

In the historical period, estimated deaths are adjusted upward or downward by actual deaths. In the forecast period, death adjustment factors, generated by the same methodology used in the calculation of birth adjustment factors, will be applied to adjust deaths accordingly.

Forecast Section

Figure 2 presents the conceptual organization of the Population Model's forecast section. While the forecast section of the model is largely similar to that of the historical section, described above, there are some differences in the calculation of the various components. Whereas in the historical section birth and death adjustment factors will be generated through the comparison of model outputs with reported data, in the forecast section these factors will be applied as controls. And whereas in the historical section it is possible to calculate net migration for an interval as the residual of the population at the interval's end after natural increase is accounted for over the interval, in the forecast section this will not be possible since it is this final population which is to be determined. Net migration for each interval will thus be projected based on the level for the previous interval, natural increase, and the effects of employment demand. The births, deaths and net migration components are discussed separately in greater detail below.

For each interval, the effects of births and deaths will be accounted for first, with births inputted as the youngest age cohort (less than five years) and each cohort then "survived" forward to produce an initial, "closed," projection which excludes the effects of net migration. The net migration component will then be added, producing a final, "open" forecast for the end of the interval. These "open" forecasts are equivalent in the model structure to Census data or interim population estimates, and will be used as a basis for forecasting in the next interval. This procedure for any given age cohort and time interval can be expressed in the equation:

$$P = P' + M + L$$

where *P* represents the "open" population forecast, *P'* represents the "closed" projection, *M* represents net migration (from prior interval rates), and *L* represents labor induced in- or out-migration.

Births and Deaths

Births are calculated in a separate model subsection within each racial/ethnic submodel. An initial births estimate will be made using the same methodology as described for the historical section above: births for each age cohort of women of child-bearing age will be calculated separately, based on the population for the cohort and its fertility rate; cohort-specific births will then be combined to equal the initial total for any given racial/ethnic group. The final total for each group will then be calculated as the product of the initial total and the group's births adjustment factor, generated in the model's historical section. Births are then divided between males and females on a 51-to-49 percent basis and incorporated in the model's main section as the youngest age cohort.

Deaths are accounted for in the model's main section. First, an initial deaths estimate will be made for each cohort based on the cohort's population at the start of the interval and the cohort-specific survival rate. For the 5 to 9 year old cohort, for example, this can be expressed in the equation:

$$_{5-9}D'_{t+5} = {}_{0-4}P_t - \left({}_{0-4}P_t \times {}_{0-4}S_t\right)$$

where *D'* represents the initial deaths estimate, *P* represents the cohort's initial population, *S* represents the cohort's five-year survival rate and *t* represents a given time interval. A final deaths estimate will be produced as a function of the initial deaths estimate and the adjustment factor:

$$_{5-9}D_{t+5} = {}_{5-9}D'_{t+5} \times a$$

where *a* represents the adjustment factor, a constant for each racial/ethnic group for all age cohorts and time intervals.

The number of deaths within a cohort will be subtracted from the cohort's initial population to produce a "closed" population projection for the cohort. This can be expressed in the equation:

$$_{5-9}P'_{t+5} = {}_{0-4}P_t - {}_{5-9}D_{t+5}$$

where *P'* is the "closed" projection, *P* is the population at the beginning of the interval and *D* is the number of deaths.

Net Migration

Net migration comprises two components and will be calculated in two steps. First, an initial estimate will be made of net migration for each age cohort, based on the net migration level for the previous interval and growth in "closed" population between the two intervals, as expressed, for any given cohort, in the equation:

$$M'_{t+5} = M_t \times \frac{P'_{t+5}}{P'_t}$$

where *M'* represents the initial net migration estimate for a given cohort, *M* represents a final Net Migration estimate, *P'* represents a "closed" population projection, and *t* represents a given time interval. Second, the initial estimate for each cohort will be adjusted based on the match between forecasted labor force supply and demand for the time period. This match is determined in the Labor Force Model. The Labor Force Model will produce outputs, at each interval, of race/sex/age-specific net migration levels, to account for the in- or out-migration of workers in search of employment. In the Population Model these figures will be added to the initial net migration estimates to produce a final net migration estimate for each age cohort:

$$M_t = M'_t + L_t$$

where *L* represents the net migration figure input from the Labor Force Model.

1.2.3. DISAGGREGATION OF SUBREGIONAL FORECASTS TO COUNTY LEVEL

County forecasts are not incorporated in the Population Model directly, but will be made by disaggregating the Model's subregional outputs to the county level based on separate projections for each county. However, for the ten counties of the NYMTC territory, separate county level cohort-survival models will be built, based upon historical data from 1990 to 2005, and forecasted to 2035 on a racial-ethnic basis. A methodology will be devised to integrate county level cohort-survival model results with subregional model forecasts, and this methodology will be described in a forthcoming technical memorandum *(TM 5.2.1)*[23].

Disaggregated county level forecasts will be developed for non-NYMTC territory counties of the Mid Hudson, New Jersey and Connecticut, using SPSS software to determine linear, cubic and quadratic equations that best fit historical trends for each county, based on annual data from the period 1970 through 2005. Projections will be modified in consultation with local planning agencies, based on local population forecasts. The limitation of this approach is that it is not directly sensitive to the factors that influence population change, but instead simply projects past patterns into the future, with adjustments based on expert opinion.

1.2.4 ENROLLMENT MODEL

The purpose of the Enrollment Model is to produce forecasts of school enrollment by location for grades K-12 and university level of all public and private institutions in the Region. The forecasts will

[23] Technical Memorandum 5.2.1A, NYMTC County Level Population and Labor Force Models, and TAZ Modeling Methodology

be prepared at the subregional level on a future 5-year interval basis for allocation to baseline 2005 enrollment estimates at the TAZ level. The Enrollment Model will be calibrated on *Public Use Micro Sample (PUMS)* data of the 2000 Census and 2005 American Community Survey. *PUMS* cross-tabulations of population by single year of age and enrollment status will be created at the sub-regional level for years 2000 and 2005. Enrollees will be estimated as a percent of all races/both sexes by 5-year population age cohort for each year, grouping the results by:

- Under 5 years: *Pre-K*
- Ages 5-9: *Kindergarten & Lower Elementary (Grades 1-4)*
- Ages 10-14: *Middle School or Upper Elementary (Grades 5-8)*
- Ages 15-19: *High School*
- Ages 20-24: *College & Graduate School*
- Ages 25+: *Other University or Technical School*

Although the Population Model age cohorts do not strictly match school-going ages by grade level, without actual annual school enrollment data by grade for all public and private institutions in the Region, it is not possible to "age" enrollment or predict school "retention" levels. Therefore, future forecasts will be related to the age structure of the population and the coarse "retention" rates derived as a percent of population enrolled by age cohort in 2000 and 2005 will be carefully examined. These rates may indicate an upward trend in enrollment in the youngest and oldest age cohorts, whereas for ages 5 to 15, nearly universal enrollment should be evident. For ages 15-19, existing conditions will be examined from the perspective of school "drop-out" rates.

The average or trended "retention" rates derived from 2000 and 2005 data will be applied to the all-races population forecasts of subregions to estimate future enrollment from 2010 to 2035. For elementary and high school enrollment, the assumption will be made that enrollment forecasts by place of residence at the subregional level will be equal to enrollment by school location. Results will be examined against available forecasts, such as the enrollment projections of the New York City Board of Education. For college and university enrollment, 2000 and 2005 enrollment by place of residence will compared against reported trends in higher educational institution enrollment, and factors will be estimated to trend residence-based enrollment against school-based enrollment to forecast higher educational enrollment by place of institution.

1.3 EMPLOYMENT MODEL

The purpose of the Employment Model is to generate annual employment forecasts by industry for each of five subregions, and subsequently by private versus public employment in the thirty-one counties of the Region. The Employment Model is based on historical data, some of which were presented in TM 1.1.5.8*: 2002-2005 County/Subregion Time Series Data for Existing Models,* of the DSF project.[24] Other data used in the model include national data provided by Global Insight, Inc. (GI),

[24] Technical Memorandum 1.1.5.8: 2002-2005 County/Subregion Time Series Data for Existing Models. October 2, 2007

and economic data from the Regional Economic Information System (REIS) of the U.S. Bureau of Economic Analysis (BEA). The model will be composed of five (5) sets of equations, comprised of approximately thirty (30) equations each, with one set for every subregion. The county forecasts will be disaggregated from the subregional forecasts for New York City, Long Island, the Mid-Hudson, northern New Jersey, and southeastern Connecticut, using historical share relationships. The equations will be derived using ordinary least squares (ols) regression analysis, a common statistical process used in econometric modeling. The data required by the model is discussed in further detail in Section 1.3.1, while the structure of the model and the methodology used to recalibrate it are explained in Section 1.3.2. A complete listing of all subregion equations that comprise the model appears in Appendix Table 2.1.

The Employment Model is critical to operation of the Best Practices Model (BPM) of the TMDI project for several reasons. As mentioned previously, the model generates annual county-level employment forecasts that provide a basis for generating work trips in the journey-to-work forecasting process. The output of the Employment Model also impacts the forecasting of future population and labor force in the Region. This relationship is portrayed in the flow chart, *Figure 3, Interrelationships of Population, Labor Force, Employment and Journey to Work Models* (before Appendix). Employment forecasts enter the Labor Force Model and set the level of demand for workers in each subregion. In turn, the labor force forecasts enter the Population Model and determine the necessary level of net in- or out-migration, in conjunction with the expected labor force participation of the resident population.

It is important to clearly understand the differences between *labor force* and *employment*. Labor force data indicate how many residents of a particular area have jobs or are unemployed, but provide no information on where the residents actually work. Employment data, by contrast, supply information on the number of persons working in an area, regardless of where the workers may actually live. Labor force forecasts tend to be driven by employment and labor force participation rates, whereas the employment forecasts are based on a range of economic variables that will be discussed below.

1.3.1 MODEL VARIABLES

Because of the range of subregions and industries that are involved, the Employment Model uses a significant number of both dependent and independent variables. Many of the variables are similar, however, and can be grouped together for the purpose of explanation. In the discussions that follow the following convention will be used for variable names:

Index	refers to:
III	Industry mnemonic
SSS	Subregion mnemonic
CCC	County mnemonic

A complete listing of all abbreviations used in the model appears in Appendix 2.8.

Dependent Variables

The dependent variables in the model are those that are to be forecasted. Each dependent variable has a unique equation associated with it in the model. For all dependent variables there exist at least 16 years of historical data on a quarterly basis which are used to develop the equations (See Section

1.3.2 for details). The dependent variables in the Employment Model are, in their most disaggregated form:

- Nonfarm payroll employment in:
    - Natural Resources & Mining
    - Construction
    - Manufacturing
    - Trade Transportation & Utilities
        - Wholesale Trade
        - Retail Trade
        - Transportation & Warehousing
        - Utilities
    - Information Services
    - Financial Activities
        - Finance & Insurance
        - Real Estate Rental & Leasing
    - Professional & Business Services
    - Education & Health Services
        - Educational Services
        - Health Care & Social Assistance
    - Leisure & Hospitality
        - Arts, Entertainment & Recreation
    - Other Services
    - Government
        - Federal Government
        - State Government
        - Local Government
- Proprietors employment
- Wage rates in:
    - Natural Resources & Mining
    - Construction
    - Manufacturing
    - Trade, Transportation & Utilities
        - Retail Trade
        - Utilities
    - Information Services
    - Financial Activities
    - Professional & Business Services
    - Education & Health Services
    - Leisure & Hospitality
    - Other Services
    - Government
        - Federal nondefense
        - State and Local
- Personal income, including:
    - Wages and salaries

Proprietors income
Other income

- Unemployment rate

The main task of the Employment Model is to forecast future nonfarm and proprietor employment levels for a measure of total employment by subregion to 2035. The other dependent variables will be forecasted because they contribute to the employment forecasting process. The forecasts of aggregate wages, personal income, and the unemployment rate are utilized in other modeling processes of the DSF project, including the Labor Force Model (Section 1.4) and the Household Model (Section 1.5). Nonfarm and proprietor employment forecasts will be presented in Technical Memorandum 5.3.1: *Regional, Subregional, County and TAZ Level Forecasts of Socioeconomic Variables from 2005 to 2035*, along with forecasts of population, labor force, household formation, enrollment, aggregate and industry-specific earnings, and employment by occupation.[25]

Nonfarm Employment

Nonfarm employment is derived from and consistent with the Series 790 Non-agricultural employment data collected by the various state Departments of Labor (DOLs) in the Region. The historical data used in the model will cover the period from 1990 through 2006 on a quarterly basis, and are categorized by industry at the subregional level, and county by public and private sector jobs. The data are quarterly average figures calculated from information collected on a monthly basis. The data cover all payroll workers who do not work on farms or are not self-employed. A summary of data trends by five year interval are presented for subregions in Table 8, while Appendix Tables 2.2 through 2.6 present the annual subregional data and Appendix Table 2.7, the annual county data.

Table 8. Total, Nonfarm & Proprietors Employment by Subregion in the New York Metropolitan Region, 1990 to 2005, (in 000s)

	1990	*1995*	*2000*	*2005*
New York City	**3,966.1**	**3,795.8**	**4,277.3**	**4360.8**
Nonfarm Employment	3,566.2	3,339.3	3,723.1	3599.4
Construction and Natural Resource	114.4	89.8	120.6	112.9
Manufacturing	265.2	207.8	176.8	114.3
Transportation, Trade, and Utilities	601.6	532.5	569.6	545.2
Information	169.2	154.4	187.3	162.9
Finance, Insurance, Real Estate and Leasing	521.8	467.2	488.8	446.2
Professional and Business Services	467.2	444.8	586.5	554.1
Educational and Health Services	476.4	551.6	620.1	678.6
Leisure and Hospitality	218.1	208.5	256.7	276.8
Other Services	124.8	122.6	147.4	153.3
Government	607.6	560.1	569.5	554.9
Proprietors	399.9	456.5	554.2	761.6

[25] Technical Memorandum 5.3.1: Regional, Subregional, County and TAZ Level Forecasts of Socioeconomic Variables from 2005 to 2035

Long Island	**1,329.8**	**1,316.4**	**1,457.6**	**1548.0**
Nonfarm Employment	1,123.6	1,093.1	1,218.0	1240.6
Construction and Natural Resource	52.5	42.8	61.0	66.6
Manufacturing	136.1	101.2	105.5	87.3
Transportation, Trade, and Utilities	264.1	255.2	273.1	271.2
Information	31.9	26.6	31.8	29.4
Finance, Insurance, Real Estate and Leasing	85.5	80.4	84.2	81.9
Professional and Business Services	114.2	128.8	155.6	158.4
Educational and Health Services	133.6	158.8	178.5	199.5
Leisure and Hospitality	81.6	77.1	86.0	95.7
Other Services	39.1	42.7	52.1	52.1
Government	185.0	179.6	190.2	198.5
Proprietors	206.2	223.3	239.5	307.4
Mid-Hudson	**1,006.9**	**978.3**	**1,080.2**	**1167.7**
Nonfarm Employment	841.9	793.7	880.5	911.3
Construction and Natural Resource	43.4	31.7	45.7	49.9
Manufacturing	103.4	79.2	74.0	61.1
Transportation, Trade, and Utilities	173.4	163.7	177.9	183.0
Information	27.8	24.2	28.4	23.6
Finance, Insurance, Real Estate and Leasing	51.3	45.1	49.3	51.7
Professional and Business Services	80.2	74.7	88.1	98.5
Educational and Health Services	121.5	136.1	158.5	167.1
Leisure and Hospitality	62.1	59.8	64.0	70.9
Other Services	27.6	29.3	35.1	37.2
Government	151.4	149.8	159.6	168.3
Proprietors	165.0	184.6	199.7	256.4
New Jersey	**3,403.9**	**3,386.4**	**3,752.6**	**3894.6**
Nonfarm Employment	2,931.1	2,896.2	3,227.6	3170.3
Construction and Natural Resource	110.8	94.0	130.1	127.2
Manufacturing	448.8	376.1	337.5	255.6
Transportation, Trade, and Utilities	677.3	654.8	708.3	687.0
Information	103.9	106.5	102.9	84.7
Finance, Insurance, Real Estate and Leasing	198.2	195.7	231.0	231.8
Professional and Business Services	373.1	395.1	508.3	493.0
Educational and Health Services	300.0	348.7	388.3	436.4
Leisure and Hospitality	168.0	175.3	200.2	214.8
Other Services	96.2	100.3	128.2	143.5
Government	454.8	449.7	492.7	496.3
Proprietors	472.8	490.1	525.0	684.8
Connecticut	**987.4**	**965.6**	**1,043.3**	**1099.6**
Nonfarm Employment	815.3	780.8	842.3	822.3
Construction and Natural Resource	34.3	25.9	31.3	31.7
Manufacturing	151.8	127.0	118.0	94.8
Transportation, Trade, and Utilities	160.0	147.4	158.5	155.6

Information	24.4	23.6	26.1	20.8
Finance, Insurance, Real Estate and Leasing	55.3	51.7	63.5	64.1
Professional and Business Services	111.0	118.6	135.3	124.2
Educational and Health Services	100.9	112.0	124.6	137.8
Leisure and Hospitality	54.4	55.5	58.3	63.2
Other Services	28.1	28.6	29.4	30.5
Government	95.2	90.5	97.1	99.5
Proprietors	172.1	184.8	201.0	255.2
Regional Total	**10,694.1**	**10,442.3**	**11,610.9**	**12070.9**
Nonfarm Employment	9,278.1	8,903.1	9,891.5	9,743.9
Construction and Natural Resource	355.4	284.2	388.7	388.3
Manufacturing	1,105.3	891.3	811.8	613.1
Transportation, Trade, and Utilities	1,876.4	1,753.6	1,887.4	1,842.0
Information	357.2	335.3	376.5	321.4
Finance, Insurance, Real Estate and Leasing	912.1	840.1	916.8	875.7
Professional and Business Services	1,145.7	1,162.0	1,473.8	1,428.2
Educational and Health Services	1,132.4	1,307.2	1,470.0	1,619.4
Leisure and Hospitality	584.1	576.1	665.2	721.4
Other Services	315.7	323.4	392.2	416.7
Government	1,494.0	1,429.7	1,509.1	1,517.6
Proprietors	1,416.1	1,539.2	1,719.4	2265.5

Source: New York, New Jersey & Connecticut State Departments of Labor and the U.S. Bureau of Economic Analysis

The variable names for US. nonfarm employment have the form *EEIII* in the model, while those for subregional nonfarm employment have *IIIEMSSS*.

Proprietors Employment

Proprietors employment, also called self-employment, represents proprietors and partners in non-limited partnerships. The historical data are derived from the Regional Economic Information System (REIS) CD-ROM, produced by the U.S. Bureau of Economic Analysis (BEA), and cover the period 1990 to 2005. The data represent annual averages of self employment at the county level summarized to the subregional level. The historical proprietor's employment is presented in Appendix Table 2.9 by county.

The variable for U.S. proprietor's employment is named PROPUS in the model, while the variable for subregional proprietors is named PROP*SSS,* and for county proprietors, PROP*CCC.*

Wage Rates

Wage rates represent average annual earnings per worker by industry and subregion. The historical data, which cover the period 1990 to 2005, come from the REIS CD-ROM (Table CA05, Personal Income & Industry Earnings). It should be noted that the CD-ROM does not contain wage rate data per se. The wage rates must be computed by dividing the total industry-specific earnings (including proprietor earnings) by the total industry-specific employment. A separate estimate of average wage and salary earnings, distinct from proprietor's income, was produced by dividing the aggregate wage and salary earnings by nonfarm payroll employment. Both total earnings by industry and total

employment by industry can be extracted from the CD-ROM. It is also important to realize that the BEA employment figures and the DOL employment figures are not equivalent because of different counting methodologies, definitions of employment and data sources.

In the model, the wage rate variables have names of the form WAGRATE*SSS.*

Personal Income

Aggregate personal income is forecasted because it often appears as an independent variable in the employment equations. In the context of the Employment Model, total personal income is considered the sum of wages and salaries, self-employment income and other income. Other income is composed of unearned income (interest, dividends, rent, etc.) transfer payments, residence adjustments (from commutation), other labor income and a subtraction for personal contributions to social insurance. The county-level historical data come from the REIS CD-ROM (Table CA05, Personal Income & Industry Earnings) and cover the period 1990 to 2005.

The various personal income variables have the following names in the model:

Total Personal Income:	PI*SSS*
Wage and Salary Income:	WAG*SSS*
Other Income:	YOTH*SSS*

Unemployment Rate

Countywide unemployment rates are also forecasted because they appear as independent variables in some of the employment forecasts. The historical unemployment rates are derived from annual average labor force statistics of the state Departments of Labor (DOLs) in the Region, collected under the Local Area Unemployment Survey *(LAUS)* program. The historical data cover the period 1990 to 2005, and data for the most recent period are provided in Technical Memorandum 1.1.5.8: *2002-2005, Series Data for Existing Models*[26].

The unemployment rate variables have names of the form RU*SSS.*

Independent Variables

The Employment Model uses a number of exogenously supplied independent variables. In addition, it is possible for a dependent variable from one equation to act as an independent variable in another. For example, in the equations for forecasting employment, the employment of another industry sector, personal income or the unemployment rate may be used as an independent variable. A subsequent discussion on *Employment Equations* will present a summary table identifying independent variables in subregional equations.

The historical and forecast data for national-level independent variables will be supplied by Global Insight, Inc. (GI), as are the regional inflation rate data. These data cover the period 1990 to 2037 on a

[26] Technical Memorandum 1.1.5.8: 2002-2005 , Series Data for Existing Models, October 2, 2007.

quarterly basis. Should the models be used for forecasting to 2050, data beyond GI's 30-year forecast horizon (2007-2037), would be estimated by Urbanomics using a linear trend assumption based upon the long term forecast of 2013 to 2037. Global Insight, Inc. is a commercial vendor of econometric services, providing economic data and widely accepted forecasts to government and businesses.

Global Insight's U.S. Long Term Trend Forecast, March 2007

GI's trend scenario is the principal long range forecast or baseline scenario. It is regarded as the best unbiased projection of where the U.S. economy is headed, with only a 10 percent chance that the realized path will lie outside this trajectory. Unlike the optimistic or pessimistic scenarios, which GI also forecasts, the baseline assumes that the national economy will grow smoothly along a full employment path, suffering no major mishaps between 2007 and 2037. Such disruptions could include excessive increases in demand, oil price shocks, or untoward swings in macroeconomic policy.

Table 9. Baseline Assumptions for National Variables in Employment Model: Global Insight, Inc. Long Term Trend Forecast, March 2007

National Variable	***Forecasted Values***						
	2010	**2015**	**2020**	**2025**	**2030**	**2035**	
Real GDP 2000$B	$12,849	$14,552	$16,604	$18,930	$21,495	$24,376	
Total Nonfarm (M)	144	148.1	153.1	160	167.1	174.9	
Pers'l Income $B	$13,619	$17,313	$22,016	$27,951	$35,316	$44,076	
Population (M)	310.3	323.7	337.2	350.8	365	379.3	
Labor Force (M)	158.2	162.9	166.4	171.9	178.2	185.1	
National Variable	***Average Annual Growth Rates***						
	2005-10	**2010-15**	**2015-20**	**2020-25**	**2025-30**	**2030-35**	**2005-35**
Real GDP	3.30%	2.70%	2.80%	2.80%	2.70%	2.70%	2.67%
CPI	2.30%	2.00%	2.00%	2.00%	2.10%	2.10%	2.11%
Nonfarm Employment							
Total Nonfarm	1.50%	0.60%	0.70%	0.90%	0.90%	0.90%	0.90%
Nat Res & Mining	0.30%	-3.60%	-1.60%	0.00%	0.10%	0.30%	-0.81%
Construction	1.80%	0.70%	1.10%	1.40%	2.50%	2.80%	1.65%
Manufacturing	-0.60%	-0.40%	-0.30%	-0.60%	-0.60%	-0.40%	-0.48%
Trade, Trans, Util	1.40%	0.30%	0.30%	0.70%	0.70%	0.50%	0.65%
Whl Trade	1.50%	0.90%	1.90%	1.90%	1.50%	0.90%	1.39%
Retail Trade	0.90%	-0.40%	-0.80%	-0.20%	0.00%	0.20%	-0.06%
Info Services	0.50%	1.20%	0.50%	1.20%	1.60%	2.00%	1.14%
Fin & Insurance	1.30%	-0.10%	-0.30%	0.20%	0.50%	0.60%	0.37%
Real Est Rtl & Lsg	1.60%	-0.40%	-0.50%	0.30%	0.40%	0.30%	0.28%
Prof & Bus Serv	4.20%	2.10%	2.60%	2.80%	2.10%	2.30%	2.54%
Ed & Health	1.80%	0.70%	0.90%	1.20%	0.90%	0.80%	1.01%
Education	1.00%	-1.70%	-0.30%	0.70%	0.70%	0.00%	0.05%
Health	2.00%	1.20%	1.00%	1.30%	0.90%	0.90%	1.18%
Leisure & Hosp'y	2.20%	0.20%	-0.20%	0.10%	0.20%	0.10%	0.43%
Arts Enter Recr	3.40%	0.60%	0.80%	1.40%	0.00%	-0.70%	0.88%
Other Services	0.40%	0.40%	0.50%	1.00%	1.20%	1.10%	0.74%
Government	0.90%	0.40%	0.40%	0.20%	0.30%	0.10%	0.38%
Federal	0.60%	-0.60%	0.60%	-0.60%	0.60%	-0.50%	0.01%
State & Local	1.00%	0.60%	0.40%	0.30%	0.20%	0.20%	0.44%
Income and Workers							
Pers'l Income	6.60%	5.40%	5.40%	5.40%	5.30%	5.00%	4.99%

Wages	6.20%	4.80%	4.60%	5.00%	5.10%	4.80%	4.63%
Propr Income	6.30%	6.00%	6.20%	6.00%	5.30%	4.80%	5.19%
Population	0.90%	0.90%	0.80%	0.80%	0.80%	0.80%	0.82%
Labor Force	1.20%	0.60%	0.40%	0.70%	0.70%	0.80%	0.72%
Employed	1.40%	0.50%	0.40%	0.70%	0.70%	0.80%	0.54%
	2010	**2015**	**2020**	**2025**	**2030**	**2035**	
Mortgage Rate	7.30%	6.90%	6.90%	6.90%	6.90%	6.90%	
T-Bill 3 Month	4.90%	4.70%	4.70%	4.70%	4.70%	4.70%	
Unemp Rate	4.40%	4.80%	4.80%	4.80%	4.80%	4.80%	

Source: Global Insight, Inc., March 2007

Between 2005 and 2010, real economic growth is expected to average 3.3 percent per year, while over the remainder of the forecast period, real growth will taper off to 2.7 to 2.8 percent yearly. Demographic forces start to slow the pace of real economic growth after 2010, with expansion of the labor force decelerating from 1.2 percent in 2005 to 2010, to 0.6 percent per year between 2010 and 2015, down again to 0.4 percent yearly between 2015 and 2020, from the retirement by the baby boom generation. After 2020, labor force growth is expected to resume a 0.7 to 0.8 percent expansion annually. Throughout the forecast period, personal income will rise ahead of inflation (CPI), advancing in nominal terms at 5 percent or greater, while the consumer price index climbs 2 percent annually.

In keeping with expected growth in output, employment expands more rapidly in the near term, by 1.5 percent annually in 2005-2010, than in the successive five year growth periods. During the initial retirement years of the baby boom generation (2010-2020), the jobs growth rate is expected to dip to 0.6 to 0.7 percent annually, rising thereafter to a 0.9 percent annual expansion from 2020 to 2035. Manufacturing employment declines throughout the forecast period, while construction rebounds after 2025. The services sector expands moderately in the national economy, with some services experiencing losses particularly in the 2010-2020 decade. Strongest gainers are in professional and business services, and the information sector, with the former consistently expanding above 2 percent annually, while the latter tends to fluctuate between one and two percent.

After a strong showing in the 2005-2010 period, job losses are expected in the 2010-2020 decade in retail trade, finance and insurance, real estate rental and leasing, and education. Although all of these industries except retail trade resume growth in the following decades, their performance is not particularly strong, falling below the total employment average. The health industry and other services, which consist primarily of personal services, grow throughout the forecast period at rates that tend to exceed the total employment average, but all below the strongest gainers. Government expansion also remains positive, with periodic federal losses offset by state and local gains, but overall, government job growth averages less than 0.5 percent annually in the national economy.

Over the next 30 years, wage rates will expand roughly 5 percent on average while inflation will advance 2 percent annually, suggesting a 3 percent real growth in earnings. Productivity growth will average 2.3 percent per year, which is higher than the 2.0 percent annual growth rate of the past 30 years. The pick-up in productivity growth, particularly over the next decade, reflects in part the expected benefits of new investment in capital goods and technology, encouraged by low interest rates. Both long term and short term interest rates will remain below their equilibrium levels by a monetary policy intended to curb inflation. The dollar continues to depreciate against foreign currencies, to slow the growth in current account deficits, while capital inflows contribute to net domestic investment over the forecast period. The growth in wage rates and productivity is also helped by the slowdown in labor force expansion.

Global Insight, Inc. forecast data from the trend scenario comprise the only exogenously supplied independent variables consisting of national data on a historical and forecast basis. A description of the use and nomenclature for each data series is described below. As previously noted, exogenously supplied independent variables consisting of local data are essentially twofold in character: (1) the historic time series of employment, proprietors, wage rates, personal income, population, unemployment and inflation by subregion, for the period 1990 to 2005; and (2) the forecasted values of selected dependent variables, such as the employment of another industry in the same or neighboring subregion, personal income, or the unemployment rate. An exception to this characterization of local variables is the historic and forecasted values of the regional inflation rate or CPI. Forecasted values for the regional CPI are given below, while the values of some historic and forecasted local variables used as exogenous inputs to the modeling process are provided in the Appendix.

Gross Domestic Product (GDP)

Historical and forecasted GDP data have been provided by Global Insight, Inc. in real (constant dollar) terms on an annual basis in chained 2000 dollars. The GDP growth rate variable will be tested in some employment equations, particularly in national market industries. No equivalent regional variable is available. The national output variable is named GDP.

Consumer Price Index (CPI)

Historical and forecasted annual CPI data have been provided by Global Insight, Inc. for both the nation and the New York-New Jersey Region. The ratio of the regional CPI to the national CPI is often used in both employment and wage rate equations as a measure of the relative cost of doing business in the Region. For some subregions, the local inflation rate -- or annual rate of change in the regional CPI -- better explains employment trends in an industry than the relative CPI.

Table 10. Baseline Assumptions for Regional Variables in Employment Model: Global Insight, Inc. Regional Forecast, March 2007

Regional Variable	*Average Annual Growth Rates*					
	2005-	**2010-**	**2015-**	**2020-**	**2025-**	**2030-**
CPINYNJ	2.5%	2.1%	2.1%	2.1%	2.2%	2.2%

Source: Global Insight, Inc., March 2007

The regional CPI variable is called CPINYNJ, while the national variable is named CPIU. The relative cost of doing business is denoted by the variable CPINYNJ/CPIU, and the annual rate of inflation by the variable, CPINYNJ/CPNYNJ\1. Note that neither variable has further indices for other levels of regional geography.

National Employment by Industry

Global Insight, Inc. provided historical and forecasted quarterly employment data by industry for the nation, from 1990 to 2037 on a NAICS (North American Industrial Classification System) basis. National employment is used as a variable in equations of the Employment Model, appearing in many industry-specific employment equations, the wage rate equations and the unemployment rate equations.

Global Insight, Inc. variable names use different industry abbreviations from those adopted for local employment variables, as noted in Appendix B. The following table lists the Global Insight national employment variables:

EME:	Total nonfarm employment
EENRM:	Natural resources & mining employment
EECON:	Construction employment
EEMFG:	Manufacturing employment
EETTU:	Transportation, trade & utilities employment
EETTR:	Transportation employment
EETRET:	Retail trade employment
EETWST:	Wholesale trade employment
ET22:	Utilities employment
EEINF:	Information employment
EEF52:	Finance & insurance employment
EEF53:	Real estate rental & leasing employment
EEPBS:	Professional & business services employment
EEE61:	Educational services employment
EEE62:	Health services employment
EELHS:	Leisure & hospitality services employment
EEOTS:	Other services employment
EEGFED:	Federal government employment
EEGSAL:	State & local government employment

Average Hourly Earnings

Global Insight, Inc. provided both historical and forecasted industry-specific data on average earnings of nonfarm employment. These data are the national comparable to subregional-level wage rate data. The ratio between the two may be used as a variable in the employment equations. National earnings data can also be used in the wage rate equations.

The variable names for the national average earnings data are as follows:

AAENRM:	Natural resources & mining earnings
AAECON:	Construction earnings
AAEMFN:	Manufacturing earnings
AAER:	Trade, transportation & utilities earnings
AAETR:	Retail trade earnings
AAET22:	Utilities earnings
AAEINF:	Information earnings
AAEFIR:	Financial activities earnings
AAEPBS:	Professional & business services earnings
AAE6162:	Education & health services earnings
AAELHS:	Leisure & hospitality earnings
AAESER:	Other services earnings
AAEGOV:	Government earnings
AAEGFD:	Federal nondefense government earnings
AAEGSL:	State & local government earnings

Note that only major industry level earnings data will be used for financial activities services and selected trade, transportation & utilities sector equations because of national data limitations.

National Personal Income

National equivalents of the personal income dependent variables were provided by Global Insight, Inc. The data include historical and forecasted national values for total personal income, wages and salaries, proprietors' income, and other income. The personal income variables will be included in the income equations where needed. The variables appear in the model as follows:

Total Personal Income:	YP
Wage and Salary Income:	WSD
Proprietors' Income:	YPROP
Other Income:	YOTH

Population

Historical population data were provided by the U.S. Bureau of the Census on annual basis for the nation and by county, aggregated to subregion, for the period 1990 to 2005. Recent and prior population estimates are contained in Technical Memorandum 1.1.5.8: *2002-2005 County/Subregion Time Series Data for Existing Models* and Technical Memorandum 1.1.1: *Population Data Collection & Analysis*.[27] Subregional population forecasts are provided to 2030 by the Transportation Models and Data Initiative (TMDI) project, as cross-accepted by NYMTC and its partners in 2004. Extrapolation of the forecasted trends to 2035 will be prepared by Urbanomics. In preparing final employment projections, subregional population forecasts will be replaced by draft level population projections of the Population Model.

The national population will occasionally be included in the model, typically as a denominator for use in measuring national per capita personal income. Forecasted national population data including armed forces overseas were provided by Global Insight, Inc. and are shown in the following table, in comparison with U.S. Bureau of the Census resident population projections. As Table 11 shows, the two series are virtually equivalent, varying only by a fraction of one percentage point in any year.

Table 11. Comparison of Global Insight, Inc. National Population Projection with U.S. Bureau of Census Projection (Middle Series)

National Variable	***Population in Millions***						
	2005	**2010**	**2015**	**2020**	**2025**	**2030**	**2035**
Global Insight	297.0	310.3	323.7	337.2	350.8	365.0	379.3
U.S. Census	295.5	308.9	322.4	335.8	349.4	363.6	377.9

Source: Global Insight, Inc., (March 2007); United States Census Bureau. Interim Projections by Age, Sex, Race and Hispanic Origin (May 11, 2004)

Subregional population variables are named POP*SSS*. The national population will appear as NP in the equations.

[27] Technical Memorandum 1.1.5.8: 2002-2005 County/Subregion Time Series Data for Existing Models. October 2, 2007 and Technical Memorandum 1.1.1: Population Data Collection & Analysis, August 31, 2000

Mortgage Rate

Historical and forecasted mortgage rate data, provided by Global Insight, Inc., may be used in some of the equations projecting construction, finance & insurance, real estate rental & leasing employment.

In the model, the variable associated with the mortgage rate will be named as RMMTGENS.

Treasury Bill Interest Rate

Historical and forecasted 3-month treasury bill interest rates are occasionally used in equations forecasting employment in the wholesale trade, retail trade, finance & insurance, real estate rental & leasing industries. When compared to mortgage rates, as in the financial sector equation, the relationship expresses the differential between long term and short term costs of financing.

The 3-month treasury bill interest rate will be represented by the variable RMGBS3NS in the model.

AAA Corporate Bond Rate

Historical and forecasted values of AAA Corporate Bond Rates were provided by Global Insight, Inc.. These rates may be used in selected employment equations as a proxy for financing costs of capital investments.

The corporate bond rate is called RMMBCAAANS.

S&P 500 Index of Common Stocks

Historical and forecasted values of the S&P 500 Index of Common Stocks may be used in equations for forecasting employment in finance & insurance. The S&P 500 is a market-value weighted index with each of 500 stocks -- chosen for market size, liquidity and industry group representation -- weighted in the index in proportion to their market value. The S&P 500 is a widely used benchmark of U.S. equity performance and a proxy for activity on the stock exchanges.

The index is represented in the model by the variable JSP500.

National Unemployment Rate

The historical and forecasted data for the national unemployment rate were provided by Global Insight, Inc.. The variable may be used in subregional unemployment rate equations.

The national unemployment rate will be denoted by the variable RUC in the model.

1.3.2 MODEL STRUCTURE

General Description

The Employment Model is a standard econometric model consisting of several hundred equations in five subregional industry models and thirty-one county disaggregation models. Appendix B provides a

specification of subregional equations used in the 2030 forecast, listed in the order solved by subregion.

Econometric modeling is a statistical technique that develops predictive mathematical models based on patterns and relationships in historical data. Econometric modeling also requires that the modeler make a number of assumptions regarding the underlying structure of the model, particularly in terms of what variables are likely to be required. Because of these assumptions and modeling processes, it should be noted that excessive reliance should not be placed on econometric equations. The regression coefficients developed from historical data cannot precisely and adequately address the interrelationships between dependent and independent variables in future years. To avoid any irregularity and/or unreasonableness in forecasts produced by econometric models, qualitative reviews are necessary.

The econometric model developed for 2030 forecasts, to be used for the 2035 forecast series, is comprised of several hundred equations that produce forecasts at the subregional level based on historical data. Each dependent variable of interest (e.g., finance and insurance employment in New York City) has a unique equation associated with it in the model. Once the structure of each equation has been satisfactorily determined based on the relevant historical data, the equations can be used to generate forecasts for all of the dependent variables. The historical data are used to determine the mathematical relationship between the historical independent variables and the historical dependent variable for each equation. By assuming that this relationship will hold true into the future, the equations can be used to forecast future values for the dependent variables based on forecasts of the necessary independent variables. By its very nature, an econometric model cannot predict future conditions that have no basis in past trends, such as the outcome of disruptive natural forces, seismic changes in public policies and regulations, or major revolutions in technology.

A regional model of comparable structure will also be developed to independently forecast the long term regional outlook in relation to the nation. The regional model will be used for informative purposes, however, and not as a control on the sum of subregional level forecasts. Once the subregional forecasts are solved, county level disaggregation models will be developed that share the industry-specific subregional forecast to component counties, based upon historical relationships. County forecasts will be expressed as the sum of private and public sector nonfarm employment. The county models will have separate equations for forecasting proprietors, in order to yield a total employment forecast of nonfarm payroll and self employment.

Each subregional model is a collection of approximately thirty (30) simultaneous[28] equations for predicting the various dependent variables of interest. Most of the subregional models are mathematically independent of one another such that the results of one have no bearing on the results of the others. The exceptions to this independence involve the employment and income equations for some suburban subregions. For example, in several suburban subregions the employment and income equations have variables for New York City. These relationships will be demonstrated in a subsequent table.

While some dependence between subregional models may be allowed, it is restricted to dependence in one direction only. For example, as mentioned above, the income and employment equations for some suburban subregions may be affected by employment in New York City. However, the employment

[28] Most of the equations within each subregional model are actually independent of one another. In some cases, the wage rate and employment equations for particular industries are dependent on one another and therefore truly simultaneous.

equations of New York City will not be affected by income and employment in those same suburban subregions. Without this simplifying restriction the five subregional models would have to be combined into larger models such that all areas dependent on others would be grouped into the same large model. Had interactions been allowed to be complex enough, it is conceivable that all subregions would have to be grouped into a single model. The number of equations would remain roughly constant but they would have a much higher degree of simultaneity because of the complex interactions between subregions. This would make the Employment Model much more complex from a computational standpoint.

The process by which independent variables are chosen for the model is called *ordinary least squares (OLS) regression analysis.* Regression analysis is a standard statistical technique for determining the "best fit" equation for a set of data points. In the case of the econometric model developed for 2030 forecasts, the data points were the historical values of the various dependent variables of interest. The historical data can be thought of as being plotted on a graph with a horizontal time axis and a vertical axis of appropriate units (e.g. employment). Regression analysis determines an equation that most closely approximates the curve defined by the plotted data points. The candidate independent variables for the equation are chosen by the modeler based upon the relationships that are believed to exist in the data. The regression analysis process identifies which of the candidate independent variables are contributing to the 'goodness of fit" of the equation and which are not. The modeler then adds and removes independent variables, changes their form from current to lagged, or from absolute to relative change, and repeats the regression analysis until satisfied with the fit of the equation.

For each equation a set of "rules" is developed concerning the coefficients for each independent variable. The sign of the coefficient indicates whether the independent variable (e.g., national construction employment) is positively or negatively associated with the dependent variable (e.g., subregional employment). For example, in the employment equations it will likely be decided that the term representing national employment in a given industry should have a positive coefficient if it appears in the equation for that subregional industry. The positive coefficient will cause an increase in subregional employment if there is an increase in national employment and a decrease if the national employment decreases. While this may not always be true it does make more intuitive sense than allowing the subregional and national employment to move in opposite directions via a negative coefficient on the national employment. It may also be necessary to require the coefficients on some variables to be less than one (1) to avoid instability problems in the model. Variables in equations that are in natural logarithmic form are not allowed to have negative coefficients in order to avoid problems with inverse logarithms. Independent variables whose coefficients fail to satisfy the set of "rules" developed for each equation, as described above illustratively, are thus dropped from the model.

The contribution made by each independent variable to the "goodness of fit" is determined by a test of significance. After each regression analysis is performed, a value called the *t-statistic* is calculated for each independent variable. The t-statistic indicates how statistically significant the variable is in terms of explaining the behavior of the dependent variable. For the Employment Model, the t-statistics *must* be either greater than +1.0 or less than -1.0, but they *should* be either greater than 1.6 or less than -1.6.

The fit of the equation can be judged in a number of ways. The most obvious is to overlay a plot of the historical data with a plot of the equivalent data as determined by the equation. By comparing the shapes of the two plots the modeler can get an overall visual impression of the fit and can identify years where the fit may need to be improved.

In addition to examining the graph, the modeler can also look at the *coefficient of determination* (commonly referred to as the "R-square") that is calculated during the regression analysis. The R-square value is a quantitative measure of goodness of fit. Its value lies between zero (0) and one (1), with one (1) indicating a perfect fit. For the Employment Model, the R-squared values for each equation should typically be greater than 0.95.

Another evaluation technique that can be used involves examining the "errors" in the fit of the equation. The difference between the actual value of a data point and the corresponding value calculated by the equation is called the *residual*. The residuals represent variation in the dependent variable that is caused by some unidentified independent variable. For equations in the Employment Model, the residuals should not be allowed to be more than five (5) percent of the corresponding historical value.

Once the "fit" of the equations is satisfactory, the equations can be used for forecasting. Future values for each of the independent variables in the model are entered and the model calculates the desired dependent variables. In some cases, the dependent variable for one equation appears as a dependent variable in another equation, and vice versa. These equations are solved simultaneously and in virtually all cases a unique solution for the two variables can be found.

Occasionally the forecasts produced by a subregional model may be clearly unrealistic, being either too high or too low. Typically, when very large changes, of say greater than 10 percent, occur in an annual forecast at the subregional level, such outcomes can be considered unrealistic. In these cases the modeler must apply expert judgment in review of the model structure and input data, including historical series, to determine what is triggering the effect. Checks will be made for instability in variable relationships, for data outliers (extreme values), and for other factors. Either some theoretically acceptable way must be found to modify the existing equations, such as expression in log-linear mode, or the model must be rebuilt. If the problem exists with the employment equations, it is often possible to include a constant adjustment factor that makes the forecast more reasonable.

It should be noted that some equations in the Employment Model will make use of special independent variables called "dummies". Dummy variables are used to improve the fit of equations when it is apparent that some unknown variable is having a deleterious effect. In some cases, an unusually large residual may appear in one year and a dummy variable will be used to eliminate it. Dummies used for this reason will appear in the model as "DUM*xx*", where *xx* corresponds to the year with the large residual (e.g. "99" for 1999). These dummy variables are treated like independent variables that have a value of one (1) for the year that corresponds to the variable name, and a value of zero (0) for all other years, past and future. This has the effect of making the variable relevant only to that one specific year, when it helps to account for effects in that year that other variables cannot explain.

The TREND variable is a special variable that is also used to improve the fit of some equations. TREND is basically a variable that is incremented by one (1) every period and can be useful when the data display a definite trend over time.

Some of the variables included in the model may be "lagged," i.e. they may represent data for the same variable from a previous year. Variables that represent lagged data are suffixed by "*x*", where *x* indicates the number of lagged time periods. The following section presents a summary table identifying lagged variables among all independent variables in each subregion's equations. An example of a lagged variable can be seen in the equation for forecasting construction employment in New York City to 2030:

CONEMNYCEQ = EXP(<#COEF1:0.390978>+<#COEF2:0.954078>*LN(CONEMNYC\4)+
<#COEF3:0.120913>*LN(EECON)+<#COEF4:1.32657>*LN(TNEMNYC\1/TNEMNYC\5)-
#COEF5:0.217743>*LN(RMMTGENS\3)+<#COEF6:7.33842>*LN(POPNYC\1/POPNYC\5)-
<#COEF7:0.0528006>*&&DUM92Q4)

The term CONEMNYC\4 is a lagged variable that indicates that the city's construction employment in any given year is based in part on the city's construction employment in four years previously. As the summary table will show, prior years' industry-specific or subregional nonfarm employment is used in most employment equations. The term POPNYC\1/POPNYC\5 uses a lagged expression of population in New York City as the numerator and denominator of a variable that represents the rate of population change. Population growth rates are often used as explanatory variables in employment equations, particularly those that model population-serving activities.

Table 12 presents the regression output for the same construction employment equation, as a means of illustrating the use of evaluation statistics. Referring to the above equation for construction employment in the New York City subregion, it should be noted that a *constant* term or intercept value (0.390978) is part of the functional form of this regression. Other equations may, or may not, include a constant term, depending upon the modeler's determination of its statistical significance.

Reading across the columns of Table 12, by the rows listing each independent variable in the equation, the *Coefficient* estimated by the regression analysis is the estimated value of the unknown percentage of each independent variable in the equation. The *Standard Error* is the error of the estimated value of the coefficient. The *T-Statistic* is a measure of the statistical significance of each estimated coefficient that reduces to the value of the coefficient divided by its standard error. Generally speaking, a t-statistic greater than two (2) in absolute value indicates that the variable in question is statistically significant in explaining changes in the dependent variable at a 95 percent confidence level.

Table 12. Illustrative Evaluation Statistics for New York City Construction Employment Equation

New York City Construction: Dependent Variable – CONEMNYC	**Coefficient**	**Standard Error**	**T-Statistic**	**Other Statistics**
For Independent Variables				
CONSTANT	0.390978	0.1504	2.599	
CONEMNYC\4	0.954078	0.05409	17.64	
EECON	0.120913	0.06243	1.937	
TNEMNYC\1/TNEMNYC\5	1.32657	0.3318	3.998	
RMMTGENS\3	-0.217743	0.04286	-5.080	
POPNYC\1/POPNYC\5	7.33842	2.726	2.692	
DUM92Q4	-0.0528006	0.02157	-2.448	
For Regression				
R-BAR SQUARED:				0.9795
DURBIN-WATSON:				1.4778
STANDARD ERROR:				0.02028
NORMALIZED:				0.004398

The *Other Statistics* pertain to overall evaluation of the regression equation. The R-Squared is a measure of how well the equation fits the data. It reduces to the explained sum of squares divided by the total sum of squares. As previously noted, the R-Squared ranges from zero (0) to one (1) where a fit close to one (1) is desired. The *R-Bar Squared* is the R-Squared corrected for degrees of freedom. As such, it is a more exacting measure of goodness of fit.

The *Durbin-Watson* statistic is a measure of first order serial correlation in the residuals. The Durbin-Watson measure ranges from zero (0) to four (4) where a value of two (2) indicates no first order serial correlation. If the regression contains lags of the dependent variable, the Durbin-Watson statistic is an unreliable measure of autocorrelation. Lastly, the *Standard Error* of the regression is a measure of the standard deviation of the calculated error term in the equation. The *Normalized Standard Error* is the standard error of the regression divided by the mean of the dependent variable. In general, the lower the standard error of the regression, the better is the equation. When estimating dependent variables with very large or very small values, the normalized standard error is especially useful.

Model Structure

Each of the five subregion models that comprise the Employment Model uses the same basic underlying structure. There are some variations to account for special situations (e.g. suburban subregions whose employment is closely tied to employment in New York City) but, for the most part, the functional format of the equations is similar. For ease in viewing the commonalties and differences in subregional equations, Table 13 is offered as a guide to the array of national, regional, local and other independent variables that may be included in an employment equation.

Table 13. Schematic of Independent Variables Used in Employment Equations, by Form of Expression

National Variables	***Local Variables***			***Other Variables***
	Regional	***Own Subregion***	***Neighboring Subregion***	
• Employment *(level, chge)*	• Inflation *(level, chge)*	• Employment *(lag, share, chge, other industry)*	• Employment *(total, or same industry)*	• Year Dummies
• Wages *(level)*	• Cost of Doing Business *(relative)*	• Wages *(level, lag, relative, real, chge)*	• Wages *(lag, relative)*	
• Population *(level)*		• Population *(level, lag, share, chge)*		• Trend Dummies

• Financial Rates *(level, relative, chge)*		• Income *(share, chge, real, lag, percapita)*		
		• Unempl't Rates *(level, lag, relative)*		

In addition to employment forecasts by industry, the subregional models will be designed to generate annual forecasts for the following variables: wage rates by industry, personal income and its components, and the unemployment rate. Separate equations disaggregate the subregional industry-specific employment to component counties and forecast the number of proprietors by county, using *ordinary least squares (OLS)* regression analysis. The functional format for each equation is discussed below.

Employment Equations

The basic form of the employment equation is a linear regression relationship that states the level of subregional employment will be a function of one or more independent variables, including national and regional measures of economic activity and competitive advantage, such that

$$III\text{EM}SSS = f\left(\text{EE}i, \frac{\text{WAGERATE}i}{\text{WSD}i}, \frac{\text{CPINYNJ}}{\text{CPIU}}, Other\right)$$

where *III* refers to the industry, EM is local employment, and *SSS* refers to the subregion. EE is national employment, WageRate/WSD is the subregion wage differential vis-à-vis the nation, and CPINYNJ/CPIU is the relative price differential of the Region to the nation. O*ther* includes other indicators of national and regional demand including:

Measure	Variable
Previous period's employment:	*III*EM*SSS*\1
Subregion unemployment rate:	RU*SSS*
Personal income:	PI*SSS*
Change in personal income:	(*PISSS* /PI*SSS*\1)
Real per capita income:	(PI*SSS*/CPINYNJ) POP*SSS*)

Change in population:	(*POPSSS/POPSSS*\1)
Mortgage rate:	RMMTGENS
Treasury bill interest rate:	RMGBS3NS

The ratios involving the wage rates and inflation should have negative coefficients. The unemployment rate, if included, should also have a negative coefficient. The previous period's employment should have a positive coefficient that is less than one (1) to avoid instability in the model. The national employment, along with any income or population measures, should also have a positive coefficient.

The requirements for the coefficients reflect the positive and negative effects that various factors will have on employment. The sensitivity of each equation to the input variables will be expressed by the value of the coefficients attached to each independent variable. Employment in an industry should be affected positively by both the national demand in the industry sector (represented by national industry employment in the model) and the subregional demand for the industry's products (as represented by a variable such as income). Industry employment will be negatively affected by factors such as the relative costs between the subregional and national economies. The wage rate and CPI ratios will represent the relative costs of doing business in the subregion.

The actual variables chosen for each equation will depend on whether the associated industry is an export industry (e.g. manufacturing) or a domestic industry (e.g. services). Export industries tend to be more strongly connected to national demand while domestic industries are more affected by local demand. Some variables are also only associated with particular industries. As examples, the mortgage rate will appear in equations for construction employment, while the difference between the mortgage rate and the treasury bill rate will appear in some of the financial employment equations.

Wage Rate Equations

The general format for the wage rate equations is:

$$\mathrm{W}AGERATEiSSS = f\left(\mathrm{WSD}i, \frac{\mathrm{CPINYNJ}}{\mathrm{CPIU}}, \frac{III\mathrm{EM}SSS}{\mathrm{EE}III}, \mathrm{WAGERATE}iSSS\backslash 1\right)$$

or, alternatively,

$$\ln\left(\mathrm{W}AGERATEiSSS\right) = f\left(\ln\left(\mathrm{WSD}i\right), \ln\left(\frac{\mathrm{CPINYNJ}}{\mathrm{CPIU}}\right), \ln\left(\frac{III\mathrm{EM}SSS}{\mathrm{EE}III}\right), \ln\left(\mathrm{WAGERATE}iSSS\backslash 1\right)\right)$$

where *i* refers to the industry and *SSS* to the subregion.

The model will assume that the wage rate in an industry is positively associated with the national average wage for that industry (WSD*i*), regional prices relative to national prices (CPINYNJ/CPIU), regional employment relative to national employment (*III*EM*SSS*/EE*III*) and, possibly, the previous period's wage rate (WAGERATE*iSSS*\1). In order to incorporate these assumptions it may be necessary to restrict the coefficients of all variables to positive values. Furthermore, if the lagged wage rate variable is used, it may be necessary to restrict its coefficient to values of less than one (1) to avoid stability problems.

Unemployment Rate Equations

The subregional unemployment rate equations have the following basic form:

$$\mathrm{RU}SSS = f\left(\frac{\mathrm{WAGEMP}SSS}{\mathrm{EME}}, \mathrm{RUC}, \mathrm{RU}SSS\backslash 1\right)$$

where *i* refers to the industry, *SSS* to the subregion.

The ratio between total nonfarm employment in a subregion and total nonfarm employment in the nation (WAGEMP*SSS*/EME) will have a negative coefficient to reflect the negative effect that an increased local share of national employment would have on the unemployment rate. The coefficients of the other two terms will be positive because of the positive association between local and national unemployment rates and the positive effect of the prior period's unemployment rate.

Income Equations

The equations for proprietors' income will be of the form:

$$\ln(\mathrm{YPROP}CCC) = f\left(\ln(\mathrm{YPROP}CCC\backslash 1), \ln(\mathrm{YENTNFADJ}), \ln\left(\frac{\mathrm{WAGEMP}SSS}{\mathrm{EME}}\right), \ln(\mathrm{PROP}CCC)\right)$$

where *C* refers to the county.

The log-log form of the equation will be used to take into account the multiplicative nature of the relationship between the variables involved.

The coefficients for all variables should be positive. Proprietors' income will be assumed to be positively related to national proprietors' income (YENTNFADJ), the ratio of subregion to national employment (WAGEMP*SSS*/EME), the number of proprietors in the county (PROP*CCC*) and the proprietors' income for the previous period (YPROP*CCC*\1). Proprietor's income by county will be summed to subregional proprietor's income (YPROP*SSS*).

The equations for other personal income will be of the form:

$$\ln(\mathrm{YOTH}SSS) = f\left(\ln(\mathrm{YOTH}SSS\backslash 1), \ln(\mathrm{YOTH}), \ln\left(\frac{\mathrm{POP}SSS}{\mathrm{NP}}\right)\right)$$

where *S* refers to the subregion.

Again, the log-log form of the equation will be used because of the multiplicative nature of the relationship.

Other personal income will be assumed to be positively related to national other personal income (YOTH), the ratio of subregional population to national population (POP*SSS*/NP) and the other personal income (YOTH*SSS*\1) of the previous period. Equations for some subregions may also include employment variables from a neighboring subregion if a significant amount of the first subregion's labor force works in the neighboring subregion. This may occur in a number of the suburban subregions whose residents work in New York City.

Total wages and salaries for each subregion will be calculated using the following formula:

$$\text{WAG}SSS = \sum_{ii} \text{EM}SSS \cdot \text{WAGERATE}SSS$$

which will simply multiply the employment by industry times the wage rate by industry and then sum the result to the subregional total.

Total personal income (PI*SSS*) will simply sum the three types of income already discussed or,

$$\text{PI}SSS = \text{WAG}SSS + \text{YPROP}SSS + \text{YOTH}SSS \ .$$

1.3.3. OCCUPATIONAL MODEL

Industry-occupation matrices will be compiled from decennial Census data of Summary File 4 (SF4) for 1990 and 2000 on a subregional basis. The matrices will cross-tabulate the major occupational classification of each industry's employment, for all industries forecasted by the Employment Model, and express the industry composition on a relative or percent share basis. Rates of change in each industry's occupational composition between 1990 and 2000 will be projected forward on an incremental basis, to produce industry-occupational matrices for 2010, 2020, 2030 and 2040. Intervening five year forecast periods will be interpolated for each industry by subregion.

The forecasted industry-occupation matrices will be applied to the forecasted industry employment to produce occupational forecasts of future employment growth on a total employment basis.

1.4 LABOR FORCE MODEL

The term "labor force" refers to workers on a resident basis, that is, residents of an area regardless of whether they are employed or unemployed in that area. By contrast, "employment" refers to jobs by location of workplace, or the number of jobs in a given area regardless of whether the people who hold those jobs also live in that area. Labor force forecasts are driven by expected growth in population, rates of labor force participation, and employment levels, while employment forecasts are based on market-driven factors. These include, at the regional level, relative competitiveness in terms of the 'cost of doing business,' productivity advantages, and local market consumption; and, at the national level, demand for output, productivity, interest and exchange rates, and inflation.

As previously noted, in Section 1.3, employment forecasts will be conducted in a separate model that drives the entire forecasting process. The labor force forecasts will depend heavily on outputs of the Population Model, as adjusted for the demand for labor as forecasted by the Employment Model. The Labor Force Model will utilize the US Bureau of Labor Statistics (BLS) Civilian Labor Force (CLF) concept from 2000 onward, while decennial Census Bureau data will be the basis of historical resident labor force before 2000. Census Bureau data will also inform the BLS current labor force information on a racial-ethnic basis, but the failure of 2000 Census enumerations to accurately capture the total resident labor force has resulted in the substitution of BLS data drawn from the *Local Area Unemployment Survey (LAUS).*

Within the process of regional transportation modeling, labor force forecasts are useful for two primary reasons. First, the size and distribution of the labor force affect the number and pattern of journey-to-work trips, which account for a large proportion of all travel within the Region, especially during peak hours. Labor force forecasts are thus necessary as a control in the process of journey-to-work forecasting. Second is the effect of labor force demand on population. Unlike at the national level,

where employment levels tend to follow population growth, at the regional level employment leads population, with the number of jobs establishing the demand for labor, which in turn affects population and migration. A growing job base can be expected to attract workers to a region, and a declining job base will result in out-migration of local workers seeking employment elsewhere. By matching the expected labor force supply to anticipated levels of employment, it is possible to account for these effects on migration.

In any forecast period, the Labor Force Model thus produces two sets of outputs: first, the initial labor force estimate and, second, the net in- or out-migration level induced by a match between labor force supply and demand for employees. Initial labor force estimates are generated, in any forecast year, based upon expected population from net natural increase and aging of the population, previous period rates of net migration, and forecasted rates of labor force participation. Induced net in- or out-migration is calculated by comparing this resulting labor force supply with the expected levels of employment as forecasted in the Employment Model. The Population Model incorporates the change to initial net migration that results from this matching process, at each five-year interval.

All outputs of the Labor Force Model are generated by sex and age-group for the population 16 years of age and over. A separate model will be run for each subregion; within each subregional model, sub-models generate outputs for each racial/ethnic group. Each model will include a historical section, covering the years 1970 through 2005, and a forecast section, covering the years 2006 through 2035. All outputs are generated on a five-year interval basis.

Reflecting these interactions between labor force supply, employment availability, and population, the Labor Force Model will be developed in conjunction with other models. It will depend on the Population Model for inputs of forecasted population by sex, race/ethnicity and age, and on the Employment Model for inputs of forecasted nonfarm employment and proprietors. The Labor Force Model, in turn, will control the results of the Population Model. In a future technical memorandum (TM 5.2.3, *Model Interoperability and Interfaces with GIS Databases),* a framework will be developed that will facilitate interoperation between the Population, Labor Force and Household Formation models.[29] A separate forecasting routine will disaggregate the subregional labor force forecasts to the county level, with the exception of NYMTC partner county labor force forecasts. Under TM 5.2.1, *NYMTC County Level Population and Labor Force Models, and TAZ Modeling Methodology*, separate Labor Force Models will be constructed for each of these counties and a method will be devised to integrate these models with the subregional Labor Force Models.[30]

1.4.1 DATA INPUTS

The subregional Labor Force Model incorporates a number of independent variables as inputs at each five-year interval:

- Population by racial/ethnic group, sex, and age-group.
- Labor Force Participation Rates by racial/ethnic group, sex, and age group.
- Unemployment Rates by racial/ethnic group.
- Net Commutation.
- Employment levels, combining nonfarm employees and proprietors.
- Work-at-Home Employment.

[29] Technical Memorandum 5.2.3, Model Interoperability and Interfaces with GIS Databases

[30] Technical Memorandum 5.2.1A NYMTC County Level Population and Labor Force Models, and TAZ Modeling Methodology, August 28, 2007

For the historical section of the model, data from 1970 through 1995 will be assumed from prior runs of the Labor Force Model, based upon data gathered under Tasks 1.1.1, *Population Data Collection And Analysis*, 1.1.2, *Employment Data Collection And Analysis*, and 1.1.3, *Labor Force Data Collection And Analysis*.[31] Where additional data collection was necessary specifically for the present task, such as labor force and unemployment levels in 2000 and 2005, data were compiled for Technical Memorandum 1.1.5.8, *2002-2005 County/Subregion Time Series Data for Existing Models*[32]. For the forecast section, some of the necessary inputs will be derived from the outputs of other models; others are the official forecasts of government sources. Where necessary, estimates will be made for some inputs based on historical data.

Tables 14 to 16 present current estimates of the resident civilian labor force, employed residents, and the unemployed by county and subregion for the period 2000 to 2005. Data were compiled from the *Local Area Unemployment Survey (LAUS),* a cooperative program of the U.S. Bureau of Labor Statistics and the state Departments of Labor. As mentioned previously, decennial Census enumerations of civilian labor force were not adopted for the year 2000 because of acknowledged deficiencies in the recorded level of participation. However, labor force participation rates by age/sex/race-ethnicity, as well as net commutation and work-at-home employment, have been compiled from the 2000 *Census of Population* and the 2005 *American Community Survey (ACS)* for the respective areas (see Tables 18 and 19). These indicators, expressed in relation to the *Census* and *ACS* civilian labor force levels, will be incorporated in the model, adjusted for relative differences in the level of civilian labor force for 2000 and 2005.

Table 14. Trends in Civilian Labor Force by County & Subregion in the New York Metropolitan Region, 2000 to 2005, (in 000s)

	Civilian Labor Force					
AREANAME	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**
NEW YORK CITY	**3,666.0**	**3,673.9**	**3,730.7**	**3,721.6**	**3,713.3**	**3,734.0**
Bronx	486.6	488.8	501.8	504.2	502.7	503.6
Kings	1,042.6	1,042.1	1,060.1	1,055.5	1,048.7	1,052.7
New York	855.3	862.2	871.8	867.8	869.0	875.3
Queens	1,064.3	1,062.5	1,074.0	1,069.4	1,066.8	1,074.6
Richmond	217.2	218.3	223.0	224.7	226.1	227.8
LONG ISLAND	**1,413.8**	**1,419.1**	**1,439.6**	**1,449.1**	**1,462.4**	**1,474.0**
Nassau	677.9	676.3	683.6	684.9	689.2	694.6
Suffolk	735.9	742.8	756.0	764.2	773.2	779.4
MID HUDSON	**1,086.7**	**1,095.6**	**1,114.0**	**1,122.5**	**1,137.0**	**1,147.3**
Dutchess	139.4	141.3	143.8	144.4	146.2	147.9
Orange	162.4	165.4	169.7	173.2	176.4	178.3
Putnam	52.4	52.9	54.0	54.6	55.6	56.2
Rockland	145.3	145.6	147.7	149.1	150.8	152.3
Sullivan	33.2	33.6	34.5	34.6	35.1	35.4
Ulster	88.7	89.0	91.2	92.0	92.1	92.2
Westchester	465.3	467.8	473.1	474.6	480.8	485.0

[31] Technical Memorandum 1.1.1., Population Data Collection And Analysis, August 31, 2000; Technical Memorandum 1.1.2 Employment Data Collection And Analysis, September 29, 2000; Technical Memorandum 1.1.3, Labor Force Data Collection And Analysis, August 31, 2000

[32] Technical Memorandum 1.1.5.8, 2002-2005 County/Subregion Time Series Data for Existing Models, October 2, 2007

NEW JERSEY	**3,395.6**	**3,410.6**	**3,452.4**	**3,456.0**	**3,443.6**	**3,478.7**
Bergen	469.2	467.0	465.9	466.3	464.8	468.9
Essex	366.1	366.1	371.1	370.4	364.7	366.3
Hudson	297.8	296.9	296.0	293.1	289.0	290.6
Hunterdon	67.9	68.7	70.1	70.8	70.8	71.2
Mercer	180.2	182.6	186.4	188.0	190.4	194.7
Middlesex	401.2	405.6	412.8	412.0	412.5	418.6
Monmouth	318.1	320.2	325.1	324.6	323.6	328.1
Morris	261.3	261.3	265.2	266.7	266.2	267.8
Ocean	232.1	236.0	242.8	245.3	246.7	250.2
Passaic	236.3	236.1	236.6	237.4	235.5	236.9
Somerset	165.5	167.7	171.3	171.6	172.7	175.4
Sussex	79.5	80.2	81.8	82.9	82.6	83.1
Union	264.7	265.5	269.4	268.9	265.9	267.3
Warren	55.7	56.7	57.9	58.0	58.2	59.6
CONNECTICUT	**994.6**	**975.9**	**993.5**	**992.1**	**990.0**	**995.3**
Fairfield	462.3	453.8	458.3	457.7	455.4	457.6
Litchfield	103.0	101.0	100.9	102.7	102.9	103.5
New Haven	429.3	421.1	434.4	431.7	431.7	434.2
REGION	**10,556.7**	**10,575.1**	**10,730.2**	**10,741.3**	**10,746.3**	**10,829.3**

Source: *LAUS* of NY/NJ/CT Departments of Labor

Table 15. Trends in Resident Employed Labor Force by County & Subregion in the New York Metropolitan Region, 2000 to 2005, (in 000s)

	Employed Labor Force					
AREANAME	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**
NEW YORK CITY	**3,453.6**	**3,451.5**	**3,422.2**	**3,413.5**	**3,451.8**	**3,518.6**
Bronx	451.8	452.7	452.9	451.5	457.1	465.9
Kings	976.1	973.8	960.3	960.5	968.6	987.4
New York	811.9	812.6	804.5	802.5	814.9	830.7
Queens	1007.7	1005.3	996.9	990.7	999.2	1018.5
Richmond	206.1	207.1	207.6	208.3	212	216.1
LONG ISLAND	**1,366.2**	**1,365.4**	**1,371.8**	**1,380.0**	**1,395.2**	**1,412.9**
Nassau	655.5	651	651.5	652.6	657.9	666.2
Suffolk	710.7	714.4	720.3	727.4	737.3	746.7
MID HUDSON	**1,050.2**	**1,054.9**	**1,064.2**	**1,071.6**	**1,086.3**	**1,100.3**
Dutchess	134.9	136.4	137.7	138.5	140.1	142.1
Orange	156.9	159.3	162.3	165.1	168.3	170.8
Putnam	50.9	51.2	51.8	52.4	53.5	54.2
Rockland	140.6	140.4	141.2	142.1	144.1	146
Sullivan	31.7	32	32.7	32.8	33.3	33.7
Ulster	85.6	85.6	87.1	87.8	87.7	88.2
Westchester	449.6	450	451.4	452.9	459.3	465.3
NEW JERSEY	**3,274.3**	**3,266.0**	**3,252.3**	**3,254.6**	**3,277.6**	**3,328.8**
Bergen	454.4	449.5	442.4	442.6	445.2	451.7
Essex	349.6	346.3	344.1	343.2	342.6	345.9
Hudson	283.4	280.3	273.2	271.3	270.9	274.9
Hunterdon	66.4	66.8	67.2	67.8	68.5	69.1

Mercer	174.3	175.8	176.7	178.3	182.2	187.2
Middlesex	387.8	389.3	389.7	389.3	393.7	401.4
Monmouth	307.9	308.1	308.0	307.1	309.1	315.1
Morris	254.4	252.7	253.0	254.5	256.8	259.2
Ocean	223.5	226.3	229.7	231.2	234.3	238.9
Passaic	225.5	223.3	219.9	220.0	220.8	224.0
Somerset	161.3	162.3	163.2	163.8	166.4	169.6
Sussex	77.2	77.3	77.8	78.6	79.2	79.9
Union	254.5	253.2	252.3	251.8	252.2	254.6
Warren	54.1	54.8	55.1	55.1	55.7	57.3
CONNECTICUT	**973.0**	**943.5**	**950.3**	**938.6**	**942.1**	**947.2**
Fairfield	453.4	439.9	439.8	434.8	435.1	437.3
Litchfield	101.1	98.0	96.6	97.5	98.1	98.8
New Haven	418.5	405.6	413.9	406.3	408.9	411.1
REGION	**10,117.3**	**10,081.3**	**10,060.8**	**10,058.3**	**10,153.0**	**10,307.8**

Source: *LAUS* of NY/NJ/CT Departments of Labor

Table 16. Trends in Resident Unemployed Labor Force by County & Subregion in the New York Metropolitan Region, 2000 to 2005, (in 000s)

	Unemployed Labor Force					
AREANAME	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**
NEW YORK CITY	**212.4**	**222.4**	**308.5**	**308.1**	**261.5**	**215.4**
Bronx	34.8	36.1	48.9	52.7	45.6	37.7
Kings	66.5	68.3	99.8	95.0	80.1	65.3
New York	43.4	49.6	67.3	65.3	54.1	44.6
Queens	56.6	57.2	77.1	78.7	67.6	56.1
Richmond	11.1	11.2	15.4	16.4	14.1	11.7
LONG ISLAND	**47.6**	**53.7**	**67.8**	**69.1**	**67.2**	**61.1**
Nassau	22.4	25.3	32.1	32.3	31.3	28.4
Suffolk	25.2	28.4	35.7	36.8	35.9	32.7
MID HUDSON	**36.5**	**40.7**	**49.8**	**50.9**	**50.7**	**47.0**
Dutchess	4.5	4.9	6.1	5.9	6.1	5.8
Orange	5.5	6.1	7.4	8.1	8.1	7.5
Putnam	1.5	1.7	2.2	2.2	2.1	2.0
Rockland	4.7	5.2	6.5	7.0	6.7	6.3
Sullivan	1.5	1.6	1.8	1.8	1.8	1.7
Ulster	3.1	3.4	4.1	4.2	4.4	4.0
Westchester	15.7	17.8	21.7	21.7	21.5	19.7
NEW JERSEY	**121.3**	**144.6**	**200.1**	**201.4**	**166.0**	**149.9**
Bergen	14.8	17.5	23.5	23.7	19.6	17.2
Essex	16.5	19.8	27.0	27.2	22.1	20.4
Hudson	14.4	16.6	22.8	21.8	18.1	15.7
Hunterdon	1.5	1.9	2.9	3.0	2.3	2.1
Mercer	5.9	6.8	9.7	9.7	8.2	7.5
Middlesex	13.4	16.3	23.1	22.7	18.8	17.2
Monmouth	10.2	12.1	17.1	17.5	14.5	13.0
Morris	6.9	8.6	12.2	12.2	9.4	8.6

Ocean	8.6	9.7	13.1	14.1	12.4	11.3
Passaic	10.8	12.8	16.7	17.4	14.7	12.9
Somerset	4.2	5.4	8.1	7.8	6.3	5.8
Sussex	2.3	2.9	4.0	4.3	3.4	3.2
Union	10.2	12.3	17.1	17.1	13.7	12.7
Warren	1.6	1.9	2.8	2.9	2.5	2.3
CONNECTICUT	**21.6**	**32.4**	**43.2**	**53.4**	**47.9**	**48.1**
Fairfield	8.9	13.9	18.5	22.9	20.3	20.2
Litchfield	1.9	3.0	4.2	5.2	4.8	4.7
New Haven	10.9	15.5	20.5	25.4	22.8	23.1
REGION	**439.4**	**493.8**	**669.4**	**682.9**	**593.3**	**521.5**

Source: *LAUS* of NY/NJ/CT Departments of Labor

Table 17. Aggregate Civilian Labor Force Participation Rates of Racial-Ethnic Population by County & Subregion in the New York Metropolitan Region, 2000 & 2005

% of Persons Aged 16+ in Civil Labor Force	*2000*			
	White nonHisp	*Black alone*	*Asian/Other*	*Hispanic*
New York City	***59.32%***	***57.40%***	***60.79%***	***53.75%***
Bronx	47.76%	55.10%	57.53%	49.01%
Kings	53.14%	58.01%	59.25%	52.59%
New York	74.35%	52.79%	63.24%	51.08%
Queens	53.71%	61.05%	60.63%	61.01%
Richmond	59.63%	62.11%	63.72%	62.46%
Long Island	***63.67%***	***64.86%***	***64.87%***	***67.78%***
Nassau	61.85%	65.37%	65.07%	67.98%
Suffolk	65.31%	64.13%	64.51%	67.59%
Mid Hudson	***64.04%***	***59.26%***	***66.41%***	***63.73%***
Dutchess	65.27%	52.04%	64.46%	57.83%
Orange	64.00%	55.78%	64.01%	62.75%
Putnam	69.46%	61.03%	71.09%	68.65%
Rockland	64.58%	66.31%	68.83%	68.60%
Sullivan	59.79%	37.90%	57.12%	52.77%
Ulster	64.93%	45.81%	60.71%	54.86%
Westchester	62.84%	61.88%	66.47%	64.71%
New Jersey	***64.12%***	***62.44%***	***67.55%***	***62.92%***
Bergen	63.53%	69.32%	65.53%	68.87%
Essex	62.26%	59.54%	66.74%	58.49%
Hudson	62.43%	60.52%	65.54%	58.76%
Hunterdon	71.05%	14.74%	69.27%	57.07%
Mercer	66.57%	60.75%	69.96%	62.84%
Middlesex	64.96%	69.61%	68.30%	67.19%
Monmouth	65.64%	61.86%	71.37%	64.69%
Morris	68.86%	69.91%	69.85%	72.33%
Ocean	54.77%	65.44%	65.18%	65.88%
Passaic	63.43%	59.22%	63.80%	60.39%
Somerset	68.89%	73.32%	71.64%	71.07%
Sussex	70.75%	74.46%	76.17%	69.24%
Union	62.00%	65.79%	68.88%	65.07%

Warren	67.75%	63.63%	62.59%	72.75%
Connecticut	***66.21%***	***65.90%***	***68.71%***	***64.36%***
Fairfield	65.71%	66.61%	68.82%	65.74%
Litchfield	69.53%	77.32%	73.30%	76.04%
New Haven	65.82%	64.95%	68.09%	61.89%
Region	***63.11%***	***59.26%***	***63.57%***	***58.26%***

Source: 2000 Census of Population, 2005 American Community Survey

Table 18. Trends in Work at Home of Resident Labor Force by County & Subregion in the New York Metropolitan Region, 2000 & 2005

	2000			***2005***			**2000-2005**	
	Workers 16 years +: Total	***Workers 16 years +: Worked at home***	***Ratio of Workers at home to Total***	***Workers 16 years +: Total***	***Workers 16 years +: Worked at home***	***Ratio of Workers at home to Total***	***Change of Workers Total***	***Change of Workers at Home***
New York City	***3,192,070***	***92,151***	***0.029***	***3,429,194***	***123,997***	***0.04***	**237,124**	**31,846**
Bronx	415,075	7,756	*0.019*	475,932	14,754	*0.03*	60,857	6,998
Kings	901,027	20,663	*0.023*	985,092	32,508	*0.03*	84,065	11,845
New York	753,114	43,853	*0.058*	780,503	46,830	*0.06*	27,389	2,977
Queens	931,709	16,673	*0.018*	992,791	25,813	*0.03*	61,082	9,140
Richmond	191,145	3,206	*0.017*	194,876	4,092	*0.02*	3,731	886
Long Island	***1,289,992***	***36,193***	***0.028***	***1,291,499***	***40,186***	***0.03***	**1,507**	**3,993**
Nassau	619,586	18,392	*0.030*	608,423	17,644	*0.03*	-11,163	-748
Suffolk	670,406	17,801	*0.027*	683,076	22,542	*0.03*	12,670	4,741
Mid Hudson	***997,717***	***35,861***	***0.036***	***1,032,220***	***46,539***	***0.05***	**34,503**	**10,678**
Dutchess	128,437	4,162	*0.032*	133,020	5,587	*0.04*	4,583	1,425
Orange	152,489	4,085	*0.027*	166,914	4,173	*0.03*	14,425	88
Putnam	48,167	1,584	*0.033*	48,135	1,252	*0.03*	-32	-332
Rockland	132,302	4,685	*0.035*	131,749	5,006	*0.04*	-553	321
Sullivan	29,544	1,090	*0.037*	34,540	1,347	*0.04*	4,996	257
Ulster	81,726	3,950	*0.048*	86,049	4,561	*0.05*	4,323	611
Westchester	425,052	16,305	*0.038*	431,813	24,613	*0.06*	6,761	8,308
New Jersey	***3,073,471***	***86,945***	***0.028***	***3,139,057***	***97,561***	***0.03***	**65,586**	**10,616**
Bergen	427,462	13,292	*0.031*	423,697	16,524	*0.04*	-3,765	3,232
Essex	328,214	9,106	*0.028*	336,256	10,424	*0.03*	8,042	1,318
Hudson	264,544	4,644	*0.018*	280,387	3,925	*0.01*	15,843	-719
Hunterdon	62,359	3,665	*0.059*	62,160	3,792	*0.06*	-199	127
Mercer	163,257	5,161	*0.032*	167,386	5,356	*0.03*	4,129	195
Middlesex	363,176	7,690	*0.021*	366,269	8,424	*0.02*	3,093	734
Monmouth	291,938	9,504	*0.033*	298,370	11,040	*0.04*	6,432	1,536
Morris	239,839	8,845	*0.037*	236,129	9,917	*0.04*	-3,710	1,072
Ocean	209,328	5,291	*0.025*	226,058	7,912	*0.03*	16,730	2,621
Passaic	210,378	4,493	*0.021*	216,796	3,252	*0.02*	6,418	-1,241
Somerset	151,284	5,438	*0.036*	157,295	5,820	*0.04*	6,011	382
Sussex	72,728	2,442	*0.034*	75,374	3,919	*0.05*	2,646	1,477

Union	238,606	5,692	*0.024*	241,001	5,543	*0.02*	2,395	-149
Warren	50,358	1,682	*0.033*	51,879	1,712	*0.03*	1,521	30
Connecticut	***901,221***	***32,219***	***0.036***	***911,177***	***31,988***	***0.04***	**9,956**	**-231**
Fairfield	419,237	18,964	*0.045*	419,902	17,636	*0.04*	665	-1,328
Litchfield	93,934	3,691	*0.039*	98,682	4,145	*0.04*	4,748	454
New Haven	388,050	9,564	*0.025*	392,593	10,207	*0.03*	4,543	643
Region	***9,454,471***	***283,369***	***0.030***	***9,803,147***	***340,271***	***0.03***	**348,676**	**56,902**

Source: 2000 Census of Population, 2005 American Community Survey

Table 19. Trends in Out Commutation of Resident Labor Force by County & Subregion in the New York Metropolitan Region, 2000 & 2005

	2000		***2005***	
	Worked in County of Residence	***Worked Outside County of Residence***	***Worked in County of Residence***	***Worked Outside County of Residence***
New York City	***1,685,614***	***1,506,456***	***1,881,344***	***1,547,850***
Bronx	168,903	246,172	206,079	269,853
Kings	431,559	469,468	501,412	483,680
New York	631,132	121,982	661,867	118,636
Queens	367,823	563,886	422,929	569,862
Richmond	86,197	104,948	89,058	105,818
Long Island	***851,534***	***438,458***	***862,983***	***428,516***
Nassau	359,698	259,888	348,626	259,797
Suffolk	491,836	178,570	514,356	168,720
Mid Hudson	***616,082***	***381,635***	***643,255***	***388,965***
Dutchess	88,963	39,474	88,325	44,695
Orange	99,901	52,588	104,488	62,426
Putnam	13,721	34,446	14,489	33,646
Rockland	72,022	60,280	78,918	52,831
Sullivan	19,922	9,622	23,453	11,087
Ulster	54,373	27,353	57,223	28,826
Westchester	267,180	157,872	276,360	155,453
New Jersey	***1,643,418***	***1,430,053***	***1,667,495***	***1,471,562***
Bergen	246,163	181,299	243,202	180,495
Essex	175,248	152,966	178,216	158,040
Hudson	121,352	143,192	127,856	152,531
Hunterdon	25,761	36,598	25,175	36,985
Mercer	112,449	50,808	113,488	53,898
Middlesex	201,811	161,365	192,657	173,612
Monmouth	175,070	116,868	182,602	115,768
Morris	138,737	101,102	137,191	98,938
Ocean	120,741	88,587	132,470	93,588
Passaic	95,790	114,588	99,943	116,853
Somerset	66,341	84,943	69,682	87,613
Sussex	29,658	43,070	33,541	41,833
Union	113,263	125,343	111,342	129,659
Warren	21,034	29,324	20,129	31,750
Connecticut	***676,977***	***224,244***	***666,319***	***244,858***
Fairfield	335,378	83,859	327,524	92,378

Litchfield	51,501	42,433	52,203	46,479
New Haven	290,098	97,952	286,593	106,000
Region	***5,473,625***	***3,980,846***	***5,721,397***	***4,081,750***

Source: 2000 Census of Population, 2005 American Community Survey

As mentioned previously, subregional Labor Force Models incorporate submodels for each racial/ethnic group. These sub-models are interdependent where necessary to aggregate racial/ethnic shares of overall employment, commutation, and work-at-home employment to subregional totals. For the sake of clarity, this aggregation process is described together with the discussion of each input, below.

Population

The Labor Force Model depends on the Population Model for inputs, at each five-year interval, of population by sex and age cohort for all persons 16 years and older, by racial/ethnic group. The eight age cohorts are determined by their differences in labor force participation, as follows:

- Age 16 - 19: Teenage workers.
- Age 20 - 24: Recent high school and college graduates.
- Age 25 - 34: Young labor force.
- Age 35 - 44: Prime labor force.
- Age 45 - 54: Middle labor force.
- Age 55 - 64: Mature labor force.
- Age 65 - 74: Early Retirement Ages.
- Age 75+ Elderly Retirement Ages

Labor Force Participation Rates

The Labor Force Participation Rate is defined as the percentage of all residents of a particular population group who are in the Civilian Labor Force, as expressed in the equation:

$$LFPR = \frac{CLF}{POP}$$

where LFPR equals Labor Force Participation Rate, CLF equals the Civilian Labor Force and POP equals the resident population. Civilian Labor Force includes both employed and unemployed workers, and excludes military personnel and all other residents who are not in the labor market. Labor Force Participation Rates for the historical section of the model were calculated by age, sex, and racial/ethnic characteristics of the subregional population, based on data from the decennial Census[33]. Rates for the forecast section were benchmarked on national forecasts prepared by the US Bureau of Labor Statistics.

Historical racial/ethnic Labor Force Participation Rates were derived for the five subregions in 1970, 1980, 1990 and 2000 from decennial Census files. Rates for 2005 through 2035 will be benchmarked

[33] It should be noted that all population figures used in the calculation of Labor Force Participation Rates are based upon the Census Bureau's STF data set, whereas the data described in the section above on population inputs rely on the Bureau's Modified Age, Race, Sex (MARS) data set for 1980 and 1990. The MARS data set incorporates adjustments to the STF data in order to improve allocation by racial/ethnic group and age. However, since the Census's labor force figures rely on unadjusted STF population data, STF data were used in the calculation of Labor Force Participation Rates for the sake of consistency.

on the differences between national and regional rates in 2000. Rates for 1990 and 2000 were derived from the Census Bureau's *Census/Equal Employment Opportunity (EEO) Special File* CD-ROMs. These files include detailed age- and sex-specific CLF figures that were aggregated by the appropriate racial/ethnic groups. Since the 1980 Census EEO file did not provide the same extensive labor force participation data by racial/ethnic, age/sex and geographic detail, and the model's need for such detail was less significant on a historical basis, both 1970 and 1980 rates were developed from decennial Census data published in the report *Characteristics of the Population* Chapter B, "General Population Characteristics." These Census sources provide county-level CLF and/or LFPR data, with limited age-, race- and sex-specific figures depending on the Census year.

There is no source for detailed intercensal labor force estimates by racial/ethnic group, sex and age for small areas comparable to the Census Bureau's annual series of population estimates. Therefore it has been necessary to estimate post-2000 rates based on the 2000 decennial Census figures, in combination with national estimates and forecasts of Labor Force Participation Rates. It should be noted, however, that the outcome of applying both the detailed 2000 Census EEO rates, and the indexed 2005 rates, will be controlled by the aggregate civilian labor force derived by subregion from the *LAUS* program, as shown previously in Table 14.

Labor Force Participation Rates for 2005, and subsequently for 2010 to 2035 by five year interval, were obtained for the nation as a whole from the Bureau of Labor Statistics, as released in a *Monthly Labor Review* publication.[34] Detailed age-sex data were provided for all races and for white non-Hispanic, Black, Asian and Hispanic groups. This racial-ethnic detail was provided on an annual basis and selected on a five year incremental basis for the eight labor force age classifications, noted by the population data, by male and female rates for white nonHispanics, Black, Asian and Hispanic groups. From 2005 to 2035, the national level estimates will be indexed on detailed 2000 subregional data for purposes of forecasting future labor force participation by race-ethnicity.

The increasing relative -- as well as absolute -- importance of minorities in the Region's labor force is considered likely to encourage greater supply, just as the declining importance of earlier dominant sources of labor has correlated with reductions in their Labor Force Participation Rates. However, the new rates are anticipated to reveal increases in labor force participation among older workers, even as the baby boom generation exits the prime-aged workforce. Appendix Tables 3.1 presents the national labor force participation rates by age/sex/race-ethnicity from 2005 to 2035.

Unemployment Rates

Unemployed persons are defined as those who are in the Civilian Labor Force but are not currently working; they are counted based on unemployment claims. Unemployment figures do not include certain groups of persons who are not employed for a variety of reasons and are not considered part of the labor force, such as the disabled or the long-term unemployed who have stopped seeking work. The unemployment rate is defined as the percentage of unemployed persons in the Civilian Labor Force:

$$UNEMP\_RATE = \frac{UNEMP}{CLF}$$

34 Toossi, Mitra. "A New Look at Long Term Labor Force Projections to 2050," *Monthly Labor Review,* November 2006 (U.S. Bureau of Labor Statistics). Detailed labor force participation rates that were summarized in this report are contained on the website: www.bls.govl, under "Detailed Data for Researchers: Labor Force."

The Labor Force Model requires Unemployment Rate inputs at each interval for each racial/ethnic group as a whole; age- and sex-specific Unemployment Rates are not required.

Historical unemployment inputs have been derived from a combination of US Census and state Department of Labor sources. The decennial Census provides county-level unemployment data by race and ethnicity for the years 1970, 1980, 1990 and 2000, although reporting is incomplete and reflects problems with racial/ethnic categorizations similar to those cited for Labor Force Participation Rates. State DOLs provide unemployment data for intercensal years, but not by racial/ethnic group. Table 16 provides trends in the unemployed by county and subregion, from 2000 to 2005.

For each Census year, subregional Unemployment Rates were derived by race-ethnicity from county-level data as the number of unemployed persons divided by the size of the CLF:

$$UNEMP\_RATE_s = \frac{\sum UNEMP_c}{\sum CLF_c}$$

where *c* denotes county-level figures and *s* denotes subregional-level figures.

Rates for 1975, 1985, 1995 & 2005

As discussed above, the source used for unemployment rates on a racial/ethnic basis is the US decennial Census. Estimates of total and unemployed civilian labor force are available from the state Departments of Labor (DOL) at the county level for all years through 2005, but not by racial/ethnic group. Subregional unemployment rates were calculated from these data using the formula above. These total subregional figures were adjusted by racial/ethnic group on the assumption that the ratio of the group-specific Unemployment Rate to that of the Civilian Labor Force as a whole would be the same as for the preceding Census year. This can be expressed as the formula:

$$UNEMP\_RATE_{ir} = UNEMP\_RATE_{it} \times \frac{UNEMP\_RATE_{cr}}{UNEMP\_RATE_{ct}}$$

where *i* refers to the intercensal year, *c* refers to the preceding Census year, *r* refers to a race-specific Unemployment Rate for the given year, and *t* refers to the Unemployment Rate for the Civilian Labor Force as a whole.

Unemployment Rates for Forecast Years

Unemployment rate estimates were prepared for the years 2010 through 2035 as part of the Employment Model. These estimates are conceptually comparable to the state DOL figures used for intercensal years; however, total and unemployed labor force figures were not developed as part of the Employment Model. It was therefore necessary to calculate subregional rates as the weighted average of the county rates.

The 2000 decennial Census rates by racial/ethnic group will then be indexed to the subregional rates to produce racial/ethnic Unemployment Rate estimates through the year 2035. Conceptually this is identical to the method described for intercensal years, above. Finally, racial/ethnic rates will be

readjusted so that the resulting total unemployment rates (based on the weighted average of the racial/ethnic rates) are reconciled to the Employment Model rates.

Employment

The employment inputs will reflect the number of available jobs in a given subregion at each five-year interval, and combine the amount of nonfarm payroll employment and the number of proprietors. Nonfarm employment includes jobs in ten major industrial sectors, and several subsectors as defined by NAICS: Natural Resources/Mining/Construction; Manufacturing; Trade, Transportation and Utilities; Information Services; Financial Activities; Professional and Business Services; Education and Health Services; Leisure and Hospitality; Other Services; and Government. Proprietors include self-employed persons, partners in non-limited partnerships, and non-farm proprietors.

The Labor Force Model will incorporate employment figures for each five-year interval from 1970 to 2005, as derived from the Employment Model. Within the Labor Force Model it will be necessary to disaggregate these figures for inclusion in each racial/ethnic sub-model. For purposes of disaggregation it will be assumed that the racial/ethnic distribution of total employment is proportional to that of the supply of local commutation-adjusted workers within each subregion, which will reflect their relative differences in unemployment. This figure will be generated in the supply-side of the Labor Force model for each racial/ethnic group based on the size of the labor force, Unemployment Rates, and net-commutation levels.

For the forecast years, the Labor Force Model will depend on the Employment Model for inputs of nonfarm employment and proprietors. The Employment Model will generate annual subregional totals for each of these groups. These totals will be disaggregated by racial/ethnic group following the same methodology used for the historical years, as described above.

Work-at-Home Employment

Table 18 presents total Work-at-Home employment by subregion for 2000 and 2005. Recent data show an escalation in the absolute and relative share of regional employment performed at home, from 3.0 to 3.5 percent between 2000 and 2005. Based on Work-at-Home employment from 1970 through 2005, future levels will be forecasted at the county-level by trend analysis. Results on a regional and subregional basis will be evaluated against current research findings at the national level. Work-at-Home employment will be incorporated in the Labor Force Model because of the necessity to exclude workers that do not generate commuting trips from transportation modeling.

For the Census years, historical Work-at-Home levels can be derived by county for each subregion from Census county-to-county journey-to-work data. Within the Labor Force Model, it will be necessary to disaggregate these total figures for inclusion in each racial/ethnic sub-model. For disaggregation purposes, it will be assumed that Work-at-Home levels are proportional to the racial/ethnic distribution of residents employed within the subregion. For the intercensal years, with the exception of 2005, Work-at-Home levels will be estimated based on the preceding and following Census years.

Total Work-at-Home employment for the forecast years will be forecasted at the county level, using *ordinary least squares regression analysis* with 1970-2005 data, a timeline and other variables such as the forecasted number of proprietors. These figures will be disaggregated by racial/ethnic group using the same methodology as for the historical period, described above.

Dual Job Rate

To account for workers holding two or more jobs within the jobs-labor force matching process in the forecast years, the Labor Force Model will apply a Dual Job Rate for each racial/ethnic group at every five-year interval. The Dual Job Rate for all forecast years will be calculated based on the average for all historical periods. The latter are calculated in the reconciliation of historical labor force and employment series as the ratio of locally-employed labor force to local trip-based employment.

Net Commutation

Net Commutation figures are input to the Labor Force Model at each five-year interval as part of the subregional labor force-employment match. They are calculated at the subregional level and are defined as the difference between the number of non-resident workers commuting into the subregion and the number of resident workers commuting out of the subregion. Net Commutation levels are positive for New York City and negative for the other subregions, reflecting the continued importance of the Manhattan CBD as an employment center. Table 19 presents trends in gross out-commutation by county and subregion of residence between 2000 and 2005. The data show an increase of one hundred thousand out-commuters in the Region over the five year period, compared to nearly a quarter million increase in intra-county work trips. Gross in-commutation data are not available for 2005, and the prior 2000 data source is acknowledged as deficient by the U.S. Bureau of the Census.

For Census years, historical Net Commutation levels of each subregion are available by county from the Census county-to-county Journey-to-Work flows of the *Census Transportation Planning Package (CTPP).* Prior versions of the Labor Force Model had net commutation levels for 1970, 1980 and 1990. Appendix Table 3.2 provides the 2000 Journey-to-Work flows. Labor Force Model net commutation numbers do not always correspond to Census numbers because of several adjustments: Census flow data reflect travel patterns of respondents during a Spring week of the decennial year. The labor force model adjusts spring travel to reflect commutation between place of work and place of residence on an annual average basis.

In addition, 2000 Journey-to-Work flows will need to be evaluated from the undercount perspective. The acknowledged deficiency in enumeration of resident labor force, by the Census Bureau, has resulted in an inaccurate representation of gross in- and out-commutation flows. Using secondary data sources, such as hub bound travel, together with the *LAUS* reported levels of civilian labor force, a tentative correction will be made to the 2000 gross commutation flows. This adjustment will be reviewed with regional transit agencies and assumed for Labor Force Model purposes until such further research is performed under separate contract to precisely model the 2000 county-to-county work trip flows.

Within the Labor Force Model it will also be necessary to disaggregate total figures for incorporation into each racial/ethnic sub-model. For the purpose of disaggregation it will be assumed that Net Commutation levels are proportional to the racial/ethnic distribution of employed workers within the subregion, as described in the formula:

$$NETCOM_R = NETCOM_{TOTAL} \times \frac{EMP_R}{EMP_{TOTAL}}$$

where *NETCOM* indicates subregional Net Commutation, *EMP* indicates subregional resident employed workers, *R* indicates a given racial/ethnic group, and *TOTAL* indicates the total for all

racial/ethnic groups. For the intercensal years, Net Commutation levels will be estimated based on the preceding and following Census years.

Total Net Commutation levels will be forecasted for each racial/ethnic group at every five-year interval in relation to two factors: first, the number of employed workers in the group at the given time period; second, historical ratios of Net Commuters to employed workers. Because of the difficulties of forecasting Net Commutation ratios by racial/ethnic group, the average historical ratio between Net Commuters and employed workers in Census years, a constant, will be used for each racial/ethnic group. This can be expressed as the equation:

$$NETCOM_i = EMPL_i \times \frac{(NETCOM_{1970} + NETCOM_{1980} + NETCOM_{1990})}{(EMPL_{1970} + EMPL_{1980} + EMPL_{1990})}$$

where *i* is the forecast year. A deflation factor will be applied to results for 1995, in order to account for the slump in New York City employment during that period.

1.4.2 METHODOLOGY

The methodology described below applies to the racial/ethnic sub-models incorporated into each subregional model. As mentioned above, each subregional model incorporates submodels for every racial/ethnic group. These sub-models are interdependent where necessary to aggregate racial/ethnic shares of overall employment, Net Commutation, and Work-at-Home employment to subregional totals. For the sake of convenience, this aggregation process is described in the above section. Thus, all figures discussed below are for individual racial/ethnic groups.

The methodology will involve a three-step process. First, an initial, unadjusted estimate will be made of Civilian Labor Force, by age and sex, and in total. This will correspond to the supply of labor that would be available based on prior-period population, natural increase, and historical rates of net migration, not modified by the anticipated demand for labor. Second, this expected supply of laborers will be matched against the expected demand for workers, input from the Employment Model, to determine if there will be a surplus or deficit of workers. Any such surplus or deficit will be assumed to induce a net in- or out-migration of an equal number of workers. Finally, this net migration figure will be disaggregated by age/sex group and added to the initial CLF figures to yield an adjusted CLF for each group. The disaggregated net migration will also become an input to the Population Model, where it will be factored up to population by application of the labor force participation rate (LFPR) and used to adjust net migration levels forecasted within that model.

Unadjusted Civilian Labor Force Forecast

Unadjusted Civilian Labor Force will be calculated separately for each age/sex group, based upon the forecasted 'closed' population and Labor Force Participation Rates for each time period. The CLF is simply the product of these two figures:

$$CLF_i = POP_i \times LFPR_i$$

where *i* represents the age/sex group. The total Civilian Labor Force is the sum of these age/sex-specific Civilian Labor Force figures:

$$CLF = \sum CLF_i$$

Labor Force-Employment Match

In the labor force-employment match, the forecasted labor supply will be compared with demand, with any difference forming the basis of an induced in or out-migration of workers. For purpose of this matching process, labor force supply will be defined as Local Employment and demand as Primary Jobs, which will be calculated as follows.

Local Employment will be defined as equal to total Civilian Labor Force after unemployed workers have been excluded and net in- or out-commuters have been accounted for; that is:

$$LOCALEMP = CLF - UNEMP + NETCOM$$

where *CLF* is the total Civilian Labor Force carried from above, *UNEMP* is the total number of unemployed workers and *NETCOM* is the net number of commuters (a positive value if there is net in-commutation and a negative value if there is net out-commutation).

The number of Primary Jobs will be calculated by first determining the level of Trip-Based Employment, which will be equal to the sum of all nonfarm employees and individual proprietors less the level of Work-at-Home employment, as follows:

$$TRIPBASED = (NONAG + PROP) - WORKATHOME$$

The number of Primary Jobs will then be calculated by excluding secondary jobs from the Trip-Based Employment. This will be done by dividing Trip-Based Employment by the Dual Job Rate (i.e., the ratio of all jobs to Primary Jobs):

$$PRIMJOB = TRIPBASED \div DUALJOB$$

The Dual Job Rate for forecast years will be calculated in the Model's historical section. The rate for each historical interval will be calculated as the ratio of Trip-Based Employment to Local Employment. The Dual Job Rate for forecast years will be calculated as the average of rates for historical years on a racial/ethnic basis by subregion.

Jobs-Labor Force Match

In the jobs-labor force match, a net in- or out-flow of workers will be induced by comparing local employment with the forecasted number of primary jobs. The level of net migration will be calculated by subtracting the former from the latter,

$$NETMIG = PRIMJOB - LOCAL\_EMP$$

resulting in a positive figure if jobs exceed labor force (generating a net in-flow of workers) and a negative figure if there are insufficient jobs for local workers (generating a net out-flow of workers).

Disaggregation of Induced Net Migration by Age & Sex

Age-group allocation of induced in- or out-migration will be based upon the age-group distribution of the initial unadjusted labor force estimate for each racial/ethnic group. In the historical period there will be no induced net migration calculation since this dynamic is incorporated in the residual of population growth and natural increase by each age-group.

Adjustment of Labor Force Net Migration to Population Net Migration

The migration of workers will bring with it an additional migration of non-workers. Thus, the labor force net-migration figures, discussed above, must be adjusted for this additional migration before incorporation into the Population. Total net-migration will be calculated based upon age/sex-specific figures for net-migration of workers and LFPR, following the equation:

$$NETMIG_{Ti} = NETMIG_{Wi} \div LFPR_i$$

where *T* denotes total net migration, *W* denotes net-migration of workers, and *i* denotes age/sex group. These age/sex-specific figures will be incorporated as inputs into the Population Model.

Additional net in-migration of dependent children of adult workers will not be included because of the lack of availability of historical data on which to base forecasts by race/ethnicity.

1.4.3 COUNTY-LEVEL DISAGGREGATION & MODELING

Separate Labor Force Models will be developed for NYMTC partner counties, and described in TM 5.2.1, *NYMTC County Level Population and Labor Force Models, and TAZ Modeling Methodology*.[35] A method will be devised to integrate these models into the subregional Labor Force Models. For all other counties in the Region, subregional totals will be disaggregated to component counties using regression analysis procedures. Historical labor force estimates of the state DOLs will be related to resident population estimates of the Census Bureau on an annual basis by county. SPSS, a software package of SPSS, Inc., provides statistical data analysis capability to fit linear, cubic and quadratic equations to the population-labor force relationship of each county. Using time as an independent variable, the best fitting form of equation will then be used to predict future population-labor force ratios on a county-by-county basis.

For county-level disaggregation, ratios will be applied to the predicted level of resident population over the forecast period, to yield initial estimates of resident civilian labor force by county. Prevailing county-to-subregion differences in resident unemployment rates will be continued over the forecast period and applied to the labor force forecasts for future estimates of employed residents. The component county estimates will then be summed to subregional totals and normalized to equal the subregional controls. County-level forecasts will be made for total and employed labor force only, with no race, age, or sex detail.

[35] Technical Memorandum 5.2.1A, NYMTC County Level Population and Labor Force Models, and TAZ Modeling Methodology, August 28, 2007.

The resulting forecasts of civilian labor force and employed resident labor force will be evaluated for reasonableness by county. Modifications will be made to smooth trends in relation to state DOL time series and forecasts will be reviewed by state and county agencies.

1.5 HOUSEHOLD FORMATION MODEL

The Household Formation and Housing Stock Preference Model represents a recalibration and extension of work performed under the Transportation Models and Data Initiative (TMDI) project. This project was a major program undertaken by NYMTC to forecast the transportation needs of the New York Metropolitan Region through the year 2020. The subsequent Demographic and Socioeconomic Forecasting project extended the forecast period for the Household Formation Model to 2025, 2030, and now 2035.

Inputs to the Household Formation Model will include results of the Population Model and historical data collected on housing for the period 1970 through 2005. In addition, the Model will incorporate state, county, and national level household data drawn from the U.S. Census sources including the decennial censuses, the 2005 American Community Survey, several Public Use Macrodata Samples (PUMS), and the March Supplement of the Current Population Survey.

Model outputs will be produced at the subregional level, aggregated by mutually exclusive racial/ethnic group, age of head of householder, household type or composition, household size, and household income. Total households will then be disaggregated to the county level.

The household model will estimate the future number of households as the product of two factors: household population by age group, and age-specific household formation rates. The latter are also known as headship rates, and represent the share of householders (or household heads) in a given age group. These age-specific household estimates will be summed to produce the total household estimate for a given subregion and racial/ethnic group. Grouping of households by type, size and income range will be performed by using a series of matrices that crosstabulate the historical and estimated rates of these parameters. The disaggregation of subregional household forecasts to the county level will be based upon the projection of county household trends, calculated as a function of forecasted household population and anticipated trends in average household size.

Under prior versions of the Household Formation Model, some model inputs were projected through the year 2000 and held constant thereafter, while others were projected through 2020. Because of the lack of new source data on which to extend these projections further, these projection periods have been retained for the current task. Thus, Household Formation Model forecasts for the year 2035 reflect the year 2000 or 2020 inputs as appropriate. For household forecasts by household income range, outputs of the Household Formation Model will be calibrated to the external forecasts of aggregate income, produced by the Employment Model. The current forecast series will add projections of housing stock preference and tenure to the household forecasts. These preference and tenure forecasts will be linked to the forecasts of household income by range.

1.5.1 DATA INPUTS

Release of the detailed tabulations from the Census 2000 long form questionnaires of Summary File 4 (SF4), and the 2005 estimates of the American Community Survey by county provide the basis for current inputs to the Household Formation Model. The 2000 Census file represents a complete cross-tabulation of household characteristics by racial-ethnic detail, while the 2005 American Community Survey provides aggregate updates of major attributes. Worked together, with inputs from the

Population Model, these data sources can provide a comprehensive and current characterization of family and nonfamily households by age of head, type and size of household, income bracket, tenure and housing preference. The following tables provide an overview of household and housing trends in the Region, as well as current patterns of household formation. The Model will contain comparable data by subregion on a racial-ethnic basis.

Table 20. Trends in Household Population & Households by County & Subregion in the New York Metropolitan Region, 2000 to 2005 *(in 000s of persons & households)*

	Household Population						Households		Ave Hsld Size	
AREANAME	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2000**	**2005**	**2000**	**2005**
NYC	**7,825.50**	**7,890.90**	**7,921.90**	**7,943.70**	**7,977.60**	**8,028.70**	**3,021.60**	**3,052.70**	**2.59**	**2.63**
Bronx	1,285.40	1,299.70	1,311.50	1,315.90	1,314.60	1,317.10	463.2	470.4	2.78	2.8
Kings	2,425.80	2,434.60	2,439.10	2,443.20	2,457.60	2,471.00	880.7	892.1	2.75	2.77
New York	1,477.20	1,498.00	1,503.20	1,515.50	1,527.50	1,544.20	738.6	738.9	2	2.09
Queens	2,202.50	2,217.20	2,221.60	2,217.90	2,224.50	2,230.50	782.7	788.2	2.81	2.83
Richmond	434.5	441.4	446.6	451.1	453.5	465.9	156.3	165.8	2.78	2.81
LONG ISLAND	**2,703.70**	**2,727.20**	**2,742.20**	**2,754.60**	**2,758.90**	**2,749.60**	**916.7**	**919.6**	**2.95**	**2.99**
Nassau	1,312.90	1,314.80	1,316.80	1,317.10	1,314.60	1,307.70	447.4	435.9	2.93	3
Suffolk	1,390.80	1,412.30	1,425.40	1,437.50	1,444.20	1,441.90	469.3	483.9	2.96	2.98
MID HUDSON	**2,097.70**	**2,124.10**	**2,144.30**	**2,162.30**	**2,172.40**	**2,184.40**	**772**	**785.8**	**2.72**	**2.78**
Dutchess	262	266.4	269.8	273.2	275.4	276	99.5	101.8	2.63	2.71
Orange	327.7	335.5	342.7	350.5	355.7	358.6	114.8	123.3	2.85	2.91
Putnam	93.6	95.3	96.5	97.5	98.2	98.3	32.7	34.5	2.86	2.85
Rockland	279.1	281.5	283.5	285.1	285.2	286.8	92.7	93.4	3.01	3.07
Sullivan	69.1	69.2	69.5	70.5	71.3	71.4	27.7	29.1	2.5	2.45
Ulster	166.5	167.1	168.8	169.9	170.1	170.7	67.5	68.3	2.47	2.5
Westchester	899.8	909.1	913.5	915.7	916.5	922.6	337.1	335.5	2.67	2.75
NEW JERSEY	**6,519.30**	**6,595.90**	**6,647.70**	**6,691.90**	**6,720.20**	**6,730.40**	**2,423.20**	**2,465.40**	**2.69**	**2.73**
Bergen	872.8	879.7	883.4	887	890.4	891	330.8	332.5	2.64	2.68
Essex	772.2	773.3	774.9	775.8	773.6	767.7	283.7	283.3	2.72	2.71
Hudson	599.5	603.5	600.3	598	595.9	593.5	230.5	229.2	2.6	2.59
Hunterdon	117.6	120.6	122.4	123.8	125	125.8	43.7	46.1	2.69	2.73
Mercer	329.7	333.6	336.4	340	343.3	344.9	125.8	127.3	2.62	2.71
Middlesex	729.3	741.1	749.2	756.8	762.8	768.5	265.8	267.8	2.74	2.87
Monmouth	605.3	612.6	618.2	623	625.1	624.8	224.2	230.5	2.7	2.71
Morris	461	464.9	469.5	474	478	480.6	169.7	172.3	2.72	2.79
Ocean	503	516.4	529.4	538.9	545.2	550.3	200.4	221	2.51	2.49
Passaic	477.7	483.3	485.7	487.4	487.6	485.7	163.9	163	2.92	2.98
Somerset	293	298.7	303.4	307.6	311.4	314.9	109	113.7	2.69	2.77
Sussex	142.5	144.9	147.1	149.3	150.4	151	50.8	55.1	2.8	2.74
Union	514.7	519.1	521.7	522.7	523	522.9	186.1	182.2	2.77	2.87
Warren	101	104.2	106	107.7	108.3	108.9	38.7	43.2	2.61	2.52
CONNECTICUT	**1,840.60**	**1,855.20**	**1,865.90**	**1,878.90**	**1,885.10**	**1,885.40**	**714.8**	**722.4**	**2.57**	**2.61**
Fairfield	864.6	871.6	875.9	880.8	883.1	882.6	324.2	324.5	2.67	2.72
Litchfield	179.7	181.9	183.9	185.4	186.6	186.8	71.6	73.6	2.51	2.54

New Haven	796.3	801.7	806.2	812.7	815.4	816	319	323.8	2.5	2.52
REGION	**20,986.90**	**21,193.20**	**21,322.10**	**21,431.30**	**21,514.20**	**21,578.50**	**7,848.30**	**7,923.70**	**2.67**	**2.72**

Source: TM 1.1.5.8, based on United States Census. Bureau

Table 20 depicts the increase in household population, the number of households and the average household size regionwide, by subregion and county over the 2000 to 2005 period. These changes will be reflected in the model calibration of the historical section, including the racial-ethnic submodels with corresponding racial-ethnic data. The extensive characterization of households revealed by 2000 Census cross-tabulations will establish the basis for patterning racial-ethnic household formation in 2005, controlled by the aggregate estimates of the American Community Survey. Tables 21 to 23 portray the relationships that will enable the Household Formation Model to generate 2005 estimates of subregional households by age of head, type and size of household, and household income bracket. These tables which present data on an all races basis are available, and will be applied on a racial-ethnic basis.

Table 21. Household Formation by Age of Head & Type of Household, by Subregion in the New York Metropolitan Region, 2000

Age of Head by Household Type	***New York City***	***Long Island***	***Mid Hudson***	***New Jersey***	***Connecti -cut***	***Region***
Total Households	***3,021,588***	***916,686***	***772,004***	***2,423,203***	***714,823***	***7,848,304***
Family households	***1,853,223***	***707,448***	***546,869***	***1,702,838***	***488,684***	***5,299,062***
Householder 15 to 24 years	64,046	6,485	10,631	33,438	11,059	125,659
Householder 25 to 34 years	354,908	90,291	78,723	267,053	75,162	866,137
Householder 35 to 44 years	483,145	191,566	150,801	462,680	132,132	1,420,324
Householder 45 to 54 years	406,695	174,056	132,267	395,120	112,573	1,220,711
Householder 55 to 64 years	259,878	113,631	83,603	251,069	71,570	779,751
Householder 65 to 74 years	169,592	81,139	56,205	173,236	49,708	529,880
Householder 75 to 84 years	92,134	42,963	28,667	100,022	30,379	294,165
Householder 85 years & over	22,825	7,317	5,972	20,220	6,101	62,435
Nonfamily households	***1,168,365***	***209,238***	***225,135***	***720,365***	***226,139***	***2,549,242***
Householder 15 to 24 years	61,993	5,247	9,534	30,864	11,226	118,864
Householder 25 to 34 years	252,415	26,887	32,905	123,365	37,254	472,826
Householder 35 to 44 years	201,285	32,156	36,280	116,336	36,314	422,371
Householder 45 to 54 years	183,010	32,380	37,562	109,655	36,132	398,739
Householder 55 to 64 years	153,395	29,107	32,669	91,253	29,149	335,573
Householder 65 to 74 years	142,382	34,795	33,015	100,900	29,563	340,655
Householder 75 to 84 years	121,551	35,503	30,532	106,643	32,640	326,869
Householder 85 years & over	52,334	13,163	12,638	41,349	13,861	133,345

Source: 2000 *Census of Population*

Table 22. Household Formation by Size of Household, by Subregion in the New York Metropolitan Region, 2000

Size of Household	***New York City***	***Long Island***	***Mid Hudson***	***New Jersey***	***Connecti -cut***	***Region***
Total Households	***3,021,588***	***916,686***	***772,004***	***2,423,203***	***714,823***	***7,848,304***

1-person household	962,624	170,273	186,010	594,166	186,083	2,099,156
2-person household	808,935	268,433	233,449	728,407	227,526	2,266,750
3-person household	486,002	165,673	131,221	420,190	119,069	1,322,155
4-person household	382,242	169,563	125,308	389,330	108,604	1,175,047
5-person household	206,080	86,483	59,214	182,255	48,957	582,989
6-person household	94,521	32,932	21,572	65,852	16,082	230,959
7-or-more person household	81,184	23,329	15,230	43,003	8,502	171,248

Source: 2000 *Census of Population*

Table 23. Household Formation by Age of Head & Income Bracket, by Subregion in the New York Metropolitan Region, 2000

Income by Age of Household Head	***New York City***	***Long Island***	***Mid Hudson***	***New Jersey***	***Connecti-cut***	***Region***
Heads < 25 years	***120,328***	***10,976***	***19,299***	***62,626***	***21,069***	***234,298***
Less than $10,000	32,316	1,527	4,696	10,673	3,975	53,187
$10,000 to $14,999	9,971	673	2,054	4,595	1,768	19,061
$15,000 to $19,999	8,481	579	1,556	4,599	2,008	17,223
$20,000 to $24,999	9,456	685	1,627	5,098	1,933	18,799
$25,000 to $29,999	8,535	692	1,604	5,281	1,782	17,894
$30,000 to $34,999	7,555	831	1,322	4,627	1,495	15,830
$35,000 to $39,999	6,576	667	1,052	4,130	1,417	13,842
$40,000 to $44,999	5,037	613	1,006	3,478	1,284	11,418
$45,000 to $49,999	4,181	635	747	3,117	895	9,575
$50,000 to $59,999	7,802	1,071	1,183	5,038	1,328	16,422
$60,000 to $74,999	7,883	1,174	1,046	4,648	1,389	16,140
$75,000 to $99,999	6,517	952	881	4,154	970	13,474
$100,000 to $124,999	2,783	421	240	1,582	474	5,500
$125,000 to $149,999	1,284	241	107	699	159	2,490
$150,000 to $199,999	1,101	99	81	546	108	1,935
$200,000 or more	850	116	97	361	84	1,508
Heads 25 to 34 years	***592,040***	***116,326***	***110,738***	***384,809***	***111,162***	***1,315,075***
Less than $10,000	81,168	4,168	7,430	22,994	6,794	122,554
$10,000 to $14,999	28,777	2,411	3,711	11,974	3,537	50,410
$15,000 to $19,999	29,695	2,698	4,612	13,601	4,559	55,165
$20,000 to $24,999	32,869	3,241	5,267	16,282	5,221	62,880
$25,000 to $29,999	36,773	3,956	6,078	18,998	5,777	71,582
$30,000 to $34,999	37,525	5,002	6,339	21,570	6,769	77,205
$35,000 to $39,999	34,031	5,544	6,606	21,769	6,491	74,441
$40,000 to $44,999	33,985	5,695	6,228	21,288	6,324	73,520
$45,000 to $49,999	27,465	5,483	5,520	19,605	5,883	63,956
$50,000 to $59,999	49,722	12,125	11,566	38,392	11,312	123,117
$60,000 to $74,999	57,520	17,050	13,538	52,427	14,989	155,524
$75,000 to $99,999	59,334	21,685	15,138	57,990	15,794	169,941
$100,000 to $124,999	31,796	12,879	8,401	30,967	7,716	91,759
$125,000 to $149,999	16,031	5,787	3,904	14,290	3,812	43,824
$150,000 to $199,999	15,937	4,981	3,150	12,465	2,972	39,505

$200,000 or more	19,412	3,621	3,250	10,197	3,212	39,692
Heads 35 to 44 years	***697,897***	***226,267***	***189,830***	***588,542***	***170,439***	***1,872,975***
Less than $10,000	93,381	6,089	8,865	27,186	7,601	143,122
$10,000 to $14,999	39,124	3,860	5,193	13,627	4,357	66,161
$15,000 to $19,999	36,858	4,129	5,447	16,269	4,165	66,868
$20,000 to $24,999	38,699	4,990	6,216	19,067	5,453	74,425
$25,000 to $29,999	39,850	5,868	6,810	22,258	6,312	81,098
$30,000 to $34,999	40,885	7,075	7,560	24,661	6,990	87,171
$35,000 to $39,999	38,402	7,624	7,999	26,338	8,162	88,525
$40,000 to $44,999	36,866	9,585	8,345	27,466	7,796	90,058
$45,000 to $49,999	30,987	9,032	8,039	25,989	7,933	81,980
$50,000 to $59,999	59,892	20,421	18,095	51,645	15,680	165,733
$60,000 to $74,999	70,472	33,390	24,105	76,253	22,404	226,624
$75,000 to $99,999	69,458	43,418	30,294	96,063	26,274	265,507
$100,000 to $124,999	38,184	28,527	18,875	60,761	15,849	162,196
$125,000 to $149,999	19,150	14,604	9,851	32,403	8,595	84,603
$150,000 to $199,999	18,309	13,718	10,195	32,876	8,998	84,096
$200,000 or more	27,380	13,937	13,941	35,680	13,870	104,808
Heads 45 to 54 years	***593,691***	***205,672***	***169,118***	***502,150***	***149,018***	***1,619,649***
Less than $10,000	73,258	5,346	6,680	21,609	6,368	113,261
$10,000 to $14,999	30,207	3,142	3,742	11,270	3,348	51,709
$15,000 to $19,999	27,094	3,065	4,522	11,056	2,985	48,722
$20,000 to $24,999	30,402	3,760	4,647	13,471	3,958	56,238
$25,000 to $29,999	31,221	4,813	5,762	15,993	4,397	62,186
$30,000 to $34,999	32,681	5,875	5,768	17,878	5,628	67,830
$35,000 to $39,999	31,507	6,125	6,027	18,979	5,773	68,411
$40,000 to $44,999	29,176	6,847	6,572	20,506	5,752	68,853
$45,000 to $49,999	27,087	6,449	6,369	18,738	5,267	63,910
$50,000 to $59,999	51,570	15,488	13,288	40,003	12,722	133,071
$60,000 to $74,999	62,309	24,796	19,040	59,262	18,026	183,433
$75,000 to $99,999	65,364	37,733	26,877	83,595	24,625	238,194
$100,000 to $124,999	37,717	28,692	20,006	58,905	16,465	161,785
$125,000 to $149,999	20,431	17,989	11,803	37,131	9,661	97,015
$150,000 to $199,999	18,472	18,569	12,630	36,376	9,640	95,687
$200,000 or more	25,195	16,983	15,385	37,378	14,403	109,344
Heads 55 to 64 years	***409,961***	***142,741***	***115,888***	***340,074***	***100,433***	***1,109,097***
Less than $10,000	63,355	5,226	7,024	21,094	6,714	103,413
$10,000 to $14,999	25,035	3,238	3,732	10,301	3,227	45,533
$15,000 to $19,999	20,251	2,931	3,624	11,589	3,288	41,683
$20,000 to $24,999	20,565	3,860	4,268	11,637	3,678	44,008
$25,000 to $29,999	21,738	4,132	4,863	13,179	3,811	47,723
$30,000 to $34,999	22,787	4,626	4,549	13,445	3,902	49,309
$35,000 to $39,999	20,607	4,991	5,056	14,090	3,752	48,496
$40,000 to $44,999	18,626	5,678	4,698	15,362	4,726	49,090
$45,000 to $49,999	16,905	4,766	4,688	14,231	4,129	44,719
$50,000 to $59,999	32,221	10,925	9,448	27,885	7,768	88,247
$60,000 to $74,999	37,835	16,445	12,116	37,570	11,238	115,204
$75,000 to $99,999	40,657	23,907	15,912	48,093	13,830	142,399

$100,000 to $124,999	25,377	17,329	11,590	34,877	9,537	98,710
$125,000 to $149,999	14,001	11,894	7,169	21,418	5,802	60,284
$150,000 to $199,999	12,669	11,639	8,049	22,284	6,091	60,732
$200,000 or more	17,332	11,154	9,102	23,019	8,940	69,547
Heads 65 to 74 years	***317,568***	***116,761***	***89,840***	***276,944***	***79,886***	***880,999***
Less than $10,000	65,761	6,991	7,833	26,336	7,074	113,995
$10,000 to $14,999	36,876	6,592	6,422	21,325	6,579	77,794
$15,000 to $19,999	24,576	7,275	5,873	20,868	5,657	64,249
$20,000 to $24,999	22,231	7,327	6,419	19,755	5,947	61,679
$25,000 to $29,999	19,225	6,995	5,484	17,896	5,306	54,906
$30,000 to $34,999	16,736	6,333	5,599	17,648	4,615	50,931
$35,000 to $39,999	14,288	6,330	5,191	16,042	5,034	46,885
$40,000 to $44,999	13,684	6,109	4,961	14,933	4,230	43,917
$45,000 to $49,999	11,534	5,578	4,030	12,908	3,417	37,467
$50,000 to $59,999	19,179	10,312	7,485	21,645	6,114	64,735
$60,000 to $74,999	20,766	11,685	7,943	24,622	7,335	72,351
$75,000 to $99,999	20,337	13,743	8,283	25,305	6,842	74,510
$100,000 to $124,999	11,826	7,531	4,899	14,352	3,967	42,575
$125,000 to $149,999	6,057	4,572	3,156	7,730	1,993	23,508
$150,000 to $199,999	5,725	4,525	2,730	7,346	2,207	22,533
$200,000 or more	8,767	4,863	3,532	8,233	3,569	28,964
Heads 75 years & Over	***290,992***	***98,595***	***77,969***	***268,889***	***83,299***	***819,744***
Less than $10,000	76,067	12,104	11,676	39,277	12,106	151,230
$10,000 to $14,999	44,431	11,148	10,778	38,273	12,475	117,105
$15,000 to $19,999	29,412	10,379	8,412	31,113	9,874	89,190
$20,000 to $24,999	23,824	8,796	6,978	25,283	8,264	73,145
$25,000 to $29,999	17,189	7,429	6,037	21,231	6,012	57,898
$30,000 to $34,999	14,077	5,750	4,529	17,089	5,000	46,445
$35,000 to $39,999	11,338	5,202	3,823	14,006	4,034	38,403
$40,000 to $44,999	9,758	4,174	3,376	11,833	3,552	32,693
$45,000 to $49,999	8,257	3,954	2,616	9,665	2,721	27,213
$50,000 to $59,999	12,716	6,273	4,333	14,299	4,536	42,157
$60,000 to $74,999	13,835	7,088	4,302	14,656	4,271	44,152
$75,000 to $99,999	11,885	6,409	4,316	13,934	4,169	40,713
$100,000 to $124,999	6,648	3,765	2,478	7,298	2,186	22,375
$125,000 to $149,999	3,268	1,861	1,333	3,592	1,177	11,231
$150,000 to $199,999	3,413	1,767	1,130	3,273	1,204	10,787
$200,000 or more	4,874	2,496	1,852	4,067	1,718	15,007

Source: 2000 *Census of Population*

Population figures by age, sex and mutually exclusive racial/ethnic group drive the household forecasts. Household population, or total persons excluding those in group quarters, is available from the American Community Survey for 2005. It can be derived from the total population forecasted by the Population Model for 2010 to 2035 by assuming constant shares of institutional population. Households are projected as a function of the forecasted population in households and projected household formation rates. Household population will be aggregated by ten-year age-of-householder on a mutually exclusive racial/ethnic basis. Household formation rates, or the share of each age cohort heading a household, will be compiled from decennial Census data, as depicted in Table 21 for all

races by household type in 2000. Household formation rate inputs will be estimated by racial/ethnic and age group for the years 2005 through 2035, based on past trends.

Inputs for estimation of household size will be produced by racial-ethnic household type in 1990 and 2000, as shown for all races in Table 22. Trend analysis will forecast changes in household size distribution through 2035. At each five year interval in the forecast period, the projected implications of household size distribution will be controlled by the forecast in household population by race-ethnicity. Inputs for estimation of household income will be compiled by age of head and race-ethnicity from Census data for the years 1990 and 2000, as shown for all races and subregions in Table 23. Trend analysis will forecast changes in household income distribution by age of head through 2005. In 2005, the projected implications of household income distribution will be controlled by the reported aggregate household income of racial-ethnic households in the American Community Survey. Adjustments will be made to the household income distribution to reflect the external aggregate income controls by subregion. Thereafter, the income distribution will be held constant and adjustments will be factored in from the Employment Model's forecast of aggregate personal income to 2035

The Region's changing racial/ethnic composition is expected to influence future patterns of household composition. These trends in part will reflect the growing proportion of immigrant households with their larger than average household sizes. As Table 20 shows, upward pressure on average household size is already apparent in the Region, throughout all subregions, after decades of decline. Trends in income distribution will also reflect racial-ethnic and immigrant influences, as growth becomes more bipolar, with gains reflected in both upper and lower income brackets. Size and income are major determinants of housing preference, expressed as changes in tenure, or ownership, and unit size. A further matrix will disaggregate projected households by housing stock preference and tenure based on income and racial/ethnic characteristics.

Table 24 depicts the overall increase of 228,200 housing units in the Region between 2000 and 2005 that accompanied the growth of 75,400 new households. The difference between households, or occupied housing units, and total housing largely reflects vacant units for sale or rent and seasonal housing. Increasing affluence in the Region has resulted in a growth in second homes, or seasonally occupied units, that will be forecasted separately as a function of income. Primary housing demand, or the preference of households for year-round occupancy by tenure and housing type (single family, townhouse, apartment, etc) will be driven by the forecasted changes in income cross-tabulated by tenure and housing type. Household size will be a determinant of housing type by unit size. Table 25 provides evidence of recent changes in housing tenure, and Table 26, by housing type.

Table 24. Trends in Housing Units by County & Subregion in the New York Metropolitan Region, 2000 to 2005 *(in 000s)*

	Housing Units						Seasonal		Change 2000-2005	
AREANAME	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2000**	**2005**	**Absolute**	**Percent**
NYC	**3,200.90**	**3,216.80**	**3,231.30**	**3,246.70**	**3,259.10**	**3,275.40**	**28.2**	**42.2**	**74.5**	**2.30%**
Bronx	490.7	492.5	494.5	496.9	498.6	502.2	1	1.7	11.5	2.30%
Kings	930.9	933.6	935.9	940.2	943.5	947.6	2.6	1.9	16.7	1.80%
New York	798.1	803.8	809.5	814.5	817.5	819.8	19.5	35.1	21.7	2.70%
Queens	817.3	820	822.5	825.1	827.3	831.8	4.6	2.9	14.5	1.80%
Richmond	164	166.9	168.8	170	172.2	174	0.5	0.7	10	6.10%
LONG ISLAND	**980.5**	**985.8**	**989.3**	**992.5**	**994.5**	**996.8**	**41.4**	**35.4**	**16.3**	**1.70%**
Nassau	458.2	458.6	458.4	458.2	458	458	3.1	1.8	-0.2	0.00%
Suffolk	522.3	527.3	530.9	534.3	536.5	538.8	38.3	33.6	16.5	3.20%

MID HUDSON	**830.7**	**837.1**	**842.4**	**848.1**	**853**	**858**	**27.7**	**27.8**	**27.3**	**3.30%**
Dutchess	106.1	107.1	108.1	109	109.9	110.7	2.4	3.3	4.6	4.30%
Orange	122.8	124.9	126.7	128.5	130	131.4	2.2	2	8.6	7.00%
Putnam	35	35.4	35.7	36	36.3	36.4	1.4	n/a	1.4	4.00%
Rockland	95	95.5	95.8	96.2	96.4	96.7	0.4	n/a	1.7	1.80%
Sullivan	44.7	45	45.4	45.8	46.3	47	13.3	13.9	2.3	5.10%
Ulster	77.7	78.2	78.6	79.2	79.8	80.5	5.2	6.2	2.8	3.60%
Westchester	349.4	351	352.1	353.4	354.3	355.2	2.7	2.4	5.8	1.70%
NEW JERSEY	**2,570.10**	**2,596.90**	**2,611.60**	**2,626.80**	**2,643.60**	**2,663.30**	**52.1**	**50.3**	**93.2**	**3.60%**
Bergen	339.8	342.3	343.3	344.2	344.7	346	1.3	0.8	6.2	1.80%
Essex	301	301.9	302.7	303.5	304.9	306.4	0.7	0.2	5.4	1.80%
Hudson	240.6	241.5	242	242.9	244.3	247.4	0.7	0.7	6.8	2.80%
Hunterdon	45	45.7	46.3	46.8	47.5	48.1	0.3	0.2	3.1	6.90%
Mercer	133.3	134.5	135.6	136.7	137.6	138.9	0.4	0.6	5.6	4.20%
Middlesex	273.6	276.1	277.5	279	280.8	282.9	0.9	0.7	9.3	3.40%
Monmouth	240.9	244.4	246.3	248.3	250.4	252.6	7.7	9.2	11.7	4.90%
Morris	174.4	177.3	178.5	180.1	181.3	182.3	1.2	1.4	7.9	4.50%
Ocean	248.7	255.2	258.7	261.9	265.4	268.8	33.2	35.4	20.1	8.10%
Passaic	170	170.1	170.3	170.6	171	171.3	0.8	0.5	1.3	0.80%
Somerset	112	115.1	116.3	117.2	118.3	119.3	0.5	0	7.3	6.50%
Sussex	56.5	57.3	58	58.6	59	59.5	3.6	n/a	3	5.30%
Union	192.9	193.3	193.4	193.5	194.2	195.1	0.5	0.5	2.2	1.10%
Warren	41.2	42.2	42.8	43.6	44.1	44.6	0.4	n/a	**3.4**	**8.30%**
CONNECTICUT	**759.5**	**763.5**	**766.4**	**769.2**	**772**	**776.2**	**11.6**	**11.1**	16.7	2.20%
Fairfield	339.5	341.4	342.8	344	345.2	346.9	3.8	4	7.4	2.20%
Litchfield	79.3	80	80.6	81.2	81.7	82.4	4.6	3.8	3.1	3.90%
New Haven	340.7	342.2	343	344	345.1	346.9	3.2	3.3	6.2	1.80%
REGION	**8,341.60**	**8,400.10**	**8,441.00**	**8,483.10**	**8,522.20**	**8,569.80**	**160.9**	**166.9**	**228.2**	**2.70%**

Source: United States Census. Bureau

Table 25. Trends in Housing Units by Tenure in the New York Metropolitan Region, 2000 & 2005 *(in 000s)*

	2000				***2005***			
	Owner occupied	**Percent**	**Renter occupied**	**Percent**	**Owner occupied**	**Percent**	**Renter occupied**	**Percent**
New York City	***912.3***	***30.2%***	***2,109.30***	***69.8%***	***1,002.30***	***33.1%***	***2,023.90***	**66.9%**
Bronx	90.7	*19.6%*	372.5	*80.4%*	98.2	*21.0%*	370	79.0%
Kings	238.4	*27.1%*	642.4	*72.9%*	265.2	*30.1%*	616.9	69.9%
New York	148.7	*20.1%*	589.9	*79.9%*	166.7	*22.8%*	564.7	77.2%
Queens	334.8	*42.8%*	447.8	*57.2%*	358.2	*45.8%*	424.3	54.2%
Richmond	99.7	*63.8%*	56.6	*36.2%*	113.9	*70.4%*	48	29.6%
Long Island	***733.6***	***80.0%***	***183.1***	***20.0%***	***766.2***	***83.2%***	***155.1***	**16.8%**
Nassau	359.3	*80.3%*	88.1	*19.7%*	362.5	*82.9%*	74.7	17.1%
Suffolk	374.4	*79.8%*	94.9	*20.2%*	403.7	*83.4%*	80.4	16.6%
Mid Hudson	***506.3***	***65.6%***	***265.7***	***34.4%***	***528.7***	***67.4%***	***255.4***	**32.6%**
Dutchess	68.6	*68.9%*	30.9	*31.1%*	74.2	*72.5%*	28.2	27.5%

Orange	77	*67.1%*	37.8	*32.9%*	82.8	*67.1%*	40.6	32.9%
Putnam	26.9	*82.3%*	5.8	*17.7%*	29.7	*86.1%*	4.8	13.9%
Rockland	66.4	*71.6%*	26.3	*28.4%*	68.6	*73.8%*	24.3	26.2%
Sullivan	18.8	*68.1%*	8.8	*31.9%*	21.3	*72.4%*	8.1	27.6%
Ulster	45.9	*68.0%*	21.6	*32.0%*	45.3	*66.2%*	23.1	33.8%
Westchester	202.7	*60.1%*	134.5	*39.9%*	206.8	*62.1%*	126.3	37.9%
New Jersey	***1,544.00***	***63.7%***	***879.3***	***36.3%***	***1,617.30***	***65.5%***	***851.8***	**34.5%**
Bergen	222.3	*67.2%*	108.5	*32.8%*	225.2	*67.8%*	107	32.2%
Essex	129.4	*45.6%*	154.3	*54.4%*	132.6	*46.7%*	151.3	53.3%
Hudson	70.7	*30.7%*	159.9	*69.3%*	75.2	*32.8%*	154	67.2%
Hunterdon	36.5	*83.7%*	7.1	*16.3%*	40.4	*87.6%*	5.7	12.4%
Mercer	84.3	*67.0%*	41.5	*33.0%*	88.4	*69.5%*	38.8	30.5%
Middlesex	177.4	*66.7%*	88.4	*33.3%*	181.9	*67.9%*	85.9	32.1%
Monmouth	167.3	*74.6%*	56.9	*25.4%*	176.1	*76.3%*	54.8	23.7%
Morris	129	*76.0%*	40.7	*24.0%*	132.4	*76.7%*	40.2	23.3%
Ocean	166.8	*83.2%*	33.6	*16.8%*	181.9	*82.3%*	39.2	17.7%
Passaic	91.2	*55.6%*	72.7	*44.4%*	92.2	*56.4%*	71.4	43.6%
Somerset	84.2	*77.2%*	24.8	*22.8%*	93.7	*82.6%*	19.8	17.4%
Sussex	42	*82.7%*	8.8	*17.3%*	46.8	*84.8%*	8.4	15.2%
Union	114.6	*61.6%*	71.5	*38.4%*	119.1	*65.3%*	63.4	34.7%
Warren	28.1	*72.6%*	10.6	*27.4%*	31.4	*72.7%*	11.8	27.3%
Connecticut	***479.6***	***67.1%***	***235.2***	***32.9%***	***504.5***	***69.8%***	***218.7***	**30.2%**
Fairfield	224.5	*69.2%*	99.7	*30.8%*	234.3	*72.2%*	90.4	27.8%
Litchfield	53.8	*75.1%*	17.8	*24.9%*	57.8	*78.3%*	16	21.7%
New Haven	201.3	*63.1%*	117.7	*36.9%*	212.4	*65.4%*	112.2	34.6%
Region	***4,175.80***	***53.2%***	***3,672.50***	***46.8%***	***4,418.90***	***55.8%***	***3,504.80***	**44.2%**

Source: United States Census. Bureau

Table 26. Trends in Housing Units by Type in the New York Metropolitan Region, 2000 & 2005 *(in 000s)*

	2000			***2005***		
	Single family unit	***Building with two or more apartments***	***Mobile home or other type of housing***	***Single family unit***	***Building with two or more apartments***	***Mobile home or other type of housing***
New York City	***536,054***	***2,661,598***	***3,260***	***555,883***	***2,716,433***	***3,096***
Bronx	55001	435144	514	53,715	448,064	432
Kings	127610	802371	885	145,726	801,264	642
New York	6424	791146	574	8,191	810,807	798
Queens	251947	564350	953	243,965	587,135	719
Richmond	95072	68587	334	104,286	69,163	505
Long Island	***815,456***	***159,039***	***5,979***	***826,501***	***164,120***	***6,216***
Nassau	367371	90274	506	363,951	92,526	1,534
Suffolk	448085	68765	5473	462,550	71,594	4,682
Mid Hudson	***515,249***	***294,670***	***20,772***	***532,939***	***305,157***	***19,882***
Dutchess	73437	28253	4413	77,969	28,019	4,724
Orange	84466	34631	3657	89,491	38,372	3,573
Putnam	30035	4646	349	32,081	4,160	138
Rockland	65355	28404	1214	66,508	28,848	1,364

Sullivan	31722	8159	4849	33,475	8,507	5,016
Ulster	54886	16686	6084	55,934	20,555	4,020
Westchester	175348	173891	206	177,481	176,696	1,047
New Jersey	***1,542,567***	***1,010,872***	***16,634***	***1,595,703***	***1,049,556***	***18,058***
Bergen	201353	137344	1123	200,278	144,588	1,167
Essex	115031	185724	256	118,900	187,321	191
Hudson	37620	202602	396	33,981	213,160	305
Hunterdon	38634	6225	173	41,901	6,084	100
Mercer	94204	38666	410	98,588	39,820	493
Middlesex	176969	94232	2436	180,043	100,693	2,122
Monmouth	180814	56775	3295	187,422	61,687	3,460
Morris	132837	40975	567	135,752	46,208	368
Ocean	209232	33606	5873	226,129	36,349	6,365
Passaic	81213	88512	323	81,730	88,213	1,372
Somerset	84073	27714	236	92,047	27,067	146
Sussex	47992	7712	824	51,296	7,558	692
Union	110745	81948	252	114,176	80,095	841
Warren	31850	8837	470	33,460	10,713	436
Connecticut	***481,546***	***274,026***	***3,893***	***498,863***	***273,574***	***3,773***
Fairfield	221046	117142	1278	226,885	118,376	1,681
Litchfield	60596	18073	598	62,851	19,211	313
New Haven	199904	138811	2017	209,127	135,987	1,779
Region	***3,890,872***	***4,400,205***	***50,538***	***4,009,889***	***4,508,840***	***51,025***

Source: United Sates Census Bureau

1.5.2 METHODOLOGY

National level household projections are produced by the US Census Bureau on a periodic basis. The Bureau produces three series, which make possible an analysis of the factors influencing the projected number of households. Series 1 uses projections of household formation rates produced through a time series analysis using a statistical function yielding a linear projection, with recent years weighted more heavily than earlier years in determining the slope. Projected changes in household formation rates are modest, especially in comparison with trends in the 1990s. Projections of rates are disaggregated by age group but not by racial/ethnic group.

Series 2 holds baseline household formation rates constant and reflects just the effect of changing age structure on household formation, for the total population. Series 3 also holds baseline household formation rates constant, but utilizes separate matrices for the different racial groups and the Hispanic population, and therefore is a reflection of the effects of both age and race.

A comparison of the different series reveals the nation's changing age structure to be the dominant influence on household formation, both in terms of numbers and composition. For example, Series 2 households in 2010 differ from Series 1 households by just 0.5%, and Series 3 differ by just 1.2%, which indicates the modest influence of shifts in family structure and racial/ethnic composition on household growth. In terms of household composition, the main driver of change is the aging of the baby-boomers. Consequently, the projections show a declining share of households with children and an increase in married couples without children. Aging baby-boomers are also expected to account for the majority of the increased share of single-person households.

In the New York Metropolitan Region, the effects of racial/ethnic restructuring are expected to be more significant than for the nation as a whole, and the Region's racial/ethnic minorities reflect a large component of foreign immigration. Because of the substantial changes in racial/ethnic composition anticipated throughout the Region, the Household Formation Model will project each racial/ethnic group separately. For each group, the following projections will be made:

- Household population
- Households by age of householder
- Households by type
- Households by size
- Households by income range
- Housing stock preference by tenure

All projections will be made at the subregional level for the period 2005 through 2035 and will be aggregated by mutually exclusive racial/ethnic group and five-year age group for the population 15 years and over.[36] Projections of households will then be disaggregated to the county level.

Projected households will be calculated as a function of projected population and household formation rates. The former will be adopted from the Population Model and reduced to reflect population in households, as discussed below. Household formation rates, also referred to as headship rates, indicate the percentage of a given population group that are classified as householders by the US Census Bureau. Because one householder is designated for each household, an estimate of the number of householders is equivalent to an estimate of the number of households.[37]

Disaggregation of households by type, size, income range and housing stock preference will be performed using a series of matrices that crosstabulate percent breakdown of age of householder by household type, household type by size, household size by income range, and household income range by housing stock preference and tenure. Household formation rates will be estimated for every five years from 2005 through 2035. For group quarters rates and all other rates, 2000 rates will be held constant, but external controls such as aggregate household population and income will be used to normalize results upward or downward.

The estimation of future household formation rates will use a methodology based on that employed by the Harvard University Joint Center for Housing Studies in their 1994 state-level household projections.[38] This involves use of a cohort model that starts with the 2000 headship rate estimates, and trends these rates forward according to the adjustments that age cohorts made in the recent past. The rate of change in headship, over a five-year period, for a given cohort as it ages in the future is assumed to be the same as the rate of change experienced by cohorts of the same age in the 1990s.

---

[36] By Census definitions a householder must be at least 15 year of age.

[37] Housing units and households, however, differ in number because of existence of unoccupied housing units for which there is no household.

[38] Working Paper W94-4, Harvard University Joint Center for Housing Studies.

Household Formation Rates

Household formation is calculated as the ratio of householders to household population for a given age cohort:

$$F_a = \frac{H_a}{P_{h \bullet a}}$$

where *F* is the household formation rate, *H* is the number of householders, P_h is household population, and *a* represents a given age cohort. Household population is calculated as total population times the percentage of population residing in households. The latter is equal to the total population minus the population residing in group quarters facilities, (e.g., dormitories, prisons, etc.).

Household formation rates will be estimated from the best available decennial census sources. Published tabulations are the preferred source, but since these are not available for geographies below the state level, PUMS will be used as a supplement. (Unadjusted PUMS cross tabulations by the required level of geographic, age and race detail may produce significant levels of sampling error for some age groups.) State-level household formation will be calculated from printed tables and state-level rates will then be factored down to the subregional level based on the PUMS subregional-to-state ratio of household formation for each racial/ethnic and age group:

$$F_a = D_{a \bullet s} \times \frac{P_{a \bullet r}}{P_{a \bullet s}}$$

where *F* indicates the estimated household formation rate, *D* indicates the decennial census household formation rate, *P* indicates the PUMS household formation rate, *r* indicates subregional level geography, *s* indicates state-level geography, and *a* indicates the particular age and racial/ethnic group.

Estimation of Other Rates

Estimation of rates necessary to disaggregate the household forecasts by type, size and income is hampered by a scarcity of published data at the adequate level of geographic and demographic detail. It will therefore be necessary to directly use the custom crosstabulations produced from the PUMS (Public Use Microdata Samples) data set for 2000. In some instances, 2005 PUMS data from the American Community Survey may be used. The approach will differ from the methodology described above for household formation rates, where PUMS crosstabs will be used to regionalize state level figures derived from published sources. For the other rates, PUMS crosstabs will be produced at the subregional level.

Estimation of Rates for 2010 through 2035

Household formation rates will be projected forward from the year 2005 through the year 2035 using a method based on the Harvard University Joint Center for Housing's cohort methodology. In this method, the future changes in household formation of a given cohort as it ages are expected to mirror those of comparable cohorts in the past. For example, the cohort of 25 to 34 year olds might show a household formation rate of 45% in 1990; in 2000, the same group, now aged 35 to 44 might show a rate of 55%, a 22% increase. The same rate of increase could be applied to the year 2000 cohort of 25 to 34 year olds to estimate the 2010 rate for 35 to 45 year olds. Comparable factors will be applied to all age groups, and the intervening years will then be interpolated.

Model Controls

Outputs of the household model will be controlled at three points:

1. Figures for total households by racial/ethnic group (as calculated based on household population and headship rates) will be controlled to match available 2000 Census enumerated figures by racial/ethnic group. This will be accomplished by adjusting 2000 household formation rates for all age groups on a proportional basis.
2. Outputs of household type by size will be controlled to match projections of household population generated in the model.
3. Projections of household size by income range will be adjusted based on external forecasts of aggregate income from the Employment Model. Aggregate money income will be forecasted for each subregion by analyzing historic trends in per capita personal income by place of residence, in the relationship between worker earnings and commuter shares, and in the proportion of money income to personal income. Personal income is a complete income concept of the US Bureau of Economic Analysis, including all earnings, property, transfer, imputed and in-kind income sources. Money income is conceptually equivalent to household income, and is a self-reported value in the decennial Census, largely reflecting earnings and transfer payments. The level of annual personal income per capita will be forecasted by subregion from the period 1970-2005, expressed in constant 2005 dollars, to 2035 using linear regression analysis. Average worker earnings by subregion, to be forecasted to 2035 by the Employment Model, will be converted to average commuter earnings by subregion and used to evaluate forecasted levels of per capita personal income. After further adjustments to per capita personal income forecasts, aggregate personal income will be computed using forecasted population levels. The relationship between money and personal income will then be applied to estimate future levels of aggregate money income.

   Though the structure of the household model will not facilitate an exact match between these two series, household model outputs will be reconciled to within plus or minus 3% of the aggregate income series for the total population in households. The aggregate income control will then be applied through a feedback process in the development of household income distribution rates for years after 2005.

1.5.3 Disaggregation to the County Level

Disaggregation of subregional household projections to the county level will be based on the projection of county household trends, which will in turn be estimated as a function of forecasted household population and anticipated trends in average household size. The long-term trend toward smaller average household size which prevailed in the post-war years has reversed itself in many areas since the 1990s. The earlier downward trend was related to demographic factors such as a decreasing fertility rates for women, the rising number of single-parent families, an increasing number of elderly-headed households without children, and a general rising affluence. The more recent upward trend can be related to the moderation of a number of the above demographic factors, a rise in housing costs in many areas, and, not least, a growing number of immigrant households, which tend to be larger on average than those of the native born population. The average household size in any given area is the result of the particular combination of a number of such factors.

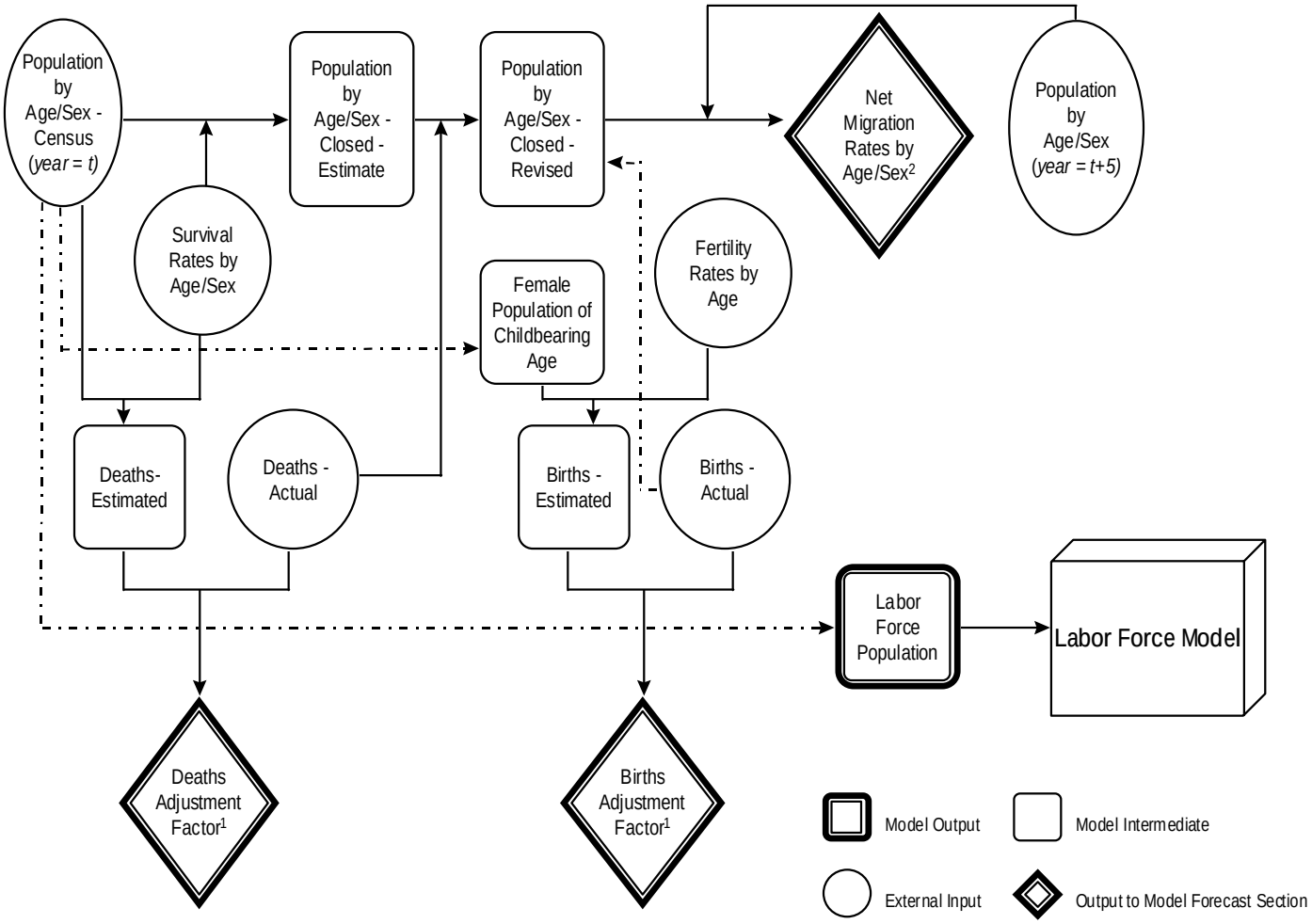
Subregional trends in average household size may not be fully reflected at the county level because of wide differences in population and housing unit composition. In an effort to best approximate likely future trends, county level estimates will be made individually based upon a combination of two approaches: 1) linear extrapolation of the 1990-2005 county-level trends, and 2) a comparative

technique in which anticipated patterns for some counties will be modified in light of the historical experience of nearby counties whose patterns they are expected to resemble in future years. County level estimates will be calibrated to subregional forecasts of households and household population.

1.6 WORK PRODUCTS

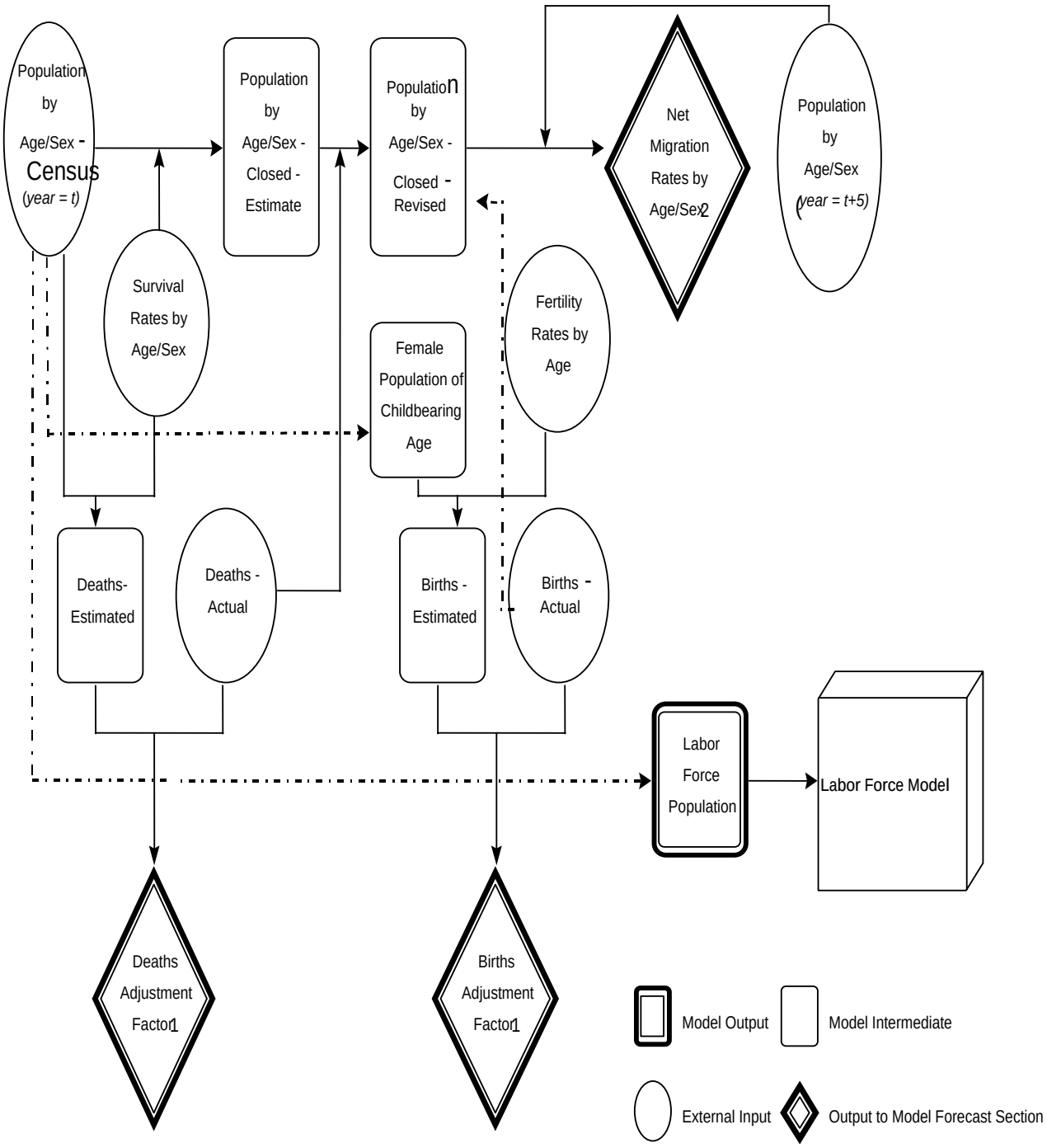
The text of this memo is contained in a Microsoft Word for Windows file, called TM 1.4.5.2text.doc. The Appendix tables are contained in a Microsoft Excel for Windows file called "TM 1.4.5.2.xls." Within this file, worksheet tabs identify each table by Appendix table number.

Figure 1. Population Model - Historical Section



[1]Output based on 1985 to 2000 time intervals only
[2]Output based on 1995 to 2000 time inteval only

Figure 2: Population Flow Chart



1Output based on 1985 to 2005 time intervals only

2Output based on 1995 to 2005 time interval only

Figure 3. Labor Force Flow Chart

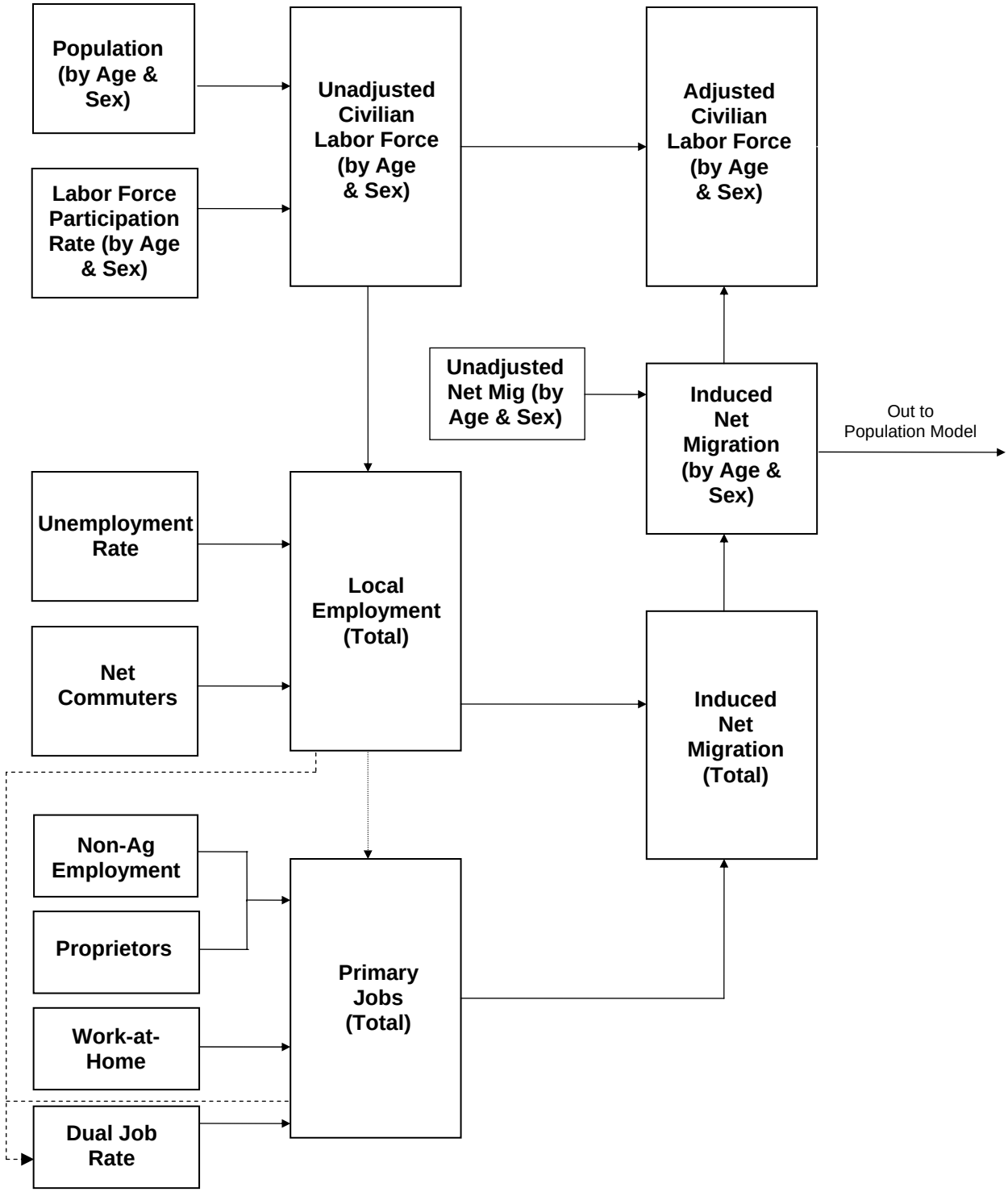


Figure 4. Employment Flow Chart

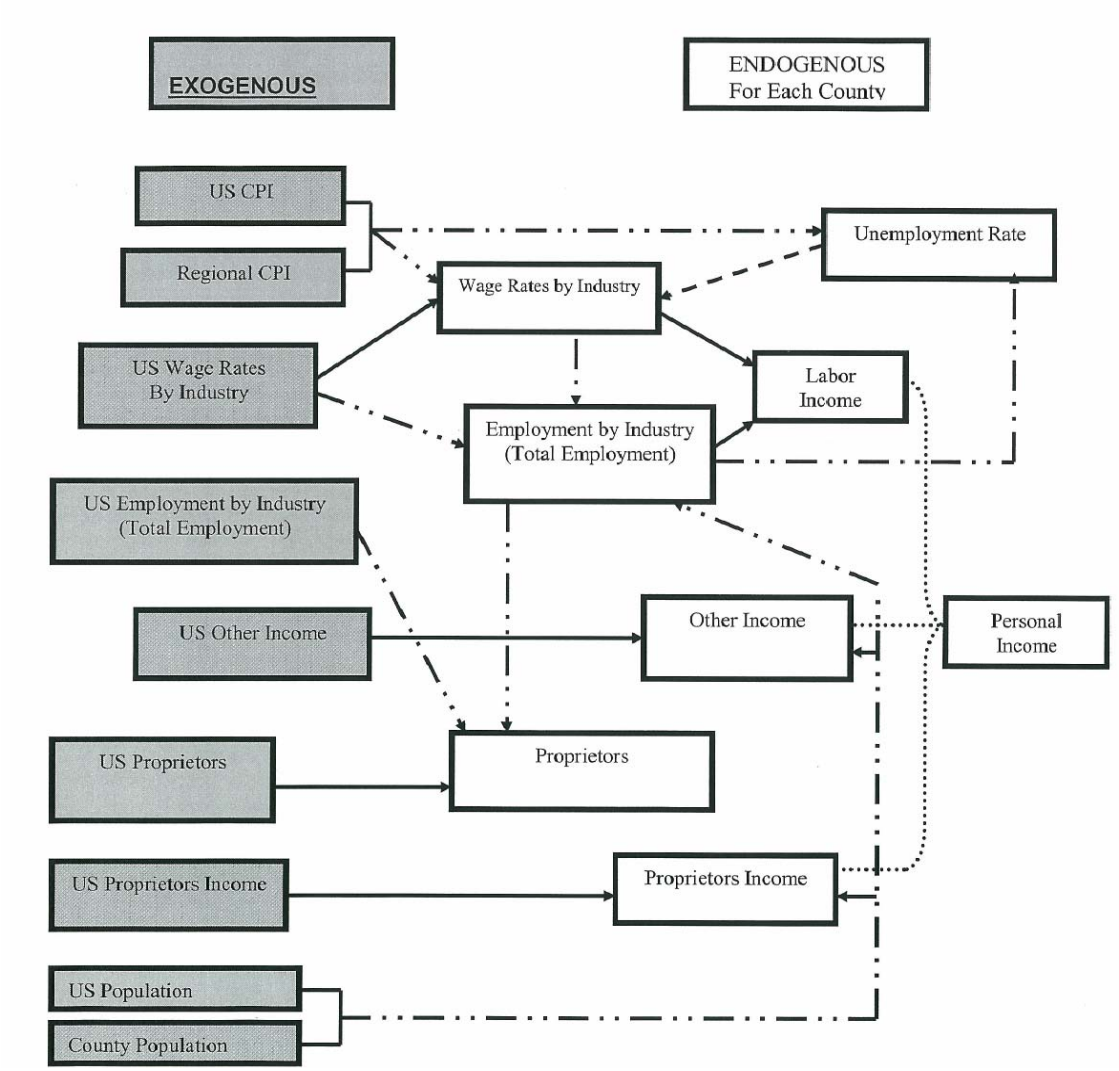


Figure 5. Household Flow Chart

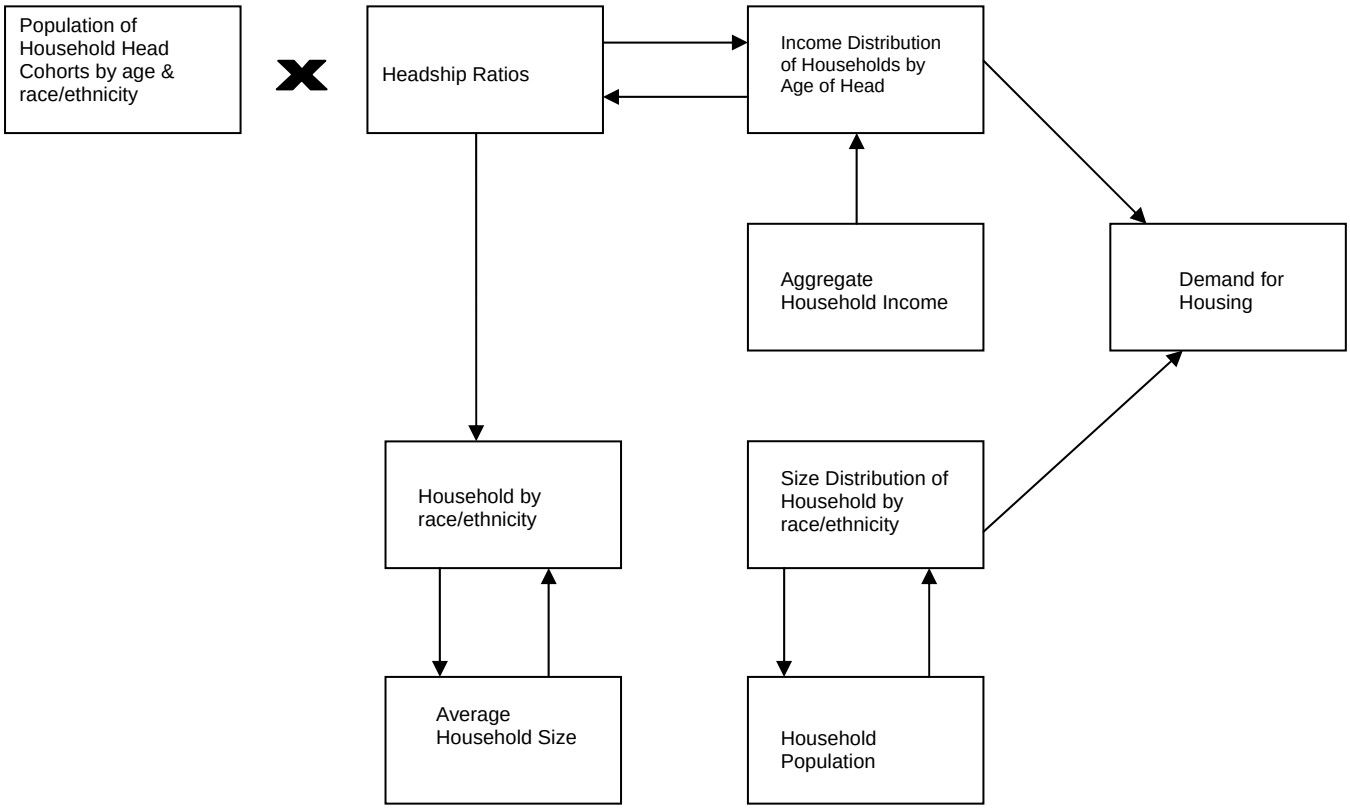


Figure 6. Enrollment Flow Chart

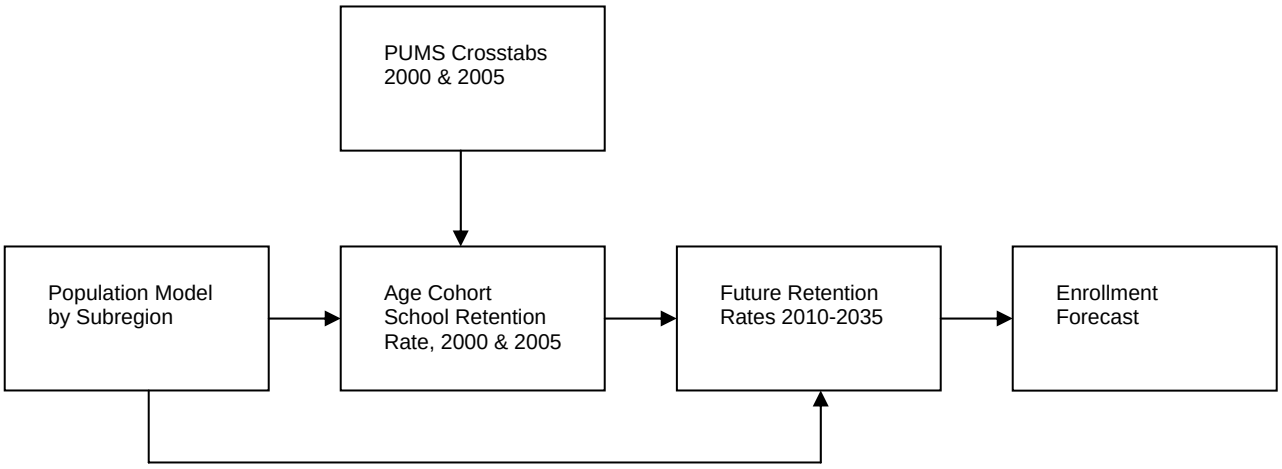


Figure 7. FORECASTING MODEL RELATIONSHIPS:

Employment, Labor Force, Population