

MapInfo Demographic Estimates and Projections

METHODOLOGY STATEMENT

This methodology statement describes the procedures used by MapInfo to create the 2004-2009 update to the U.S. demographic estimates and projections database. The MapInfo Estimates and Projections for the U.S. are updated annually. The reference date for the data is always July 1, which is considered the midpoint for the reference year. The reference date should be considered a point of “annual average” not strictly, for example, the population number for a location on July 1.

The MapInfo Estimates and Projections database contains over 1,800 variables for nine layers of geography, including a national layer. This methodology statement describes procedures used to produce the basic variable streams – for example, population detail and household income – as well as the processes used to validate the data and assure data quality.

MapInfo personnel responsible for producing this data update combine over 25 years of experience in producing demographic estimates and projections for the U.S. and Canada. The methodologies used to develop and update the U.S. estimates and projections build on this expertise using a combination of traditional demographic techniques as well as innovative processes, which take advantage of proprietary resources.

Overview of Demographic Methods

The basic methodology uses a combined top-down / bottom-up approach. The top-down phase develops national, state, and county estimates and projections that become “control totals” for the sub-county estimates and projections. Resources used in the top-down phase include the Census Bureau’s national population projections by age, sex, race, and Hispanic origin, Census Bureau state population projections, and the Census Bureau county estimates program. We build a complete cohort-component model at the state and county level for each race/Hispanic origin group by age and sex. The models contain our assumptions for fertility, mortality, internal migration, and international migration based on race-specific fertility schedules and life tables from the National Center for Health Statistics together with migration estimates from the U.S. Bureau of the Census.

The bottom-up phase of the estimation and projection methodology begins with the most recent decennial census, in this case the Census 2000 block and block group level data. In addition to census data the bottom-up process utilizes comprehensive consumer list resources, US Postal Service delivery statistics, and selected state and local estimates from the Census Bureau's State Data Center network. The objective of the bottom-up phase is to estimate households for the current year and project five years out assessing the impacts of differential sub-county demographic change across the primary units of sub-county geography – Block Groups, Census Tracts, Zip Codes, as well as county subdivisions (e.g. minor civil divisions) as delineated by the Census Bureau.

The estimates and projections are organized into the following data streams:

- Population and Households
- Population Detail
- Household Characteristics
- Income by Age
- Income Detail
- Consumer Expenditure Potential
- Retail Sales Potential
- Wealth and Financial Assets
- Housing Unit Characteristics
- Census-based Update Variables

The methodologies used to develop each data stream are described below. In addition, we provide notes on some of the special processes we have implemented to ensure internal consistency and external validity of the each data stream.

Population and Households

Our standard estimation procedure begins with an analysis of the latest demographic estimates and projections of the U.S. Bureau of the Census. For example, the bureau produces county-level population estimates by age, sex, race, and Hispanic origin. These estimates form an annual time-series for each post-census year, from 2000 to the most current year. The bureau also produces a parallel time-series of the

components of population change for all counties. The components include: births, deaths, net domestic migration, net international migration, net federal movements which track military populations and other federal employees, and changes to the group quarters population. These time-series estimates form the starting point for the county-level cohort-component estimation and projection model.

The cohort component model makes use of the Census 2000 county-level age distributions of the population by race, sex, and Hispanic origin. This starting point, in essence, recalibrates the model to be consistent with the new census. Additional resources for the cohort-component procedure include life tables (showing morality schedules and life expectancy by age, race and sex) from the National Center for Health Statistics (NCHS), part of the Centers for Disease Control and Prevention. Our fertility assumptions by race and Hispanic origin are developed from resources of the National Vital Statistics System as well as those of the Census Bureau.

The Census Bureau has relationships with producers of independent demographic estimates at the county and sub-county level. The State-Federal Cooperative Program for Population Estimates includes academic experts as well as State Demographers who head the Census Bureau's State Data Centers in each state. Our procedures involve comparing locally produced estimates and projections to determine which series most closely track our own growth assumptions and the results of our cohort component projection.

Through a modified cohort-component method we advance the Census Bureau's county population estimates to the current year and to the projection year five years out. The cohort-component model is age-sex-specific for each race and Hispanic origin group. A cohort-component model essentially takes a population by age and sex and "ages" it through time based on the addition of births, the subtraction of deaths, and the addition of a positive or negative net migration amount. We allocate births to race/Hispanic/sex-specific groups based on county-level fertility trends consistent with the Census Bureau's middle-series national fertility rates. Similarly we allocate deaths to race/Hispanic/sex and age groups based on group-specific mortality rates from NCHS life tables.

Net migration estimates and projections are developed from an analysis of historical trends reflected in the Census Bureau's county estimates. We also take into account county growth trends in the econometric model projections of Woods & Poole Economics, Inc., a Washington, DC-based forecasting firm. The Woods & Poole model projects total population based on an analysis of local employment and

economic trends modified by the county-to-county commuting patterns from the decennial census. Thus we use a combination of explicit and implicit migration estimates to inform our assumptions about net migration for counties. Age/sex-specific migration rates from the Census Bureau's Current Population Survey are analyzed to estimate the age/sex-specific impacts of a county's net migration trend.

At the state and national level we make use of Census Bureau's latest middle-series national population projections. However, our assumption with respect to international migration falls between the Census Bureau's middle and high series. In general, the Census Bureau official projections have tended to understate the impacts of international immigration. While we have made extensive use of Census Bureau estimates, projections, and survey data, we have supplemented those with our own judgments as to their impacts on small area population change.

For example, the predominant component of population change for most areas is migration. The Census Bureau's county estimates incorporate special tabulations of Internal Revenue Service data (maintaining strict confidentiality) showing county-to-county migration patterns. Knowing where individual tax forms are filed from year to year provides a powerful basis for estimating the net effects of migration. The Census Bureau can also estimate the effect of international migration, including undocumented immigration, using files from the Immigration and Naturalization Service and other sources. Despite these rich sources of information available to Census Bureau researchers, we compare results from our county-level estimates to independent State Data Center projections for selected states. State Data Centers often use alternative indicators of migration, such as driver's license registrations. Special attention is given to those states and counties for which the Census Bureau results diverge significantly from our state and local sources.

The result of this "top-down" process (national level down to county level) provides current-year and five-year control totals for the nation, the 50 states (plus the District of Columbia) and 3,141 counties for total population by age, sex, race, and Hispanic origin. The control totals are used in the next steps of the process to guarantee that all estimates and projections from sub-county levels of geography add up to county and higher levels of geography.

Census 2000 Results and Post-Census Geographic Changes

We incorporated the Census 2000 population enumeration on the new geographies beginning with the 2002 estimates and 2007 projections. Subsequently, this process involved advancing in time the population control totals from April 1, 2000 (Census Day) to July 1, 2004 and subsequently to July 1, 2009 – the reference days for these estimates and projections. At the national and state levels we evaluated the Census Bureau’s population projections in light of the new census data. Our procedures included matching our county level cohort model to new controls by race, Hispanic origin, sex, and age at the national and state levels. In essence, we integrated our cohort component models at the county, state, and national levels, taking into account differential growth patterns as suggested by the Census 2000 results. Following the 2002-2007 update, the MapInfo methodology provides Census 2000-consistent annual updates.

Most users understand that no census is perfect and a degree of under-enumeration is expected. Therefore, our population estimates will differ from the census results in part because we have implemented national control totals that overcome to a degree the problem of under-enumeration. Thus, while the Census Bureau decided not to produce a set of census results adjusted for net undercount, the MapInfo series takes the undercount into account in setting top-down control totals. We have also made limited “corrections” to some of the known census errors in group quarters locations.

The creation of Broomfield County in Colorado occurred after the 2000 Census. The MapInfo 2004-2009 Estimates and Projections include this new county, along with consequent changes to geographic codes down to block group. Also, code changes in Virginia, associated with the merging of Clifton Forge (formerly a county-equivalent independent city) into Allegheny County, are incorporated into this update.

The “Bottom Up” Phase

The next phase of our procedure involves a “bottom-up” household level approach. In a nutshell, the key to this phase is to determine the current number of households at the smallest levels of geography – Block Group, Census Tract, and Zip Code. Then, the Block Group estimate is “rolled-up” to the county level and made consistent (through iterative proportional fitting techniques) with the results of the “top-down” phase.

We use four primary data inputs at small spatial scales to estimate current households – 1990 Census, 2000 Census, a comprehensive consumer household list, and U.S. Postal Service delivery statistics. The comprehensive list of over 115 million U.S. households is called TotalSource from the Atlanta, Georgia-based firm Equifax. The entire list is geocoded and summarized to census block group, census tract, and Zip Code geographies. The list itself is rigorously maintained and updated twice a year by Equifax’s Consumer Information Services Group. Occupied households are verified against transaction databases.

The geocoded household list data are analyzed over time determine changes in occupied households and small area growth trends. Because of the release of the Census 2000 data, we were able to benchmark the aggregated list data to the census results on the new census geography. Benchmarking is a process that determines how closely the list data matches the census enumeration.

It should be noted that the Census 2000 published data contain some significant location errors with respect to Group Quarters populations. We took special pains to properly locate some of the larger Group Quarters facilities, primarily correctional institutions, in order to correct errors in the published census data. However, this process remains incomplete and will continue as more census errata are released. For the post-censal years we have in place procedures to update major group quarters facilities such as military barracks (primarily affected by base closings) and prisons (primarily affected by new prison construction).

Our household estimation method also uses U.S. Postal Service delivery statistics for Zip Codes and Carrier Routes. The monthly-updated active residential deliveries by Zip Code and Carrier Route are allocated to block groups and census tracts based on correspondence tables we develop from Geographic Data Technology’s block-to-Zip Code and block-to-CR correspondence files. The GDT files are based on U.S. Postal Service delivery-weighted centroids for each census block. As a point of comparison with the consumer list data, the USPS data provide another indicator of “occupied households” (as opposed to vacant housing units) by the fact that residents collect their mail regularly and are, therefore, considered “active residential deliveries.” The Census 2000 enumeration also provided an opportunity to benchmark postal delivery statistics in order to determine where they are most accurate.

The combination of USPS delivery statistics, geographically aggregated consumer list data, and the Census 2000 enumeration provide a powerful basis for summarizing household counts for small areas of geography covering the entire U.S. However, we do not take any of these inputs at face value. We apply a series of checks and demographic controls to our data at several points in the process. For example, the postal delivery statistics condition our estimates in areas where residential post office (RPO) deliveries are common, that is, where people primarily pick up their mail at the Post Office. By developing estimates for RPO Zip Codes, the data help maximize mailing list matches and analysis. In this way, census statistics, postal deliveries, and our own cohort component analysis condition the estimates in areas where the household database may fall short of or over-estimate a realistic current-year household estimate.

Once a “total households” estimate is established for an area, we use estimates of updated average household size, group quarters, and total population controls to create a consistent set of basic block group-level variables upon which to build the next set of current-year estimates and five-year projections.

By combining the “top down” and “bottom up” phases, we create a geographically and demographically consistent database of core variables – total population, households, group quarters population, household population, and average household size. These variables form the base upon which we build estimates and projections of population and household characteristics.

Population Detail

Population Detail is a variable set that includes five-year age groups for total population, female population, and male population. It also includes “school ages” (pre-school, elementary, middle school, high school) or “marketing” age groups for total population, population by sex, and population by race and Hispanic origin. Estimates for all geographies are provided for the current year and the five-year projection.

- **Race and Hispanic Origin**

The MapInfo estimates and projections by race and Hispanic origin variables conform to the new Census 2000 conceptualization of race and Hispanic origin. That is, the race variables include a category for “Some Other Race” and the individual race categories are “single race alone.” Individuals of two or more races are combined in the MapInfo data with the “Some Other Race” variable. In the past, we

deferred to the Office of Management and Budget (OMB) race categories, which in turn were reflected in Census Bureau's estimates and projections programs. The OMB standard does not include a "some other race" category and, therefore, comparisons with census results are more difficult. The main points to understand are the following:

1. The old OMB mandate (implemented in MapInfo data from 1998 through 2001) was four race groups cross-tabulated by Hispanic origin. The four race groups are: White, Black, American Indian or Alaska Native, and Asian or Pacific Islander. The Census Bureau, subsequent to the primary census tabulation, merged the "some other race" responses (historically permitted in the census) with the four main race groups to conform to the OMB directives. The merged data were published after the census and were called the MARS data for Modified Age Race Sex data. This permitted government agencies and others to use baseline statistics from the Census Bureau that conformed to OMB directives, primarily to administer the many anti-discrimination and affirmative action programs of the federal government.
2. The new Census 2000 convention permitted multiple-race responses. As in prior censuses an "other race" response (now, single race "other race") is also captured. The MapInfo estimates and projections conform to the Census 2000 tabulations and include an "other race" categories. However, in order to maintain "cell size" and the integrity of the race projections, MapInfo combines Hawaiian / Other Pacific Islander with Asian and persons of two or more races with the "Other Race" category. The race question has always been a matter of self-identification but this is the first time multiple responses were permitted. In order to use the new information on race, we modified our process of calibrating national, state, and county control totals to conform to the new race categories. We maintain an "other race" category in order to maintain consistency with Census 2000 in the future. The resulting categories, therefore, allow for qualified comparisons to Census 2000 race categories:

- White alone
- Black or African American alone
- American Indian / Alaska Native alone
- Asian, Hawaiian, Other Pacific Islander alone
- Some Other Race (includes persons of two or more races)

3. The MapInfo Estimates and Projections of the Hispanic population are consistent with the estimates and projections by race except that the race categories are combined in order to maintain reasonable “cell size” and integrity of the projection for small areas. Hispanic origin, of course, is not a matter of race but a matter of ethnic origin more broadly speaking. The Hispanic Detail variables in the MapInfo Estimates and Projections are as follows:

- Hispanic
 - White Hispanic
 - All Other Races Hispanic
- Non-Hispanic
 - White, non-Hispanic
 - All Other Races, non-Hispanic

In order to generate the Population Detail variables each race/sex group is “advanced” – or moved forward in time demographically – to current and projected years while maintaining the county-level control totals. Our estimation process, therefore, takes into account a detailed structure of the population by age, sex, race, and Hispanic origin. We use an iterative proportional fitting technique to guarantee that the small area estimates sum to control totals at all higher geographic levels. The final step involves systematically “sprinkling” the integer results such that differences due to rounding are eliminated in the data.

Household Characteristics

MapInfo produces a national set of household projections by age of householder, race/Hispanic origin of householder, and household type (family, non-family) to form the basis for our household detail variables. We project households at the national level from a detailed set of household headship rates applied against projections of the civilian, non-institutional household population by age. The assumptions behind this method take into account trends in age-specific household headship rates as well as trends in the distribution of the civilian non-institutional household population by age/race/Hispanic origin. Our U.S. national household projections become the control totals for estimates and projections of households by age at the state and county level.

Our projections of households by age at the county level are based on our population projections by age conditioned by our estimates of the group quarters population and trends in average household size. In this way, our county estimates of households by age are consistent with local trends as well as with state and national household trends.

For the smallest levels of geography, such block groups, census tracts, and Zip Codes, our estimates of total households are distributed by age in a manner similar to the estimation procedure for population. The household procedure uses a headship rate method applied to the underlying population shifts by age. The 2000 Census distribution of households by age is “advanced” to the current year, and subsequently projected out five years, such that the county control totals are maintained. As with the population procedure, households by age are subjected to an iterative proportional fitting routine, also known as a demographic “rake,” that assures that the household totals by age for all small areas of geography conform to corresponding totals at all higher geographic levels.

Household Income

The first step in the MapInfo method for estimating and projecting household income is to adjust 2000 Census income distributions to current dollars. We use the Consumer Price Index to inflate 1999 dollars from the 2000 Census to current dollars. The process involves moving a distribution of households by income group for each householder age group to a new distribution of households by income group given in current year dollars. These detailed distributions at the block group, census tract and county levels of geography form the starting point for estimation and projection.

In the next step, we use trends in aggregate income provided by Washington, DC-based Woods & Poole Economics, Inc. to establish county and metropolitan area level control totals for average household income. The Woods & Poole regional econometric models take into account employment and earnings trends in 13 major industry groups. The Woods & Poole model uses the concept of aggregate “personal income,” which includes: wages and salaries, other labor income (primarily employer contributions to private pensions), proprietors’ income (income of the self-employed), dividend income, rental income, personal interest income, government transfer payments (including Social Security, Medicaid, and veterans’ benefits), and employees contributions to social insurance. We have, however, made adjustments

to these components in order to provide household income estimates more in line with the Census Bureau's concept of money income. These county-level adjustments provide a basis for using Census Bureau money income estimates by householder age at the national level (Current Population Survey) to constrain income by demographic as well as geographic variables. The primary benefit of incorporating the Woods & Poole household income projection is that regional and local trends in income generation are taken into account. These trends provide a geographic specificity to income estimates and projections while the Census Bureau data provide the demographic benchmarks.

As discussed above, starting with the MapInfo 2003 update (and continuing) are income distributions in "real" and "current" dollars. While the income by age of household distributions are provided only in current dollars, the overall distributions are given in "real" 1999 dollars (for 2004 and 2009) in order to permit trending of the income data in "real" terms which factors out the effects of inflation. Other the other hand, the current dollar series provide a basis for comparing markets with respect to income growth as the combined effect of both real income change and inflationary effects.

Income distributions for all households and by age of householder are derived from a model that advances in time each distribution in accordance with trends in mean and aggregate household income at the census tract and county level. At the census tract level an input to the income model is based on summarized mortgage application data. As discussed below, our estimates of Home Value use U.S. government reports of mortgage applications (a cumulative total of 14 million mortgages) to determine current distributions of Home Value for local areas. The same data source provides summary data on household income at the Census Tract level. This data provides an up-to-date, small area input to our income modeling process.

We extend the upper ends of the income distribution for all households beyond published Census Bureau figures using mathematical techniques constrained by our estimates and projections of aggregate income above \$200,000. While the extension of the income distributions is theoretically possible and consistent with higher-level controls, data users should necessarily apply an extra degree of caution in interpreting such information. In general, there is a higher degree of certainty associated with the distributions below \$200,000 and with the summary measures of income – mean and median household income.

Measures of per capita income are provided in two variables: (1) Per Capita Income is based on aggregate household income divided by the total population, and (2) Income Per Household Member is based on aggregate income divided by the household population.

Consumer Expenditure Potential

The traditional approach to “market potential” estimation involves assigning dollar estimates of household demand for consumer goods and services. Our approach to this type of market potential estimate makes use of the most recent four years of Consumer Expenditure Survey data from the U.S. Bureau of Labor Statistics (through the 2001 survey for the 2004 update) – the most extensive national survey of consumer expenditures available – and a powerful cluster analysis of U.S. neighborhoods, PSYTE US, the MapInfo proprietary geo-demographic cluster system.

The MapInfo estimates and projections of Consumer Expenditure Potential begin with an analysis of the Interview and Diary portions of the Consumer Expenditure Survey. The interview schedule covers the larger expenditures households make on an infrequent basis (e.g. refrigerators) as well as those made on a regular monthly basis (e.g. insurance payments). The diary portion covers expenditures made frequently, such as grocery items and drug store purchases. We use four years’ worth of CE data to achieve a sample size of 80,000+ households in an integrated dataset for diary and interview-based expenditures. The survey analysis establishes the relationship between detailed consumer expenditures and key demographic variables such as age, income, consumer unit size, family type, owner-renter status, and metropolitan residence.

The principal challenge of this type of exercise is to translate statistical relationships established at the national level to dollar estimates and projections for small areas of geography. Analysts run the risk of committing an “ecological fallacy” if results from a national survey are applied directly to smaller geographic areas. Our approach to this problem is to use results from our block group and neighborhood classification system – PSYTE US.

In essence, our method is a two-step process:

1. First we develop a typical, annual household “budget” for each PSYTE US cluster at the national level based on the patterns of expenditures found among Consumer Expenditure

Survey respondents. The survey respondents are matched to each PSTYE US Advantage cluster based on the key demographic variables that drive both cluster membership and household expenditures: age of householder, income, household size, family type, race/Hispanic origin, metropolitan area status, and homeownership.

2. Second, we apply the typical household “budget” (expressed as a percent of total annual expenditures) to our estimates of the dollars available for household expenditures at the block group level. Each cluster-specific household “budget” is applied to all corresponding block groups classified by PSYTE US cluster. Spending levels are adjusted for each block group within each cluster according to its “fit” with the national cluster profile. Finally, the estimates are controlled by our estimate of dollars available for consumer spending in the block group.

The result of this process is a set of estimates and projections of consumer market potential for over 350 goods and services for all levels of geography. These variables can be considered traditional demand-side market potential estimates. For example, custom trade areas can be drawn to represent geographic markets, which in turn represent a dollar potential of annual spending in a given category. Users may calculate the average market potential per household (mean) or create an index of market potential to compare several trade areas.

The five-year projections make the assumption that spending patterns remain constant while underlying demographic shifts, including population growth/decline and real income growth, drive changes in levels of category spending. As in our income variables, consumer expenditure dollars are held constant in the projection so that differences can be attributed to changes in income levels as well as underlying demographic change.

Retail Sales Potential

Retail Sales Potential provides estimates and projections of consumer potential organized by store type. As such, the Retail Sales Potential estimates are also “demand side” estimates. They are explicitly not derived from actual retail sales estimates. (Note: For “supply side” estimates of Retail Sales, see the

MapInfo Business Summary Data.) The variables in the Retail Sales Potential series are based on the variables in the latest Consumer Expenditure Potential series and are, therefore, consistent with the larger CEP data series from category to category and across the various geographies.

The store-type classification was developed from the new North American Industry Classification System (NAICS), which will replace the familiar SIC coding system. This system will provide consistency among the United States, Canada, and Mexico with respect to business activity statistics. The MapInfo Retail Sales Potential store categories reflect the new retail sector categories in the NAICS together with the consumer expenditure categories of the Bureau of Labor Statistics' Consumer Expenditure Survey.

Users will note that there is a natural overlap or double counting in data generated using store types as many consumer categories can be purchases in multiple store types. For example, a refrigerator may be purchased in an appliance store or a department store. For a complete listing of the consumer expenditure variables included in each store type, please see the following table. This table contains a listing of the items typically sold by each store type. The items are correlated with there corresponding Consumer Expenditure Potential codes and descriptions.

RSP Variable	Retail Sales Potential Variable List by Store Type
AU4411	Automobile dealers
AU11000	Cars and trucks, new (net outlay)
AU14000	Cars and trucks, used (net outlay)
AA12001	Automobile maintenance and repairs, total
AU4412	Other motor vehicle dealers
AU17000	New motorcycles, motor scooters, mopeds (net outlay)
AU18000	Used motorcycles, motor scooters, mopeds (net outlay)
AU19000	Motorized camper (net outlay)
AU44131	Automotive parts, accessories, and repair stores
AA12101	Coolant, additives, brakes/transmission fluid
AA12103	Parts, equipment, accessories
AA12104	Audio equipment excluding labor
AA12105	Body work, painting, upholstery
AA12106	Clutch & transmission repair
AA12107	Drive shaft & rear-end repair
AA12108	Brake work, excluding brake adjustment
AA12109	Steering or front end repair
AA12110	Cooling system repair
AA12111	Motor tune-up

	AA12112	Lubrication & oil changes
	AA12113	Front end alignment, wheel balance & rotation
	AA12114	Shock absorber replacement
	AA12115	Brake adjustment
	AA12117	Vehicle air conditioner repair
	AA12118	Exhaust system repair
	AA12119	Electrical system repair
	AA12120	Motor repair & replacement
	AA12121	Vehicle accessories including labor
	AA12122	Vehicle audio equipment including labor
AU44132	Tire dealers	
	AA12102	Tires, purchased, replaced, installed
	AA12116	Tire & other repair work
BU44411	Home centers	
	HO11100	Floor coverings, total
	HO12201	Paint/wallpaper and supplies - own home
	HO12202	Painting/papering equip. - own home
	HO12203	Panel/siding, etc. supplies - own home
	HO12204	Roofing/gutters mat. /equip. - own home
	HO12205	Patio, masonry, etc. mat. - own home
	HO12206	Plumbing supplies/equip. - own home
	HO12207	Electrical, heating, cooling supplies - own home
	HO12208	Insulation, other improvements, repair, maintenance - own home
	HO12209	Landscaping materials - own home
	HO19004	Power tools
	HO19005	Non-power tools
	HO19006	Grills and other Outdoor equipment
	HO19007	Fresh flowers or potted plants
	HO19008	Smoke alarms & detectors, purchases & rentals of - own home
	HO23105	Lawn-mowing equipment & other yard machinery
BU44412	Paint & wallpaper stores	
	HO12201	Paint/wallpaper and supplies - own home
	HO12202	Painting/papering equip. - own home
BU44413	Hardware stores	
	HO19004	Power tools
	HO19005	Non-power tools
	HO12206	Plumbing supplies/equip. - own home
	HO12207	Electrical, heating, cooling supplies - own home
BU44421	Outdoor power equipment stores	
	HO23105	Lawn-mowing equipment & other yard machinery
BU44422	Nursery & garden centers	
	HO19007	Fresh flowers or potted plants
	HO23110	Gardening/lawn care services

CA44811	Men's clothing stores	
	AP11000	Men's apparel, 16 and over, total
CA44812	Women's clothing stores	
	AP13300	Women's apparel, 16 and over, total
CA44813	Children's & infants' clothing stores	
	AP12200	Boys apparel, 2 to 15, total
	AP14000	Girls apparel, 2 to 15, total
CA4482	Shoe stores	
	FO11000	Footwear, total
CA44831	Jewelry stores	
	JE10000	Jewelry items, total
EA44311	Appliance, television, & other electronics stores	
	AL11001	Major appliances, total
	AL21405	Small electrical kitchen appliances
	AL21200	Portable heating & cooling equipment
	AL21305	Telephones & accessories
	AL21310	Telephone answering devices
	AL31100	Color TV console & combinations of TV; large screen color TV projection equipment..
	AL31200	Color TV (portable & table models)
	AL31300	VCR, video disc player, video camera, & camcorder
	AL31400	Video cassettes, tapes, & discs
	AL31500	TV video game hardware/software and computer game software
	AL31600	Radios
	AL31700	Tape recorders & players
	AL31800	Sound components, component systems, & compact disc sound systems
EA44312	Computer & software stores	
	AL21403	Computers, computer systems, & related hardware - home use
	AL21406	Computer software & accessories - home use
	AL21409	Computer repair services
	AL31500	TV video game hardware/software and computer game software
EA44313	Camera & photographic supplies stores	
	PH10000	Photographic, total
FB44511	Supermarkets & other grocery stores (except convenience)	
	GR11000	Food items at grocery stores, total
	HL12001	Health and Beauty Aid, total
	HL12100	Personal care products, total
FB44512	Convenience and specialty food stores	
	FS12001	Food at convenience/specialty stores

FB4453	Beer, wine, & liquor stores
	LI10000 Alcoholic beverages, total
FU4421	Furniture stores
	FU10001 Furniture, total
FU4422	Home furnishings stores
	HF10000 Home Furnishings, Cookware, and Houseware Items, total
GA44711	Gasoline stations with convenience stores
	AA11000 Gasoline and oil, total
	FS12001 Food at convenience/specialty stores
GA44719	Other gasoline stations
	AA11000 Gasoline and oil, total
GR452	General merchandise stores
	AL20411 Computing Equipment, Telephones, and Small Appliances, total
	AL31300 VCR, video disc player, video camera, & camcorder
	AL31400 Video cassettes, tapes, & discs
	AL31500 TV video game hardware/software and computer game software
	AL31600 Radios
	AL31700 Tape recorders & players
	AL31800 Sound components, component systems, & compact disc sound systems
	AL32001 Compact discs, tapes, needles, or records not from a club
	AP11000 Men's apparel, 16 and over, total
	AP12200 Boys apparel, 2 to 15, total
	AP13300 Women's apparel, 16 and over, total
	AP14000 Girls apparel, 2 to 15, total
	EN13500 Toys, games, hobbies, tricycles
	FO10000 Footwear, total
	HF10000 Home Furnishings, Cookware, and Houseware Items, total
	HL12001 Health and Beauty Aid, total
	HL12100 Personal care products, total
	JE10000 Jewelry items, total
	AL11001 Major appliances, total
	AL21405 Small electrical kitchen appliances
	AL21200 Portable heating & cooling equipment
	AL21305 Telephones & accessories
	AL21310 Telephone answering devices
	AL31100 Color TV console & combinations of TV; large screen color TV projection equipment...
	AL31200 Color TV (portable & table models)
	AL31300 VCR, video disc player, video camera, & camcorder
	AL31400 Video cassettes, tapes, & discs
	AL31500 TV video game hardware/software and computer game software
	AL31600 Radios
	AL31700 Tape recorders & players
	AL31800 Sound components, component systems, & compact disc sound systems

GR4521	Department stores (exc. leased depts.)
AL20411	Computing Equipment, Telephones, and Small Appliances, total
AL31300	VCR, video disc player, video camera, & camcorder
AL31400	Video cassettes, tapes, & discs
AL31500	TV video game hardware/software and computer game software
AL31600	Radios
AL31700	Tape recorders & players
AL31800	Sound components, component systems, & compact disc sound systems
AL32001	Compact discs, tapes, needles, or records not from a club
AP11000	Men's apparel, 16 and over, total
AP12200	Boys apparel, 2 to 15, total
AP13300	Women's apparel, 16 and over, total
AP14000	Girls apparel, 2 to 15, total
EN13500	Toys, games, hobbies, tricycles
FO10000	Footwear, total
HF10000	Home Furnishings, Cookware, and Houseware Items, total
HL12001	Health and Beauty Aid, total
HL12100	Personal care products, total
JE10000	Jewelry items, total
FU10001	Furniture, total
HP44611	Pharmacies & drug stores
HL13000	Prescription and OTC drugs and medical supplies, total
HL12001	Health and Beauty Aid, total
HL12100	Personal care products, total
HP44619	Other health & personal care stores
HL12200	Personal care services for females, including haircuts
HL12300	Personal care services for males, including haircuts
MR45391	Pet & pet supplies stores
EN12100	Pets, pet supplies and medicines for pets
EN12200	Pet food
RE	Restaurant Expenditures
FS11000	Full service restaurants (excludes alcoholic beverages)
SP45111	Sporting goods stores
SP10000	Sporting goods, total
SP45112	Hobby, toy, & game stores
EN13500	Toys, games, hobbies, tricycles
SP45113	Sewing, needlework, & piece goods stores
AP15000	Sewing materials for making clothes
AP16000	Sewing notions, patterns
SP45114	Musical instrument & supplies stores
MU10000	Music items, total

SP4512	Book, periodical, & music stores
	EN14600 Books not through book clubs
	EN14200 Newspapers, non-subscriptions
	EN14400 Magazines, non-subscription
	AL32001 Compact discs, tapes, needles, or records not from a club
SP45122	Video Tape Stores, Retail
	VD10000 Rental of video cassettes, tapes, & discs
GAF	GAF stores (General Merchandise, Apparel, and Furniture)
	Consists of the following RSP Codes:
	GR452 General merchandise stores
	AP11000 Men's apparel, 16 and over, total
	AP13300 Women's apparel, 16 and over, total
	AP12200 Boys apparel, 2 to 15, total
	AP14000 Girls apparel, 2 to 15, total
	FU4421 Furniture stores

Household Wealth (Net Worth) and Financial Assets

The Survey of Consumer Finances from the Federal Reserve Board is the only ongoing, comprehensive national survey that tracks household-level assets, liabilities and consequent net worth. We use the detailed demographic and financial information from this survey to establish relationships between demographic characteristics of households and net worth. The analysis also provides estimates of ownership and relative value of financial assets. Our model shows a clear relationship between mean income for specific demographic groups and net worth.

We apply the model to smaller geographies, conditioned by mean household income, in a manner analogous to our estimates of consumer expenditures as described in the previous section. By repeating the national survey analysis for each the PSYTE US Advantage geo-demographic clusters, we generate a wealth or net worth profile and a financial asset profile for each cluster. Each cluster's demographic composition provides the link to small geographic entities such as block groups. However, as income and home value at the block group level can condition estimates of wealth and financial assets within each cluster, we adjust wealth averages and distributions according to income and home value at the block group level. We use block group level income estimates alone to condition the distributions of Financial Assets.

The results are presented as mean and median wealth estimates as well as distributions of wealth. Similarly, estimates of financial assets are presented as means, medians, and distributions. The following are definitions of these concepts and their component parts:

Wealth is household net worth, or the difference between assets and liabilities at the household level. Assets include vehicles, primary residence, investment real estate, business assets, and a residual category of non-financial assets, plus financial assets. **Financial Assets** include transaction accounts (e.g. checking accounts), certificates of deposit, savings bonds, bonds, stocks, mutual funds, retirement accounts, cash value of life insurance, and a residual category of other managed assets and other financial assets. The concept of financial assets is a subset of the components of household wealth.

Components of Financial Assets (FA)

- Transaction Accounts
- CDs
- Savings Bonds
- Bonds
- Mutual Funds (excluding money market accounts)
- Retirement accounts
- Cash Value of life insurance
- Other Managed assets
- All other financial assets

Components of Non-Financial Assets

- Vehicles
- Primary Residence
- Investment Real estate
- Business Assets
- Other Non-financial assets

Financial Assets + Non-Financial Assets = Total Assets

Components of Liabilities

- Home Mortgage
- Home Equity
- Lines of Credit (secured by Home)
- Installment Loans
- Other Lines of Credit
- Credit Card Balance
- All other Debt

Total Assets – Liabilities = Net Worth

Home Value

The MapInfo approach to estimating a current-year home value distribution involves using a unique resource: complete US government records of all mortgages consummated over the most recent three period (though 2002 in the 2004 update) – representing more than 14 million records of recent home sales. This database includes all mortgages approved and accepted for single-family housing. Through a proprietary process we have attached census tract codes to these records and created summary distributions at the tract level. Our process involves, first, adjusting the 2000 distribution of home values for specified owner-occupied housing to current dollars. (“Specified” owner-occupied properties are a subset of all owner-occupied properties being those owner-occupied housing units with no more than 10 acres of land, no commercial space in the building, and no mobile homes. Condominiums are also not included in these estimates.)

Next, we calculate current means and medians, as well as distributions, of home value for all U.S. census tracts from the mortgage database adjusted to estimate home value. Our model is applied against estimates of “specified” owner-occupied housing derived as a subset of our household estimates. A final step distributes the census tract results to block groups based on the dollar-adjusted 1990 distributions. The block group distributions are shifted in accordance with changes in the average home value for each block group relative to its parent census tract.

The basic assumption of this approach is that current home values are best reflected in the most recent sale prices, values, and mortgages extended in small areas such as census tracts. Results of our home value methodology are given for the current year only and are not available as a five-year projection.

Census Update Variables

A set of variables referred to as “Census Update Variables” is derived from the 2000 Census and applied to the current year estimates of the population by selected age groups. These variables include: Population Aged 25+ by Educational Attainment, Population Aged 15+ by Marital Status, and Population Aged 16+ by Occupational Group. Also available in this set are estimates of owner and renter occupied housing units by the number of units in the structure. The base data for this variable set is the 2000 Census and current year update demographics.

In essence, the percent distribution of the population, as it was at the Block Group level in 2000, is applied to the new, current year base population. While this is a so-called “straight line” projection method, users should realize that the distributions will shift as they are rolled up to all higher geographies. The distributions shift, for example, at the county level according to the differential population growth of the sub-county geographies. If higher educational attainment areas grew faster than lower educational attainment areas, then the county may show a higher pattern of educational attainment. Similarly, Zip Code summary data may shift if differential growth patterns are exhibited across Block Groups that make up the Zip Code.

New (continuing from 2003) MapInfo “census update” include:

- Owner-occupied housing units by number of units in structure
- Renter-occupied housing units by number of units in structure
- Index of Relative Potential for Vacancy
- Index of Relative Potential for Rental Vacancy
- Index of Relative Potential for Seasonal Vacancy

Housing tenure by number of units in structure is based on the corresponding Census 2000 variables and updated household estimates. The relative potential indexes are based on the vacancy variables in Census 2000. The vacancy index can be interpreted in terms of the percent increase in Total Households that would be needed to fully occupy the existing stock of housing units. The Index of Relative Potential for Seasonal Vacancy is similarly interpreted and is, in that sense, an indicator of the influence of seasonality on the number of households in an area.

Daytime Population

The 2004 Estimates include a set of variables called “Daytime Population.” Daytime Population has two components: At-Home Population and At-Work Population (total employees). The At-Home Population is the current estimate of the number of persons aged 16+ that are not in the labor force and, therefore, presumed to be at home during the day. The At-Work Population is based on the MapInfo Business Summary Data, which contains estimates of the number of persons who work in the given Block

Group. The addition of the At-Home Population and the At-Work Population gives the estimate of the number of persons in the Block Group during day.

A detailed description of the Business Summary Data methodology, including the development of estimates of employees by SIC and NAICS by Block Group, is provided with the Business Summary Data package.

A Note on Data Sprinkling

Any set of data that involves summations within and across geographic units, as well as within and across demographic categories necessarily involves numerical rounding due to ratio adjustments such as those used in iterative proportional fitting. MapInfo has developed procedures to eliminate differences in summations due to rounding by systematically “sprinkling” the differences across subcategories within a data series in a manner that preserves the data distribution while simultaneously providing exact demographic and geographic summations.

Conclusion

Demographic estimates and projections developed by MapInfo provide a sound basis for market analysis and business planning. Users should keep in mind, however, that a degree of uncertainty exists in any comprehensive set of demographic estimates and projections. Variation can be greater across smaller geographic units and in areas undergoing rapid demographic change. Nevertheless, our methodology as described in this document provides a consistent framework and a set of demographic estimates and projections that can be used confidently for making reasonable comparisons across the entire U.S.

We are interested in any comments readers may have on this methodology statement. Please address comments to:

Thomas G. Exter, Ph.D.
Chief Demographer
MapInfo Corporation
26 Wellington St., East
Toronto, Ontario
Canada M5E 1S2

Email: tom_exter@mapinfo.com