Water Demand and Demographics: An Exploratory Analysis

Prepared by

William F. Gayk
California State University, Fullerton
Center for Demographic Research

Prepared for

National Water Research Institute

April 2004

INTRODUCTION

Water agencies routinely forecast water demand to more effectively deliver and plan for a dependable water supply and the necessary capital improvements to ensure its reliability. There are a variety of models and techniques that are used to produce a forecast. Some of the more common models include regression models, econometric models, time series, and even artificial intelligence (Brekke, et. al., 2002; Kiefer and Kocik, 1999; Jain and Ormsbee, 2002; Billings and Jones, 1996). The model may have a simple structure that forecasts future water demand by applying the current ratio of water use per capita to a forecast of population or a more elaborate structure that forecasts water demand by user sectors or "sectoral disaggregations" (Billings and Jones, 1996). Most of these models incorporate one or more "predictor" variables that have a known or assumed relationship with water consumption. A wide variety of variables are incorporated into these models as explanatory variables. Some of the models, especially those focusing on short-term forecasts, will include meteorological variables, such as temperature and rainfall, and hydrologic conditions, such as groundwater withdrawals and stormwater runoff (Megic, et. al. 2004). Other models that focus on the long-term changes in the size and composition of the customer base, or "rate payers", will incorporate socio-economic variables. The most common of these socio-economic variables include population, housing units or households by type (single-family and multi-family), persons per household, income, and employment by industry (Billings and Jones, 1996; Metropolitan Water District of Southern California, 2000). These variables are assumed to have a direct or indirect relationship with water consumption. For

example, in the Metropolitan Water District of Southern California's MWD-Main Water Use Forecasting System, water demand per single-family residential household is a function of both the number of single-family households and the average household size of these households and their income (Metropolitan Water District of Southern California, 2000). Some of the relationships between these variables and water consumption are as follows:

> *Population* – Changes in the population are a direct measure of the changes in the size of the number of people who will be consuming water for bathing, toilet use, cooking, drinking etc. As population expands or contracts, water demand increases or decreases.
>
> *Housing Units or Households by Type*– Like population, housing units are a direct measure of the residential customer base or the number of rate payers. However, the type of units will have an influence on demand. Single-family dwelling units require more water for lawn and garden uses as opposed to higher density multiple-family dwelling units that may have, at most, a common area. These may often be treated as an "end use" variable or a customer sector.
>
> *Persons per Household* – this is a measure of population density. As the demand by sub areas within a service area becomes important, the distribution and the density of the population differentiates the distribution of the population.
>
> *Income* – This is a measure of economic resources and life-style, both of which can influence water consumption. In the first place, income will influence the number and types of amenities a population will possess such as automobiles, motor homes, swimming pools and spas. In addition, income can be related to activities such as washing cars, bathing habits, etc. Finally, there is an income effect meaning that the relative cost of water decreases as income increases.
>
> *Employment by Industry* – Like dwelling units, employment differentiates the industrial from the residential sector, but may also differentiate within the sector. For example, producers of computer chips consume vast amounts of water in the manufacturing process, while wholesale trade companies will generally use minimal amounts of water.

These socio-economic factors are all important for forecasting water consumption, but there are other key population or demographic characteristics that, in all likelihood, may also influence water consumption and for that reason demand.

Foremost among these demographic characteristics is the age structure of the population and, as important in a multi-cultural society, the racial and ethnic composition of the population. These factors have been suggested but little research is available that has focused on the influence of such demographic characteristics on the consumption of water. If such factors strongly influence water consumption, then the inclusion of such variables in the models could help to better specify the models. Let us begin by asking "Why are these characteristics important?"

Social scientists, both academic and popular, and marketing researchers have all paid close and particular attention to age cohorts and racial/ethnic groups as they have strongly influenced so many American institutions and consumer practices. Age is important for two reasons. First as people age, they pass through various stages of life that are characterized by different physical and social needs (Fuchs, 1983). This is clearly obvious when contrasting those of an infant to those of an adult. Age is also symptomatic of one's generation. Karl Mannheim, a German Sociologist, noted that each generation is influenced by the critical events during one's youth that then form into their view of the world (Mannheim, 1952). These world-views will influence one's values and how they work, recreate consume, and behave, which includes their civic engagement. It is this latter that is particularly important for this discussion, because water can be looked at from the perspective of a consumer product, that is it is a product that is consumed and that different segments of the population will consume it at different rates. Edmunds and Turner (2002) note the importance of the post-World War II generation to the rise of modern consumerism.

According to Strauss and Howe (1991), there are three broad generational cohorts into which most living American adults fall. The first includes those that were born or raised during the Great Depression through the early years of World War II. Today, they are generally between the ages of 60 and 80 years. Their formative years were marked by hardship and sacrifice, first a depression that lasted for more than ten years followed by nearly five years of war. Their consumption habits tend to be conservative and frugal, always saving for that inevitable rainy day. Strauss and Howe (1991, 1997) have labeled this group the "Silent" generation.

The second of the large adult generations is the "Baby Boom" or "Boomer" generation. Although the precise births years of the Boomer generation are debated, it is generally a label affixed to the cohort of persons born immediately following World War II up to 1964 with the introduction and widespread use of the "pill." Nearly 83 million persons in the United States are part of the part of the Baby Boom generation. Even in California, with nearly 20 years of unprecedented levels of immigration that have reshaped the racial and ethnic composition, Baby Boomers still represent the single largest generation. The "Baby Boom" generation is perhaps the first generation to be labeled as such and is characterized as a major social force. For years, journalists, historians, and social scientists have written and lectured on the "psyche," culture and influence of the Boomer generation. So pervasive has the attention been on this generation that the terms "Baby Boom," "Baby Boomer," and "Boomer" are all part of our everyday vernacular. They grew up in an age of economic expansion, economic prosperity, educational opportunity, "suburbanization," housing opportunities, and expanding public infrastructure among others. "Baby boomers" dominate markets by

their shear size, have set many new trends and redefined numerous norms and practices in this country. Many expect that as these boomers move into their senior years they will also redefine what it means to be an older American. They are expected to be healthier than their predecessors, meaning their lifestyles will be more active, they will be able and may be willing to work beyond the traditional retirement ages, and they will maintain their high levels of consumption.

The generation following the Baby Boomers has been labeled both "Generation X" or "Gen X" (Coupland, 1991) and "Generation 13" or "13thGen" (Strauss and Howe, Howe and Strauss, 1993). Like the previous generations, the exact years of birth are not precise but generally run from the early 1960s to the early 1980s. Generation X was raised in an era marked by war, civil strife, political scandals, energy crises, failures in public education, Americans held hostage, declining worker benefits and, perhaps most important, a period of relative economic stagnation. This generation has been characterized as cynical, yet pragmatic. It has a preference for brand names and is willing to go into debt to consume: it carries the highest per-capita consumer debt of all adult generations. This is the first generation to have a curriculum that includes environmental topics throughout their education.

The oldest of Generation Y, or the "Millennials," as they are labeled by Howe and Strauss (200), are now just reaching adulthood. These are persons born during the 80s and beyond. Although this generation has not fully reached adulthood, it points to the growing importance of race and ethnicity: nearly one out of every three persons in this age cohort is other than a non-Hispanic White. This is particularly so in states such as California, Florida, New York, Arizona, Illinois and Texas that are experiencing

significant changes in their racial and ethnic compositions due to immigration. Much of this growth has been among the Latino and Asian populations. In fact, Latinos have now surpassed African-Americans as the second largest minority group in the United States, while Asians are the fastest growing minority group. These people are bringing with them cultural practices from their homelands that are having noticeable effects on domestic markets. Latinos, in particular, are having a profound influence on water consumption. Williams, et. al. (2001) found in a survey of Tucson area residents that Hispanics are four times as likely to drink bottled water than tap or filtered water. Baldassare (2002) found similar results in California. In a statewide survey, 55 percent of the Latinos indicated that they drink bottled water compared to 30 percent of the non-Hispanic Whites. Latinos were also more likely to state that pollution of drinking water is a problem. Similarly, Williams et. al. (2001) found that Latinos were less likely to report that the smell and taste of water was acceptable. These beliefs and practices on the part of Latinos are rooted in the poor quality of municipal water in their native lands.

The research reported below is an exploratory analysis that focuses on the effects of age and ethnicity on water consumption. Water use is another form of consumption. Like food and clothing, it is a necessity but also like food and clothing, it is a commodity. It follows, like other products, that its consumption will be influenced by lifestyle and cultural practices that are tied to age and ethnicity of the consumer. The essential purpose of this exploratory study is to address the question of whether age and ethnicity should be considered as variables in models that forecast future water demand, especially models that forecast residential use.

METHODOLOGY

Data

In order to undertake this analysis, it was critical to look at the relationship between water users and the volume of their use in relationship to their demographic characteristics. Such data generally does not exist in a level of detail that would contain information on both water consumption and the demographic characteristics of the customer households. Recognizing the unavailability of such data and the great expense necessary to develop such data, a geographic based data set containing both 2000 Census and water customer consumption information was utilized as a more expedient and efficient approach. The census block was selected as the "unit of analysis" since it is the smallest level of census geography that would include enough detailed demographic information for this study. In order to take this approach it was also necessary to obtain residential customer records. The Irvine Ranch Water District (IRWD) made a customer-based file available for the study. IRWD serves approximately 75,000 residential customers in central Orange County, covering all of the City of Irvine and parts of the cities of Costa Mesa, Lake Forest, Newport Beach and Tustin. The IRWD customer file included the customer's address, type of account, type of user and the amount of water consumed in CCFs (100 cubic feet). To match the customer information with census information, it was necessary to first determine the 2000 Census version of census tract and census block location of each customer. In order to determine the location of each, the IRWD customer file was geo-coded to census geography using Geographic Information System (GIS) technology. The geo-coding process involves matching the address on the customer file to a unique address range in a geographic reference file that

is associated with a census tract and block. Where a customer record is successfully matched, relevant geographic (latitude and longitude) and census geography is assigned to the record. A multi-stage process using two geo-coding software engines was employed to maximize the "match rate" (the percentage of records which match successfully during geo-coding). By utilizing more than one geo-coding software engine, the numbers of non-geo-coded addresses are minimized because of the different address matching protocols used in each program. This also holds true when utilizing multiple reference databases, if the address cannot be found in one database, there is a possibility that it could be found in another. The MapMarker Plus and ArcView geocoding engines were used for the geocoding. The reference databases include 2000 U.S Census TIGER/Line files, enhanced 2000 U.S. Census TIGER Street Geometry and related information from GDT's Dynamap/2000® file as well as the Thomas Bros. Street Centerline file.

The Irvine Ranch Water District provided water consumption by address for a total of 73,963 residential customer records. MapMarker Plus was initially used as the geo-coding engine. Using enhanced 2000 U.S. Census TIGER Street Geometry as the reference database, MapMarker Plus can locate addresses with street level precision. Equally important, the software can return Census Block IDs as output, or partial Census information, depending on the quality of the match.

The residential water consumption file was initially geo-coded using the strictest match settings in MapMarker Plus. This pass requires exact matches on address number, street name, city, and ZIP Code. All records geo-coded during this pass are at street level accuracy. The first pass resulted in a 63% match rate. A second pass was performed on

only unmatched records in the residential water consumption file. In this pass, ZIP Codes were relaxed. Relaxing criteria allows the software to search a wider area for a match while increasing the match rate. Each match in this pass was still at street level accuracy. The second pass resulted in an increased match rate of 28% for unmatched records; the total match rate for the residential water consumption file in the second pass was 91%. All records geo-coded in MapMarker Plus were automatically assigned Census Block IDs based on their street address position. In an effort to further maximize the match rate, residential water consumption records that were unmatched in MapMarker Plus were exported and geo-coded using ArcView's geo-coding engine. The 2000 U.S. Census TIGER/Line files were used as the reference database in this phase. This pass was also at street level accuracy. The third pass again resulted in an increased match rate for unmatched records; the match rate for the entire residential water consumption file after three passes was 97%. The ArcView geocoding engine does not automatically assign block information to records, as does MapMarker Plus. Hence, for those records geo-coded in ArcView, a spatial join between 2000 Block boundary information and residential water consumption data was performed. The file was then merged with the original data. Finally, auditing of a representative sample of the geo-coded data was performed. Samples of the geo-coded residential water consumption records were reviewed against 2000 U.S. Census address and block information. If errors were found, the records were moved to their correct geographic location, and block IDs updated. Latitude and longitude were updated as well.

The IRWD residential customer file was then summarized by census block to produce a data set with 2000, 2001 and 2002 water use, number of customers and number

of customer units. Two intermediary files were created that included a listing of each of the census tracts and blocks with IRWD residential customers and census tracts and block groups with IRWD residential customers. The block file was matched with 2000 Census Summary File 1 data and the block group file was matched with the 2000 Census Summary File 3 data. Population, age, housing, household and income information was selected from these census files and then merged with the block level IRWD customer file information.

Water Consumption

Residential water demand is generally expressed in two broad measures: use per housing unit (or customer) or use per capita. Each of these measures has its merits. Much of the residential water is consumed by activities of persons or tied to the number of persons in a household. For example, nearly 60 percent of the end-use of water consumption for single-family residential customers is tied directly to the number of individuals in the household (AWWA, 1993). These uses include activities such as toilet flushing, showering and bathing, washing clothes and dishes and faucet use. On the other hand, the customer is normally a household that would include the composite of all the population based water uses and those that would be tied to the household as a whole. Thus for single-family residential users, these could include lawn and gardening watering, evaporative coolers, humidifiers, swimming pools, and driveway and sidewalk cleaning. For this analysis it was decided to use both types of measures thereby providing an opportunity to assess the importance of the demographic variables on both per capita and household consumption.

There was another methodological issue that had to be resolved related to the measure of use. The data set contained water consumption data for the years 2000, 2001 and 2002. This presented a variety of ways of creating the numerator. These included measures for each individual year, an average of all three years, a sum of all three years, or one of the years. A major concern was change within the water district. Several areas within the IRWD service area are "high growth" areas. Several blocks showed major changes in the level of water consumption between 2000 and 2002. Several of these blocks contained no population or housing at the time of the 2000 Census. It was possible to assign estimates for those areas based on known land use characteristics and records of added housing, but at a considerable effort. For this analysis, it was decided that the most prudent approach was to only include those blocks that had census counts for 2000 and to base the demand measure on 2000 consumption. Two measures of demand were calculated for each block: "2000 Water Use per 2000 Housing Units" and "2000 Water Use per Capita."

After eliminating those blocks that did not have 2000 Census counts, the file required some additional "scrubbing" before it could be used for analyses. Several blocks had very high ratios of 2000 water use per housing unit and 2000 water use per capita. These outliers were also eliminated from the data set. In almost every instance, the blocks that were removed from the data set had less than five dwelling units according to the 2000 Census.

Explanatory Variables

In order to answer the basic question of "Does the age structure of the population and the racial/ethnic composition of the population influence water demand?" the analysis could have focused on simply the age composition and the racial and ethnic composition of the population. A more convincing answer to the question needed to include analysis comparing the different groups of socio-economic variables. Therefore, two groups of explanatory variables were selected. The first group included those socio-economic variables that are commonly incorporated into residential water demand models. We relied heavily on the Metropolitan Water District of Southern California's Demand Forecast model to identify these variables. These include persons-per-household, single-family detached units, and household income. In addition, the year structure was built was also included as an indirect measure of water conservation measures. It is assumed that newer homes will incorporate more types of water conservation technologies because these homes are more likely to be subject to more water conservation measures required through building codes. The second group included the population and ethnic variables. It was intended that the analysis would first address if the population variables influence water consumption and if they do, then how do they measure up against the more traditional variables.

The persons-per-household is a block level measure produced by the Census Bureau by dividing the residential population by the total number of households in each block. The percentage of single-family detached units was calculated for each block by dividing the number of single-family dwellings by the total count of all housing units. Household income is measured as the median household income at the block group level.

The lowest geography for which census income information is available is the census block group. A block group is a combination of two or more contiguous blocks. An assumption was made that the census blocks that comprise a census block group would have similar income levels because of their close geographic proximity. Year structure built is a block level census measure that reflects the median year the housing in a block were constructed.

Three age-related variables were created and labeled as follows for the analysis: "Percent Silent Generation," "Percent Baby Boomers," and "Percent Generation X." "Silent Generation" was defined as the population 55 years and older as of the 2000 Census, even though this variable includes some individuals who are categorized by Howe and Strauss as in the "G.I. Generation", that is persons born prior to 1925. "Baby Boomers" were defined operationally as the population between the ages of 35 and 54 in the year 2000. Percent Generation X was defined as the percent of population between the ages of 18 and 34.

Two race and ethnicity variables were used for this study. The first, the "Percent Population Hispanic or Latino" was chosen for this analysis, because of previous research indicating that this group does have different water consumption patterns. In addition, the "Percent Population Asian" was included because several areas in IRWD contain large numbers of Asians. Asians were individuals who did not classify themselves as Hispanic and were classified as "Population of one race: Asian Alone" in the 2000 Census. The denominator for both of the racial and ethnic variables was the total population.

For this exploratory analysis One-Way Analysis of Variance was selected to answer the basic question of whether the age and race and ethnicity variables influence water consumption. This technique determines if the means between different groups are significantly different. The basic approach was to determine if the average consumption of water was different given different concentrations of the demographic groups. In other words, if census blocks with a high concentration of Baby Boomers have a higher average use of water than census blocks with a low concentration of Baby Boomers, it would suggest that the particular age concentration does influence water consumption.

Each of the blocks was therefore classified into groups depending on the measured value of each of the explanatory variables. In order to classify the blocks into distinct groups, the Z-scores of the explanatory variables were calculated and plotted into histograms. Using the histograms, the blocks were placed into one of three 'natural' groups of the blocks based on the particular variables. For example, if a block had an average percent of Hispanics or Latinos, it was classified in the "Middle Value" group on the Percent Hispanic or Latino variable; if it had a below average percent of Hispanic or Latino, it was classified in the "Low Value" group on the Percent Hispanics or Latino variable; and if it had an above average percent of Hispanics or Latinos, it was classified in the "High Value" group on the Percent Hispanics or Latino variable.

FINDINGS

The overall results of the One-Way Analysis of Variance are presented in Table 1. To answer the first question of whether age and ethnicity make a difference in terms of water consumption, the answer is "yes." When water use is measured on a per capita basis, there was a significant difference in water use at the block level differentiated by the demographic variables as measured by the F statistic. Specifically, the average or mean water use per capita varies significantly for the Percent Silent Generation, Percent Generation X, Percent Baby Boomers, and Percent Hispanic. Water consumption differentiated by the Percent Asian was marginally significant for per capita water consumption but not significant for per dwelling water consumption.

Table 1
2000 Water Use per Capita and Dwelling Unit by Selected Explanatory Variable

Explanatory Variables	2000 CCF per 2000 Population		2000 CCF per 2000 Dwelling Units	
	F	Sig.	F	Sig.
Traditional Variables				
Persons per Household	0	.954	18.5	.000
Year Housing Built	2.9	.054	2.5	.081
Percent Single-Family Detached	7.3	.001	22.6	.000
Median Household Income	11.9	.000	17.9	.000
Demograhic Variables				
Percent Generation X	21.5	.000	15.4	.000
Percent Baby Boomers	18.1	.000	7.3	.001
Percent Silent Generation	9.0	.000	4.8	.008
Percent Hispanic or Latino	8.4	.000	10.4	.000
Percent Asian	2.4	.091	0.3	.709

The F Score was higher for two of the age-based variables, Percent Baby Boomer and Percent Generation X, than the traditional explanatory variables when the measure of consumption is population based. Only one of the traditional explanatory variables, Median Household Income, had a higher F Score, than any of these two population based

variables. Percent Hispanic and Percent Silent Generation had higher values of the F statistic than all but one of the traditional explanatory variables.

Interestingly and not surprising, when the denominator for the measure of consumption is housing units, the F Scores increased for the household variables such as owner occupied, median income and single-family attached. However, the demographic variables still remained strong in differentiating between average water use. All of the values of the F statistic are significant with the exception of Percent Asian, as was noted above.

The pattern of average use per capita broken down by each of the explanatory variables is displayed in Table 2. Generally speaking, for the traditional variables, the average Water Use per 2000 Population Average is in the expected direction. Specifically, water use increases as the percent Owner Occupied, Single-Family Detached, and Median Household Income increases. With Persons per Household and Year Housing Built, the average water use decreases and then increases slightly.

Table 2
Average 2000 CCF per 2000 Population by Explanatory Variables

Traditional Variables	**Lower Value**	**Middle Value**	**Higher Value**
Percent Single-Family Detached	44.5	67.2	80.5
Median Household Income	52.2	61.4	99.1
Persons per Household	71.2	68.6	70
Year Housing Built	81.7	63	66.1
Demograhic Variables			
Percent Generation X	133.1	62.9	58.2
Percent Baby Boomers	83	63	148.9
Percent Silent Generation	107.8	62.1	84.7
Percent Hispanic or Latino	88.3	58	61.6
Percent Asian	81.6	66.3	61.9

For the demographic variables, there are some conflicting patterns of water use. As the Percent Generation X population increases, water use declines by almost 50

percent from 133.1 CCF per 2000 Population to 58.2 CCF per 2000 Population. This may be tied to the fact that they are younger and may not have the same number of end uses as other age groups and more away-from-residence activities. As the percentage of Baby Boomers increases water use decreases, but then increases as the percentage of Baby Boomers increases. There is to some extent a multicollinearity effect here. As the number of Baby Boomers declines, the percentage of either or both Generation X and Silent Generation population increases. The average water use for both Generation X and the Silent Generation is likely influencing the average water use in these blocks. It should be pointed out however, at the extremes, water use increases as the percentage of Baby Boomers increases, as would be expected. Blocks that had the higher concentrations of Baby Boomers had higher levels of water use. Likewise, at the extremes for the Silent Generation, water use decreases as the percentage of this population increases. The same pattern exists for Hispanics. Blocks with the lowest concentration of Hispanics had higher levels of water use and for Asians the pattern was the opposite.

The average water per household consumption for the traditional explanatory variables follows in the expected pattern. For each of the variables, average water consumption per 2000 Dwelling Units increases as the value of the explanatory variables increases. For the demographic variables, the patterns are similar to those with water use per population. The actual pattern of average use by concentrations of Generation X population is as expected; a decline in water consumption as the concentration increases. With the other variables, the patterns between areas with the lowest and highest concentrations is as expected. Again, due to multicollinearity, areas with a mid-level

concentration are confounded. It should also be pointed out that the average use for all age groups with a mid-level value is about the same, further suggesting the percentages by these generation cohorts are interdependent.

Table 3
Average 2000 CFS per 2000 Dwelling Units by Explanatory Variables

Traditional Variables	Lower Value	Middle Value	Higher Value
Percent Single-Family Detached	103.2	181.9	235.5
Median Household Income	124.1	177.6	264.9
Persons per Household	236.6	269.2	308.1
Year Housing Built	126.8	195.4	258.0
Demograhic Variables			
Percent Generation X	302.1	188.3	155.4
Percent Baby Boomers	222.0	183.7	297.8
Percent Silent Generation	251.8	186.1	192.8
Percent Hispanic or Latino	238.1	167.0	172.5
Percent Asian	203.1	188.6	191.6

Questions can be raised regarding the relationships between the demographic and the traditional independent variables. For example, Hispanics and Latinos generally have a larger number of people residing in a household, and Boomer households are generally thought to be located primarily in suburban neighborhoods thick with single-family housing. This then raises questions if the relationship between the demographic variables and water consumption are spurious.

Table 4 presents a summary of several tables that measured the association between the two groups of variables. These tables are presented in Appendix I along with the summary statistics. For this part of the analysis, Gamma was used because it measures both the strength of the relationship between the two variables and, based on this measure, the approximate statistical significance of the relationship can be determined.

Table 4
Gamma Coefficients between Demographic and Traditional Variables

	Percent Generation X	Percent Baby Boomers	Percent Silent Generation	Percent Hispanic or Latino	Percent Asian
Percent Single-Family Detached	-0.38	0.14	-.02*	-.06*	0.12
Median Household Income	-0.38	0.21	-.06*	-0.33	0.24
Persons per Household	-0.34	-.01*	-0.18	.01*	0.21
Year Housing Built	0.32	-.07*	-0.52	-.07*	0.32

* p >.05, not significant

Of the five demographic variables, two of them were statistically associated with each of the four traditional variables, Percent Generation X and Percent Asian. For this analysis we can dismiss the Percent Asian since it was not associated with water consumption. In the case of the Percent Generation X, which was associated with levels of water consumption, these results cannot be dismissed. Recalling that this population was based on the number of persons between the ages of 18 and 34 years, these relationships do suggest that the patterns are related to their stage of the "life-cycle." Younger adults are more likely to live in multi-family dwellings, have lower incomes, be members of smaller households, and live in newer housing if they are owners. In this particular water district, it could also be noted that there was a large number of multi-family dwellings constructed at the time of the 2000 Census.

There is a positive relationship between the proportion of Baby Boomers and the Percent Single-Family Detached housing, but this relationship is not particularly strong; a much stronger relationship was expected. Part of the explanation for the weak relationship could be the measurement of the variable. Since it was based on population, the areas that had the highest percent of Baby Boomers had a lower percent of other age groups. Thus, the higher percentages would probably represent areas that were populated

with smaller adult households. The positive relationship between Percent Boomers and Median Household Income is expected because of the relationship between age and when one peaks in their career. These are individuals who are reaching the pinnacles of their career ladders.

One of the strongest relationships found here is between the Percent Silent Generation and Year Housing Built. This is more than likely an artifact of when people in this age group purchased the house rather than the degree to which they embrace water savings technology. Of course, the relationship with Persons per Household is as expected.

The expected relationship between the Percent Hispanic or Latino and Persons per Household was not found. This is an anomaly because it is well documented that Hispanics or Latinos do have larger households. This anomaly may be due to the areas covered by the Irvine Ranch Water District in that it does not include many areas with large concentrations of Hispanics or Latinos. In addition, these Hispanics and Latinos residing within the Irvine Ranch Water District are likely to live in areas with more multiple-family dwellings that also have larger concentrations of Generation X population. The smaller Generation X households may be pushing the average household size down in these areas. This is also consistent with the relationship between Percent Hispanic or Latino and Median Household Income. As a group, they would not have the income necessary to purchase housing, which would cluster them in areas with multiple-family dwellings.

Returning to the question of whether there may be spurious relationships between these demographic variables and water consumption, there is not a clear answer one way

or the other based on the results just discussed above. Additional analysis was conducted to evaluate the combined effect of the demographic variables and the traditional variables using Two-Way Analysis of Variance on the per capita water consumption. Here we focused on the nine statistically significant associations discussed above. The summary of the results is presented in Table 5 and the details in Appendix II. Table 5 indicates whether each of the main effects or interaction effects were statistically significant (MD=main effects demographic variables, MT=main effects traditional variables, I=interaction effects).

Table 5
Effects Within and Between Demographic and Traditional Variables

	Percent Generation X	Percent Baby Boomers	Percent Silent Generation	Percent Hispanic or Latino
Average CFS per 2000 Population				
Percent Single-Family Detached	MD,MT,I	MD,MT,I		
Median Household Income	MD,MT,I	MD,MT,I		MD,MT
Persons per Household	MD		MD,MT,I	
Year Housing Built	MD,MT,I		MD,MT,I	

These results suggest that the relationships between water consumption and the traditional explanatory variables when combined with the demographic variables are complicated. In most of the combinations, the main effects of each of the predictor variables, both traditional and demographic, are statistically significant and there is an interaction effect between the two predictor variables. The two exceptions are the combination of Percent Generation X and Persons per Household and Percent Hispanic or Latino and Median Household Income. In the former, only Percent Generation X was statistically significant and in the latter, both Percent Hispanic or Latino and Median Household Income were both statistically significant, but there was no interaction effect.

Generally, the patterns followed those as displayed in Table 3. For example, using Generation X and Single-Family housing, generally the per capita water consumption in an area decreased as the percentage of Generation X population increased and per capita water consumption increased as the percent of Single-Family housing increased. Thus, average water consumption is lowest in areas with a higher percentage of Generation X population and a lower percentage of Single-Family housing, and is highest in areas with a lower percentage of Generation X population and a higher percentage of Single-Family housing. One of the more significant exceptions to this general pattern is the combination of Percent Baby Boomers and Single-Family Dwellings (See Appendix II). Average water consumption decreases in areas where the Percentage of Single-Family dwellings is lower as the percentage of Baby Boomers increases. However, per capita water consumption increases in areas where the percentage of Single-Family dwellings is higher as the percentage of Baby Boomers increases. In fact, the combination of a high percent of Baby Boomers along with a high percent of Single-Family dwellings is associated with the second highest per capita water consumption (218.6 CCF) while the lowest is associated with a higher level of Baby Boomers and a low percentage of Single-Family dwellings (36 CCF) among the various combinations of traditional and demographic variables. The highest level (286 CCF) occurred in areas with a higher level of Baby Boomers and a high level of Median Household Income.

DISCUSSION

Before we begin the discussion, there are several cautions that need to be made. The first is that the data is aggregated data by census block. There is always the potential problem of an "ecological fallacy." This is attributing a pattern that is found in the aggregate to individuals. Thus, even though we find that areas with high concentrations of Generation Xers have lower water use, it does not establish that if a person in that area is in the Generation X cohort, they use less water. But since the purpose is to look at ways of forecasting future demand by a geographic area (normally a water district) we are looking for relationships with water consumption at the aggregate level not attempting to explain individual water consumption.

It needs to be noted that these data are also cross-sectional. They represent an empirical relationship at one time. This is important because the water consumption patterns may change over time. Also, the fact that it appears that age is important, it would be valuable to determine the relative importance of age as a measure of life cycle versus membership in a cohort or generation. It is essential to look at such relationships longitudinally. A final caution is that this study only included one water agency. It does contain a certain level of socio-economic diversity on the one hand, but it does tend to be more affluent and ethnically homogeneous than other areas of Orange County.

The purpose of this exploratory study was to address the question of "Does the age structure of the population and the racial/ethnic composition of the population influence water demand?" A secondary question was if the answer to the first question was "yes" then "Do these variables do as well as traditional variables in explaining different levels of water consumption?" The results do suggest that age structure of the population and the racial and ethnic composition do influence water consumption. At

this time, the answer is a qualified answer. First of all, these variables seemed to have much more influence when water consumption is measured on a per capita or population basis. The age structure, as measured by the percentage of the population in the Generation X and Baby Boomer age cohorts, strongly differentiates average water use. The same is the case for the percentage of the population that is Hispanic or Latino. The average water use was only weakly differentiated by the percentage of the population that is Asian. It needs to be pointed out that the census category of "Asian" is made up of many different groups who are not only culturally different, but have had very different life experiences. When water consumption is based on a per dwelling unit basis, average water consumption does differ by the concentrations of the population by age and Hispanic population, but only the percentage of the population in Generation X is as strongly associated with average water use as the more traditional explanatory variables.

These preliminary results also indicate that the relationships between the demographic variables and water consumption are not spurious because of their relationships with traditional variables. In fact, the results indicate that levels of water consumption in an area are influenced by combinations of these variables working in concert. Thus, areas with high concentrations of Baby Boomers and single-family dwellings were associated with some of the higher levels of water consumption. Because of the nature of census blocks, it is most probable that these Baby Boomers were living in the single-family dwellings. Income in combination with the percentage of Baby Boomers was also associated with higher levels of per capita water consumption.

Certainly further research is warranted and is needed, but these results do suggest that further consideration of race and ethnicity are worth pursuing. It may even be an

understatement for at least two reasons. The first is that the water consumption levels of areas with high concentrations of Baby Boomers were higher. In needs to be noted that, as these percentages increased, the percentage of other age groups decreased. This suggests that larger households that may be associated with Baby Boomers or single-family residences may not be intervening variables. Therefore, suggesting that the water consumption patterns of Baby Boomers holding all other things constant may be higher. This could have major implications for future water demand, because as this group ages, the overall average water consumption can change. Further research on age and ethnicity is also worth pursuing because many regional agencies are routinely producing population projections broken out by these characteristics. Thus, for many water agencies, these data may be available and they may help to improve the models used to forecast water demand.

REFERENCES

American Water Works Association, 1993. Evaluating Urban Water Conservation Programs: A Procedures Manual. Denver, CO: AWWA.

Baldassare, Mark "PPIC Statewide Survey: Special Survey on Californians and the Environment", June 2002 Public Policy Institute of California. San Francisco.

Brekke, Levi et. al. "Suburban Water Demand Modeling Using Stepwise Regression" Journal of American Water Works Association 94:10 October 2002, 65-75.

Billings, R. Bruce and Jones, C. Vaughan 1996. Forecasting Urban Water Demand. Denver: American Water Works Association.

Edmunds, June and Bryan S. Turner Generations, Culture and Society Buckingham, UK and Philadelphia, PA: Open University Press.

Fuchs, Victor R. How We Live. 1983 Cambridge, Massachusetts and London, England: Harvard University Press.

Gardyn, Rebecca "Habla English?" American Demographics, April 2001 54-57.

Jain, Ashu and Ormsbee, Lindell E. "Short-term Water Demand: Forecast Modeling Techniques-Conventional Methods versus AI" Journal of American Water Works Associations 94:7 July 2002, 64-72.

Mannheim, Karl "The Problem of Generations" in Collected Works of Karl Mannheim, pp. 276-320. London: Routledge.

Megic, Brian J. MacIntyre, David F., Aikens, Alan W. and Christina Beatham "The Effects of Urbanization on Water Supply" Florida Water Resources Journal February, 2004: 31-36.

Metropolitan Water District of Southern California "The Regional Urban Water Management Plan for the Metropolitan Water District of Southern of California" December 2000.

Raymond, Joan "The Joy of Empty Nesting" American Demographics, May 2000 48-54.

Stein Wellner, Alison "The Forgotten Baby Boom" American Demographics, Feb 2001 46-51.

Weiss, Michael J. "Great Expectations" American Demographics, May 2003 pp 27-35.001) 11 (6) Pp. 510-521.

Williams B.L. and Florez, Y and Pettygrove, S. "Inter- and Intra-ethnic variations in water intake, contact, and source estimates among Tucson Residents: Implications for Exposure Analysis," Journal of Exposure Analysis & Environmental Epidemiology, December 2001, 510-521.

APPENDIX I

Cross Tablulations of Traditional Variables with Demographic Variables

	Percent Asian							
Household Income	Lower Value		Middle Value		Higher Value		Total	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Lower Value	69	18.8%	59	8.4%	25	8.3%	153	11.2%
Middle Value	231	62.8%	478	67.9%	177	59.0%	886	64.6%
Higher Value	68	18.5%	167	23.7%	98	32.7%	333	24.3%
Total	368	100.0%	704	100.0%	300	100.0%	1372	100.0%

	Percent Baby Boomers							
Household Income	Lower Value		Middle Value		Higher Value		Total	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Lower Value	24	24.2%	122	10.3%	7	8.6%	153	11.2%
Middle Value	53	53.5%	786	66.1%	46	56.8%	885	64.6%
Higher Value	22	22.2%	282	23.7%	28	34.6%	332	24.2%
Total	99	100.0%	1190	100.0%	81	100.0%	1370	100.0%

	Percent Generation X							
Household Income	Lower Value		Middle Value		Higher Value		Total	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Lower Value	11	7.5%	64	7.0%	78	24.9%	153	11.2%
Middle Value	78	53.1%	622	68.2%	186	59.4%	886	64.6%
Higher Value	58	39.5%	226	24.8%	49	15.7%	333	24.3%
Total	147	100.0%	912	100.0%	313	100.0%	1372	100.0%

	Percent Hispanic							
Household Income	Lower Value		Middle Value		Higher Value		Total	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Lower Value	36	7.5%	64	10.5%	53	18.8%	153	11.2%
Middle Value	288	60.0%	393	64.4%	205	72.7%	886	64.6%
Higher Value	156	32.5%	153	25.1%	24	8.5%	333	24.3%
Total	480	100.0%	610	100.0%	282	100.0%	1372	100.0%

	Percent Silent Generation							
Household Income	Lower Value		Middle Value		Higher Value		Total	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Lower Value	16	10.6%	117	11.5%	20	10.0%	153	11.2%
Middle Value	84	55.6%	671	65.7%	131	65.5%	886	64.6%
Higher Value	51	33.8%	233	22.8%	49	24.5%	333	24.3%
Total	151	100.0%	1021	100.0%	200	100.0%	1372	100.0%

Percent Single-Family Detached Units	Percent Asian: Lower Value		Middle Value		Higher Value		Total	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Lower Value	59	15.9%	141	19.6%	52	17.1%	252	18.1%
Middle Value	161	43.3%	244	33.9%	75	24.7%	480	34.4%
Higher Value	152	40.9%	334	46.5%	177	58.2%	663	47.5%
Total	372	100.0%	719	100.0%	304	100.0%	1395	100.0%

Percent Single-Family Detached Units	Percent Baby Boomers: Lower Value		Middle Value		Higher Value		Total	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Lower Value	35	35.4%	204	16.8%	13	15.9%	252	18.1%
Middle Value	26	26.3%	420	34.7%	32	39.0%	478	34.3%
Higher Value	38	38.4%	588	48.5%	37	45.1%	663	47.6%
Total	99	100.0%	1212	100.0%	82	100.0%	1393	100.0%

Percent Single-Family Detached Units	Percent Generation X: Lower Value		Middle Value		Higher Value		Total	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Lower Value	20	13.2%	108	11.6%	124	39.2%	252	18.1%
Middle Value	57	37.7%	309	33.3%	114	36.1%	480	34.4%
Higher Value	74	49.0%	511	55.1%	78	24.7%	663	47.5%
Total	151	100.0%	928	100.0%	316	100.0%	1395	100.0%

Percent Single-Family Detached Units	Percent Hispanic: Lower Value		Middle Value		Higher Value		Total	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Lower Value	59	12.0%	132	21.3%	61	21.5%	252	18.1%
Middle Value	190	38.7%	191	30.8%	99	34.9%	480	34.4%
Higher Value	242	49.3%	297	47.9%	124	43.7%	663	47.5%
Total	491	100.0%	620	100.0%	284	100.0%	1395	100.0%

Percent Single-Family Detached Units	Percent Silent Generation: Lower Value		Middle Value		Higher Value		Total	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Lower Value	27	17.2%	181	17.4%	44	22.0%	252	18.1%
Middle Value	63	40.1%	350	33.7%	67	33.5%	480	34.4%
Higher Value	67	42.7%	507	48.8%	89	44.5%	663	47.5%
Total	157	100.0%	1038	100.0%	200	100.0%	1395	100.0%

Persons per Household	Percent Asian						Total	
	Lower Value		Middle Value		Higher Value			
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Lower Value	132	35.8%	161	22.6%	31	10.3%	324	23.4%
Middle Value	180	48.8%	406	56.9%	168	55.6%	754	54.5%
Higher Value	57	15.4%	146	20.5%	103	34.1%	306	22.1%
Total	369	100.0%	713	100.0%	302	100.0%	1384	100.0%

Persons per Household	Percent Baby Boomers						Total	
	Lower Value		Middle Value		Higher Value			
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Lower Value	33	33.7%	259	21.5%	32	0.4%	324	23.4%
Middle Value	49	50.0%	666	55.4%	37	45.1%	752	54.4%
Higher Value	16	16.3%	277	23.0%	13	15.9%	306	22.1%
Total	98	100.0%	1202	100.0%	82	100.0%	1382	100.0%

Persons per Household	Percent Generation X						Total	
	Lower Value		Middle Value		Higher Value			
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Lower Value	43	28.5%	129	14.1%	152	48.3%	324	23.4%
Middle Value	74	49.0%	550	59.9%	130	41.3%	754	54.5%
Higher Value	34	22.5%	239	26.0%	33	10.5%	306	22.1%
Total	151	100.0%	918	100.0%	315	100.0%	1384	100.0%

Persons per Household	Percent Hispanic						Total	
	Lower Value		Middle Value		Higher Value			
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Lower Value	119	24.4%	148	24.1%	57	20.3%	324	23.4%
Middle Value	255	52.3%	337	54.8%	162	57.7%	754	54.5%
Higher Value	114	23.4%	130	21.1%	62	22.1%	306	22.1%
Total	488	100.0%	615	100.0%	281	100.0%	1384	100.0%

Persons per Household	Percent Silent Generation						Total	
	Lower Value		Middle Value		Higher Value			
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Lower Value	30	19.2%	204	19.8%	90	45.5%	324	23.4%
Middle Value	65	41.7%	590	57.3%	99	50.0%	754	54.5%
Higher Value	61	39.1%	236	22.9%	9	4.5%	306	22.1%
Total	156	100.0%	1030	100.0%	198	100.0%	1384	100.0%

Year Housing Built	Percent Asian						Total	
	Lower Value		Middle Value		Higher Value			
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Lower Value	172	46.2%	190	26.4%	74	24.3%	436	31.3%
Middle Value	168	45.2%	380	52.9%	138	45.4%	686	49.2%
Higher Value	32	8.6%	149	20.7%	92	30.3%	273	19.6%
Total	372	100.0%	719	100.0%	304	100.0%	1395	100.0%

Year Housing Built	Percent Baby Boomers						Total	
	Lower Value		Middle Value		Higher Value			
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Lower Value	27	27.3%	389	32.1%	20	24.4%	436	31.3%
Middle Value	44	44.4%	591	48.8%	50	61.0%	685	49.2%
Higher Value	28	28.3%	232	19.1%	12	49.2%	272	19.5%
Total	99	100.0%	1212	100.0%	82	134.6%	1393	100.0%

Year Housing Built	Percent Generation X						Total	
	Lower Value		Middle Value		Higher Value			
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Lower Value	47	31.1%	323	34.8%	66	20.9%	436	31.3%
Middle Value	88	58.3%	478	51.5%	120	38.0%	686	49.2%
Higher Value	16	10.6%	127	13.7%	130	41.1%	273	19.6%
Total	151	100.0%	928	100.0%	316	100.0%	1395	100.0%

Year Housing Built	Percent Hispanic						Total	
	Lower Value		Middle Value		Higher Value			
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Lower Value	152	31.0%	166	26.8%	118	41.5%	436	31.3%
Middle Value	251	51.1%	309	49.8%	126	44.4%	686	49.2%
Higher Value	88	17.9%	145	23.4%	40	14.1%	273	19.6%
Total	491	100.0%	620	100.0%	284	100.0%	1395	100.0%

Year Housing Built	Percent Silent Generation						Total	
	Lower Value		Middle Value		Higher Value			
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Lower Value	16	10.2%	317	30.5%	103	51.5%	436	31.3%
Middle Value	68	43.3%	528	50.9%	90	45.0%	686	49.2%
Higher Value	73	46.5%	193	18.6%	7	3.5%	273	19.6%
Total	157	100.0%	1038	100.0%	200	100.0%	1395	100.0%

APPENDIX II

Average Water Consumption per Capita

Percent Single-Family Detached Units	Percent Generation X			
	Lower Value	Middle Value	Higher Value	Total
Lower Value	51.5	52.2	36.7	44.5
Middle Value	101.8	64.2	58.0	67.2
Higher Value	179.2	64.4	92.9	80.5
Total	133.1	62.9	58.3	69.4

Percent Generation X F= 7.8
Percent Single-Family Detached Units F= 14.6
Interaction F= 4.7

Median Household Income	Percent Generation X			
	Lower Value	Middle Value	Higher Value	Total
Lower Value	105.7	57.8	40.1	52.2
Middle Value	93.2	60.6	50.7	61.4
Higher Value	196.5	70.5	115.9	99.1
Total	134.9	62.9	58.3	69.5

Percent Generation X F= 9.4
Median Household Income F= 16.9
Interaction F= 4.8

Persons per Household	Percent Generation X			
	Lower Value	Middle Value	Higher Value	Total
Lower Value	166.0	61.9	52.1	71.2
Middle Value	130.3	62.2	60.3	68.5
Higher Value	97.3	64.7	80.2	70.0
Total	133.1	62.8	58.4	69.5

Percent Generation X F= 16.7
Persons per Household F= 0.6 not significant
Interaction F= 1.7 not significant

Year Built	Percent Generation X			
	Lower Value	Middle Value	Higher Value	Total
Lower Value	223.1	63.6	69.2	81.6
Middle Value	83.8	61.3	54.6	62.9
Higher Value	139.4	67.0	56.1	66.1
Total	133.1	62.9	58.3	69.4

Percent Generation X F= 20.4
Year Built F= 12.7
Interaction F= 8

Percent Single-Family Detached Units	Percent Baby Boomers			
	Lower Value	Middle Value	Higher Value	Total
Lower Value	46.3	44.7	36.0	44.5
Middle Value	84.7	62.5	114.2	67.2
Higher Value	115.5	69.6	218.6	80.5
Total	83.0	63.0	148.9	69.4

Percent Baby Boomers F= 9.1
Percent Single-Family Detached Units F= 15.3
Interaction F= 4.8

Median Household Income	Percent Baby Boomer			
	Lower Value	Middle Value	Higher Value	Total
Lower Value	45.6	47.3	160.7	52.2
Middle Value	54.8	61.6	65.6	61.4
Higher Value	191.5	73.4	286.0	99.2
Total	82.3	62.9	150.0	69.5

Percent Baby Boomers F= 18.5
Median Household Income F= 34.8
Interaction F= 4.8

Persons per Household	Percent Baby Boomer			
	Lower Value	Middle Value	Higher Value	Total
Lower Value	119.1	52.3	174.5	71.2
Middle Value	49.6	65.5	149.5	68.6
Higher Value	115.6	66.7	84.2	70.0
Total	83.8	62.9	48.9	69.5

Percent Silent Generation F= 12.6
Persons per Household F= 2 not significant
Interaction F= 3.5

Persons per Household	Percent Silent Generation			
	Lower Value	Middle Value	Higher Value	Total
Lower Value	208.2	49.8	74.0	71.2
Middle Value	91.5	63.8	81.9	68.6
Higher Value	74.4	69.1	62.3	70.0
Total	107.3	62.3	77.4	69.5

Percent Silent Generation F= 14.8
Persons per Household F= 4.5
Interaction F= 6.7

Year Built	Percent Silent Generation			
	Lower Value	Middle Value	Higher Value	Total
Lower Value	380.0	64.6	87.7	81.7
Middle Value	76.3	62.0	58.9	63.0
Higher Value	75.4	59.2	157.0	66.1
Total	106.8	62.3	77.2	69.4

Percent Generation X F= 39.6

Year Built F= 35.3

Interaction F= 20.8

Median Household Income	Percent Hispanic or Latino			
	Lower Value	Middle Value	Higher Value	Total
Lower Value	69.4	50.2	43.0	52.2
Middle Value	71.4	53.8	62.0	61.4
Higher Value	125.7	71.9	99.4	99.1
Total	88.9	57.9	61.6	69.5

Percent Hispanic or Latino F= 4

Median Household Income F= 6.4

Interaction F= 1.2 not significant