ACRP

Project Number 03-36

Using Disaggregated Socioeconomic Data in Air Passenger Demand Studies

Final Report

Appendix D

Detailed Case Study Analysis Results

6/28/18

TRANSPORTATION RESEARCH BOARD
NAS-NRC

Prepared by:

GRA, Incorporated

With:

Atlanta Analytics, LLC

and

Aviation System Consulting, LLC

Table of Contents

List of Figures and Tables

List of Figures and Tables (cont.) Page

List of Figures and Tables (cont.)

Page

List of Figures and Tables (cont.) Page

List of Figures and Tables (cont.)

Page

List of Figures and Tables (cont.)

Page

List of Figures and Tables (cont.)

Page

D.1 Introduction

This appendix to the project Final Report presents a more detailed analysis and discussion of the case study regression modeling for each of the seven case study airports and the case study airport system that is addressed in Section 4 of the Final Report. It also presents a more detailed analysis of the Baltimore-Washington region that is also included in Section 4 of the Final Report. Along with the analyses presented in the Final Report, these additional analyses contributed to the research results and conclusions that are presented in the Final Report.

Section D.2 of the appendix presents the regression analysis of the seven individual case study airports and the three-airport Baltimore-Washington metropolitan region. For each case study subject, data is provided to present, using figures and tables, the aviation activity and the socioeconomic environment in the regions served by the airports during the 1990 to 2010 sample period. Following the documentation of airport and regional activity, the individual airport regression results and regression model performance are presented.

Section D.3 presents a more detailed analysis of the Baltimore-Washington region which used an alternative disaggregated household income variable and a more complex model structure with several additional variables, including the evolution of the modeling approach and specifications that were used in the analysis.

D.2 Case Study Regression Models

a. Variables Included in Case Study Regression Models

As discussed in Appendices A and B, the demand for air travel is a derived demand, arising from the desire of individuals and groups to travel in pursuit of business or personal activities in locations relatively distant from their homes. Because of this, demand for air travel is customarily modeled as a function of socioeconomic factors, including population levels, employment levels, household and regional incomes, and other expressions or measurements of regional economic activity. These socioeconomic factors are an expression of the economic activity in a region that leads to business air travel to and from the region as well as the existence of household incomes that enable personal travel.

Six aggregate socioeconomic variables were selected for use as independent variables in the regressions for eight case study examples, serving as representatives or measures of the socioeconomic activity that "drives" air travel demand (and "drives" demand for other forms or modes of transportation). These six candidate independent socioeconomic variables are:

- Regional Population
- Regional Total Employment
- Regional Total Earnings (reported in $2013, converted to $2015 using the Consumer Price Index)(footnote: this is applied to all monetary variables, does not matter for model performance)
- Regional Wages and Salaries ($2015)
- Gross Regional Product ($2015) (GRP)
- Average Regional Household Income ($2015)

These data are available from 1969 to the present[1], and for each of the case study regions this was a period during which these variables were generally rising steadily from year to year. Population tended to rise most slowly, with employment rising more rapidly, reflecting the fact that it was during this period that an increasing share of adult women joined the labor force. The measures of economic activity – earnings, incomes and GRP – grew more rapidly, reflecting the greater pace of economic growth as productivity gains augmented the increases that could be associated with a larger population and labor force. For each of the regions, increasing productivity means that not only did the population and the size of the employed population increase over time, so did the output and earnings associated with each worker, on average.

Since these aggregate socioeconomic variables measure different aspects of the economic development occurring in a region, they are highly correlated over time, which creates obstacles to using two or more of them as independent variables in multivariate regressions. For each of

[1] These data were purchased from Woods and Poole, a company that compiles and processes demographic and socioeconomic data and projections for the United States, the individual states and metropolitan areas, and other regions. Woods and Poole data products are commonly used by airports and airport consultants as a source for demographic and socioeconomic data that is used in air passenger demand and other analyses.

the regions served by the case study airports or airport systems, these correlation patterns are reported as part of the case study results.[2]

The purpose of each of the eight case study examples is to assess the effect of adding a disaggregated socioeconomic variable as a new independent explanatory variable to the regression estimates. In the case study analyses, we use as the disaggregated socioeconomic variable the percentage of households in each study region that are making $100,000 or more per year (in 2009 dollars). This variable was chosen because the analysis of passenger survey results indicates that on average the propensity of individuals to travel by air increases as their household income increases. Because of this, changes in the share of regional households with higher incomes could help explain changes in airport enplanements from year to year, in combination with the aggregate socioeconomic variables and other variables traditionally used in air passenger enplanements regressions.

The baseline and disaggregated alternative case study regressions will also include the annual average oil prices[3] (adjusted for inflation and expressed in 2015 dollars) as a national factor that influences both passenger demand for air travel (through its effect on real consumer incomes as well as its possible impact on airline fares) and the provision of flights by airlines (through its effects on air carrier operating costs). Typically we would expect the parameter estimate for the influence of the oil price variable on annual airport enplanements to be negative, although this may not be true in all cases, such as for airports serving oil producing regions.

Over the sample period (1990 to 2010), oil prices have been volatile and have reached both extremely high and extremely low levels. To illustrate this (and to introduce the indexing approach that will be used to present the socioeconomic data for individual regions on a single scale for comparative purposes), Figure D-1 shows the annual oil price series from 1990 to 2010, with the real price in the year 2000 set to an index value of 100. For comparative purposes, the figure also includes the series of annual enplanement values for each of the case study airports or airport systems, also with the year 2000 value for each airport also set to 100. The figure depicts the greater volatility and range of variation of the real oil price series, compared to case study airport annual O&D enplanements, which increased during the years 1990 to 2000, when oil prices were relatively low. In the remaining years of the sample period, case study airport enplanements grew more slowly or even declined, in a time period when commercial aviation was strongly affected by the events of 9/11 and the volatility of oil prices after 2003.

---

[2] In statistics, this regression problem is termed "multicollinearity," and one of its consequences is the inflation of the variances of the coefficient estimates, making it less likely that the coefficient estimates are statistically significant. As described in Kennedy 1985, "The higher the correlation between the independent variables (the more severe the multicollinearity), the less information [is] used by the Ordinary Least Squares estimator *(that is, the regression*) to calculate the parameter estimates, and thus the greater the variances." (Kennedy 1985)
[3] The Oil Price variable is the per barrel price of the U.S. Crude Oil Composite Acquisition Cost to U.S. Refiners, a monthly series (with annual average) reported by the Energy Information Administration. https://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PET&s=R0000____3&f=M

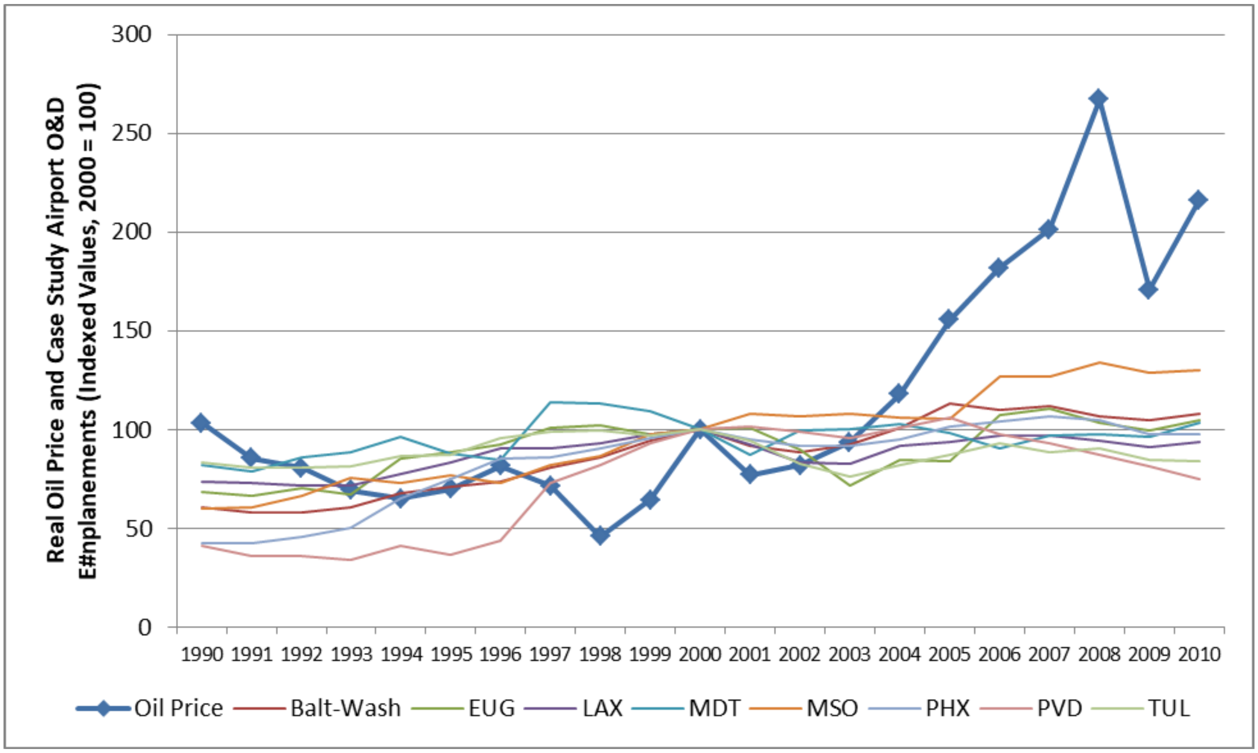

Figure D-1 Indexed Real Oil Prices and Case Study Airport Annual O&D Enplanements, 1990 to 2010, 2000=100

The value of displaying these series in indexed form is that it permits a focus on the behavior over time of the series and allows those behaviors to be compared with one another, even though they differ significantly in their magnitudes. The actual magnitudes are reported below for the individual airports and airport system.

The next subsections present the results of the eight case studies.

b. Eugene Airport (EUG)

Figure D-2 shows the indexed values for the socioeconomic variables within the Eugene, Oregon, MSA, from 1990 to 2010 (with the value for 2000 set to 100). This region is served by Eugene Airport (EUG), currently categorized by the FAA as a small hub airport. For comparative purposes, the figure also shows the indexed series of annual DB1B passenger enplanements at EUG from 1990 to 2010.

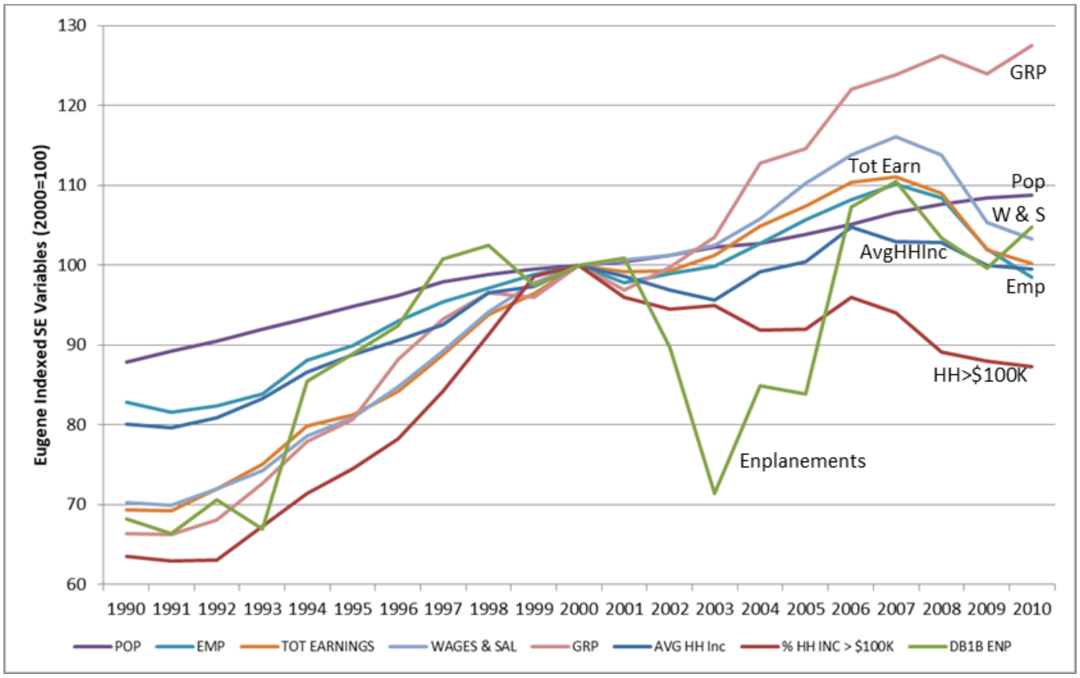

Figure D-2 Eugene, OR MSA Indexed Socioeconomic Variables, 1990 to 2010, 2000=100

While all the socioeconomic variables experienced a gradual rising trend over the 1990 to 2010 period, there are some differences. As can be seen in the 2010 data points, in comparative terms Eugene's population, employment, and average household income have grown most slowly, since these reflect demographic processes and dynamics rather than strictly economic ones. Growing somewhat more rapidly are the economic variables related to regional output (RGP) and regional earnings (Total Earnings and Wages and Salaries). Somewhat more variable in the later years of the period is the percentage of Eugene MSA households with incomes exceeding $100,000, which declined in the years following 2000.

Annual passenger enplanements at EUG have been much more volatile than the socioeconomic series, generally rising in the years prior to 2002, then sharply declining and then growing until the year 2008 and the beginning of the financial crisis.

Table D-1 reports values for the variables for the Eugene MSA from 1990 to 2010. Also shown are annual enplanements at EUG for those years.

Table D-1 Values of Socioeconomic Variables, Eugene, OR MSA (1990-2010)

	Year	Eugene Airport (EUG)								
		Annual O&D Enplanements	Real Oil Price	Regional Population (000s)	Regional Employment (000s)	Regional Total Earnings ($M)	Regional W&S Earnings ($M)	Regional Product (GRP) ($M)	Regional Avg HH Income	% Regional HHs with Income > $100K
	1990	231,011	$40.16	284.3	154.5	$5,043.6	$3,524.0	$7,367.5	$61,422	9.1%
	1991	224,742	$33.20	288.5	152.2	$5,034.9	$3,503.0	$7,355.1	$61,048	9.0%
	1992	238,949	$31.33	292.5	153.5	$5,230.5	$3,610.8	$7,567.0	$62,050	9.0%
	1993	226,580	$26.96	297.3	156.4	$5,458.9	$3,723.9	$8,065.0	$63,830	9.6%
	1994	289,359	$25.31	301.8	164.3	$5,802.0	$3,941.0	$8,656.6	$66,358	10.2%
	1995	300,909	$27.16	306.7	167.6	$5,901.8	$4,058.2	$8,963.2	$68,029	10.7%
	1996	312,900	$31.73	311.0	173.4	$6,121.2	$4,250.6	$9,801.6	$69,430	11.2%
	1997	341,124	$27.80	316.6	177.9	$6,453.4	$4,476.5	$10,354.4	$70,964	12.1%
	1998	346,878	$17.87	319.6	181.1	$6,818.7	$4,720.4	$10,726.2	$73,988	13.1%
	1999	329,900	$25.03	321.8	184.3	$7,009.5	$4,906.8	$10,664.2	$74,655	14.1%
	2000	338,636	$38.86	323.5	186.6	$7,272.4	$5,015.4	$11,111.7	$76,687	14.3%
	2001	341,576	$29.99	324.9	182.4	$7,206.6	$5,044.5	$10,765.2	$75,600	13.7%
	2002	303,450	$31.83	327.5	184.6	$7,221.2	$5,076.9	$11,081.0	$74,296	13.5%
	2003	241,784	$36.38	330.8	186.2	$7,361.4	$5,136.3	$11,497.0	$73,333	13.6%
	2004	287,417	$45.87	332.3	191.7	$7,623.5	$5,306.5	$12,528.1	$76,001	13.1%
	2005	283,763	$60.44	335.8	197.1	$7,805.9	$5,526.1	$12,727.6	$76,991	13.2%
	2006	363,164	$70.72	339.9	201.7	$8,026.9	$5,706.8	$13,552.7	$80,342	13.7%
	2007	374,231	$78.22	344.8	205.5	$8,077.0	$5,823.4	$13,764.1	$78,899	13.5%
	2008	350,070	$103.82	348.2	202.3	$7,923.4	$5,706.3	$14,029.0	$78,817	12.8%
	2009	337,124	$66.42	350.9	190.4	$7,410.3	$5,282.5	$13,770.9	$76,691	12.6%
	2010	354,543	$83.99	351.9	183.7	$7,285.3	$5,179.6	$14,170.1	$76,272	12.5%
Used to assess out-of-sample forecasts	*2011*	*384,447*	*$101.87*	*353.5*	*185.6*	*$7,354.2*	*$5,212.9*	*$14,409.9*	*$75,813*	*12.1%*
	2012	*391,283*	*$100.93*	*354.5*	*185.8*	*$7,565.0*	*$5,278.2*	*$14,336.3*	*$76,672*	*12.4%*
	2013	*428,316*	*$100.49*	*355.7*	*187.2*	*$7,701.2*	*$5,368.4*	*$14,145.1*	*$75,865*	*12.5%*
	2014	*440,853*	*$92.02*	*358.3*	*191.4*	*$7,899.9*	*$5,569.7*	*$14,498.5*	*$78,256*	*12.5%*
	2015	*435,582*	*$48.39*	*360.9*	*194.3*	*$8,084.9*	*$5,778.8*	*$14,821.9*	*$79,118*	*12.5%*

The correlation between these socioeconomic variables is reported in Table D-2 below. The socioeconomic variables are very strongly and positively correlated with one another over the sample period. The correlations between the Eugene MSA socioeconomic variables and either oil prices or annual O&D enplanements at EUG are also positive but less strong.

Table D-2 Correlations among Socioeconomic Variables, Eugene, OR MSA (1990-2010)

Eugene Airport (EUG)	Population	Emp	Total Earnings	Wages & Salaries	GRP	Avg HH Inc	% > $100K	DB1B Enp	Real Oil Price
Population	1								
Employment	0.923	1							
Total Earnings	0.941	0.987	1						
Wages & Salaries	0.949	0.989	0.997	1					
GRP	0.990	0.938	0.944	0.956	1				
Avg HH Income	0.938	0.974	0.984	0.979	0.939	1			
% > $100K	0.801	0.879	0.911	0.888	0.778	0.928	1		
DB1B Enplanements	0.733	0.745	0.713	0.714	0.733	0.807	0.704	1	
Real Oil Price	0.723	0.642	0.612	0.659	0.772	0.580	0.293	0.448	1

The next four figures provide information on the distribution of age and household income in the Eugene, Oregon MSA between 1990 and 2010. The region saw an increase in the percentage of older residents (those over 50 years of age) over the case study years, as shown in Figure D-3, while Eugene's under 20 population was slightly declining over the period.

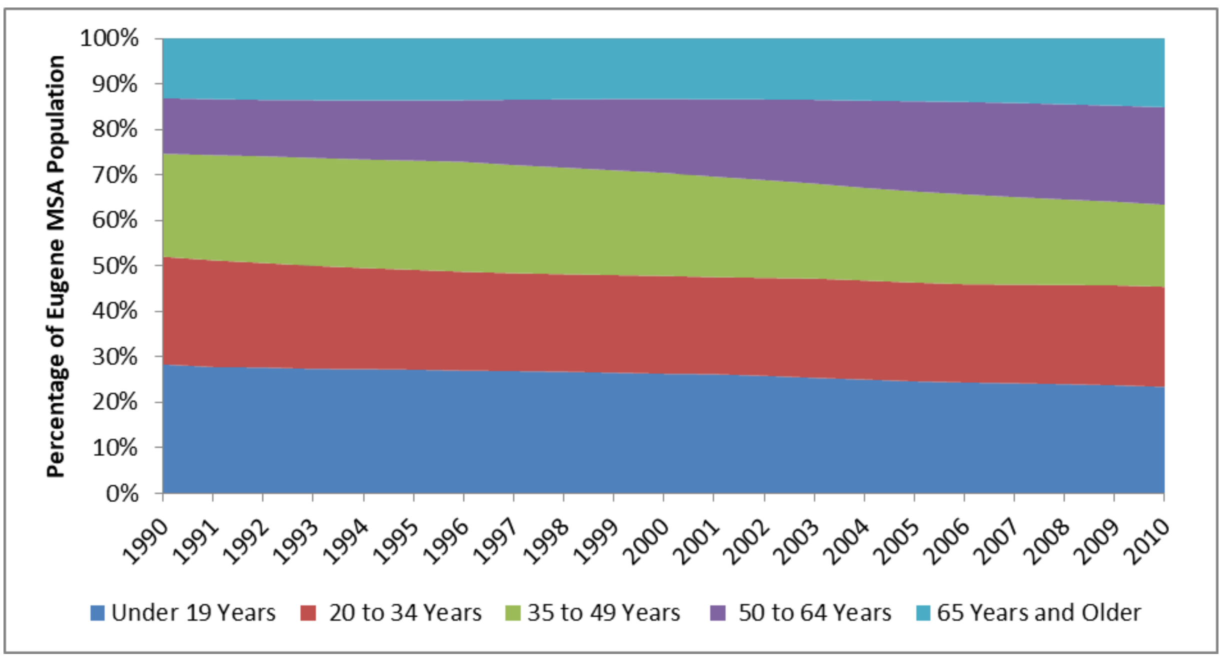

Figure D-3 Population Age Distribution, Eugene, OR MSA, 1990 to 2010

Figure D-4 shows the distribution of Eugene's adult population, those 20 and older, since few air travel decisions would be made by residents younger than 20. During the case study period Eugene's population aged 50 to 64 increased from just over 15 percent of the adult population in the MSA to around 25 percent by 2010, while the share of adults aged 35 to 49 declined from more than 30 percent to less than 25 percent.

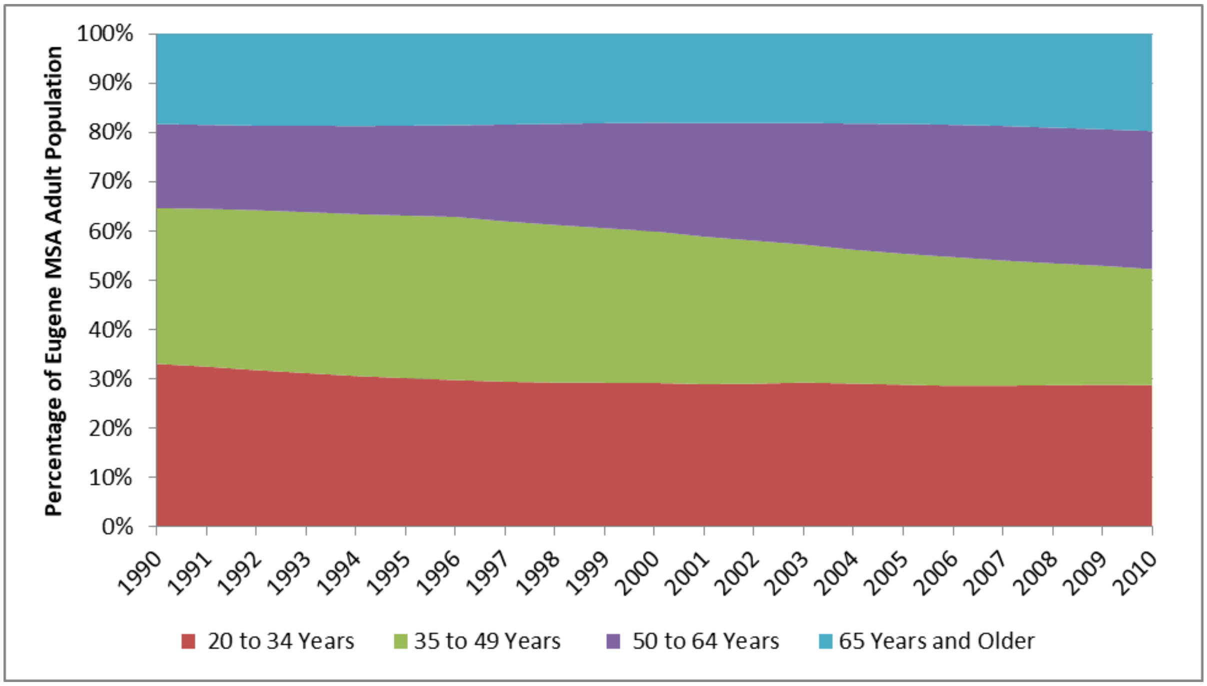

Figure D-4 Adult Population Age Distribution, Eugene, OR MSA, 1990 to 2010

Figure D-5 shows household income distributions for the residents of the Eugene MSA over the case study sample period. The figure shows a period of increase in the proportion of Eugene households with incomes exceeding $100,000 beginning in the late 1990s, which stabilized in the subsequent years, and as shown in the analysis of aggregate socioeconomic trends over the case study sample period for Eugene (and other locations), average household incomes also increased over the period. Changes in income distributions are of special interest because the analysis of passenger survey data indicated that there is an increasing propensity to take air trips as household incomes increase, which suggests that changes in enplanements from year to year may be related to changes in household income distributions as well as changes in household income averages.

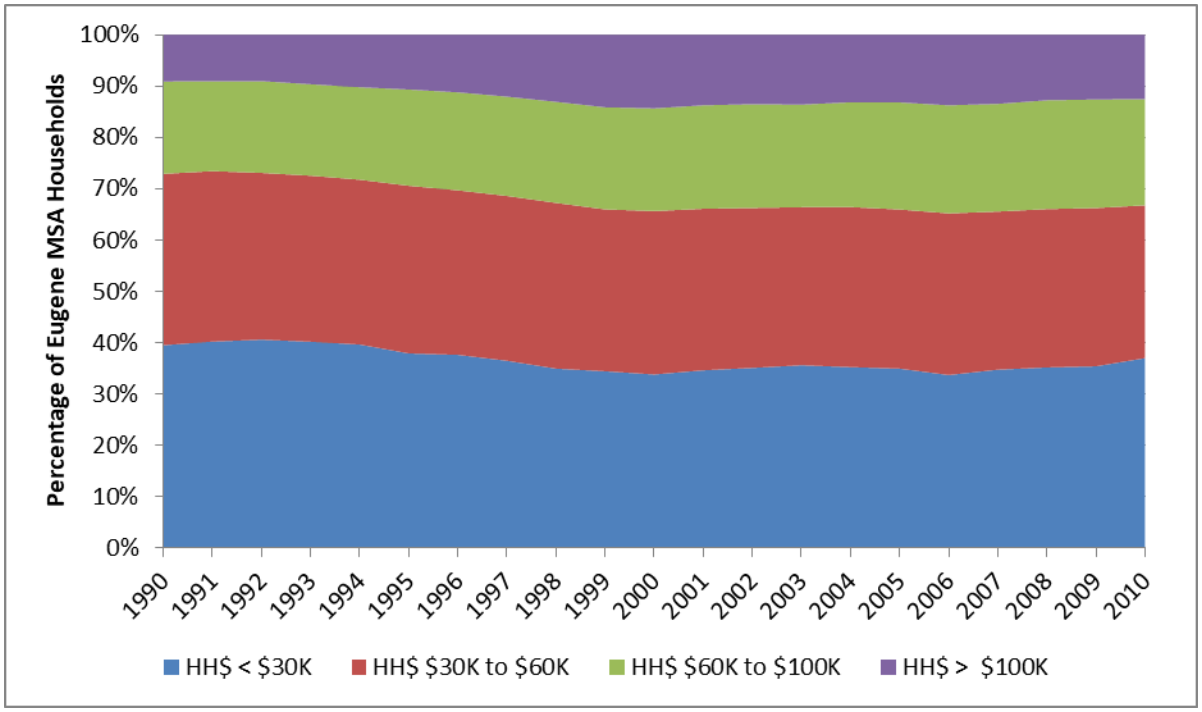

Figure D-5 Household Income Distribution, Eugene, OR MSA, 1990 to 2010

Figure D-6 reports the annual changes in the percentage of Eugene MSA households with annual incomes of $100,000 (in 2009 dollars), shown as the uppermost band in Figure D-5.

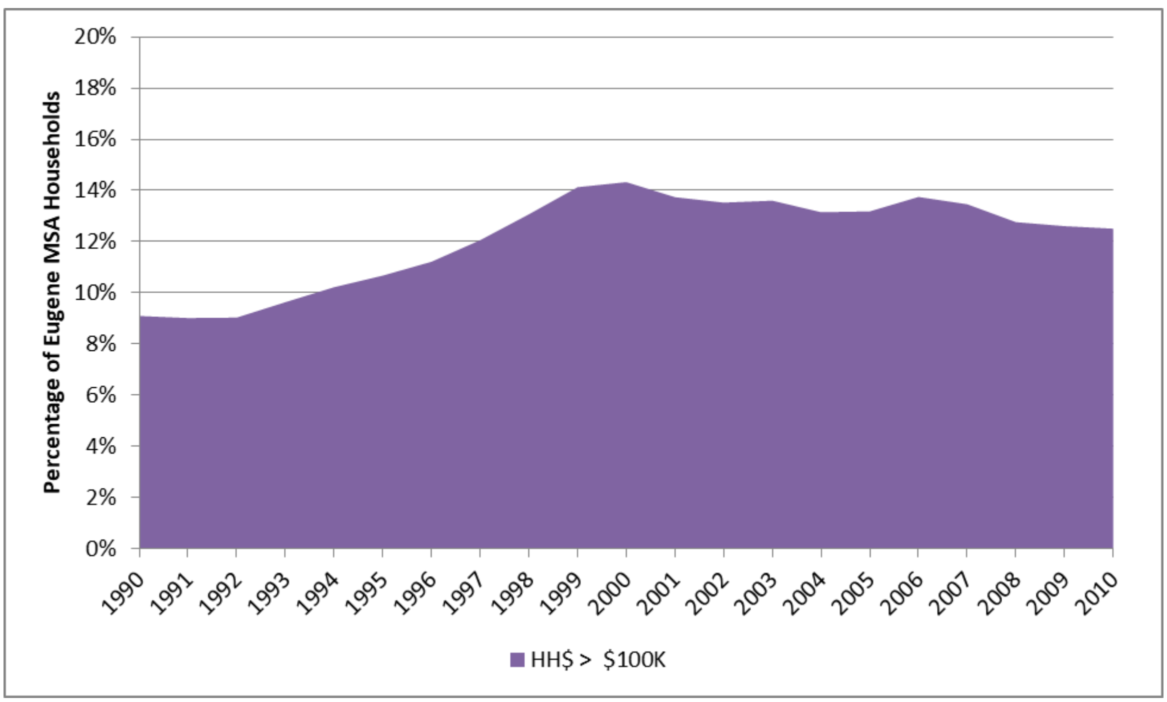

Figure D-6 Percentage of Households with Incomes Exceeding $100,000, Eugene, OR MSA, 1990 to 2010

This percentage rose steadily from 1992 to 2000, from about nine percent of the households in the Eugene MSA to over 14 percent. The percentage then stabilized and declined slightly over the second half of the case study sample period.

The aggregate and disaggregated socioeconomic data for the Eugene MSA were used to estimate a range of models for DB1B O&D enplanements at EUG from 1990 to 2010. The baseline regressions expressed EUG enplanements as a function of the (national) oil price series and one of the six aggregate socioeconomic variables. Results for these baseline regressions modeling annual passenger O&D enplanements at Eugene Airport (EUG) from 1990 to 2010 are shown in the first six columns of Table D-3. These results are summarized below.

- The models had similar degrees of goodness of fit, with adjusted R-squared statistics ranging from 0.672 to 0.769, indicating that the baseline models regressing annual enplanements on the annual real oil price and one of six aggregate socioeconomic variables for the Eugene MSA accounted for a large share of the observed variation in the annual enplanements series.
- With the exception of the baseline regression using Eugene MSA population as the aggregate socioeconomic explanatory variable, the regression constant terms were significant at the 0.05 level or better, indicating that it is unlikely that the true regression constant term is zero, given the sample.
- The coefficient on the oil price variable, which in a log linear regression represents the elasticity of changes in annual EUG enplanements with respect to changes in the oil price, is relatively small (and negative) in all baseline regressions (ranging between -0.037 and -0.123) and somewhat statistically significant only in the baseline regression using Eugene MSA GRP as the aggregate socioeconomic explanatory variable (at the 0.05 level of significance). This indicates that over the sample period there was on average little to no responsiveness of annual air passenger enplanements to changes in oil prices. Because the model estimates are log linear regressions, the parameter estimates represent elasticities of annual EUG enplanements with respect to changes in the price of oil, a commodity that is important both as an input to airline operations and a factor affecting household and business firm budgets. While the magnitude of the estimated parameter is small – the estimate of -0.12 in the GRP equation (number 5) indicates that a one percent increase in the price of oil will reduce annual passenger enplanements at EUG by the modest amount of 0.12 percent, it also suggests that large changes in oil prices like those that were experienced between 2006 and 2010 can have meaningful impacts on annual EUG enplanements. For example, a 20 percent change in oil prices would be associated with a more meaningful change in EUG passenger enplanements of around 2.5 percent.
- The influence of the aggregate socioeconomic variables on annual enplanements at EUG is positive and strongly statistically significant. The range of coefficient (elasticity) estimates exhibits the same pattern as that described for the baseline regressions covering the Baltimore-Washington CSA, with the largest elasticity estimate (2.50, significant at the 0.005 level) coming for Eugene MSA population, a smaller value for the regression estimate using Eugene MSA total employment (1.57, also significant at the 0.005 level), and then estimates of around 0.80 for Total Earnings, Wages and Salaries, and GRP (all also significant at the 0.005 level). The elasticity estimate for Average Real Household

Income in the Eugene MSA was in the middle range of these, estimated at 1.72, and also significant at the 0.005 level.

The EUG regression equation substituting the annual series of the percentage of Eugene MSA households with incomes of $100,000 or more (a disaggregated socioeconomic variable) for the aggregate socioeconomic variables used in the six baseline regressions is very similar to the results from those baseline regressions. The estimated constant term of 13.97, is much larger that the constant term that resulted in the six baseline regressions (reflecting the much smaller magnitude of the "share of households" socioeconomic variables, which is reported in percentages), and it is significant at the 0.005 significance level. The estimated coefficient for the Oil Price variable is unexpectedly positive (0.049) and is not statistically significant. The coefficient estimate for the disaggregated socioeconomic variable introduced into the model specification, 0.73, is comparable to the estimates in the baseline equations for the individual socioeconomic variables. This estimate is significant at the 0.005 level of significance. Finally, the adjusted R-squared value for this specification is 0.687, a value similar to those reported for the baseline estimations. The disaggregated socioeconomic variable does not appear to bring new information to the estimations for EUG annual O&D enplanements in this simple specification.

Table D-3 Baseline and Other Regression Results for Models of Annual Enplanements, Eugene Airport, 1990 to 2010

Independent Variable	Statistic	Regression						
		1	2	3	4	5	6	7
Constant	Coefficient	-1.45	4.68	5.22	5.71	5.81	-6.45	13.97
	t-Statistic	-0.476	2.782 **	2.976 **	3.579 ***	4.442 ***	-1.894 *	29.280 ***
Oil Price	Coefficient	-0.11	-0.06	-0.04	-0.06	-0.12	-0.05	0.05
	t-Statistic	-1.455	-0.931	-0.543	-0.824	-1.748 *	-0.856	0.906
Pop	Coefficient	2.50						
	t-Statistic	4.485 ***						
Emp	Coefficient		1.57					
	t-Statistic		4.467 ***					
Total Earnings	Coefficient			0.86				
	t-Statistic			3.983 ***				
Wages and Salaries	Coefficient				0.84			
	t-Statistic				4.067 ***			
Gross Regional Product	Coefficient					0.78		
	t-Statistic					4.892 ***		
Average HH Income	Coefficient						1.72	
	t-Statistic						5.467 ***	
<u>Disaggregated SE Variable</u>: % HH with Income > $100K	Coefficient							0.73
	t-Statistic							4.209 ***
	R Squared	0.736	0.734	0.705	0.711	0.758	0.792	0.718
	Adj R Squared	0.707	0.705	0.672	0.679	0.732	0.769	0.687

* Coefficient estimate significant at 0.05 level ** at 0.01 level *** at 0.001 level

The second approach to comparing the contribution to model performance of the disaggregated socioeconomic variable – the percentage of Eugene households with relatively high incomes – to the baseline results involves adding the disaggregated variable as an additional independent regressor variable to individual aggregate socioeconomic variables used in the baseline regressions. Results from these regressions for EUG and the Eugene MSA are reported in Table D-4. When the disaggregated socioeconomic variable is added as a new variable to the baseline regression specifications, many of the relationships that were estimated in the baseline equations no longer hold. The equation constant estimates are different across the equations, possibly a reflection of the much smaller magnitude of the values that make up the disaggregated socioeconomic variable.

- For these regressions, the adjusted R-squared estimates are slightly greater or about the same as those reported for the baseline regression equations, suggesting that inclusion of the disaggregated socioeconomic variable in the models provided no new information for modeling the 1990 to 2010 O&D passenger enplanements at EUG.
- The regression constant estimation (-42.55) in only one of the specifications (that using the Eugene MSA Average Real Household Income series as the aggregate socioeconomic variable) is statistically significant, at the 0.005 level of significance or better. Constant estimates in the other five equations are not statistically significant.
- Of these specifications combining one of the aggregate socioeconomic variables with the disaggregated socioeconomic variable, only one has an estimated coefficient for the Oil Price variable that is statistically significant. This is again the specification using the Eugene MSA Average Real Household Income as the aggregate socioeconomic variable. This estimate is -0.193 and is significant at the 0.005 level of significance. The estimate for the Oil Price coefficient is not statistically significant in any of the others, and in two cases the estimate is positive rather than the expected negative value.
- Three of the coefficient estimates for the aggregate socioeconomic variables are statistically significant, at the 0.05 level of significance or better. These are the specifications that use the Eugene MSA Population (1.84), the Eugene MSA Gross Regional Product (0.79),and the MSA's Average Real Household Income (4.72). This third estimated value is surprisingly large. The remaining estimates were all small in magnitude and not statistically significant.
- The coefficient estimates for the disaggregated socioeconomic variable include only one specification that is statistically significant, that for the specification that uses the Eugene MSA Average Real Household Income as the aggregate socioeconomic variable (with an estimate of -1.48, significant at the 0.01 level of significance). Coefficient estimates in the remaining specifications on the disaggregated socioeconomic variable are not statistically significant.

As noted above, for these regressions, the adjusted R-squared estimates are about the same as those reported for the baseline regression equations, suggesting that inclusion of the disaggregated socioeconomic variable in the models provided on a modest amount of new information for modeling the 1990 to 2010 O&D passenger enplanements at EUG.

Table D-4 Regression Results for Models of Annual Enplanements, Eugene Airport, 1990 to 2010, that Include both Aggregate and Disaggregated Socioeconomic Variables

Independent Variable	Statistic	Regression					
		1	2	3	4	5	6
Constant	Coefficient	2.75	5.41	11.75	10.18	5.69	-42.55
	t-Statistic	0.801	0.801	1.205	1.100	1.289	-2.984 **
Oil Price	Coefficient	-0.068	-0.054	0.027	-0.001	-0.125	-0.193
	t-Statistic	-0.689	-0.546	0.229	-0.008	-1.175	-2.602 **
Pop	Coefficient	1.835					
	t-Statistic	1.434					
Emp	Coefficient		1.452				
	t-Statistic		1.271				
Total Earnings	Coefficient			0.219			
	t-Statistic			0.229			
Wages and Salaries	Coefficient				0.389		
	t-Statistic				0.410		
Gross Regional Product	Coefficient					0.793	
	t-Statistic					1.885 *	
Average HH Income	Coefficient						4.715
	t-Statistic						3.965 ***
Disaggregated SE Variable: % HH with Income > $100K	Coefficient	0.226	0.061	0.546	0.399	-0.012	-1.475
	t-Statistic	0.583	0.111	0.681	0.491	-0.028	-2.588 **
	R Squared	0.744	0.735	0.718	0.718	0.758	0.842
	Adj R Squared	0.698	0.688	0.668	0.669	0.716	0.814

* Coefficient estimate significant at 0.05 level ** at 0.01 level *** at 0.001 level

To assess the forecasting performance of these models, we limit the analysis to consideration of a single pair of regressions, those that use the Eugene MSA Gross Regional Product as the aggregate socioeconomic variable. As mentioned above, it is not possible to obtain the Woods and Poole projections that would have been available for the years 2011 through 2015, which would have been included as part of the Woods and Poole 2012 data release. To create a test projection of values for Eugene MSA GRP for those years, we use the percentage annual growth terms from the Woods and Poole 2016 data release for the years 2015 through 2019. With those percentage values for annual GRP growth, notional projected values for Eugene GRP can be calculated for 2011 through 2015. To create notional projected values for the disaggregated socioeconomic variable, the percentage of Eugene MSA households with 2009 dollar denominated income exceeding $100,000 (in 2009 dollars), a similar procedure was used.

Forecast values for the Oil Price dependent variable are also needed to calculate out of sample projections. These were obtained for the years 2011 to 2015 from the data reported in the FAA's 2012 Annual Aerospace Forecast document. These national values are used for each of the case study airports out of sample forecast exercises.

Four out of sample model forecasts were calculated for the baseline model using Eugene MSA GRP and the model that also includes the disaggregated socioeconomic variable. The first three are based on the out of sample forecasts described above. For the first forecast, the projected values – based on the percentage increases projected for the years 2015 through 2019 – for GRP, and the percentage of Eugene household incomes exceeding $100,000 in the case of the model using the disaggregated socioeconomic variable as input, are used as described above. For the second, the values for percentage annual increases over the projection period (2011 to 2015) are increased by 50 percent, indicating a more rapidly growing Eugene economy. For the third EUG projection, the values for percentage annual increases are reduced by 50 percent, representing a projection for a more slowly growing economic environment in the Eugene MSA.[4] For the fourth projection, the actual values observed in 2011 through 2015 for GRP and the share of Eugene MSA households with incomes exceeding $100,000 are used to drive the "forecast" of EUG annual O&D enplanements.

Using this approach to simulating a range of forecast inputs that would have been available to analysts in 2011 or 2012, we can compare the forecasting performance of the baseline regressions for EUG annual O&D enplanements based on the annual oil price variable and the Eugene MSA GRP, with that of the regression that includes input from the disaggregated socioeconomic variable. The independent variable data and projections are shown in Table D-5. The projections for Eugene MSA variables reported in the Woods and Poole 2016 data have annual GRP growing at a rate of 2.2 percent, and the percentage of Eugene MSA households with incomes exceeding $100,000 growing at about 1 percent initially, growing to 2 percent over a two year period. The annual percentage changes used to calculate the Oil Price projected

[4] For example, if for the first projection the independent variables "driving" the enplanements model are project to grow one percent annually (for Eugene GRP) and 0.5 percent annually (for the percentage of Eugene households with incomes exceeding $100,000), then the projected rates of growth for the second EUG projection are 1.5 percent (for Eugene GRP) and 0.75 percent ((for the percentage of Eugene households with incomes exceeding $100,000). In this example, the third projection for the Eugene independent variables would be 0.50 percent annual growth for Eugene GRP and 0.25 percent annual growth for the percentage of Eugene households with incomes exceeding $100,000 (the disaggregated socioeconomic variable.

values for 2011 through 2015 are those implied by the Oil Price projections reported in the 2012 FAA Aerospace Forecast, which relies on projected values for the years following 2010.

Table D-5 Inputs to Out-of-Sample EUG O&D Enplanement Projections

EUG Inputs to Out-of-Sample Enplanement Forecasts	EUG					
	Sources	2011	2012	2013	2014	2015
Oil Price Forecast	*2012 FAA Aero Forecast*	$84.32	$92.71	$97.16	$100.65	$104.64
GRP Forecast 1 (Medium)	*W&P 2016 %*	$14,488.1	$14,801.0	$15,111.5	$15,422.0	$15,733.6
GRP Forecast 2 (High)	*1.5 x F'cast 1*	$14,647.1	$15,121.6	$15,597.4	$16,078.2	$16,565.5
GRP Forecast 3 (Low)	*0.5 x F'cast 1*	$14,329.1	$14,483.8	$14,635.7	$14,786.1	$14,935.5
HHInc% Forecast 1 (Medium)	*W&P 2016 %*	12.6%	12.8%	13.0%	13.3%	13.6%
HHInc% Forecast 2 (High)	*1.5 x F'cast 1*	12.7%	12.9%	13.3%	13.8%	14.2%
HHInc% Forecast 3 (Low)	*0.5 x F'cast 1*	12.6%	12.6%	12.8%	12.9%	13.1%
	2010					
Actual Oil Price	$83.99	$101.87	$100.93	$100.49	$92.02	$48.39
Actual GRP	$14,170.1	$14,409.9	$14,336.3	$14,145.1	$14,498.5	$14,821.9
Actual HHInc%	12.5%	12.1%	12.4%	12.5%	12.5%	12.5%

The forecasting or projection results for the baseline regression equation for annual enplanements at EUG are shown in Figure D-7.

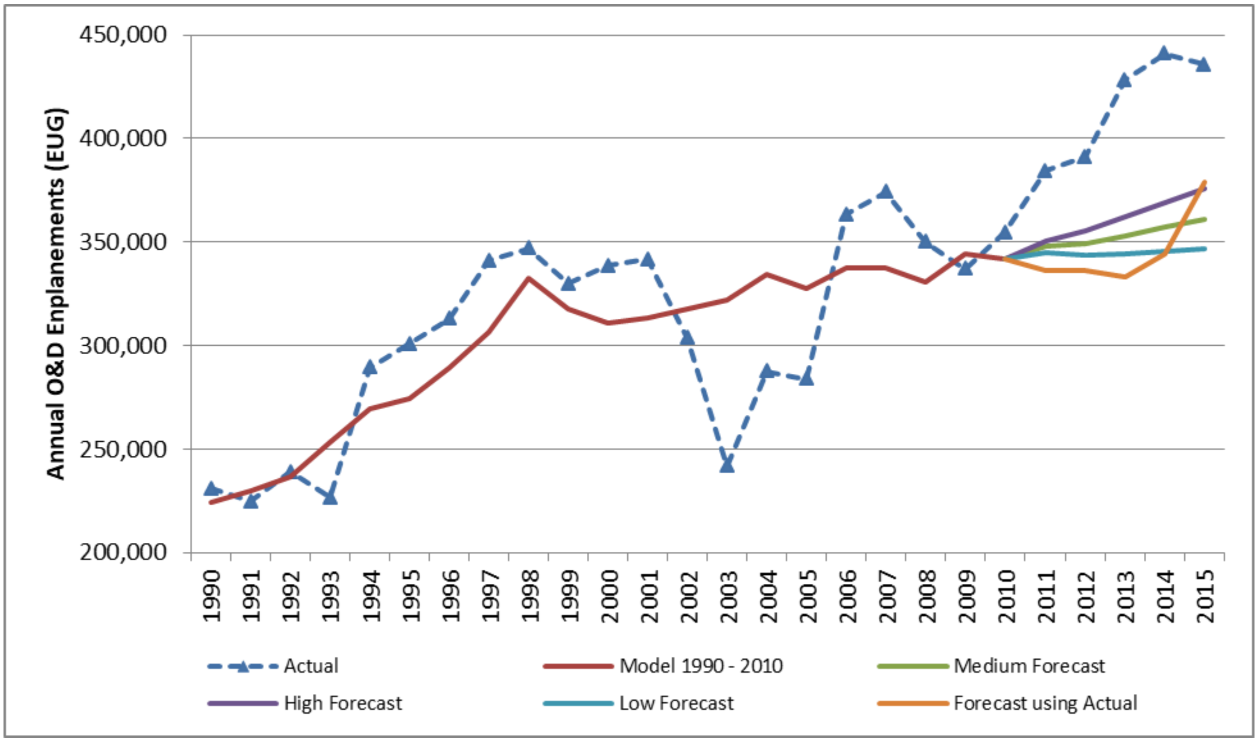

Figure D-7 Baseline Model Forecasts of EUG Annual Enplanements, 2011 to 2015, Using Eugene MSA Gross Regional Product, 1990 to 2010

The results for the regression equation using the disaggregated socioeconomic variable along with Eugene MSA GRP are shown in Figure D-8.

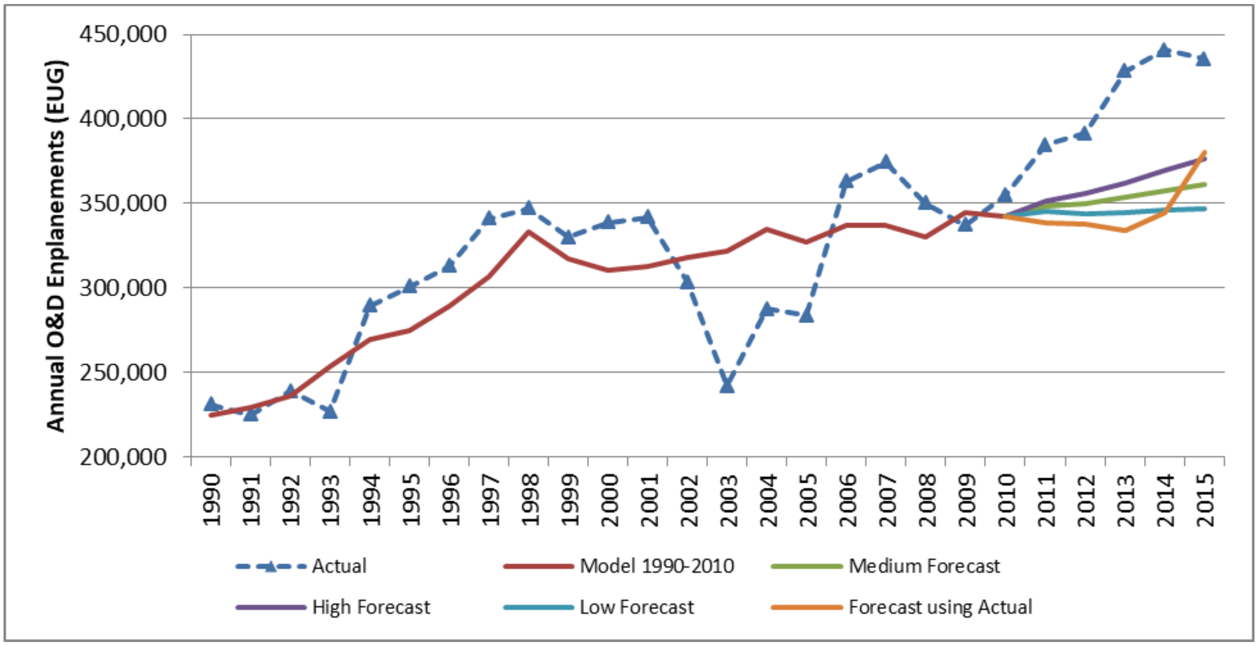

Figure D-8 Disaggregated SE Variable Model Forecasts of EUG Annual Enplanements, 2011 to 2015, Using Eugene MSA GRP and Share of Eugene Households with Incomes exceeding $100,000, 1990 to 2010

Actual enplanements at EUG increased much more robustly than was forecast by either model. In fact, the two forecast charts look remarkably similar, a reflection of the fact that the estimated baseline and disaggregated socioeconomic variable regression equations using Eugene MSA GRP are very similar.[5] The accuracy of the projections driven by the actual values taken by the Oil Price, Eugene GRP, and the Share of Households with Income exceeding $100,000 in the years 2011 through 2015 is also modest, with the projected enplanement counts undershooting the actual history over those years. Tables D-6 to D-8 quantify this, reporting the percentage error for each of the forecasts and the Root Mean Squared Error (RMSE) for each forecast, a calculation which can be used to compare the accuracy of the baseline model forecasts and the forecasts calculated from the model also using the disaggregated socioeconomic variable.

[5] The baseline regression equation is ln(EUG_Enp) = 5.81 – 0.123ln(OilPrice) + 0.78ln(GRP), with an adjusted R squared of 0.732,while the estimated equation that includes the disaggregated socioeconomic variable is ln(EUG_Enp) = 5.69 – 0.125ln(OilPrice) + 0.79ln(GRP) – 0.01ln(%HHInc>$100K) with an adjusted R squared of 0.716, so the explanatory contribution from the disaggregated socioeconomic variable is very modest for the Eugene region.

Table D-6 Percentage Forecast Errors, Baseline Model Projections of EUG Annual O&D Enplanements, 2011 to 2015

EUG (using GRP)	Percentage Forecast Error (EUG Baseline Regression with Oil Price and GRP)			
	Medium	High	Low	Actual
2010	--	--	--	--
2011	-9.5%	-8.8%	-10.3%	-12.5%
2012	-10.7%	-9.2%	-12.2%	-14.1%
2013	-17.5%	-15.5%	-19.6%	-22.2%
2014	-19.0%	-16.3%	-21.6%	-21.9%
2015	-17.1%	-13.7%	-20.4%	-13.0%

The negative percentage errors in the tables reflect the fact that the model projections of EUG enplanements consistently underestimated the number of annual passenger, even for the "forecast" based on the values that actually occurred for the Oil Price and Eugene MSA GRP. As can be seen in the figures and the tables of percentage errors, the model forecasts using these actual values did turn up in 2015, although not enough to catch up to the growth that actually took place in EUG enplanements. This upturn in 2015 is a reflection of the sharp drop in the actual 2015 Oil Price variable, from $92.02 in 2014 to $48.39 in 2015, which, combined with the negative regression coefficient on the Oil Price variable, resulted in an increased forecast for 2015 EUG annual enplanements.

Table D-7 Percentage Forecast Errors, Disaggregated SE Model Projections of EUG Annual O&D Enplanements, 2011 to 2015

EUG (using GRP)	Percentage Forecast Error (EUG Regression including Disaggregated SE Variable)			
	Medium	High	Low	Actual
2010	--	--	--	--
2011	-9.5%	-8.7%	-10.2%	-11.9%
2012	-10.6%	-9.1%	-12.1%	-13.7%
2013	-17.5%	-15.4%	-19.5%	-22.0%
2014	-18.9%	-16.2%	-21.5%	-21.8%
2015	-17.0%	-13.6%	-20.4%	-12.7%

Table D-8 reports the Root Mean Squared Error (RMSE) of the different forecasts for the two model specifications. This statistic can be used to compare the accuracy of two or more of these forecasts, with a smaller RMSE value indicating a more accurate forecast. As can be seen, the inclusion of the additional disaggregated socioeconomic variable in the model specification provided only very modest improvement in the accuracy of the five year forecasts, based on the percentage change in the RSME compared to those based on the baseline specification.

Table D-8 Root Mean Squared Error (RMSE) for EUG Baseline and Disaggregated SE Variable Models (Annual EUG O&D Enplanements, 2011 to 2015)

EUG (using GRP)	Root Mean Squared Error		% Difference BL v Alt
	Baseline	Disagg Var	
Medium Forecast	65,207	64,977	0.4%
High Forecast	55,697	55,431	0.5%
Low Forecast	74,596	74,399	0.3%
Actual Values	73,422	72,403	1.4%

c. Missoula International Airport (MSO)

The collection of indexed socioeconomic variables from 1990 to 2010 for the Missoula, Montana, MSA is shown in Figure D-9.

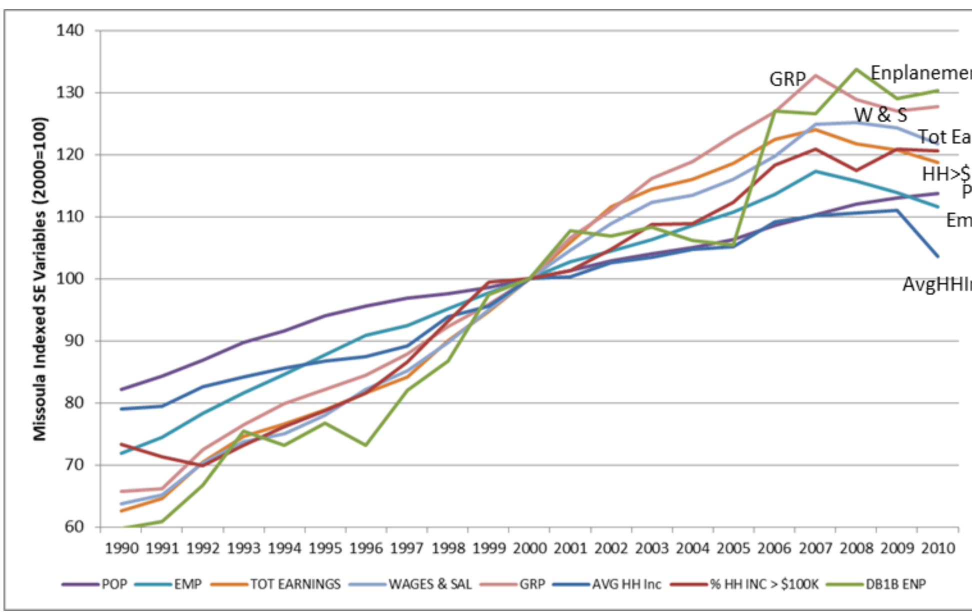

Figure D-9 Missoula, MT MSA Indexed Socioeconomic Variables, 1990 to 2010, 2000=100

These variables, also indexed with their 2000 values set to 100, are also highly correlated, each series gradually increasing over time. Differences in paces of growth also exist, with the demographic variables – population and employment – growing more modestly, and economic variables growing more strongly, reflecting the influence of productivity improvements on economic activity over time. The region is served by Missoula International Airport (MSO), a non-hub airport.

Table D-9 reports the values taken by the regional and other variables for the Missoula MSA from 1990 to 2010.

Table D-9 Values of Socioeconomic Variables, Missoula, MT MSA (1990-2010)

	Year	Missoula International Airport (MSO)								
		Annual O&D Enplanements	Real Oil Price	Regional Population (000s)	Regional Employment (000s)	Regional Total Earnings ($M)	Regional W&S Earnings ($M)	Regional Product (GRP) ($M)	Regional Avg HH Income	% Regional HHs with Income > $100K
	1990	130,830	$40.16	79.1	47.6	$1,593.7	$1,114.2	$2,475.0	$64,241	8.4%
	1991	133,287	$33.20	81.1	49.3	$1,646.1	$1,140.9	$2,492.2	$64,619	8.2%
	1992	146,307	$31.33	83.5	51.9	$1,793.2	$1,230.0	$2,724.9	$67,092	8.0%
	1993	165,260	$26.96	86.2	54.0	$1,899.6	$1,289.4	$2,877.2	$68,369	8.4%
	1994	160,251	$25.31	88.0	56.0	$1,951.6	$1,311.7	$3,008.7	$69,596	8.8%
	1995	168,200	$27.16	90.4	58.1	$2,008.4	$1,363.6	$3,090.0	$70,485	9.1%
	1996	160,367	$31.73	91.9	60.2	$2,077.0	$1,436.2	$3,178.7	$71,094	9.4%
	1997	179,571	$27.80	93.2	61.3	$2,140.8	$1,489.2	$3,308.3	$72,496	10.0%
	1998	189,840	$17.87	93.8	63.0	$2,291.4	$1,570.0	$3,476.4	$76,320	10.7%
	1999	213,497	$25.03	94.8	64.7	$2,412.6	$1,661.4	$3,608.1	$77,679	11.5%
	2000	219,030	$38.86	96.2	66.2	$2,545.9	$1,749.3	$3,764.3	$81,274	11.5%
	2001	236,103	$29.99	97.4	68.0	$2,694.4	$1,828.4	$4,012.8	$81,472	11.7%
	2002	234,140	$31.83	99.0	69.1	$2,841.0	$1,905.5	$4,180.6	$83,378	12.1%
	2003	237,269	$36.38	100.0	70.4	$2,913.5	$1,964.5	$4,371.5	$84,118	12.5%
	2004	232,491	$45.87	100.9	72.0	$2,954.0	$1,983.5	$4,476.8	$85,126	12.5%
	2005	230,916	$60.44	102.3	73.4	$3,018.0	$2,029.9	$4,629.7	$85,481	12.9%
	2006	278,082	$70.72	104.4	75.2	$3,119.0	$2,093.5	$4,775.8	$88,731	13.6%
	2007	277,284	$78.22	106.1	77.6	$3,158.0	$2,184.8	$4,993.8	$89,535	13.9%
	2008	293,024	$103.82	107.7	76.6	$3,100.0	$2,189.6	$4,852.5	$89,825	13.5%
	2009	282,671	$66.42	108.7	75.4	$3,073.2	$2,173.4	$4,780.8	$90,261	13.9%
	2010	285,537	$83.99	109.4	73.9	$3,024.4	$2,128.3	$4,809.0	$84,172	13.9%
Used to assess out-of-sample forecasts	*2011*	*291,585*	*$101.87*	*110.1*	*75.1*	*$3,052.2*	*$2,134.4*	*$4,958.6*	*$85,424*	*14.1%*
	2012	*297,329*	*$100.93*	*111.1*	*75.6*	*$3,085.3*	*$2,171.8*	*$4,975.8*	*$89,018*	*14.4%*
	2013	*303,965*	*$100.49*	*111.8*	*76.6*	*$3,080.5*	*$2,195.7*	*$5,002.7*	*$85,463*	*14.5%*
	2014	*345,608*	*$92.02*	*112.7*	*77.7*	*$3,166.6*	*$2,265.9*	*$5,073.4*	*$87,138*	*14.5%*
	2015	*359,069*	*$48.39*	*113.5*	*78.9*	*$3,268.5*	*$2,321.0*	*$5,164.4*	*$88,377*	*14.7%*

As reported in Table D-10, the 1990 to 2010 values for the socioeconomic data variables are very strongly correlated, with all correlation coefficients exceeding 0.9. The correlations between the Missoula MSA socioeconomic variables and annual O&D enplanments at MSO are also strongly correlated, but oil prices, while they are also positively correlated with the other reported variables, are not as strongly correlated.

Table D-10 Correlations among Socioeconomic Variables, Missoula, MT MSA (1990-2010)

	Population	Emp	Total Earn	Wages & Salaries	GRP	Avg HH Inc	% > $100K	DB1B Enp	Real Oil Price
Population	1								
Employment	0.984	1							
Total Earn	0.968	0.990	1						
Wages & Salaries	0.980	0.991	0.996	1					
GRP	0.979	0.990	0.995	0.997	1				
Avg HH Inc	0.962	0.987	0.991	0.991	0.985	1			
% > $100K	0.969	0.979	0.983	0.990	0.986	0.981	1		
DB1B Enplanements	0.972	0.966	0.968	0.980	0.974	0.970	0.975	1	
Real Oil Price	0.732	0.693	0.696	0.734	0.749	0.696	0.732	0.770	1

The Missoula, Montana MSA saw only modest change in its population distribution over the case study period. Those 19 years old or younger and those aged between 20 and 34 combined to make up 56 percent of Missoula's population in 1990, a total which decreased slightly by 2010 to 51 percent, as shown in Figure D-10.

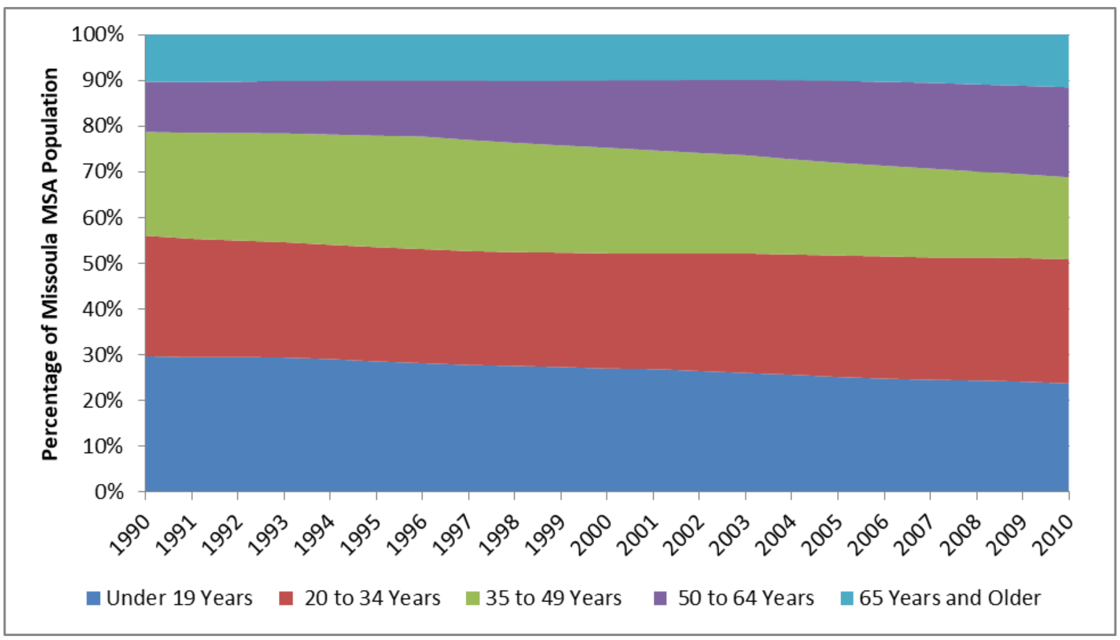

Figure D-10 Population Age Distribution, Missoula, MT MSA, 1990 to 2010

Figure D-11 shows the distribution of Missoula's adult population, those 20 and older. During the case study period Missoula's adult population distribution changed little, although those between 50 and 64 years of age did increase from around 15 percent of the adult population to nearly 25 percent by 2010.

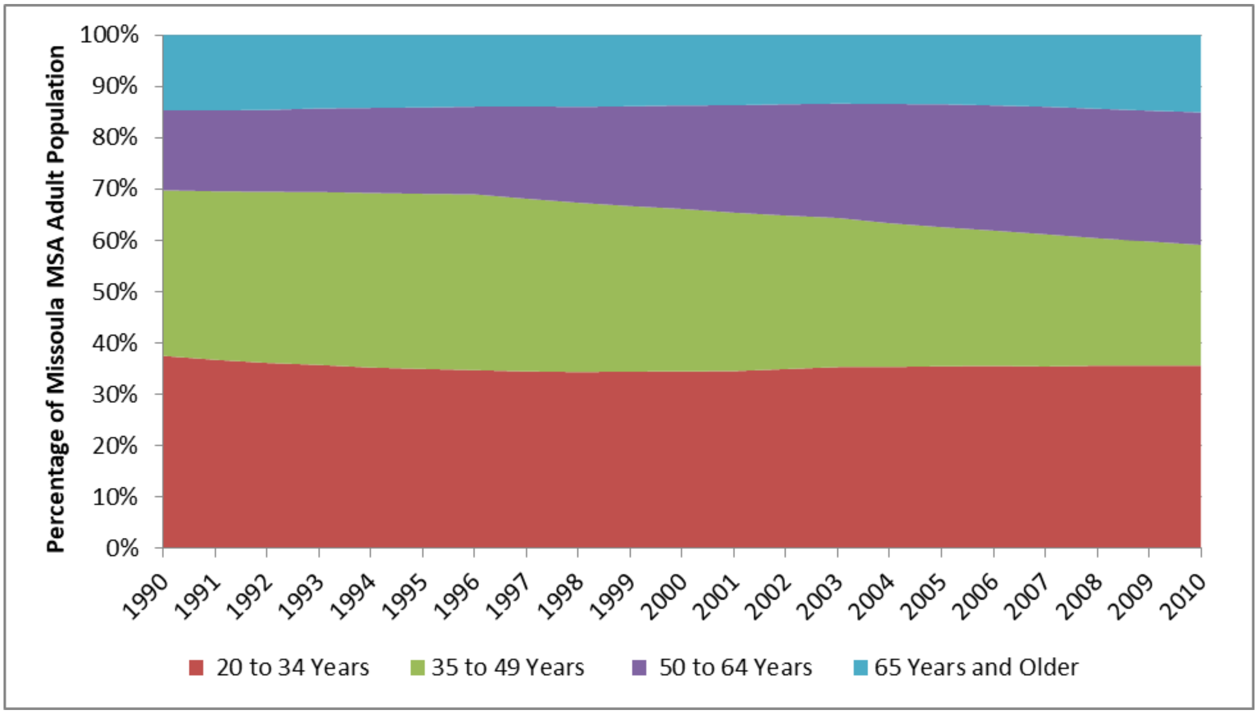

Figure D-11 Adult Population Age Distribution, Missoula, MT MSA, 1990 to 2010

Figure D-12 shows household income distributions for the residents of the Missoula MSA over the case study sample period, 1990 to 2010.

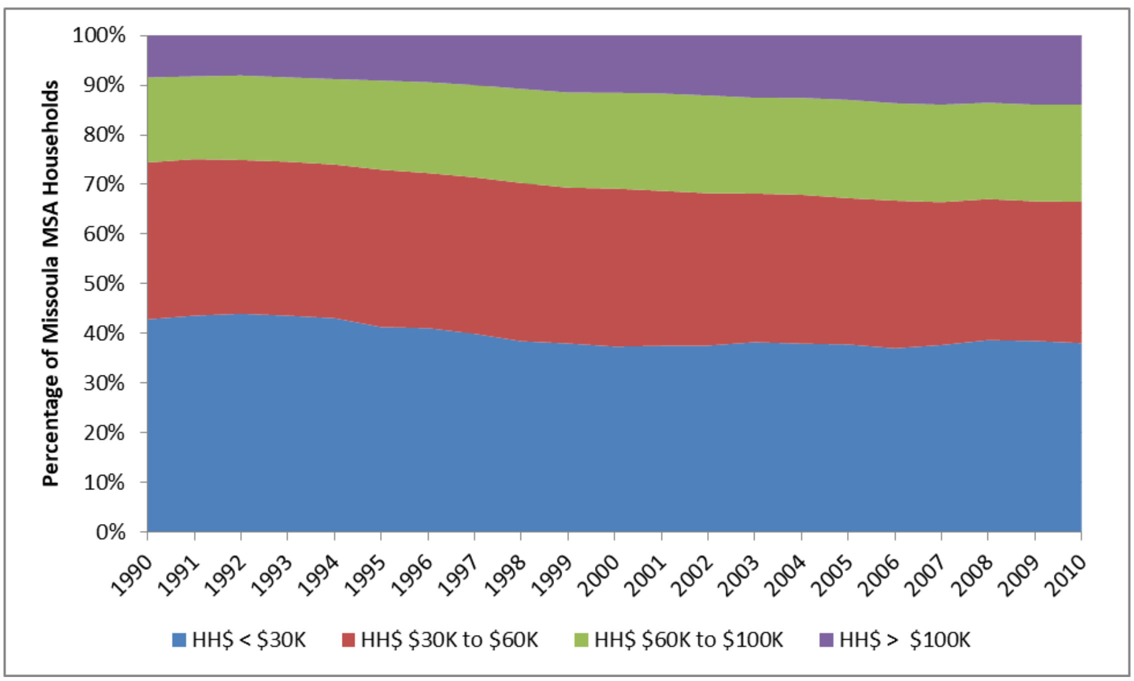

Figure D-12 Household Income Distribution, Missoula, MT MSA, 1990 to 2010

The figure shows a gradual increase in the proportion of Missoula households with incomes exceeding $100,000 over the sample period. As shown in the analysis of aggregate socioeconomic trends over the case study sample period for Missoula (and other locations), average household incomes also increased over the period, declining somewhat in the final years.

Figure D-13 reports specifically the changes in the percentage of Missoula MSA households with annual incomes of $100,000 (in 2009 dollars), a group shown as the uppermost band in Figure D-12. This percentage rose steadily from 1992 to 2000, from about eight percent of the households in the Eugene MSA to nearly 14 percent by 2010.

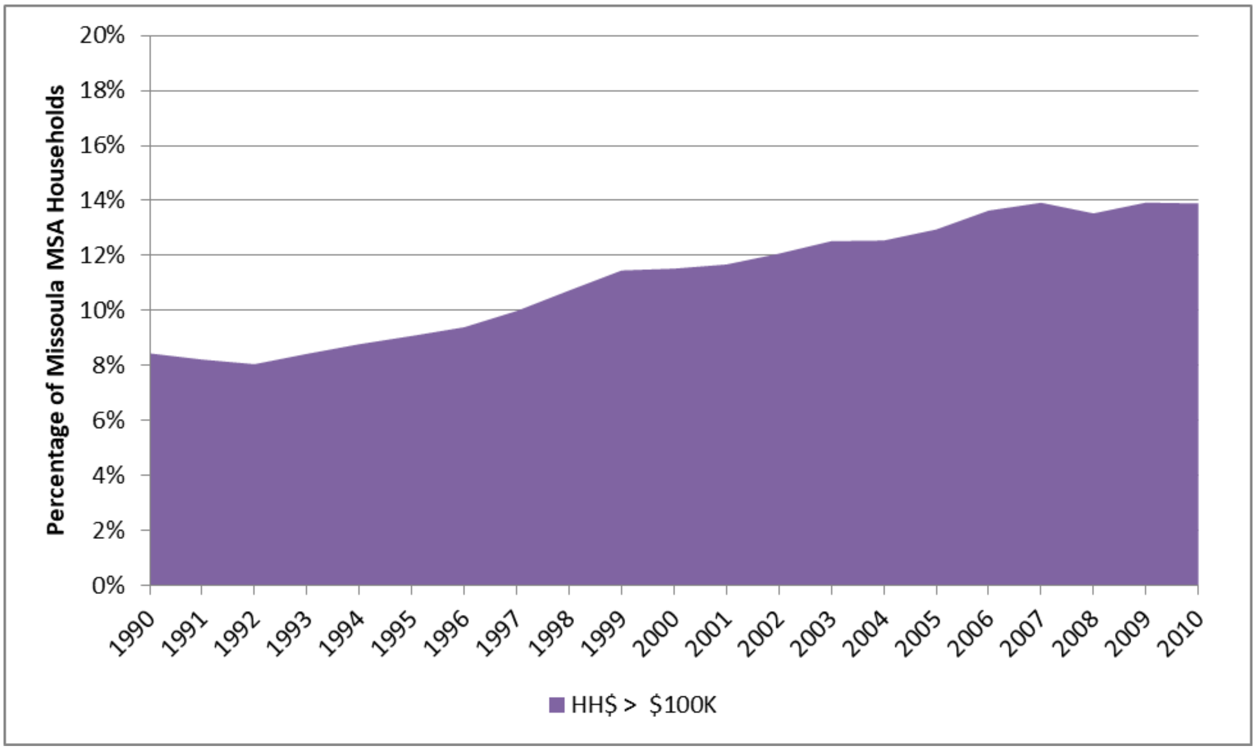

Figure D-13 Percentage of Households with Incomes Exceeding $100,000, Missoula, MT MSA, 1990 to 2010

Characteristics of the results for the six baseline regressions modeling annual passenger enplanements at Missoula International Airport (MSO) from 1990 to 2010 are reported in Table 5.11 and include:

- The model estimates based on the individual regional aggregate socioeconomic variables had similarly high goodness of fit to the actual MSO annual enplanements series, with adjusted R-squared statistics ranging from 0.943 to 0.960, indicating that the baseline models regressing annual enplanements on the annual real oil price and one of the six aggregate socioeconomic variables for the Missoula MSA accounted for nearly all of the observed variation in the annual MSO enplanements series.

- The estimated regression constant terms varied across the baseline models, although the estimates for four of the models -- those using Missoula MSA Employment, Total Earnings in the MSA, Annual Wages and Salaries in the MSA, and Gross Regional Product were roughly comparable, ranging from 3.35 to 5.61. All of these constant estimates were highly significant at the 0.005 significance level (or better). The estimated constant term for the model using MSA population was estimated at 0.49, and was not statistically significant at customary levels of significance. The constant estimate for the models using Missoula MSA Average Real Household Income took a negative value of -11.75, which was statistically significant at the 0.005 level.
- The coefficients estimated in the baseline MSO models for the oil price variable, which in a log linear regression represents the elasticity of changes in annual MSO enplanements with respect to changes in the oil price, were relatively small in all baseline regressions, and only one of them (Missoula MSA employment) was statistically significant at accepted levels for significance. The coefficient estimates were also positive, indicating that higher oil prices were associated with higher annual enplanements at MSO. While this is not the expected outcome in general, a positive relationship is not surprising for the Missoula region, which has been increasingly involved in oil extraction. Nevertheless, the estimated elasticities were modest and not statistically different from zero.
- The influence of the individual aggregate socioeconomic variables on annual enplanements at MSO is positive and strongly statistically significant across all of the baseline models. The range of coefficient (elasticity) estimates exhibits the same pattern as that described for the baseline regressions covering the Baltimore-Washington CSA, with the largest elasticity estimate (2.55, significant at the 0.005 level) occurring for Missoula MSA population, a smaller value for the regression estimate using Missoula MSA total employment (1.54, also significant at the 0.005 level), and then estimates between 1.03 and 1.08 for Missoula MSA Total Earnings, Wages and Salaries, and GRP (all also significant at the 0.005 level). The elasticity estimate for Average Real Household Income in the Missoula MSA was somewhat greater than the estimates for these three, estimated at 2.12, and is also significant at the 0.005 level.

The baseline MSO regression equation 7, in which the annual series of the percentage of Missoula MSA households with incomes of $100,000 or more (a disaggregated socioeconomic variable) replaces the aggregate socioeconomic variables used in the six baseline regressions is very similar to the results from those baseline regressions. The estimated constant term of 14.95 is significant at the 0.005 significance level. The (positive) estimated coefficient for the Oil Price variable, 0.015, is comparable to the low end of the coefficient values in the six baseline regressions for MSO, and like those baseline estimates, it is not statistically significant. The coefficient estimate for the disaggregated socioeconomic variable introduced into the model specification, 1.26, is also comparable to the baseline estimates for the individual socioeconomic variables. This estimate is significant at the 0.005 level of significance. Finally, the adjusted R-squared value for this specification is 0.951, a value similar to the equally high adjusted R-squared values that occurred in the baseline estimations. The disaggregated socioeconomic variable does not appear to bring new information to the estimations for MSO annual O&D enplanements in this simple specification.

Table D-11 Baseline and Other Regression Results for Models of Annual Enplanements, Missoula International Airport, 1990 to 2010

Ind. Variable	Statistic	Regression						
		1	2	3	4	5	6	7
Constant	Coefficient	0.49	5.61	4.03	4.12	3.35	-11.75	14.95
	t-Statistic	0.611	13.856 ***	8.174 ***	9.745 ***	6.223 ***	-7.312 ***	43.103 ***
Oil Price	Coefficient	0.03	0.06	0.05	0.02	0.01	0.03	0.02
	t-Statistic	0.676	1.754 *	1.480	0.605	0.254	0.923	0.380
Pop	Coefficient	2.55						
	t-Statistic	12.901 ***						
Emp	Coefficient		1.54					
	t-Statistic		13.313 ***					
Total Earnings	Coefficient			1.03				
	t-Statistic			14.077 ***				
Wages and Salaries	Coefficient				1.08			
	t-Statistic				16.225 ***			
Gross Regional Product	Coefficient					1.08		
	t-Statistic					14.172 ***		
Average HH Income	Coefficient						2.12	
	t-Statistic						14.091 ***	
<u>Disaggregated SE Variable</u>: % HH with Income > $100K	Coefficient							1.26
	t-Statistic							12.332 ***
	R Squared	0.953	0.951	0.949	0.964	0.951	0.949	0.956
	Adj R Squared	0.948	0.945	0.943	0.960	0.945	0.943	0.951

* Coefficient estimate significant at 0.05 level ** at 0.01 level *** at 0.001 level

When the disaggregated socioeconomic variable is added as a new variable to the baseline regression specifications, many of the relationships that were present in the baseline equations no longer hold as strongly. The equation constant estimates are slightly different across the equations, possibly a reflection of the much smaller magnitude of the values that make up the disaggregated socioeconomic variable. These results are reported in Table D-12.

- For these regressions, the adjusted R-squared estimates are high and similar to those reported for the baseline regression equations, suggesting that inclusion of the disaggregated socioeconomic variable in the models provided no new information for modeling the 1990 to 2010 O&D passenger enplanements at MSO. The range of adjusted R squared values runs from 0.949 to 0.962.
- With the exception of the regression using the Missoula MSA Average Real Household Income as the aggregate socioeconomic variable, the regression constants in the equations of the specifications are statistically significant at the 0.1 level of significance or better.
- For each of the specifications combining one of the aggregate socioeconomic variables with the disaggregated socioeconomic variable, coefficient estimates for the Oil Price variable are slightly smaller in magnitude (but still positive), compared to the baseline regressions that used only aggregate socioeconomic variables. None of the estimates are statistically significant.
- The coefficient estimates for the aggregate socioeconomic variables are all significant at the 0.05 level of statistical significance or better. The magnitude of the estimates is somewhat smaller that the estimates in the corresponding baseline specifications
- The coefficient estimates for the disaggregated socioeconomic variable are positive as expected, but only two of them, in the specifications using Missoula MSA Population and Missoula MSA Employment, have statistically significant estimates at the 0.05 level or better. With respect to magnitude, it is as if in the Missoula case, the influence of socioeconomic characteristics on the MSO annual O&D enplanements that is estimated in the baseline specifications is in effect split between the aggregate socioeconomic variable and the disaggregated socioeconomic variable when we come to the second set of estimates.

Table D-12 Regression Results for Models of Annual Enplanements, Missoula International Airport, 1990 to 2010, that Include both Aggregate and Disaggregated Socioeconomic Variables

Ind. Variable	Statistic	Regression					
		1	2	3	4	5	6
Constant	Coefficient	7.16	9.51	7.67	4.94	7.11	-2.99
	t-Statistic	4.309 ***	4.309 ***	2.738 **	1.645	2.388 *	-0.436
Oil Price	Coefficient	0.008	0.035	0.033	0.017	0.005	0.020
	t-Statistic	0.227	0.949	0.912	0.529	0.143	0.583
Pop	Coefficient	1.403					
	t-Statistic	3.063 **					
Emp	Coefficient		0.911				
	t-Statistic		2.491 *				
Total Earnings	Coefficient			0.692			
	t-Statistic			2.611 **			
Wages and Salaries	Coefficient				1.004		
	t-Statistic				3.352 **		
Gross Regional Product	Coefficient					0.734	
	t-Statistic					2.646 **	
Average HH Income	Coefficient						1.430
	t-Statistic						2.613 **
Disaggregated SE Variable: % HH with Income > $100K	Coefficient	0.612	0.540	0.432	0.098	0.421	0.429
	t-Statistic	2.707 **	1.792 *	1.319	0.275	1.285	1.308
	R Squared	0.968	0.961	0.957	0.964	0.957	0.958
	Adj R Squared	0.962	0.954	0.949	0.958	0.950	0.951

* Coefficient estimate significant at 0.05 level ** at 0.01 level *** at 0.001 level

To assess the forecasting performance of these models, we limit the analysis to consideration of a single pair of regressions, those that use the Missoula MSA Gross Regional Product as the aggregate socioeconomic variable. As mentioned above, it is not possible to obtain the Woods and Poole projections that would have been available for the years 2011 through 2015, which would have been included as part of the Woods and Poole 2012 data release. To create a test projection of values for Missoula MSA GRP for those years, we use the percentage annual growth terms from the Woods and Poole 2016 data release for the years 2015 through 2019. With those percentage values for annual GRP growth, notional projected values for Missoula GRP can be calculated for 2011 through 2015. To create notional projected values for the disaggregated socioeconomic variable, the percentage of Missoula MSA households with 2009 dollar denominated income exceeding $100,000 (in 2009 dollars), a similar procedure was used.

Forecast values for the Oil Price dependent variable are also needed to calculate out of sample projections. These were obtained for the years 2011 to 2015 from the data reported in the FAA's 2012 Annual Aerospace Forecast document. These national values are used for each of the case study airports out of sample forecast exercises.

Four out of sample model forecasts were calculated for the baseline model using Missoula MSA GRP and the model that also includes the disaggregated socioeconomic variable. The first three are based on the out of sample forecasts described above. For the first forecast, the projected values – based on the percentage increases projected for the years 2015 through 2019 – for GRP, and the percentage of Missoula household incomes exceeding $100,000 in the case of the model using the disaggregated socioeconomic variable as input, are used as described above. For the second, the values for percentage annual increases over the projection period (2011 to 2015) are increased by 50 percent, indicating a more rapidly growing Missoula economy. For the third MSO projection, the values for percentage annual increases are reduced by 50 percent, representing a projection for a more slowly growing economic environment in the Missoula MSA. For the fourth projection, the actual values observed in 2011 through 2015 for GRP and the share of Missoula MSA households with incomes exceeding $100,000 are used to drive the "forecast" of MSO annual O&D enplanements.

Using the projections published by Woods and Poole for 2015 and subsequent years for simulating a range of forecast scenarios based on inputs that would have been available to analysts in 2011 or 2012, we can compare the forecasting performance of the baseline regressions for MSO annual O&D enplanements using the annual oil price variable and the Missoula MSA GRP, with those of the regression parameters based additionally on input from the disaggregated socioeconomic variable. The independent variable data and projections are shown in Table D-13. The projections for the Missoula MSA variables reported in the Woods and Poole 2016 data have annual GRP growing at a rate of around 1.6 percent, and the percentage of Missoula MSA households with incomes exceeding $100,000 growing at about 1 percent initially, growing to 2 percent after three years. The annual percentage changes used to calculate the Oil Price projected values for 2011 through 2015 are those implied by the Oil Price projections reported in the 2012 FAA Aerospace Forecast, which relies on projected values for the years following 2010. Thus we have four "forecast" scenarios for the two regression specifications for annual MSO O&D enplanements:

- Projected values for the socioeconomic independent variables based on the annual percentage changes in the variables, as reported by Woods and Poole for the five years beginning in 2015,
- Projected values based on annual percentage changes that are 1.5 times these values,
- Projected values based on annual percentage changes that are half these values, and]
- Projected values that equal the actual values for the independent variables that occurred in the years 2011 through 2015.

Table D-13 Inputs to Out-of-Sample MSO O&D Enplanement Projections

MSO Inputs to Out-of-Sample Enplanement Forecasts	MSO					
	Sources	2011	2012	2013	2014	2015
Oil Price Forecast	*2012 FAA Aero Forecast*	$84.32	$92.71	$97.16	$100.65	$104.64
GRP Forecast 1 (Medium)	*W&P 2016 %*	$4,892.2	$4,973.7	$5,053.4	$5,132.9	$5,211.9
GRP Forecast 2 (High)	*1.5 x F'cast 1*	$4,933.8	$5,057.2	$5,178.6	$5,300.8	$5,423.3
GRP Forecast 3 (Low)	*0.5 x F'cast 1*	$4,850.6	$4,891.0	$4,930.2	$4,969.0	$5,007.2
HHInc% Forecast 1 (Medium)	*W&P 2016 %*	14.0%	14.2%	14.4%	14.6%	14.9%
HHInc% Forecast 2 (High)	*1.5 x F'cast 1*	14.1%	14.4%	14.6%	15.0%	15.5%
HHInc% Forecast 3 (Low)	*0.5 x F'cast 1*	14.0%	14.0%	14.1%	14.3%	14.4%
	2010					
Actual Oil Price	$83.99	$101.87	$100.93	$100.49	$92.02	$48.39
Actual GRP	$4,809.0	$4,958.6	$4,975.8	$5,002.7	$5,073.4	$5,164.4
Actual HHInc%	13.9%	14.1%	14.4%	14.5%	14.5%	14.7%

The forecasting or projection results for the baseline regression equation for annual enplanements at MSO are shown in Figure D-14 below.

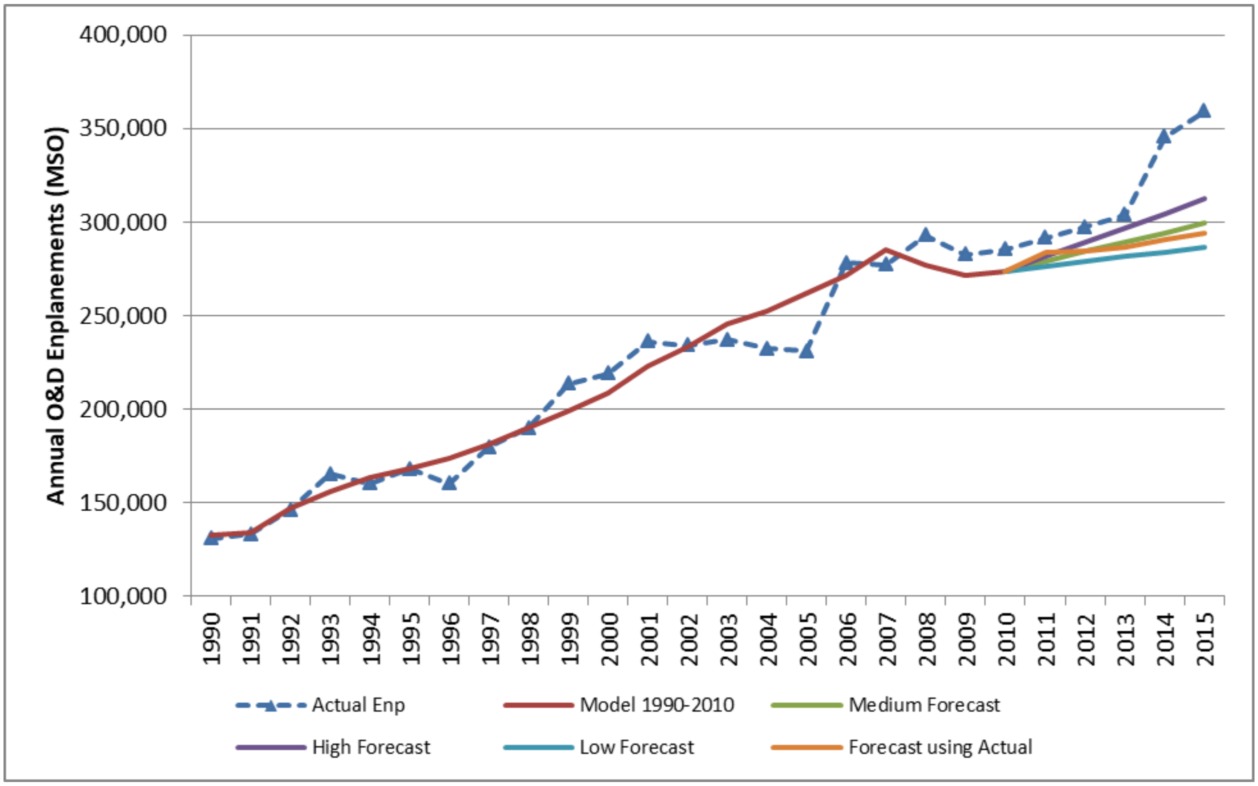

Figure D-14 Baseline Model Forecasts of MSO Annual Enplanements, 2011 to 2015, Using Missoula MSA Gross Regional Product, 1990 to 2010

The results for the regression equation using the disaggregated socioeconomic variable along with Missoula MSA GRP are shown in Figure D-15.

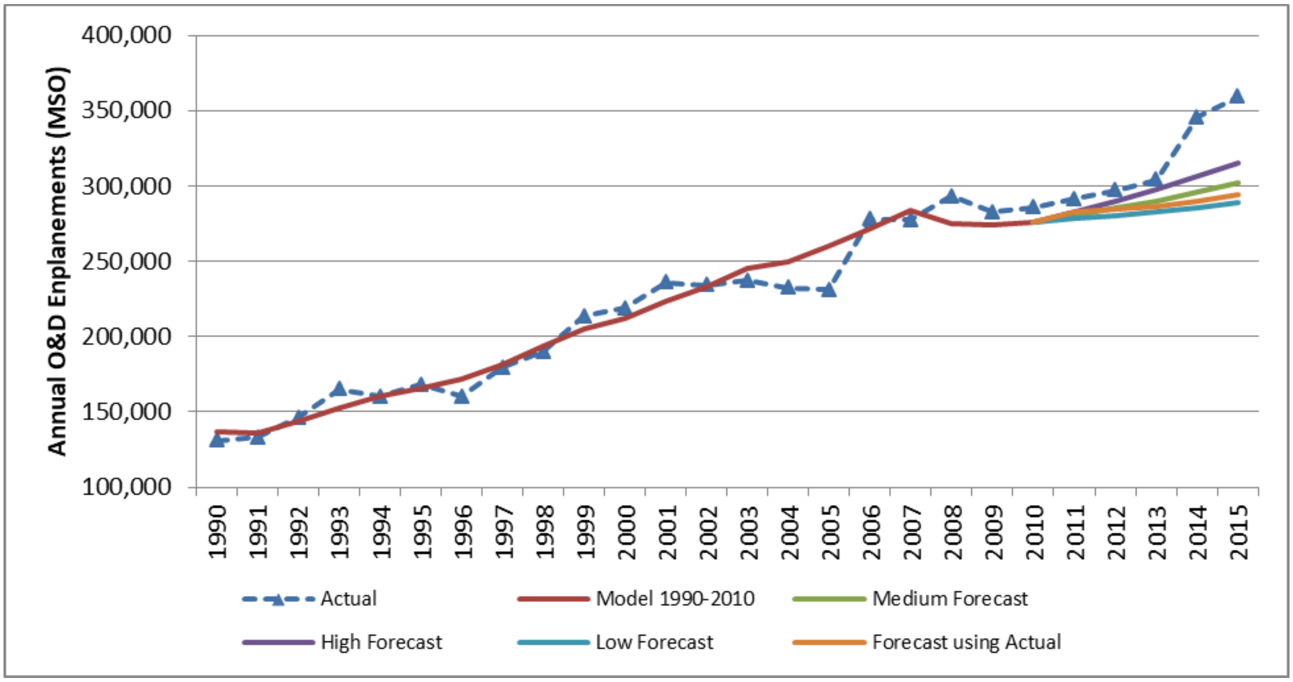

Figure D-15 Disaggregated SE Variable Model Forecasts of MSO Annual Enplanements, 2011 to 2015, Using Missoula MSA GRP and Share of Missoula Households with Incomes exceeding $100,000, 1990 to 2010

Enplanements at MSO increased more strongly than was forecast by either model and under any of the forecast scenarios for the independent variables. In fact, the two forecast charts look remarkably similar, suggesting that even though the forecast models differ in some respects between the two specifications, they are using similar information from the available data.[6] This is also suggested by the high R-squared scores for both model specifications, without and with the disaggregated socioeconomic variable. The accuracy of the projections driven by the actual values taken by the Oil Price, Missoula MSA GRP, and the Missoula Share of Households with Income exceeding $100,000 in the years 2011 through 2015 is also modest, with the projected enplanement counts undershooting the actual history over those years. The following tables D-14 through D-16 quantify this, reporting the percentage error for each of the forecasts and the Root Mean Squared Error (RMSE) for each forecast, a calculation which can be used to compare the accuracy of the baseline model forecasts and the forecasts calculated from the model also using the disaggregated socioeconomic variable.

---

[6] The baseline regression equation is ln(MSO_Enp) = 3.35 + 0.01ln(OilPrice) + 1.08ln(GRP),), with an adjusted R squared of 0.945, while the estimated equation that includes the disaggregated socioeconomic variable is ln(MSO_Enp) = 7.11 + 0.005ln(OilPrice) + 0.73ln(GRP) + 0.42ln(%HHInc>$100K) with an adjusted R squared of 0.950, so for MSO the explanatory contribution from the Oil Price variable is (counterintuitively) positive but very small in magnitude.

Table D-14 Percentage Forecast Errors, Baseline Model Projections of MSO Annual O&D Enplanements, 2011 to 2015

MSO (using GRP)	Percentage Forecast Error (Baseline Regression with Oil Price and GRP)			
	Medium	High	Low	Actual
2010	--	--	--	--
2011	-4.3%	-3.4%	-5.2%	-2.8%
2012	-4.4%	-2.7%	-6.1%	-4.3%
2013	-4.8%	-2.3%	-7.3%	-5.8%
2014	-14.8%	-11.8%	-17.8%	-16.0%
2015	-16.7%	-13.0%	-20.2%	-18.0%

The percentage errors in the tables reflect the fact that the model-based projections of MSO enplanements consistently underestimated the number of annual passenger, even for the "forecast" based on the values that actually occurred for the Oil Price and Missoula MSA GRP. As can be seen in the figures and the tables of percentage errors, the model forecasts using these actual values do not differ from those based on the projected values for the independent variables, in spite of the significant drop in Oil Prices in 2015, from $92.02 in 2014 to $48.39 in 2015. This is because the estimated parameter for the Oil Price independent variable in the MSO regression estimations using GRP as the aggregate socioeconomic variable is very small in magnitude, so the change in actual oil prices between 2014 and 2015 exerts little influence on the forecast of MSO enplanements for both the baseline model and the alternative model.

Table D-15 Percentage Forecast Errors, Disaggregated SE Model Projections of MSO Annual O&D Enplanements, 2011 to 2015

MSO (using GRP)	Percentage Forecast Error (Specification including Disaggregated SE Variable)			
	Medium	High	Low	Actual
2010	--	--	--	--
2011	-3.8%	-3.0%	-4.6%	-3.0%
2012	-4.0%	-2.4%	-5.6%	-4.4%
2013	-4.5%	-2.1%	-6.9%	-5.9%
2014	-14.4%	-11.4%	-17.3%	-16.1%
2015	-15.9%	-12.1%	-19.6%	-18.1%

The final MSO analytic table reports the Root Mean Squared Error (RMSE) of the different forecasts for the two model specifications. This statistic can be used to compare the accuracy of two or more of these forecasts, with a smaller RMSE value indicating a more accurate forecast. As can be seen, the inclusion of the additional disaggregated socioeconomic variable in the model specification provided only very modest improvement in the accuracy of the five year forecasts, compared to those based on the baseline specification.

Table D-16 Root Mean Squared Error (RMSE) for MSO Baseline and Disaggregated SE Variable Model Forecasts of Annual MSO O&D Enplanements, 2011 to 2015)

MSO (using GRP)	Root Mean Squared Error		% Difference BL v Alt
	Baseline	Disagg Var	
Medium Forecast	36,751	35,200	4.2%
High Forecast	28,522	26,877	5.8%
Low Forecast	44,891	43,407	3.3%
Actual Values	39,466	39,628	-0.4%

d. Harrisburg International Airport (MDT)

The case study indexed socioeconomic variables from 1990 to 2010 for the Harrisburg, Pennsylvania MSA are shown in Figure D-16. These variables, indexed with their 2000 values set to 100, are also highly correlated series, each gradually increasing over time. Differences in paces of growth exist for Harrisburg as well, with the more demographic variables – population and employment – growing more modestly, and economic variables growing more strongly, reflecting the influence of productivity improvements on economic activity over time. The Harrisburg region is served by a small hub, Harrisburg International Airport (MDT).

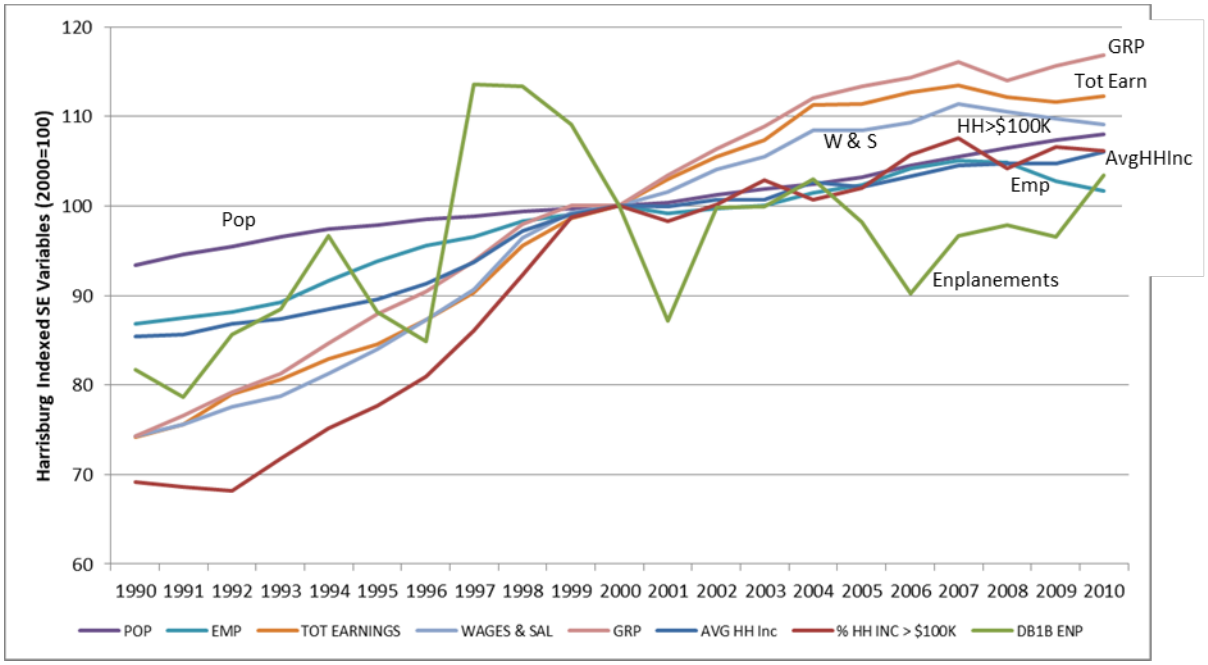

Figure D-16 Harrisburg, PA MSA Indexed Socioeconomic Variables, 1990 to 2010, 2000=100

Table D-17 reports values of the socioeconomic variables for the Harrisburg MSA from 1990 to 2010.

Table D-17 Values of Socioeconomic Variables, Harrisburg, PA MSA (1990-2010)

	Year	Harrisburg International Airport (MDT)								
		Annual O&D Enplanements	Real Oil Price	Regional Population (000s)	Regional Employment (000s)	Regional Total Earnings ($M)	Regional W&S Earnings ($M)	Regional Product (GRP) ($M)	Regional Avg HH Income	% Regional HHs with Income > $100K
	1990	514,185	$40.16	475.9	327.4	$14,276.1	$10,568.7	$21,385.2	$84,711	12.5%
	1991	494,718	$33.20	482.0	329.9	$14,548.8	$10,748.2	$22,060.7	$84,997	12.4%
	1992	538,808	$31.33	486.5	332.6	$15,186.8	$11,037.9	$22,807.5	$86,185	12.3%
	1993	556,868	$26.96	491.9	336.6	$15,505.9	$11,207.6	$23,404.4	$86,738	13.0%
	1994	607,961	$25.31	496.2	345.7	$15,941.0	$11,553.7	$24,370.8	$87,735	13.6%
	1995	554,327	$27.16	498.8	353.8	$16,256.1	$11,945.4	$25,341.3	$88,899	14.1%
	1996	534,030	$31.73	502.0	360.4	$16,789.0	$12,412.1	$26,061.1	$90,630	14.6%
	1997	714,189	$27.80	503.6	364.4	$17,380.2	$12,891.4	$27,021.7	$92,951	15.6%
	1998	712,940	$17.87	506.5	370.8	$18,380.3	$13,712.6	$28,226.1	$96,422	16.7%
	1999	686,385	$25.03	508.2	373.5	$18,967.6	$14,122.6	$28,806.9	$98,238	17.9%
	2000	629,003	$38.86	509.5	377.2	$19,237.2	$14,221.5	$28,802.8	$99,205	18.1%
	2001	548,079	$29.99	511.4	374.0	$19,819.2	$14,442.8	$29,802.8	$99,100	17.8%
	2002	627,848	$31.83	515.6	376.0	$20,291.6	$14,807.0	$30,632.1	$99,896	18.1%
	2003	628,824	$36.38	519.2	377.6	$20,650.7	$15,010.5	$31,369.6	$99,861	18.6%
	2004	647,651	$45.87	521.9	382.9	$21,400.3	$15,415.8	$32,279.6	$101,881	18.2%
	2005	617,432	$60.44	525.8	386.1	$21,430.0	$15,415.6	$32,662.2	$101,334	18.5%
	2006	567,441	$70.72	532.3	392.9	$21,671.2	$15,543.2	$32,942.1	$102,506	19.2%
	2007	608,349	$78.22	537.3	396.4	$21,829.2	$15,840.7	$33,445.9	$103,642	19.5%
	2008	615,269	$103.82	542.3	395.6	$21,583.8	$15,723.3	$32,851.2	$103,958	18.9%
	2009	607,005	$66.42	546.9	387.9	$21,476.6	$15,607.8	$33,319.1	$103,952	19.3%
	2010	650,801	$83.99	550.3	383.8	$21,589.6	$15,522.5	$33,658.6	$105,231	19.2%
Used to assess out-of-sample forecasts	*2011*	*626,567*	*$101.87*	*551.9*	*387.2*	*$21,876.9*	*$15,647.5*	*$33,650.1*	*$105,736*	*19.0%*
	2012	*637,949*	*$100.93*	*554.8*	*389.2*	*$22,010.6*	*$15,838.5*	*$33,776.3*	*$106,915*	*19.6%*
	2013	*630,798*	*$100.49*	*557.9*	*393.6*	*$22,506.2*	*$16,113.2*	*$34,426.0*	*$106,153*	*19.6%*
	2014	*642,726*	*$92.02*	*560.8*	*397.6*	*$22,978.1*	*$16,516.0*	*$35,139.0*	*$107,646*	*19.6%*
	2015	*595,245*	*$48.39*	*565.0*	*403.8*	*$23,381.9*	*$16,826.6*	*$35,840.5*	*$108,259*	*19.7%*

As reported in Table D-18, the 1990 to 2010 values for the socioeconomic data variables are very strongly correlated, with all values exceeding 0.9. This strong correlation across these socioeconomic factors suggests that the individual variables represent alternative perspectives on a commonly evolving socioeconomic environment. The correlations between the Harrisburg MSA socioeconomic variables and either oil prices or annual O&D enplanements at MDT are also positive but less strong.

Table D-18 Correlations among Socioeconomic Variables, Harrisburg, PA MSA (1990 – 2010)

	Population	Emp	Total Earn	Wages & Salaries	GRP	Avg HH Inc	% > $100K	DB1B Enp	Real Oil Price
Population	1								
Employment	0.918	1							
Total Earn	0.936	0.972	1						
Wages & Salaries	0.928	0.981	0.997	1					
GRP	0.956	0.976	0.996	0.994	1				
Avg HH Inc	0.939	0.972	0.989	0.994	0.987	1			
% > $100K	0.906	0.974	0.980	0.990	0.977	0.991	1		
DB1B Enplanements	0.427	0.539	0.483	0.518	0.499	0.531	0.546	1	
Real Oil Price	0.792	0.635	0.665	0.642	0.675	0.659	0.596	0.030	1

The Harrisburg, Pennsylvania MSA saw only modest change in its population's age distribution over the case study period. Those 19 years old or younger and those aged between 20 and 34 combined to make up 50 percent of Harrisburg's population in 1985, a total which decreased slightly to 44 percent by 2010, as shown in Figure D-17.

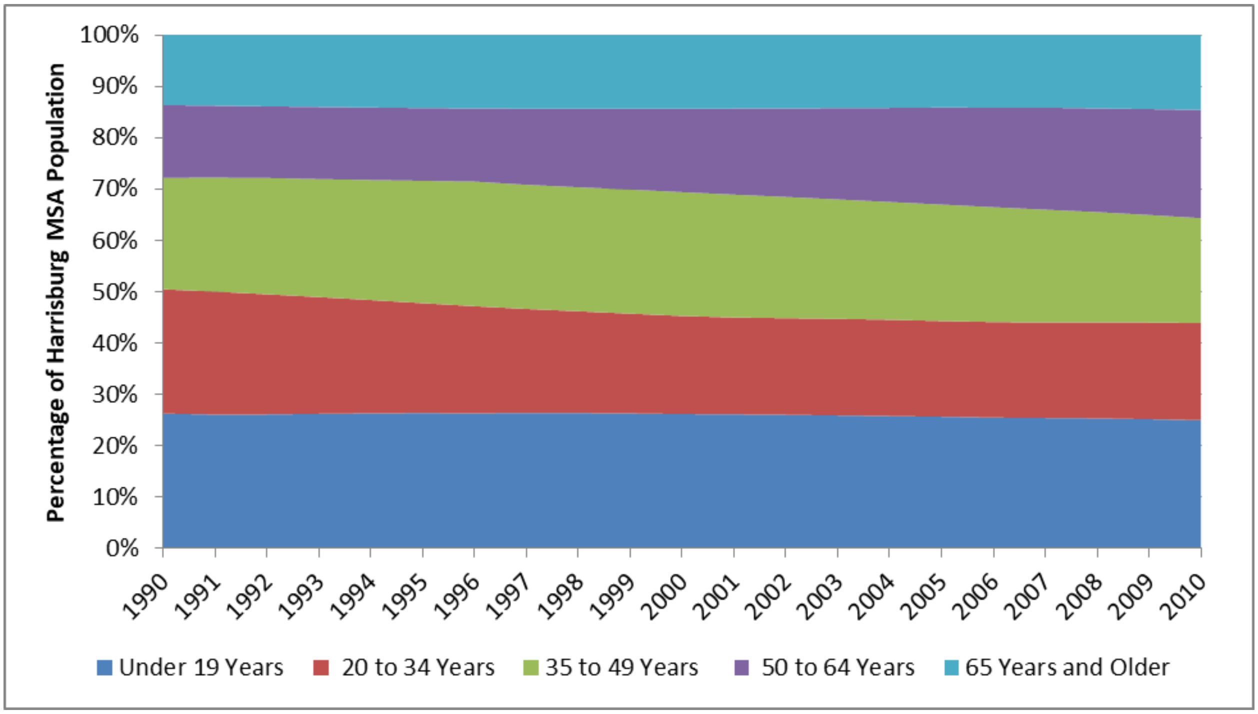

Figure D-17 Population Age Distribution, Harrisburg, PA MSA, 1990 to 2010

Figure D-18 shows the distribution of Harrisburg's adult population, those 20 and older. It may be worthwhile to exclude the youngest cohort, since few air travel decisions would be made by residents younger than 20. During the case study period Harrisburg's adult population distribution changed little, although those between 50 and 64 years of age did increase from around 20 percent of the adult population to nearly 35 percent by 2010.

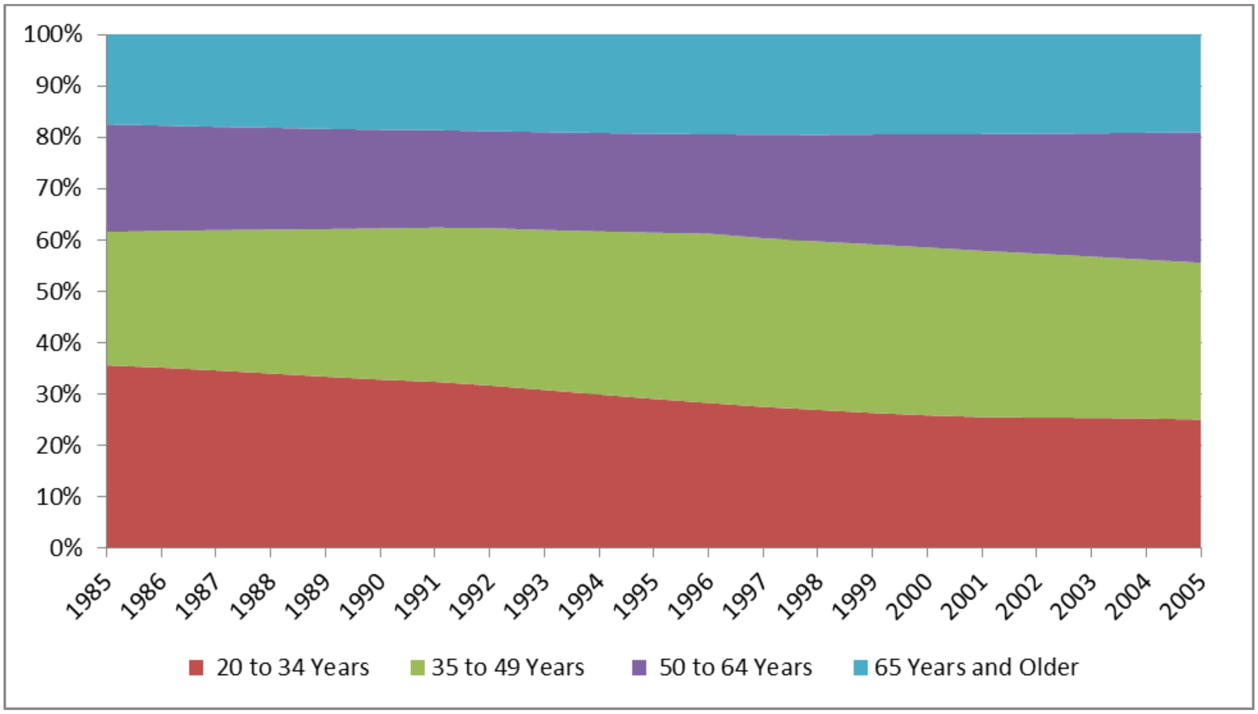

Figure D-18 Adult Population Age Distribution, Harrisburg, PA MSA, 1990 to 2010

Figure D-19 shows household income distributions for the residents of the Harrisburg MSA over the case study sample period. The figure shows a gradual increase in the proportion of Harrisburg households with incomes exceeding $100,000 over the sample period, from around 12 percent in 1990, when the income distribution data is first available, to nearly 20 percent by 2010. As shown in the analysis of aggregate socioeconomic trends over the case study sample period for Harrisburg (and other locations), average household incomes also increased over the period.

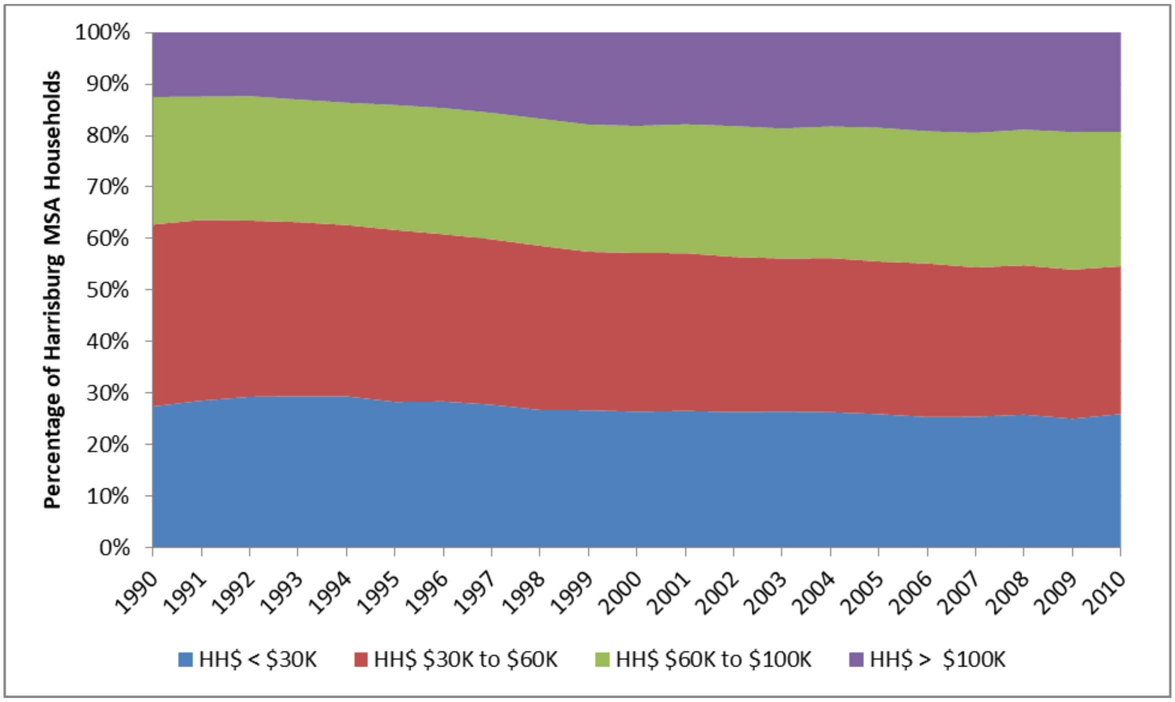

Figure D-19 Household Income Distribution, Harrisburg, PA MSA, 1990 to 2010

Figure D-20 reports specifically the changes in the percentage of Harrisburg MSA households with annual incomes of $100,000 (in 2009 dollars) (a group shown as the uppermost band in Figure D-19 above). This percentage rose steadily from 1992 to 1999, from about 12 percent of Tulsa MSA households to 18 percent by 1999. The percentage then flattened out through 2010, with periods of growth and decline.

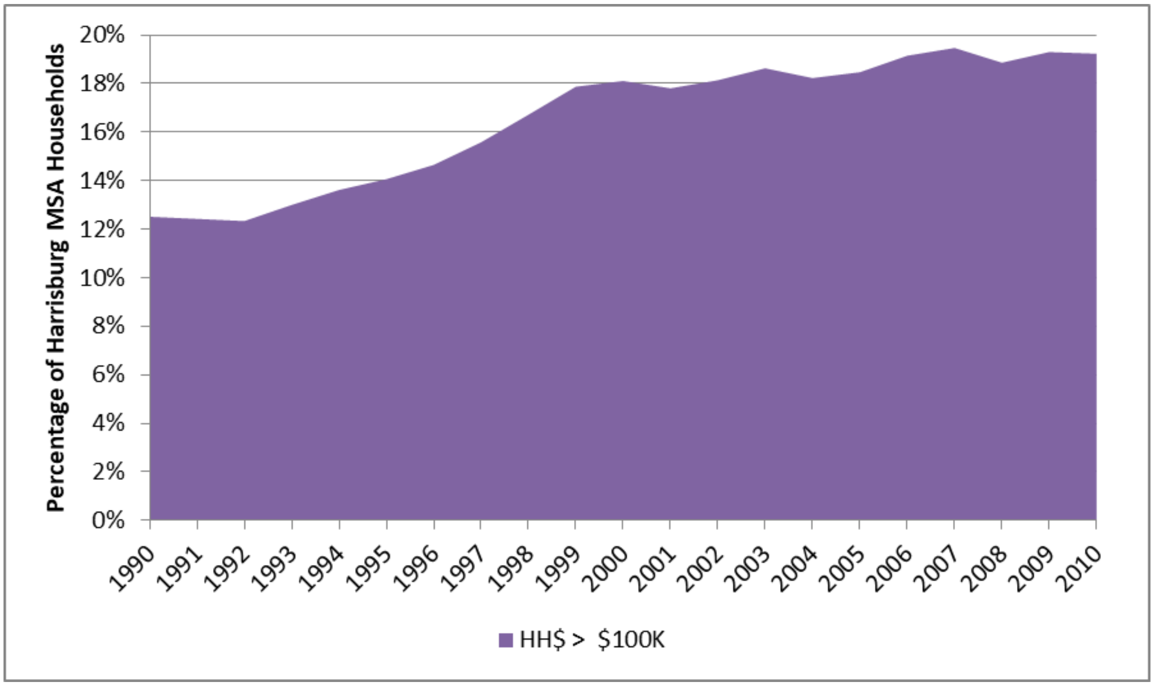

Figure D-20 Percentage of Households with Incomes Exceeding $100,000, Harrisburg, PA, 1990 to 2010

Characteristics of the results for the six baseline regressions modeling annual passenger enplanements at Harrisburg International Airport (MDT) from 1990 to 2010 include:

- The six baseline model estimates had similar goodness of fit to the actual MDT annual enplanements series, with adjusted R-squared statistics ranging from 0.416 to 0.484, indicating that the baseline models regressing annual enplanements on the annual real oil price and one of the six aggregate socioeconomic variables for the Harrisburg MSA accounted for less than half of the observed variation in the annual MDT enplanements series.
- The estimated regression constant terms varied across the baseline models, although the estimates for the models using Total Earnings in the Harrisburg MSA, Annual Wages and Salaries in the MSA, and Gross Regional Product were roughly comparable, ranging from 6.88 to 7.21. All of these estimates were at the 0.005 significance level (or better). The estimated constant term for the model using MSA annual employment was estimated at 4.39, and was statistically significant at the 0.05 level. Constant estimates for the models using Harrisburg MSA population and Average Real Household Income took negative values and were not statistically for customary test values.
- The coefficients estimated in the baseline MDT models for the oil price variable, which in a log linear regression represents the elasticity of changes in annual MDT enplanements with respect to changes in the oil price, were relatively small (and negative, as would be expected) in all baseline regressions. The estimated coefficient values were within a close range of one another, ranging between -0.121 and -0.189), and all were statistically significant only at the 0.005 level. This indicates that over the sample period

there was on average modest responsiveness of annual air passenger enplanements to changes in oil prices, and that these levels of response are precisely estimated.

- The influence of the individual aggregate socioeconomic variables on annual enplanements at MDT is positive and strongly statistically significant across all of the baseline models. The range of coefficient (elasticity) estimates exhibits the same declining pattern as that described for the baseline regressions covering the Baltimore-Washington CSA, with the largest elasticity estimate (2.96, significant at the 0.005 level) occurring for Harrisburg MSA population, a smaller value for the regression estimate using Harrisburg MSA total employment (1.58, also significant at the 0.005 level), and then estimates between 0.67 and 0.69 for Total Earnings, Wages and Salaries, and GRP (all also significant at the 0.005 level). The elasticity estimate for Average Real Household Income in the Harrisburg MSA was somewhat greater than the estimates for these three, estimated at 1.11, and also significant at the 0.005 level.

The MDT regression equation substituting the series of annual data on the percentage of Harrisburg MSA households with incomes of $100,000 or more (a disaggregated socioeconomic variable) for the aggregate socioeconomic variables used in the six baseline regressions is very similar to the results from those baseline regressions. The estimated constant term of 14.72, is much larger that the constant term that resulted in the six baseline regressions, reflecting the smaller magnitude of the values taken by the "share of households" socioeconomic variable, which affects the magnitude of the constant term estimate, and it is significant at the 0.005 significance level. The estimated coefficient for the Oil Price variable, -0.111, is comparable to the coefficient values in the six baseline regressions for MDT, and the estimate is statistically significant at the 0.01 level of significance. The coefficient estimate for the disaggregated socioeconomic variable introduced into the model specification, 0.56, is also comparable to the baseline estimates for the individual socioeconomic variables. This estimate is significant at the 0.005 level of significance. Finally, the adjusted R-squared value for this specification is 0.453, a value similar to those found for the baseline estimations. The disaggregated socioeconomic variable does not appear to bring new information to the estimations for MDT annual O&D enplanements in this simple specification.

Table D-19 Baseline and Other Regression Results for Models of Annual Enplanements, Harrisburg International Airport, 1990 to 2010

Ind. Variable	Statistic	Regression 1	2	3	4	5	6	7
Constant	Coefficient	-4.50	4.39	7.21	7.21	6.88	-2.03	14.72
	t-Statistic	-1.170	2.264 *	5.150 ***	5.406 ***	4.939 ***	-0.648	44.059 ***
Oil Price	Coefficient	-0.19	-0.12	-0.13	-0.12	-0.13	-0.14	-0.11
	t-Statistic	-3.572 ***	-2.848 **	-2.866 **	-2.889 **	-3.035 **	-3.216 **	-2.692 **
Pop	Coefficient	2.96						
	t-Statistic	4.626 ***						
Emp	Coefficient		1.58					
	t-Statistic		4.603 ***					
Total Earnings	Coefficient			0.67				
	t-Statistic			4.362 ***				
Wages and Salaries	Coefficient				0.69			
	t-Statistic				4.579 ***			
Gross Regional Product	Coefficient					0.67		
	t-Statistic					4.626 ***		
Average HH Income	Coefficient						1.38	
	t-Statistic						4.900 ***	
Disaggregated SE Variable: % HH with Income > $100K	Coefficient							0.56
	t-Statistic							4.595 ***
	R Squared	0.519	0.518	0.474	0.503	0.510	0.535	0.508
	Adj R Squared	0.465	0.464	0.416	0.448	0.455	0.484	0.453

* Coefficient estimate significant at 0.05 level ** at 0.01 level *** at 0.001 level

When the disaggregated socioeconomic variable is added as a new variable to the baseline regression specifications, many of the relationships that were estimated in the baseline equations no longer hold. These results are reported in Table D-20. The equation constant estimates are different across the equations, possibly a reflection of the much smaller magnitude of the values that make up the disaggregated socioeconomic variable.

- For these regressions, the adjusted R-squared estimates are about the same as those reported for the baseline regression equations, suggesting that inclusion of the disaggregated socioeconomic variable in the models provided no new information for modeling the 1990 to 2010 O&D passenger enplanements at MDT.
- Regression constants in only three of the specifications are statistically significant at the 0.1 level of significance or better, the equation that includes the Harrisburg MSA Total Earnings aggregate socioeconomic variable and the equation that uses the Harrisburg MSA Gross Regional Product variable.
- For each of the specifications combining one of the aggregate socioeconomic variables with the disaggregated socioeconomic variable, the coefficient estimate for the Oil Price variable is similar in magnitude (and sign) with the estimates for the Oil Price variable in the baseline regressions described above. These estimates range from -0.110 to -0.161, and all are statistically significant estimates at the 0.05 significance level or better.
- The coefficient estimates for the aggregate socioeconomic variables are, with the exception of the specifications using the Harrisburg MSA Average Real Household Income, all small in magnitude, and none of the estimates are statistically significant.
- The coefficient estimates for the disaggregated socioeconomic variable are positive as expected, with the exception of the regression specification using the Average Real Household Income, and none of the estimates are statistically significant. For these regressions, the adjusted R-squared estimates are about the same as those reported for the baseline regression equations, ranging from 0.384 to 0.453. This suggests that inclusion of the disaggregated socioeconomic variable in the models provided no new information for modeling the 1990 to 2010 O&D passenger enplanements at MDT.

Table D-20 Regression Results for Models of Annual Enplanements, Harrisburg International Airport, 1990 to 2010, that Include both Aggregate and Disaggregated Socioeconomic Variables

Ind. Variable	Statistic	Regression					
		1	2	3	4	5	6
Constant	Coefficient	4.33	9.36	14.83	11.43	10.35	-16.02
	t-Statistic	1.090	1.090	1.872 *	1.069	1.434	-0.665
Oil Price	Coefficient	-0.161	-0.118	-0.110	-0.118	-0.126	-0.156
	t-Statistic	-2.833 **	-2.716 **	-2.199 *	-2.468 *	-2.593 **	-2.900 **
Pop	Coefficient	1.618					
	t-Statistic	1.263					
Emp	Coefficient		0.826				
	t-Statistic		0.625				
Total Earnings	Coefficient			-0.010			
	t-Statistic			-0.014			
Wages and Salaries	Coefficient				0.302		
	t-Statistic				0.308		
Gross Regional Product	Coefficient					0.378	
	t-Statistic					0.606	
Average HH Income	Coefficient						2.529
	t-Statistic						1.276
Disaggregated SE Variable: % HH with Income > $100K	Coefficient	0.295	0.280	0.570	0.319	0.256	-0.487
	t-Statistic	1.207	0.595	0.977	0.398	0.491	-0.586
	R Squared	0.547	0.522	0.508	0.509	0.516	0.542
	Adj R Squared	0.467	0.438	0.421	0.422	0.430	0.461

* Coefficient estimate significant at 0.05 level ** at 0.01 level *** at 0.001 level

To assess the forecasting performance of these models, we limit the analysis to consideration of a single pair of regressions, those that use the Harrisburg MSA Gross Regional Product as the aggregate socioeconomic variable. As mentioned above, it is not possible to obtain the Woods and Poole projections that would have been available for the years 2011 through 2015, which would have been included as part of the Woods and Poole 2012 data release. To create a test projection of values for Harrisburg MSA GRP for those years, we use the percentage annual growth terms from the Woods and Poole 2016 data release for the years 2015 through 2019. With those percentage values for annual GRP growth, notional projected values for Harrisburg GRP can be calculated for 2011 through 2015. To create notional projected values for the disaggregated socioeconomic variable, the percentage of Harrisburg MSA households with 2009 dollar denominated income exceeding $100,000 (in 2009 dollars), a similar procedure was used.

Forecast values for the Oil Price dependent variable are also needed to calculate out of sample projections. These were obtained for the years 2011 to 2015 from the data reported in the FAA's 2012 Annual Aerospace Forecast document. These national values are used for each of the case study airports out of sample forecast exercises.

Four out of sample model forecasts were calculated for the baseline model using Harrisburg MSA GRP and the model that also includes the disaggregated socioeconomic variable. The first three are based on the out of sample forecasts described above. For the first forecast, the projected values – based on the percentage increases projected for the years 2015 through 2019 – for GRP, and the percentage of Harrisburg household incomes exceeding $100,000 in the case of the model using the disaggregated socioeconomic variable as input, are used as described above. For the second, the values for percentage annual increases over the projection period (2011 to 2015) are increased by 50 percent, indicating a more rapidly growing Harrisburg economy. For the third MDT projection, the values for percentage annual increases are reduced by 50 percent, representing a projection for a more slowly growing economic environment in the Harrisburg MSA. For the fourth projection, the actual values observed in 2011 through 2015 for GRP and the share of Harrisburg MSA households with incomes exceeding $100,000 are used to drive the "forecast" of MDT annual O&D enplanements.

Using the projections published by Woods and Poole for 2015 and subsequent years for simulating a range of forecast scenarios based on inputs that would have been available to analysts in 2011 or 2012, we can compare the forecasting performance of the baseline regressions for MDT annual O&D enplanements using the annual oil price variable and the Harrisburg MSA GRP, with those of the regression parameters based additionally on input from the disaggregated socioeconomic variable. The independent variable data and projections are shown in Table D-21 below. The projections for the Harrisburg MSA variables reported in the Woods and Poole 2016 data have annual GRP growing at a rate of around 2.4 percent, and the percentage of Harrisburg MSA households with incomes exceeding $100,000 growing at about 1.8 percent across the forecast period. The annual percentage changes used to calculate the Oil Price projected values for 2011 through 2015 are those implied by the Oil Price projections reported in the 2012 FAA Aerospace Forecast, which relies on projected values for the years following 2010. Thus we have the four "forecast" scenarios for the two regression specifications for annual MDT O&D enplanements:

- Projected values for the socioeconomic independent variables based on the annual percentage changes in the variables, as reported by Woods and Poole for the five years beginning in 2015 ,
- Projected values based on annual percentage changes that are 1.5 times these values,
- Projected values based on annual percentage changes that are half these values, and]
- Projected values that equal the actual values for the independent variables that occurred in the years 2011 through 2015.

Table D-21 Inputs to Out-of-Sample MDT O&D Enplanement Projections

MDT Inputs to Out-of-Sample Enplanement Forecasts	Sources	2011	2012	2013	2014	2015
Oil Price Forecast	*2012 FAA Aero Forecast*	$84.32	$92.71	$97.16	$100.65	$104.64
GRP Forecast 1 (Medium)	*W&P 2016 %*	$34,464.8	$35,272.5	$36,085.3	$36,906.1	$37,736.0
GRP Forecast 2 (High)	*1.5 x F'cast 1*	$34,867.9	$36,093.5	$37,341.2	$38,615.3	$39,917.7
GRP Forecast 3 (Low)	*0.5 x F'cast 1*	$34,061.7	$34,460.8	$34,857.9	$35,254.3	$35,650.7
HHInc% Forecast 1 (Medium)	*W&P 2016 %*	19.6%	19.9%	20.3%	20.6%	21.0%
HHInc% Forecast 2 (High)	*1.5 x F'cast 1*	19.7%	20.3%	20.8%	21.4%	22.0%
HHInc% Forecast 3 (Low)	*0.5 x F'cast 1*	19.4%	19.6%	19.8%	19.9%	20.1%
	2010					
Actual Oil Price	$83.99	$101.87	$100.93	$100.49	$92.02	$48.39
Actual GRP	$33,658.6	$33,650.1	$33,776.3	$34,426.0	$35,139.0	$35,840.5
Actual HHInc%	19.2%	19.0%	19.6%	19.6%	19.6%	19.7%

The forecasting or projection results for the baseline regression equation for annual enplanements at MDT are shown in Figure D-21 below.

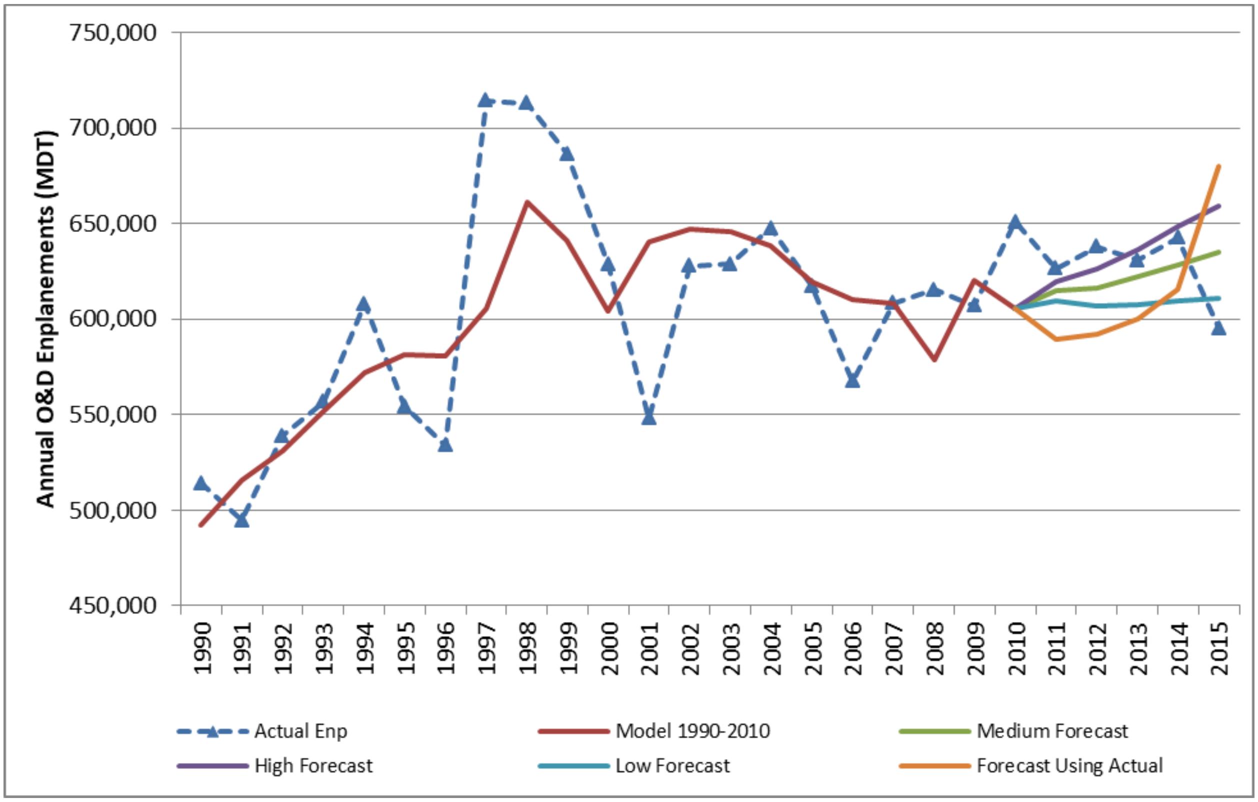

Figure D-21 Baseline Model Forecasts of MDT Annual Enplanements, 2011 to 2015, Using Harrisburg MSA Gross Regional Product, 1990 to 2010

The results for the regression equation using the disaggregated socioeconomic variable along with Harrisburg MSA GRP and the Oil Price independent variable are shown in Figure D-22 below.

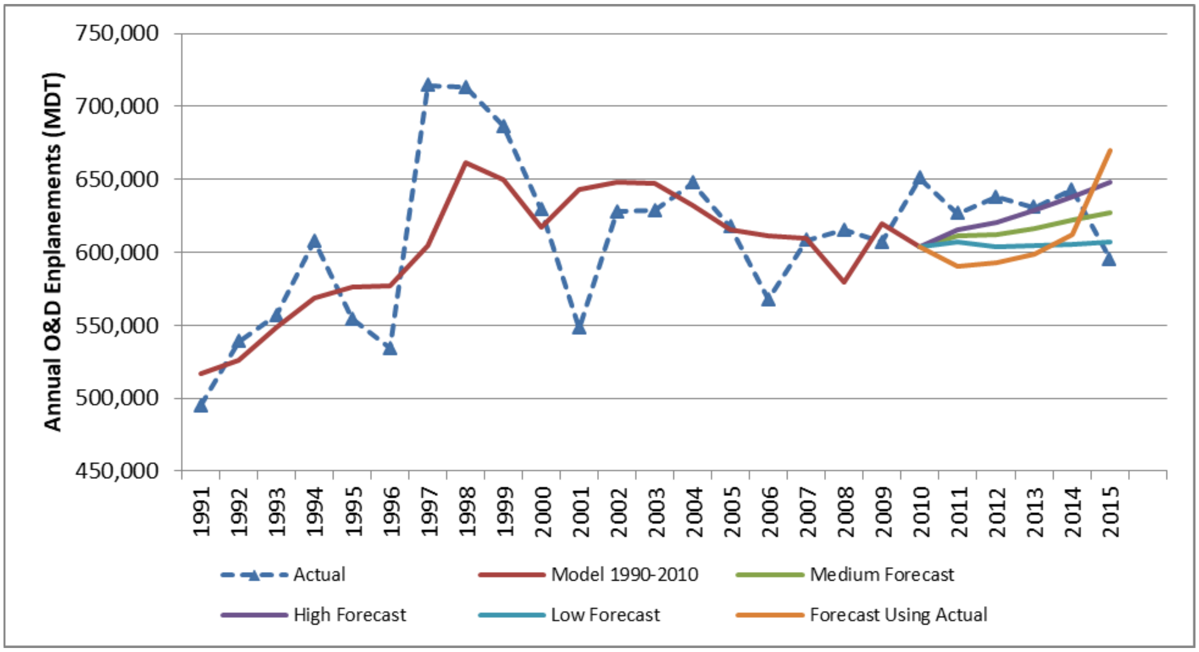

Figure D-22 Disaggregated SE Variable Model Forecasts of MDT Annual Enplanements, 2011 to 2015, Using Harrisburg MSA GRP and Share of Harrisburg Households with Incomes exceeding $100,000, 1990 to 2010

Enplanements at MDT resembled those projected in the high forecast scenario for both the baseline and alternative model specifications, except for 2015, when actual enplanements declined sharply from the prior year. This decline in enplanements is in contrast to the increase in enplanements predicted by the model forecasts that used actual independent variable values for 2011 to 2015 as inputs. These forecasts resulted from the sharp decline in actual oil prices between 2014 and 2015, combined with the negative coefficient estimate for the oil price variable in both the baseline and alternative model specifications.[7] The following tables D-22 through D-24 quantify this, reporting the percentage error for each of the forecasts and the Root Mean Squared Error (RMSE) for each forecast, a calculation which can be used to compare the accuracy of the baseline model forecasts and the forecasts calculated from the model also using the disaggregated socioeconomic variable.

[7] The baseline regression equation is ln(MDT_Enp) = 6.88 – 0.13ln(OilPrice) + 0.767ln(GRP), with an adjusted R squared of 0.455,while the estimated equation that includes the disaggregated socioeconomic variable is ln(MDT_Enp) = 10.35 – 0.13ln(OilPrice) + 0.38ln(GRP) + 0.26ln(%HHInc>$100K) with an adjusted R squared of 0.430.

Table D-22 Percentage Forecast Errors, Baseline Model Projections of MDT Annual O&D Enplanements, 2011 to 2015

MDT using GRP	Percentage Forecast Error (Baseline Regression with Oil Price and GRP)			
	Medium	High	Low	Actual
2010	--	--	--	--
2011	-1.9%	-1.1%	-2.7%	-5.9%
2012	-3.4%	-1.9%	-4.9%	-7.2%
2013	-1.4%	0.9%	-3.7%	-4.9%
2014	-2.2%	0.8%	-5.2%	-4.2%
2015	6.6%	10.8%	2.6%	14.2%

The predominantly negative percentage errors in the tables reflect the fact that the model-based projections of MDT enplanements consistently underestimated the number of annual passenger, even for the "forecast" based on the values that actually occurred for the Oil Price and Harrisburg MSA GRP. The model forecasts, both baseline and alternative, also do not capture the sharp drop in MDT enplanements in 2015. When the actual values for the independent variables are used for the forecasts, the significant drop in the actual oil price in 2015, from \$92.02 in 2014 to \$48.39 in 2015, combined with the negative regression coefficient on the Oil Price variable, resulted in a sharply increased forecast for 2015 MDT annual enplanements. This forecasted increase was at odds with the actual decline in MDT enplanements in 2015, leading to a positive forecast error, with the forecast enplanement values exceeding the actual ones.

Table D-23 Percentage Forecast Errors, Disaggregated SE Model Projections of MDT Annual O&D Enplanements, 2011 to 2015

MDT using GRP	Percentage Forecast Error (Specification including Disaggregated SE Variable)			
	Medium	High	Low	Actual
2010	--	--	--	--
2011	-2.4%	-1.8%	-3.1%	-5.8%
2012	-4.0%	-2.8%	-5.3%	-7.0%
2013	-2.3%	-0.3%	-4.2%	-5.0%
2014	-3.2%	-0.7%	-5.7%	-4.8%
2015	5.3%	8.8%	1.9%	12.5%

The final table for the regression analysis of MDT reports the Root Mean Squared Error (RMSE) of the different forecasts for the two model specifications. This statistic can be used to compare the accuracy of two or more of these forecasts, with a smaller RMSE value indicating a more accurate forecast. As can be seen, the inclusion of the additional disaggregated socioeconomic variable in the model specification provided only very modest improvement in the accuracy of the five year forecasts, compared to those based on the baseline specification.

Table D-24 Root Mean Squared Error (RMSE) for MDT Baseline and Disaggregated SE Variable Model Forecasts of Annual MDT O&D Enplanements, 2011 to 2015)

MDT (using GRP)	Root Mean Squared Error		% Difference BL v Alt
	Baseline	Disagg Var	
Medium Forecast	22,097	22,551	-2%
High Forecast	29,535	25,352	14%
Low Forecast	24,999	27,250	-9%
Actual Values	49,684	46,568	6%

e. Tulsa International Airport (TUL)

The indexed socioeconomic data series for the Tulsa, Oklahoma MSA over the years 1990 to 2010, shown in Figure D-23, indicates relatively strong growth in economic variables, relative to the more demographic variables such as MSA population and employment, followed by much more volatile behavior in the series in the later years of the sample period. This is likely related to the importance of the volatile oil and gas extraction industries for the Oklahoma economy. Tulsa air passengers are served by the third small hub airport in the case study group, Tulsa International Airport (TUL).

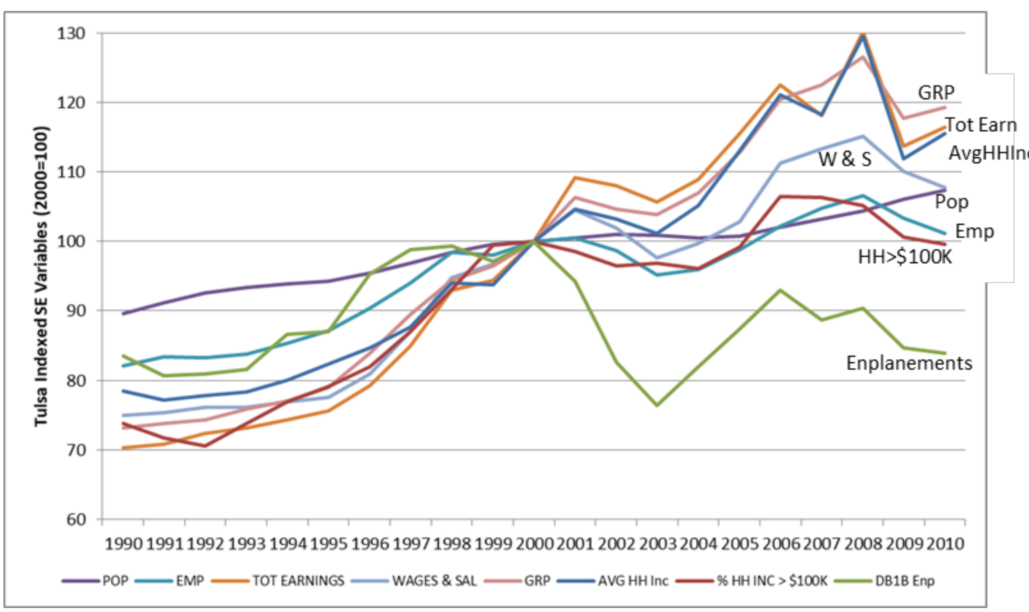

Figure D-23 Tulsa, OK MSA Indexed Socioeconomic Variables, 1990 to 2010, 2000=100

Table D-25 reports the values taken by the socioeconomic and other variables for the Tulsa MSA from 1990 to 2010.

Table D-25 Values of Socioeconomic Variables, Tulsa, OK MSA (1990-2010)

	Year	Tulsa International Airport (TUL)								
		Annual O&D Enplanements	Real Oil Price	Regional Population (000s)	Regional Employment (000s)	Regional Total Earnings ($M)	Regional W&S Earnings ($M)	Regional Product (GRP) ($M)	Regional Avg HH Income	% Regional HHs with Income > $100K
	1990	1,426,236	$40.16	763.1	430.8	$17,625.2	$12,856.5	$26,906.9	$74,712	11.4%
	1991	1,378,398	$33.20	774.8	437.7	$17,795.4	$12,961.8	$27,118.7	$74,122	11.1%
	1992	1,382,199	$31.33	785.8	438.2	$18,239.0	$13,125.5	$27,347.9	$74,944	10.9%
	1993	1,392,500	$26.96	795.1	444.9	$18,571.7	$13,208.6	$28,090.0	$75,362	11.4%
	1994	1,478,474	$25.31	800.5	453.9	$18,885.4	$13,373.6	$28,587.7	$77,103	11.9%
	1995	1,485,519	$27.16	805.7	466.0	$19,283.7	$13,590.3	$29,359.7	$79,062	12.2%
	1996	1,628,550	$31.73	816.3	481.9	$20,154.1	$14,149.8	$31,123.7	$81,032	12.6%
	1997	1,687,476	$27.80	829.4	499.8	$21,517.1	$15,159.9	$33,101.1	$83,653	13.4%
	1998	1,695,089	$17.87	843.1	523.1	$23,475.2	$16,500.0	$34,780.1	$88,949	14.4%
	1999	1,658,675	$25.03	854.6	522.6	$23,751.2	$16,744.1	$35,348.4	$89,007	15.3%
	2000	1,707,647	$38.86	861.2	533.3	$25,235.1	$17,347.0	$36,891.9	$94,408	15.4%
	2001	1,610,385	$29.99	867.6	541.0	$27,480.6	$18,188.4	$39,202.0	$98,193	15.2%
	2002	1,409,709	$31.83	874.8	533.8	$27,284.2	$17,825.1	$38,809.5	$97,153	14.9%
	2003	1,305,129	$36.38	877.6	519.8	$26,814.1	$17,195.6	$38,957.1	$95,393	14.9%
	2004	1,400,564	$45.87	878.0	523.7	$27,591.6	$17,486.1	$40,127.1	$98,558	14.8%
	2005	1,492,890	$60.44	882.9	538.9	$29,288.3	$18,075.7	$42,563.0	$104,530	15.3%
	2006	1,586,708	$70.72	894.0	557.3	$31,082.7	$19,568.4	$45,471.0	$111,679	16.4%
	2007	1,513,964	$78.22	906.4	572.7	$30,207.3	$20,059.9	$46,411.7	$110,474	16.4%
	2008	1,544,204	$103.82	916.5	582.5	$33,021.2	$20,442.0	$48,278.9	$119,403	16.2%
	2009	1,446,627	$66.42	929.8	564.4	$29,097.7	$19,509.4	$44,573.0	$106,506	15.5%
	2010	1,431,896	$83.99	939.9	553.2	$29,790.2	$19,128.4	$45,431.5	$109,223	15.4%
Used to assess out-of-sample forecasts	*2011*	*1,381,097*	*$101.87*	*945.9*	*558.8*	*$32,139.7*	*$19,924.6*	*$47,252.8*	*$114,993*	*15.4%*
	2012	*1,360,202*	*$100.93*	*952.8*	*573.0*	*$34,222.9*	*$20,723.5*	*$49,707.2*	*$122,478*	*15.7%*
	2013	*1,361,934*	*$100.49*	*962.4*	*581.5*	*$34,700.4*	*$20,993.8*	*$50,288.6*	*$120,378*	*15.8%*
	2014	*1,411,683*	*$92.02*	*969.2*	*592.3*	*$35,963.7*	*$21,613.9*	*$52,057.9*	*$123,785*	*15.8%*
	2015	*1,395,482*	*$48.39*	*976.4*	*601.4*	*$36,768.4*	*$22,969.2*	*$53,577.9*	*$124,598*	*15.9%*

In spite of this wider spread in growth outcomes for the demographic and more economic variables for the Tulsa MSA, the series themselves are strongly correlated for the Tulsa MSA as well. The correlations between the Tulsa MSA socioeconomic variables and oil prices are also positive and relatively strong, but those between the socioeconomic factors for the Tulsa MSA and annual O&D enplanements at TUL are also positive but less strong.

Table D-26 Correlations among Socioeconomic Variables, Tulsa, OK MSA (1990 – 2010)

	Population	Emp	Total Earn	Wages & Salaries	GRP	Avg HH Inc	% > $100K	DB1B Enp	Real Oil Price
Population	1								
Employment	0.936	1							
Total Earn	0.932	0.940	1						
Wages & Salaries	0.947	0.975	0.983	1					
GRP	0.954	0.951	0.992	0.988	1				
Avg HH Inc	0.916	0.928	0.991	0.972	0.990	1			
% > $100K	0.919	0.976	0.949	0.973	0.951	0.928	1		
DB1B Enplanements	0.141	0.383	0.115	0.208	0.137	0.127	0.330	1	
Real Oil Price	0.687	0.630	0.755	0.714	0.783	0.826	0.605	-0.107	1

The Tulsa, Oklahoma MSA saw only modest change in its population's age distribution over the case study period. Those 19 years old or younger and those aged between 20 and 34 combined to make up 59 percent of Tulsa's population in 1985, a total which decreased slightly to around 48 percent by 2010, as shown in Figure D-24.

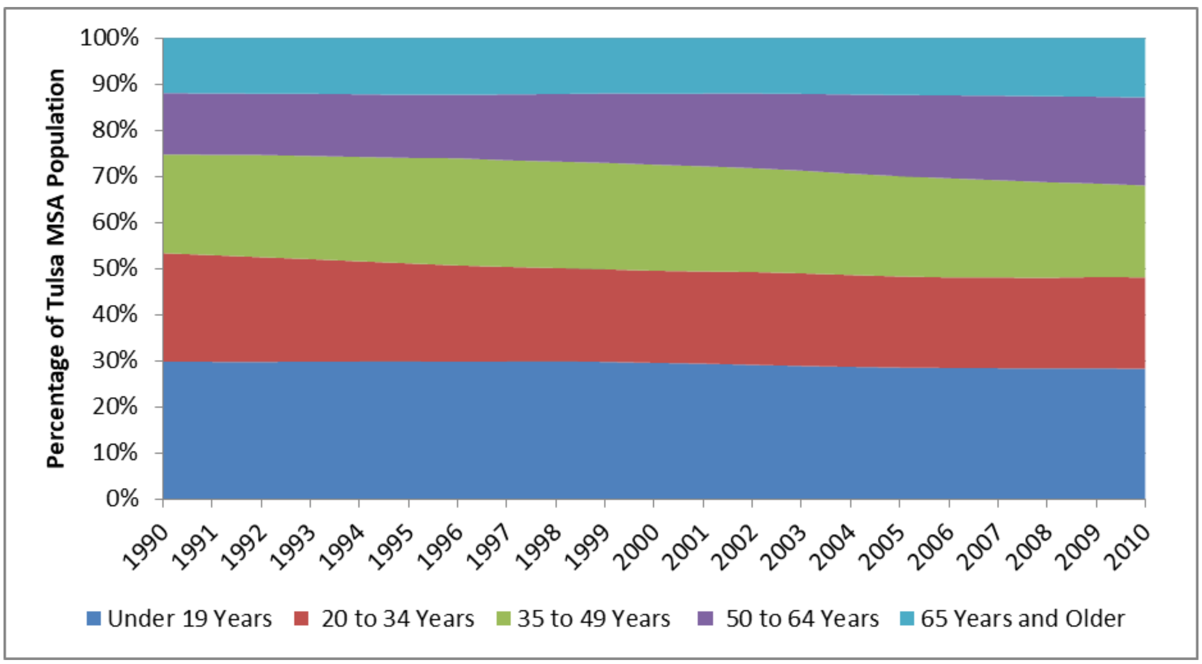

Figure D-24 Population Age Distribution, Tulsa, OK MSA, 1990 to 2010

Figure D-25 shows the distribution of Tulsa's adult population, those 20 and older. It may be worthwhile to exclude the youngest cohort, since few air travel decisions would be made by residents younger than 20. During the case study period Tulsa's adult population distribution changed, with those between 20 and 34 years of age declining from around 34 percent of the adult population to around 28 percent in 2010.

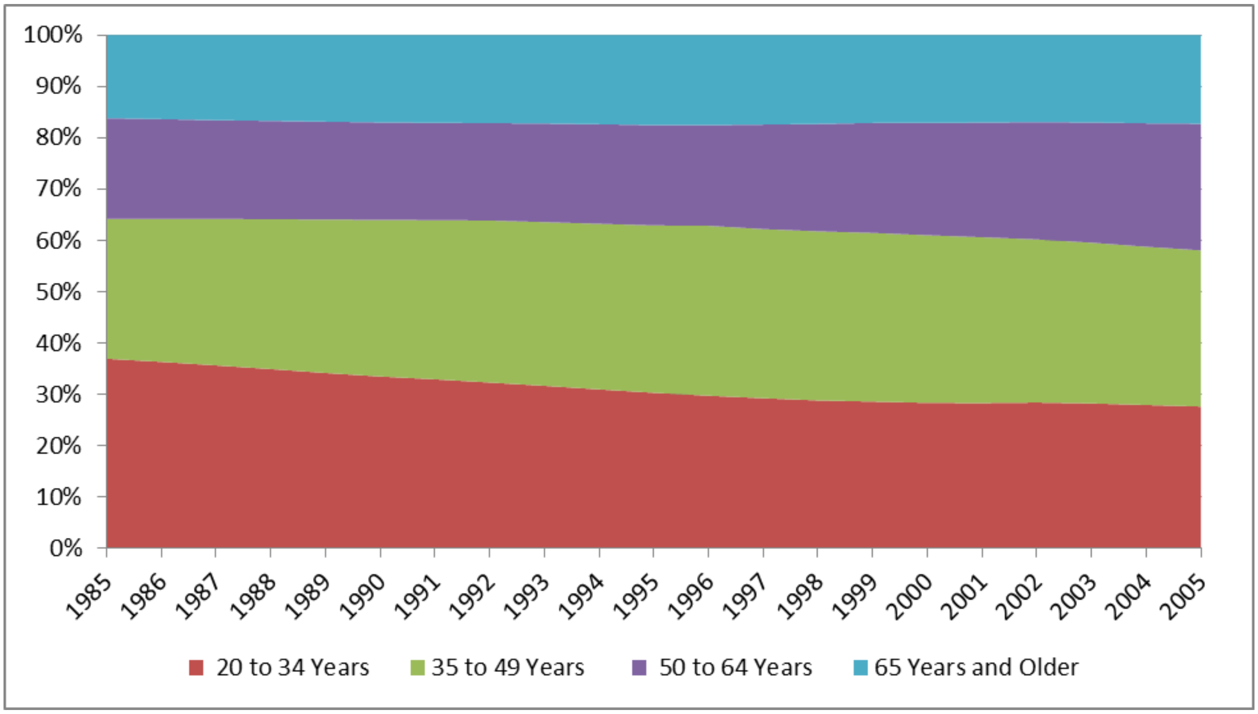

Figure D-25 Adult Population Age Distribution, Tulsa, OK MSA, 1990 to 2010

Figure D-26 shows household income distributions for the residents of the Tulsa MSA over the case study sample period. The figure shows a gradual fall in the proportion of Tulsa households with incomes of $60,000 or less over the sample period, from nearly 70 percent in 1990, when the income distribution data is first available, to around 60 percent by 2010. As reported in the analysis of aggregate socioeconomic trends over the case study sample period for Tulsa (and other locations), average household incomes also increased over the period.

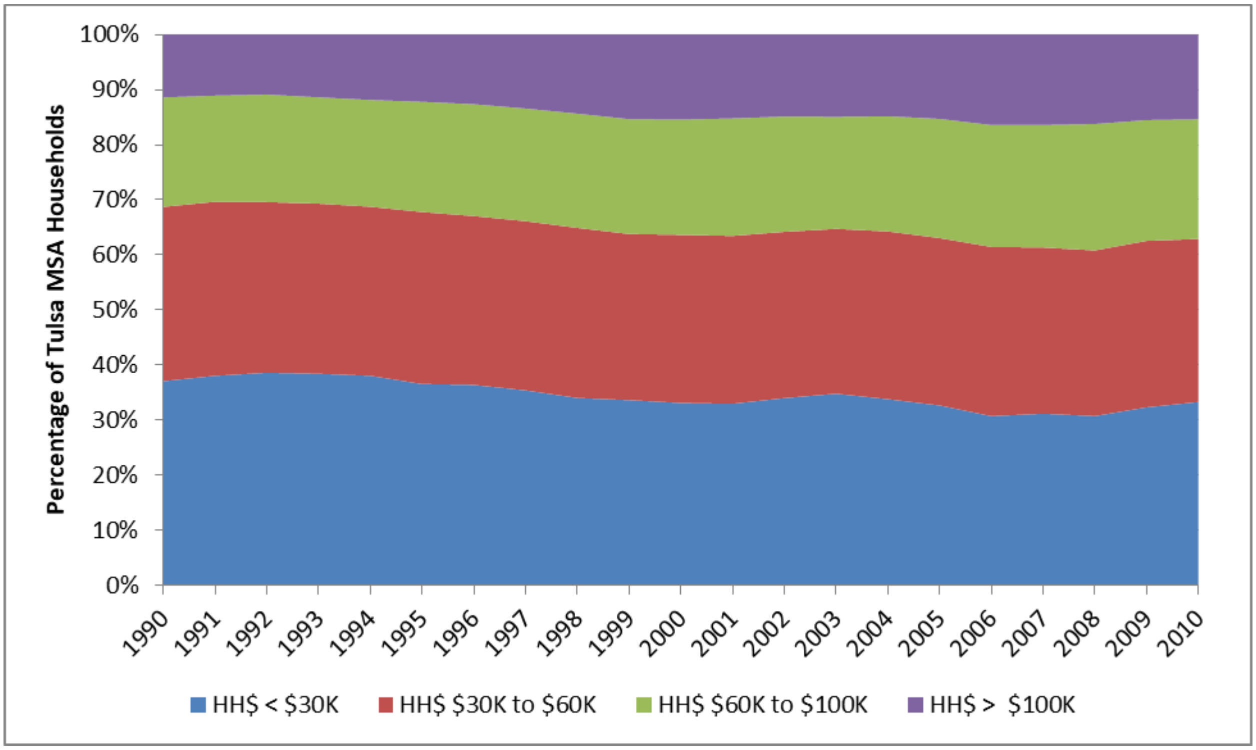

Figure D-26 Household Income Distribution, Tulsa, OK MSA, 1990 to 2010

Figure D-27 reports specifically the changes in the percentage of Tulsa MSA households with annual incomes of $100,000 (in 2009 dollars), a group shown as the uppermost band in Figure D-26 above. This percentage rose steadily from 1992 to 1999, from about 11 percent of Tulsa MSA households to over 15 percent by 1999. The percentage then flattened out through 2010, with periods of growth and decline.

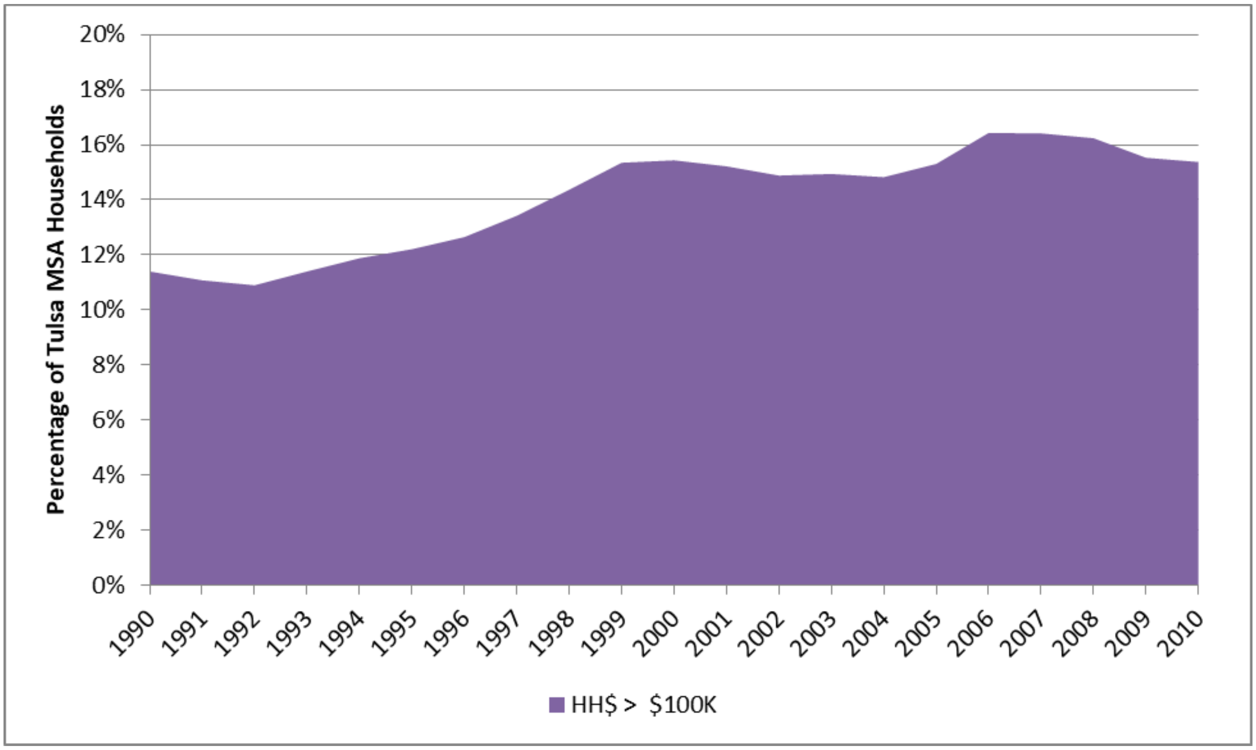

Figure D-27 Percentage of Households with Incomes Exceeding $100,000, Tulsa, OK, 1990 to 2010

Characteristics of the results for the six baseline regressions modeling annual passenger enplanements at Tulsa International Airport (TUL) from 1990 to 2010 are reported in Table D-27. These characteristics include:

- The model estimates based on the individual regional aggregate socioeconomic variables had very poor results for goodness of fit to the actual TUL annual enplanements series, with adjusted R-squared statistics ranging from 0.043 to 0.254, indicating that the baseline models regressing annual enplanements on the annual real oil price and one of the six aggregate socioeconomic variables for the Tulsa MSA accounted for no more than a quarter of the observed variation in the annual TUL enplanements series. In spite of this general characteristic, many of the individual coefficient estimates were statistically significant.
- The estimated regression constant terms were similar across the baseline models, a difference with the baseline estimation profiles of other case study airports. Across the six baseline regressions, each using an individual aggregate socioeconomic variable, the constant term estimates for the TUL annual enplanements series ranged between 10.18 and 12.41, and all of these estimates were significant at the 0.005 significance level, or better.
- While the coefficients estimated in the baseline TUL models for the oil price variable, which in a log linear regression represents the elasticity of changes in annual TUL enplanements with respect to changes in the oil price, were relatively small in all baseline regressions, the estimated coefficients in each equation were statistically significant at the

0.05 significance level (with the exception of the regression using the Tulsa MSA population as the aggregate socioeconomic variable, which was significant at the 0.1 significance level. The coefficient estimates for the six equations ranged from -0.80 to -0.123, the tightest range of values among all the case study analyses.

- The influence of the individual aggregate socioeconomic variables on annual enplanements at TUL is positive and statistically significant across all of the baseline models. The range of coefficient (elasticity) estimates exhibits the same pattern as that described for the baseline regressions covering the Tulsa MSA, with the largest elasticity estimate (0.63, significant at the 0.1 level) occurring in the regression specification using Tulsa MSA population, a similar value for the regression estimate using Tulsa MSA total employment (0.60, also significant at the 0.05 level), and then estimates between 0.21 and 0.32 for Tulsa MSA Total Earnings, Wages and Salaries, and GRP (all also significant at the 0.05 level). The elasticity estimate for Average Real Household Income in the Tulsa MSA was slightly greater than the estimates for these three, estimated at 0.40, and is also significant at the 0.05 level.

The first comparison of the effects of using disaggregated socioeconomic variables to model annual O&D enplanements at TUL is to replace the individual aggregated socioeconomic variables used individually in the baseline regressions with a disaggregated socioeconomic variable – the percentage in each year of Tulsa MSA households with real household incomes (in 2009 dollars) exceeding $100,000. This estimation resulted in a model constant value of 15.31, which was statistically significant at the 0.001 level of significance. As with the other case study airports and airport systems, this value is larger than the constant values estimated in the six baseline regressions. The estimated coefficient on the Oil Price variable, -0.090, is of similar magnitude to the Oil Price coefficients estimated in the baseline equations, and like most of those, is statistically significant at the 0.05 significance level.. The coefficient estimated for the disaggregated socioeconomic variable is also comparable to those estimated for each of the individual aggregate variables, and like them is statistically significant, at the 0.05 level of significance. The goodness of fit of the overall model to the actual data is also low, with an adjusted R-squared statistic of 0.243, a value comparable to the higher adjusted R-squared scores reported for the baseline regressions.

Table D-27 Baseline and Other Regression Results for Models of Annual Enplanements, Tulsa International, 1990 to 2010

Ind. Variable	Statistic	Regression						
		1	2	3	4	5	6	7
Constant	Coefficient	10.28	10.83	12.41	11.47	11.78	10.15	15.31
	t-Statistic	3.863 ***	9.025 ***	11.507 ***	9.058 ***	9.465 ***	5.400 ***	41.074 ***
Oil Price	Coefficient	-0.08	-0.10	-0.09	-0.10	-0.11	-0.12	-0.09
	t-Statistic	-1.568	-2.482 *	-1.783 *	-2.118 *	-2.037 *	-2.214 *	-2.312 *
Pop	Coefficient	0.63						
	t-Statistic	1.515						
Emp	Coefficient		0.60					
	t-Statistic		2.909 **					
Total Earnings	Coefficient			0.21				
	t-Statistic			1.775 *				
Wages and Salaries	Coefficient				0.32			
	t-Statistic				2.254 *			
Gross Regional Product	Coefficient					0.27		
	t-Statistic					2.043 *		
Average HH Income	Coefficient						0.40	
	t-Statistic						2.217 *	
Disaggregated SE Variable: % HH with Income > $100K	Coefficient							0.39
	t-Statistic							2.815 **
	R Squared	0.139	0.329	0.167	0.233	0.201	0.220	0.318
	Adj R Squared	0.043	0.254	0.074	0.148	0.112	0.133	0.243

* Coefficient estimate significant at 0.05 level ** at 0.01 level *** at 0.001 level

The second comparison of the contribution to model performance of the disaggregated socioeconomic variable – the percentage of Tulsa households with relatively high incomes – is made by adding the disaggregated variable as an additional independent regressor added to the baseline regressions. Results from these regressions are reported for all the case study airports in Table D-28. For TUL and the Tulsa MSA, there is remarkable similarity among the six regression specifications that include the disaggregated socioeconomic variable.

- For each of the model specifications, there is an improvement in the model goodness of fit to the actual data, compared to the baseline equations, with the adjusted R-squared scores rising to values between 0.210 and 0.466.
- The regression constant estimates are all positive and statistically significant at the 0.1 significance level, and usually better. In magnitude they are usually about double the estimates in the baseline equations.
- Although the Oil Price coefficient estimates were significant in each of the baseline regressions, this is true for only one of the equations (using Tulsa MSA Employment as the aggregate socioeconomic variable) once the disaggregated socioeconomic variable has been added to the specification. In the remaining equations, the estimates are not statistically significant and are positive in two cases.
- The estimated coefficients on the aggregate socioeconomic variables used in each equation are negative in five of the specifications, and the positive estimate, on Tulsa MSA Employment, is not statistically significant. Three of the negative estimates are statistically significant at the 0.05 significance level.
- Finally, for five of the six specifications the coefficient estimates for the disaggregated socioeconomic variable introduced into these regressions are positive as expected and statistically significant at the 0.05 level of significance or better. In the model estimated with the Tulsa MSA Employment variable as the aggregate socioeconomic variable, the coefficient estimated for the disaggregated socioeconomic variable is positive but not statistically significant. The results suggest that the model estimates "split" the socioeconomic influences on TUL O&D passenger enplanements between the aggregate socioeconomic variable and the disaggregated variable.

For the TUL and Tulsa MSA data, the addition of the disaggregated socioeconomic variable does improve overall model fit (from the poor fits indicated by the low R-squared results from the baseline equations) a great deal. While the Oil Price variable and the individual aggregate socioeconomic variables are consistently significant in the baseline regressions, and replacing the aggregate socioeconomic variables with the disaggregated one gives similar results, adding the disaggregated socioeconomic variable as a third independent regressor disrupt this consistent outcome while improving overall model fit.

Table D-28 Regression Results for Models of Annual Enplanements, Tulsa International Airport, 1990 to 2010, that Include both Aggregate and Disaggregated Socioeconomic Variables

Ind. Variable	Statistic	Regression					
		1	2	3	4	5	6
Constant	Coefficient	26.10	11.31	25.57	24.50	25.67	24.59
	t-Statistic	1.694	1.694	6.917 ***	3.947 ***	5.007 ***	3.313 **
Oil Price	Coefficient	-0.041	-0.099	0.006	-0.036	0.022	-0.003
	t-Statistic	-0.946	-2.329 *	0.119	-0.690	0.332	-0.036
Pop	Coefficient	-1.481					
	t-Statistic	-1.938 *					
Emp	Coefficient		0.539				
	t-Statistic		0.600				
Total Earnings	Coefficient			-0.859			
	t-Statistic			-2.786 **			
Wages and Salaries	Coefficient				-0.813		
	t-Statistic				-1.483		
Gross Regional Product	Coefficient					-0.861	
	t-Statistic					-2.026 *	
Average HH Income	Coefficient						-0.746
	t-Statistic						-1.252
Disaggregated SE Variable: % HH with Income > $100K	Coefficient	0.879	0.043	1.411	1.182	1.305	0.966
	t-Statistic	3.048 **	0.073	3.654 ***	2.136 *	2.770 **	2.001 *
	R Squared	0.458	0.329	0.546	0.411	0.468	0.396
	Adj R Squared	0.362	0.210	0.466	0.308	0.374	0.290
	* Coefficient estimate significant at 0.05 level ** at 0.01 level *** at 0.001 level						

To assess the forecasting performance of these models, we limit the analysis to consideration of a single pair of regressions, those that use the Tulsa MSA Gross Regional Product as the aggregate socioeconomic variable. As mentioned above, it is not possible to obtain the Woods and Poole projections that would have been available for the years 2011 through 2015, which would have been included as part of the Woods and Poole 2012 data release. To create a test projection of values for Tulsa MSA GRP for those years, we use the percentage annual growth terms from the Woods and Poole 2016 data release for the years 2015 through 2019. With those percentage values for annual GRP growth, notional projected values for Tulsa GRP can be calculated for 2011 through 2015. To create notional projected values for the disaggregated socioeconomic variable, the percentage of Tulsa MSA households with 2009 dollar denominated income exceeding $100,000 (in 2009 dollars), a similar procedure was used.

Forecast values for the Oil Price dependent variable are also needed to calculate out of sample projections. These were obtained for the years 2011 to 2015 from the data reported in the FAA's 2012 Annual Aerospace Forecast document. These national values are used for each of the case study airports out of sample forecast exercises.

Four out of sample model forecasts were calculated for the baseline model using Tulsa MSA GRP and the model that also includes the disaggregated socioeconomic variable. The first three are based on the out of sample forecasts described above. For the first forecast, the projected values – based on the percentage increases projected for the years 2015 through 2019 – for GRP, and the percentage of Tulsa household incomes exceeding $100,000 in the case of the model using the disaggregated socioeconomic variable as input, are used as described above. For the second, the values for percentage annual increases over the projection period (2011 to 2015) are increased by 50 percent, indicating a more rapidly growing Tulsa economy. For the third TUL projection, the values for percentage annual increases are reduced by 50 percent, representing a projection for a more slowly growing economic environment in the Tulsa MSA. For the fourth projection, the actual values observed in 2011 through 2015 for GRP and the share of Tulsa MSA households with incomes exceeding $100,000 are used to drive the "forecast" of TUL annual O&D enplanements.

Using the projections published by Woods and Poole for 2015 and subsequent years for simulating a range of forecast scenarios based on inputs that would have been available to analysts in 2011 or 2012, we can compare the forecasting performance of the baseline regressions for TUL annual O&D enplanements using the annual oil price variable and the Tulsa MSA GRP, with those of the regression parameters based additionally on input from the disaggregated socioeconomic variable. The independent variable data and projections are shown in Table D-29 below. The projections for the Tulsa MSA variables reported in the Woods and Poole 2016 data have annual GRP growing at a rate of around 2.1 percent over the 2011 to 2015 forecast period, and the percentage of Tulsa MSA households with incomes exceeding $100,000 growing at about 1.8 percent initially, growing to 2 percent after three years. The annual percentage changes used to calculate the Oil Price projected values for 2011 through 2015 are those implied by the Oil Price projections reported in the 2012 FAA Aerospace Forecast, which relies on projected values for the years following 2010. Thus we have four "forecast" scenarios for the two regression specifications for annual TUL O&D enplanements:

- Projected values for the socioeconomic independent variables based on the annual percentage changes in the variables, as reported by Woods and Poole for the five years beginning in 2015,
- Projected values based on annual percentage changes that are 1.5 times these values,
- Projected values based on annual percentage changes that are half these values, and]
- Projected values that equal the actual values for the independent variables that occurred in the years 2011 through 2015.

Table D-29 Inputs to Out-of-Sample TUL O&D Enplanement Projections

TUL Inputs to Out-of-Sample Enplanement Forecasts	TUL					
	Sources	2011	2012	2013	2014	2015
Oil Price Forecast	*2012 FAA Aero Forecast*	$84.32	$92.71	$97.16	$100.65	$104.64
GRP Forecast 1 (Medium)	*W&P 2016 %*	$46,402.6	$47,364.2	$48,325.4	$49,290.9	$50,262.7
GRP Forecast 2 (High)	*1.5 x F'cast 1*	$46,888.1	$48,345.6	$49,817.2	$51,310.2	$52,827.6
GRP Forecast 3 (Low)	*0.5 x F'cast 1*	$45,917.1	$46,392.8	$46,863.6	$47,331.7	$47,798.3
HHInc% Forecast 1 (Medium)	*W&P 2016 %*	15.6%	15.9%	16.2%	16.5%	16.9%
HHInc% Forecast 2 (High)	*1.5 x F'cast 1*	15.8%	16.2%	16.7%	17.1%	17.6%
HHInc% Forecast 3 (Low)	*0.5 x F'cast 1*	15.5%	15.7%	15.8%	15.9%	16.1%
	2010					
Actual Oil Price	$83.99	$84.32	$92.71	$97.16	$100.65	$104.64
Actual GRP	$45,431.5	$47,252.8	$49,707.2	$50,288.6	$52,057.9	$53,577.9
Actual HHInc%	15.4%	15.4%	15.7%	15.8%	15.8%	15.9%

The forecasting or projection results for the baseline regression equation for annual enplanements at TUL are shown in Figure D-28 below.

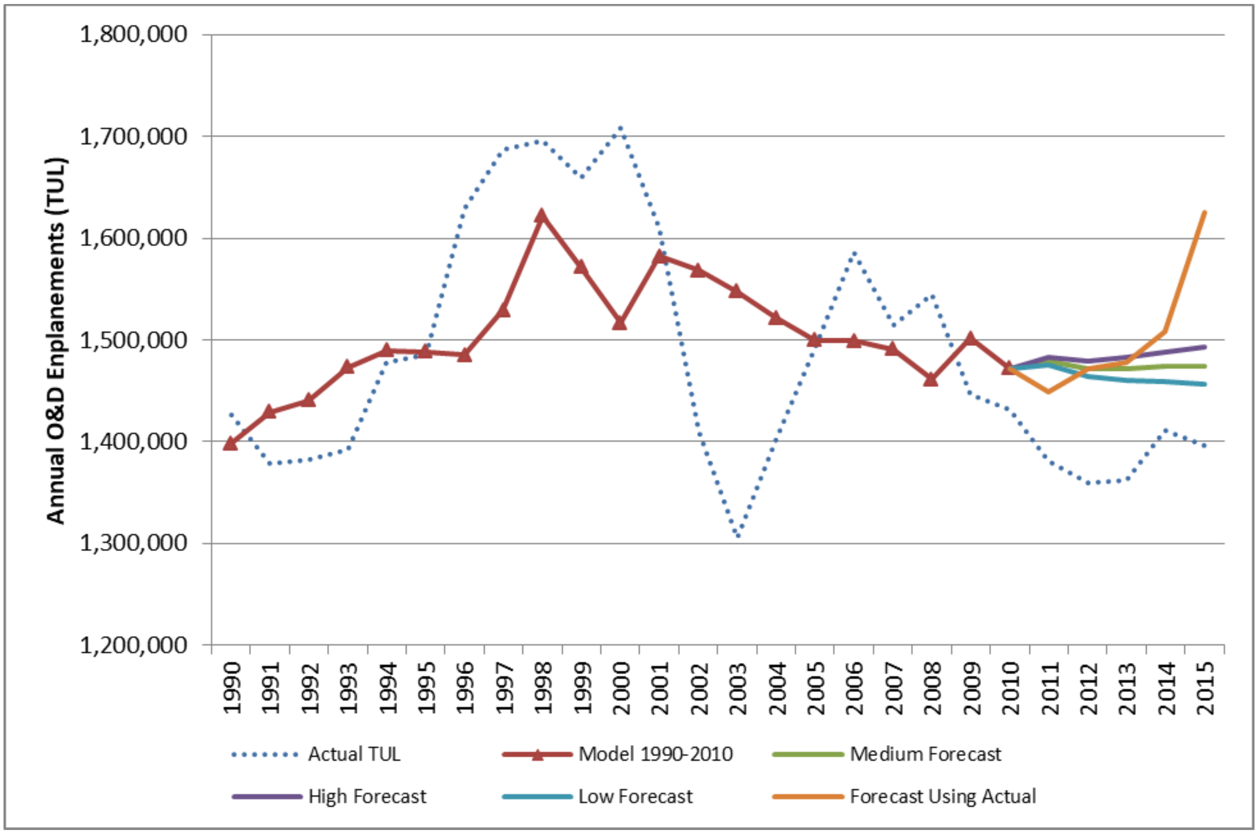

Figure D-28 Baseline Model Forecasts of TUL Annual Enplanements, 2011 to 2015, Using Tulsa MSA Gross Regional Product, 1990 to 2010

The results for the regression equation using the disaggregated socioeconomic variable along with Tulsa MSA GRP are shown in Figure D-29 below.

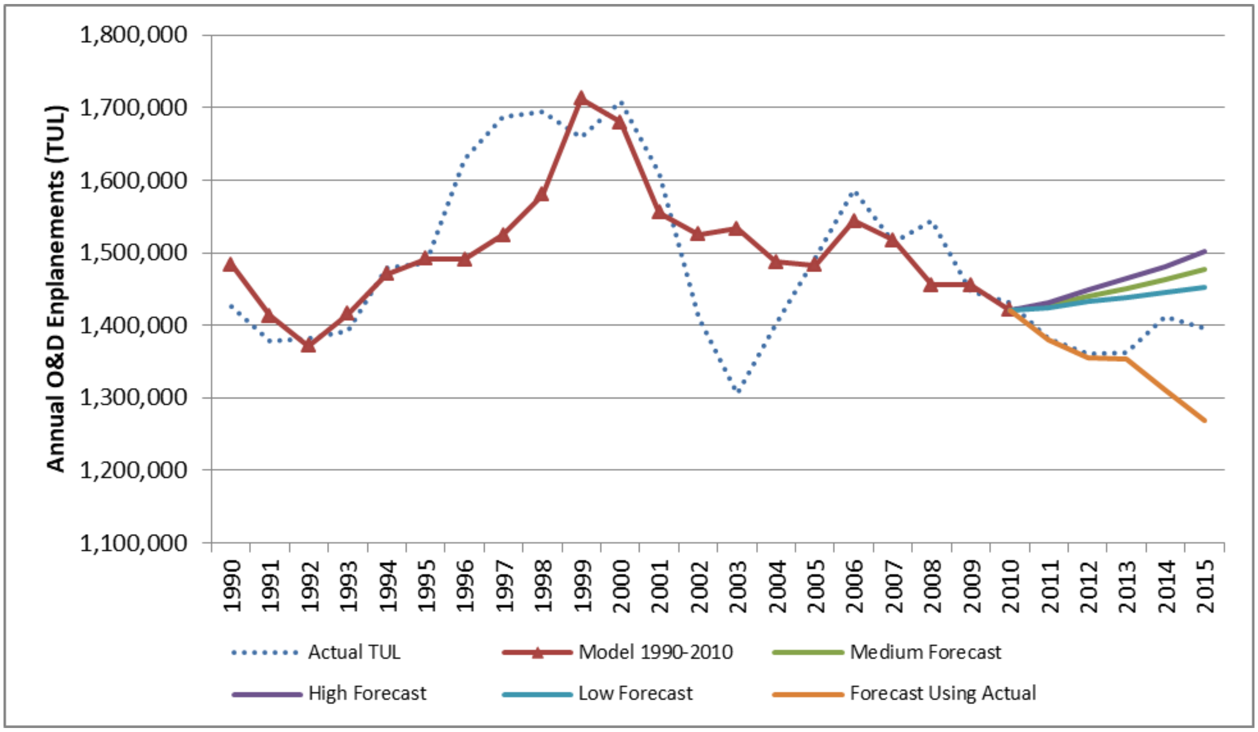

Figure D-29 Disaggregated SE Variable Model Forecasts of TUL Annual Enplanements, 2011 to 2015, Using Tulsa MSA GRP and Share of Tulsa Households with Incomes exceeding $100,000, 1990 to 2010

Enplanements at TUL declined over the 2011 to 2015 forecast period, which is at odds with nearly all the forecast equations. For the baseline specifications, shown in Figure D-28, the forecasts using projected values for the independent variables were relatively flat, and the forecast scenario using the actual independent values over the forecast period shows a forecast of strongly increased TUL enplanements, resulting from the stronger growth in actual Tulsa MSA GRP and the sharp fall in oil prices between 2014 and 2015.[8] The forecasts from the model using the Tulsa MSA disaggregated socioeconomic variable along with Tulsa GRP projects some increase in TUL enplanements over the forecast period, as can be seen in Figure D-29. When the actual values for the independent variables are used in this alternative model, the forecast enplanements are very accurate through 2013 before sharply underforecasting TUL enplanements for 2014 and 2015. This is a consequence of the fall in oil prices over these years combined with the positive coefficient estimate on the oil price variable in the alternative model.

The following tables D-30 to D-32 quantify this, reporting the percentage error for each of the forecasts and the Root Mean Squared Error (RMSE) for each forecast, a calculation which can be used to compare the accuracy of the baseline model forecasts and the forecasts calculated from the model also using the disaggregated socioeconomic variable.

[8] The baseline regression equation is ln(TUL_Enp) = 11.78 – 0.109ln(OilPrice) + 0.27ln(GRP),), with an adjusted R squared of 0.112, while the estimated equation that includes the disaggregated socioeconomic variable is ln(TUL_Enp) = 25.67 + 0.022ln(OilPrice) - 0.86ln(GRP) + 1.30ln(%HHInc>$100K) with an adjusted R squared of 0.374.

Table D-30 Percentage Forecast Errors, Baseline Model Projections of TUL Annual O&D Enplanements, 2011 to 2015

TUL using GRP	Percentage Forecast Error (Baseline Regression with Oil Price and GRP)			
	Medium	High	Low	Actual
2010	--	--	--	--
2011	7.1%	7.4%	6.8%	4.9%
2012	8.2%	8.8%	7.7%	8.2%
2013	8.1%	8.9%	7.3%	8.5%
2014	4.4%	5.4%	3.3%	6.8%
2015	5.7%	7.0%	4.4%	16.4%

The positive percentage errors in the tables reflect the fact that the model-based projections of TUL enplanements consistently overestimated the number of annual passenger, especially for the baseline "forecast" derived from use of the values that actually occurred for the Oil Price and Tulsa MSA socioeconomic variables. However, in the case of the alternative model forecast using actual oil price and socioeconomic variable values over the forecast period, the forecasts undershoot the actual TUL enplanements, especially in the last two years of the forecast period.

Table D-31 Percentage Forecast Errors, Disaggregated SE Model Projections of TUL Annual O&D Enplanements, 2011 to 2015

TUL using GRP	Percentage Forecast Error (Specification including Disaggregated SE Variable)			
	Medium	High	Low	Actual
2010	--	--	--	--
2011	3.4%	3.7%	3.1%	-0.1%
2012	5.9%	6.5%	5.3%	-0.3%
2013	6.6%	7.5%	5.6%	-0.6%
2014	3.6%	4.9%	2.4%	-7.2%
2015	5.9%	7.7%	4.1%	-9.1%

The final table for the TUL case study results reports the Root Mean Squared Error (RMSE) of the different forecasts for the two model specifications. This statistic can be used to compare the accuracy of two or more of these forecasts, with a smaller RMSE value indicating a more accurate forecast. Although forecasts from the alternative model using both the aggregate and the disaggregated socioeconomic variables have lower RMSE values than the forecasts from the baseline model, the prior figures and tables indicate that neither model provides an accurate forecast of annual enplanements at TUL over the forecast period.

Table D-32 Root Mean Squared Error (RMSE) for TUL Baseline and Disaggregated SE Variable Model Forecasts of Annual TUL O&D Enplanements, 2011 to 2015

TUL (using GRP)	Root Mean Squared Error		% Difference BL v Alt
	Baseline	Alternative	
Medium Forecast	94,079	72,038	23%
High Forecast	104,440	86,068	18%
Low Forecast	84,182	58,826	30%
Actual Values	135,787	72,748	46%

f. Providence T. E. Green Airport (PVD)

Indexed annual socioeconomic data from 1990 to 2010 are shown for the Providence, Rhode Island MSA in Figure D-30. This region is served by a medium hub, T.F. Green Airport (PVD). The data for the Providence MSA exhibit similar growth patterns to those occurring in the other case study MSAs. Again, demographic variables grew more slowly than did more economic variables, reflecting the contributions of productivity growth to economic variables along with the effects of growth in population and employment levels. For the Providence area, for example, gross regional product, indexed to 100 for 2000, grew to an indexed level of over 117 by 2010, having begun the sample period at an indexed level of around 78, rising steadily, and then declining in the wake of the Financial Crisis. In contrast, population growth in the Providence MSA has been modest, starting from an indexed level of 95 in 1990, reaching the index baseline level of 100 in 2000, and only reaching an indexed value less than 105 by 2010.

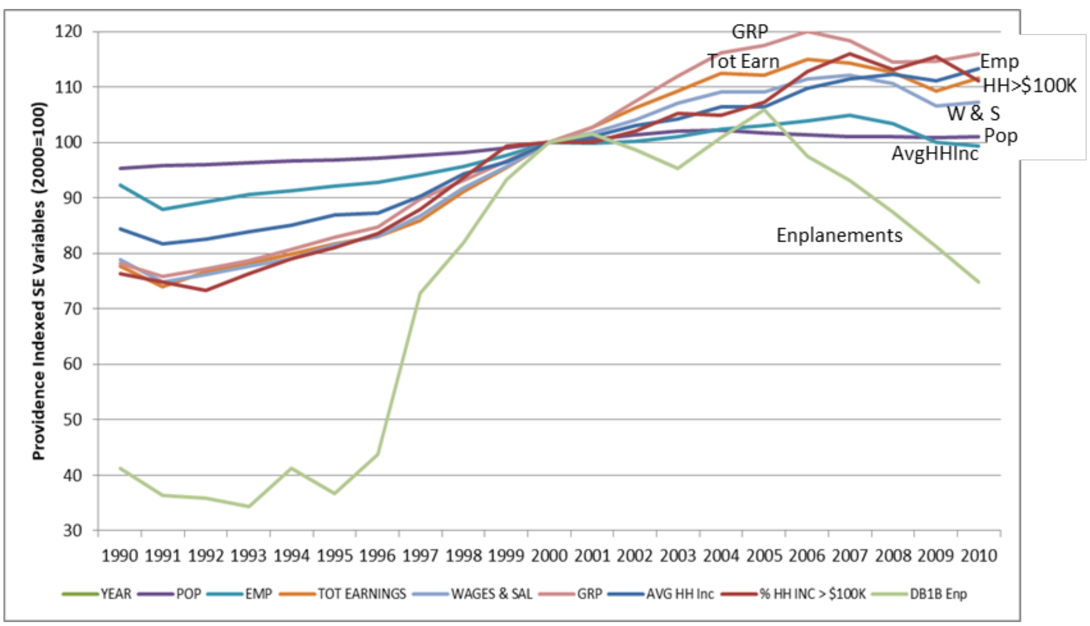

Figure D-30 Providence, RI MSA Indexed Socioeconomic Variables, 1990 to 2010, 2000=100

Table D-33 reports the reported value of the socioeconomic variables for the Providence MSA in five year intervals from 1990 to 2010.

Table D-33 Values of Socioeconomic Variables, Providence, RI MSA (1990-2010)

	Year	T. F. Green Airport Providence (PVD)								
		Annual O&D Enplanements	Real Oil Price	Regional Population (000s)	Regional Employment (000s)	Regional Total Earnings ($M)	Regional W&S Earnings ($M)	Regional Product (GRP) ($M)	Regional Avg HH Income	% Regional HHs with Income > $100K
	1990	1,148,711	$40.16	1,513.2	786.5	$32,180.5	$24,170.6	$49,758.8	$82,772	14.7%
	1991	1,009,743	$33.20	1,519.2	749.1	$30,690.8	$22,910.5	$48,290.6	$80,133	14.8%
	1992	996,377	$31.33	1,522.4	760.3	$31,783.6	$23,353.0	$49,103.1	$81,044	14.8%
	1993	954,062	$26.96	1,528.1	772.2	$32,416.7	$23,801.1	$50,111.4	$82,369	15.8%
	1994	1,144,668	$25.31	1,532.1	778.5	$33,101.6	$24,221.5	$51,379.7	$83,497	16.6%
	1995	1,022,280	$27.16	1,535.8	784.8	$33,876.1	$24,985.2	$52,763.5	$85,283	17.2%
	1996	1,218,725	$31.73	1,541.7	791.0	$34,422.6	$25,432.0	$53,893.1	$85,724	18.0%
	1997	2,027,004	$27.80	1,548.9	802.6	$35,624.9	$26,583.1	$57,069.5	$88,554	19.1%
	1998	2,276,180	$17.87	1,558.4	815.5	$37,788.8	$28,107.0	$59,257.5	$92,546	20.5%
	1999	2,596,493	$25.03	1,571.6	831.9	$39,591.9	$29,307.1	$61,487.1	$94,750	21.8%
	2000	2,780,558	$38.86	1,586.1	852.1	$41,459.5	$30,626.3	$63,613.4	$98,149	22.1%
	2001	2,823,261	$29.99	1,595.4	851.1	$42,557.6	$31,129.3	$65,320.9	$99,181	21.4%
	2002	2,742,317	$31.83	1,608.7	853.9	$44,020.1	$31,871.5	$68,285.2	$101,124	21.3%
	2003	2,652,426	$36.38	1,617.6	860.7	$45,272.4	$32,792.9	$71,196.1	$102,247	21.7%
	2004	2,803,479	$45.87	1,620.6	872.1	$46,604.2	$33,400.9	$73,888.2	$104,513	21.4%
	2005	2,943,633	$60.44	1,613.4	878.5	$46,518.4	$33,403.1	$74,729.4	$104,368	22.0%
	2006	2,709,914	$70.72	1,607.6	884.9	$47,647.7	$34,116.0	$76,332.0	$107,683	22.9%
	2007	2,587,998	$78.22	1,602.6	893.5	$47,423.6	$34,317.1	$75,290.0	$109,430	22.8%
	2008	2,429,732	$103.82	1,601.5	881.7	$46,695.5	$33,868.2	$72,786.6	$110,166	21.5%
	2009	2,259,474	$66.42	1,601.0	852.1	$45,336.0	$32,627.9	$72,939.4	$108,982	21.3%
	2010	2,079,410	$83.99	1,602.2	846.2	$46,290.6	$32,845.4	$73,810.5	$111,245	20.8%
Used to assess out-of-sample forecasts	*2011*	*2,026,427*	*$101.87*	*1,601.1*	*855.7*	*$46,491.5*	*$33,078.9*	*$73,604.1*	*$110,851*	*19.7%*
	2012	*1,916,271*	*$100.93*	*1,603.4*	*860.1*	*$47,314.1*	*$33,460.8*	*$74,491.0*	*$112,508*	*19.6%*
	2013	*2,000,796*	*$100.49*	*1,605.5*	*876.1*	*$47,774.1*	*$34,044.4*	*$74,943.6*	*$110,670*	*19.6%*
	2014	*1,874,954*	*$92.02*	*1,609.4*	*891.7*	*$49,176.6*	*$35,195.6*	*$77,186.9*	*$113,986*	*19.6%*
	2015	*1,848,714*	*$48.39*	*1,615.5*	*903.8*	*$50,094.5*	*$36,189.9*	*$78,644.6*	*$115,020*	*19.8%*

As with the other case study regions, the socioeconomic measures for the Providence MSA were strongly correlated in their growth patterns over the sample period. Table D-34 reports the correlations for Providence, with nearly all of the socioeconomic variables positively correlated with others at levels of 0.95 or greater. The correlations between the Providence MSA socioeconomic variables and annual O&D enplanements at PVD are also positive and nearly as strong, with slightly weaker correlations between Providence MSA socioeconomic variables and real oil prices.

Table D-34 Correlations among Socioeconomic Variables, Providence, RI MSA (1990 – 2010)

	Population	Emp	Total Earn	Wages & Salaries	GRP	Avg HH Inc	% > $100K	DB1B Enp	Real Oil Price
Population	1								
Employment	0.950	1							
Total Earn	0.979	0.976	1						
Wages & Salaries	0.978	0.984	0.998	1					
GRP	0.970	0.966	0.996	0.992	1				
Avg HH Inc	0.942	0.947	0.985	0.981	0.984	1			
% > $100K	0.940	0.956	0.978	0.981	0.978	0.989	1		
DB1B Enplanements	0.912	0.913	0.881	0.903	0.863	0.828	0.861	1	
Real Oil Price	0.559	0.650	0.692	0.675	0.705	0.763	0.712	0.364	1

The Providence, Rhode Island MSA saw only modest change in its population's age distribution over the case study period. Those 19 years old or younger made up about 26 percent of the population over the case study period, and there was a slight reduction in the share of those aged between 20 and 34, from 25 percent to around 20 percent. There was a modest increase in the share of those aged between 50 and 64, from about 13 percent to around 20 percent by 2010, as shown in Figure D-31.

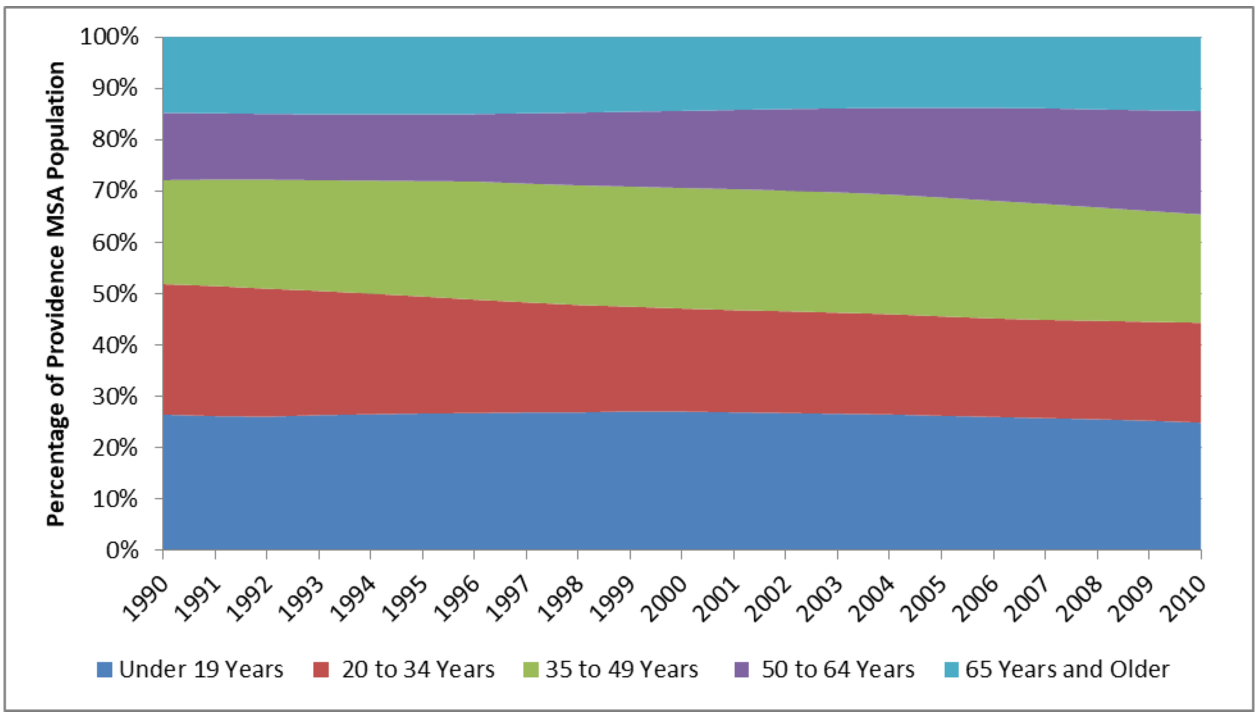

Figure D-31 Population Age Distribution, Providence, RI MSA, 1990 to 2010

Figure D-32 shows the distribution of Providence's adult population, those 20 and older. During the case study period Providence's adult population distribution changed, with those between 20 and 34 years of age declining from around 35 percent of the adult population to around 26 percent in 2010, although the proportion aged 65 or older was stable at around 20 percent of the Providence adult population..

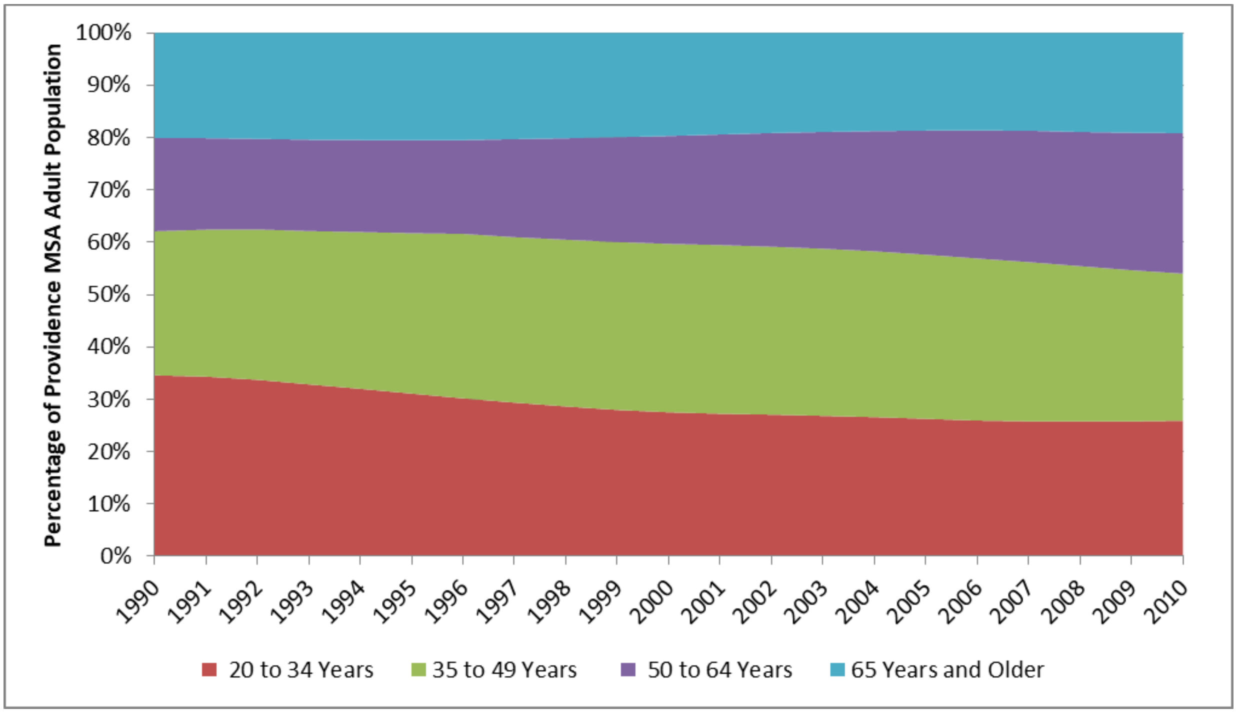

Figure D-32 Adult Population Age Distribution, Providence, RI MSA, 1990 to 2010

Figure D-33 depicts household income distributions for the residents of the Providence MSA over the case study sample period. The figure shows that a steady proportion of just less than 30 percent of Providence households had incomes of $30,000 or less over the sample period, from around 15 percent in 1990.

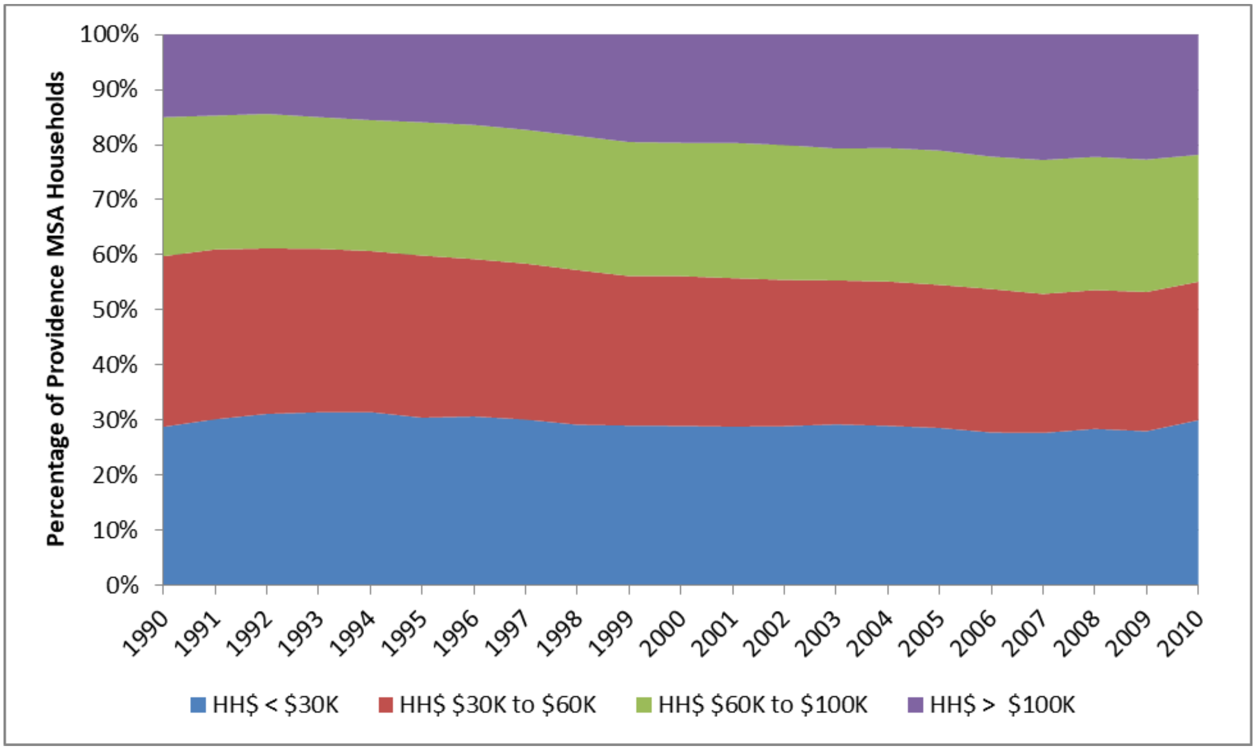

Figure D-33 Household Income Distribution, Providence, RI MSA, 1990 to 2010

Figure D-34 reports specifically the changes in the percentage of Providence MSA households with annual incomes of $100,000 (in 2009 dollars), a group shown as the uppermost band in Figure D-33 above. This percentage has risen from about 15 percent of Providence MSA households to around 22 percent by 2010.

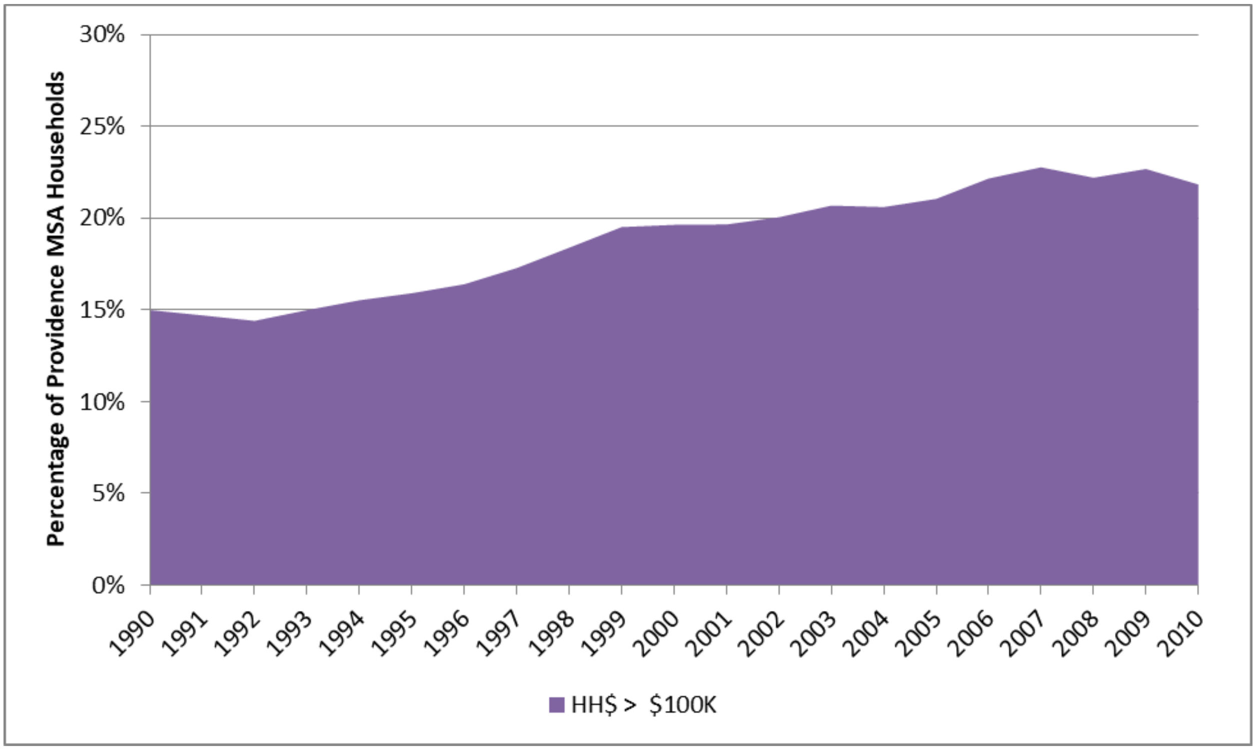

Figure D-34 Percentage of Households with Incomes Exceeding $100,000, Providence, RI 1990 to 2010

Characteristics reported below in Table D-35 of the results for the six baseline regressions modeling annual passenger enplanements at T. F. Green Airport in Providence, RI (PVD) from 1990 to 2010 include:

- The six baseline model estimates of PVD annual enplanements had similar strong goodness of fit to the actual annual enplanements series, with adjusted R-squared statistics ranging from 0.797 to 0.909, indicating that the baseline models regressing annual enplanements on the annual real oil price and one of the six aggregate socioeconomic variables for the Providence MSA accounted for nearly all of the observed variation in the annual PVD enplanements series.
- The estimated regression constant terms varied across the baseline models, although the estimates for the models using Total Earnings in the Providence MSA, Annual Wages and Salaries in the MSA, and Gross Regional Product were comparable in magnitude, ranging from -20.49 to -18.71. All of these estimates were at the 0.005 significance level (or better). The estimated constant term for the model using MSA annual employment was estimated at -46.21, and was also statistically significant at the 0.005 level. The constant estimate for the log linear model using Providence MSA population took a value of -123.16 and was statistically significant at the 0.005 level, while the baseline model using Average Real Household Income took a negative value of -39.18, and was also statistically significant at the 0.005 significance level.
- The coefficients estimated in the baseline PVD models for the Oil Price variable, which in a log linear regression represents the elasticity of changes in annual PVD

enplanements with respect to changes in the oil price, were, like the PVD estimates, a bit larger in magnitude than those seen for other case study airports (and also negative, as would be expected) in all baseline regressions. The estimated coefficient values ranged between -0.184 (in the log linear regression using Providence MSA Population as the independent aggregate socioeconomic variable) and -0.491 (in the regression using Providence MSA Average Real Household Income), and all were statistically significant at the 0.005 level (except for the estimate in the MSA Population based equation, in which the Oil Price coefficient estimate was significant at the 0.05 level).

- The influence of the individual aggregate socioeconomic variables on annual enplanements at PVD, based on the coefficient estimates, is positive and strongly statistically significant across all of the baseline models. The range of coefficient (elasticity) estimates exhibits a pattern somewhat similar to that described for the baseline regressions covering the Baltimore-Washington CSA, with the largest coefficient estimate (18.79, significant at the 0.005 level) occurring for Providence MSA population, a smaller value for the regression estimate using Providence MSA total employment (9.20, also significant at the 0.005 level), and then estimates between 3.16 and 3.53 for Total Earnings, Wages and Salaries, and GRP (all also significant at the 0.005 level). The coefficient estimate for Average Real Household Income in the Providence MSA was slightly higher than these three, estimated at 4.83, and also significant at the 0.005 level.

The first comparison of the effects of using disaggregated socioeconomic variables to model annual O&D enplanements at PVD is to replace the individual aggregated socioeconomic variables used individually in the baseline regressions with a disaggregated socioeconomic variable – the percentage in each year of Providence MSA households with real household incomes (in 2009 dollars) exceeding $100,000. This estimation resulted in a model constant value of 21.15, which was statistically significant at the 0.001 level of significance. As with the other case study airports and airport systems, this value is considerably larger than the constant values estimated in the six baseline regressions. The estimated coefficient on the Oil Price variable, -0.354, is of similar magnitude to the Oil Price coefficients estimated in the baseline equations, and like most of those, is statistically significant at the 0.005 significance level.. The coefficient estimated for the disaggregated socioeconomic variable is also comparable to those estimated for each of the individual aggregate variables, and like them is statistically significant, at the 0.001 level of significance. The goodness of fit of the overall model to the actual data is also high, with an adjusted R-squared statistic of 0.830, a value comparable to the adjusted R-squared values for the baseline regressions.

Table D-35 Baseline and Other Regression Results for Models of Annual Enplanements, Providence T. F. Greene Airport, 1990 to 2010

Ind. Variable	Statistic	Regression						
		1	2	3	4	5	6	7
Constant	Coefficient	-123.16	-46.21	-18.71	-20.49	-18.95	-39.18	21.15
	t-Statistic	-8.188 ***	-9.434 ***	-6.533 ***	-8.284 ***	-6.201 ***	-7.236 ***	26.287 ***
Oil Price	Coefficient	-0.18	-0.32	-0.37	-0.35	-0.40	-0.49	-0.35
	t-Statistic	-1.842 *	-3.811 ***	-3.939 ***	-4.583 ***	-3.928 ***	-4.201 ***	-3.625 ***
Pop	Coefficient	18.79						
	t-Statistic	9.058 ***						
Emp	Coefficient		9.20					
	t-Statistic		12.113 ***					
Total Earnings	Coefficient			3.26				
	t-Statistic			11.133 ***				
Wages and Salaries	Coefficient				3.53			
	t-Statistic				13.610 ***			
Gross Regional Product	Coefficient					3.16		
	t-Statistic					10.512 ***		
Average HH Income	Coefficient						4.83	
	t-Statistic						9.660 ***	
Disaggregated SE Variable: % HH with Income > $100K	Coefficient							3.21
	t-Statistic							10.572 ***
	R Squared	0.817	0.915	0.886	0.919	0.858	0.828	0.847
	Adj R Squared	0.797	0.905	0.873	0.909	0.843	0.809	0.830

* Coefficient estimate significant at 0.05 level ** at 0.01 level *** at 0.001 level

The second comparison of the contribution to model performance of the disaggregated socioeconomic variable – the percentage of Providence households with relatively high incomes – is made by adding the disaggregated variable as an additional independent regressor added to the baseline regressions. Results from these regressions are reported for all the case study airports in Table D-36 below. For PVD and the Providence MSA, there is remarkable similarity among the six regression specifications that include the disaggregated socioeconomic variable.

- For each, there is a slight improvement in the model goodness of fit to the actual data, with the adjusted R-squared scores rising slightly to values between 0.822 and 0.917.
- In four of the six regressions adding the disaggregated socioeconomic variable to the baseline regression specifications, the regression constant is negative, and of these two are statistically significant estimates at the 0.1 level of significance, and two are not statistically significant. Two of the regression specifications have constant term estimates that are positive, and both of these are not statistically significant. This is in contrast to the constant term estimate in the baseline regressions, all of which were negative and statistically significant. .
- The coefficient estimates for the Oil Price variable are negative in all six regressions, and are similar in magnitude to those estimated in the baseline regressions for PVD. All of the estimates are significant at the 0.01 significance level or better.
- The coefficients on the aggregate socioeconomic variables used in each equation are positive but smaller than the estimates seen in the baseline equations and are, with the exception of the estimated coefficient on the Providence MSA Average Real Household Income, statistically significant at the 0.1 level or better.
- Finally, for four of the six specifications the coefficient estimates for the disaggregated socioeconomic variable introduced into these regressions are positive as expected and statistically significant at the 0.1 level of significance or better. The two specifications using the Providence MSA Total Earnings aggregate socioeconomic variable and the Wages and Salaries aggregate socioeconomic variable have positive coefficient estimates that are not statistically significant.

For the PVD and Providence MSA data, the addition of the disaggregated socioeconomic variable does not improve overall model fit very much. While the Oil Price variable and the individual aggregate socioeconomic variables are consistently significant in the baseline regressions, and replacing the aggregate socioeconomic variables with the disaggregated one gives similar results, adding the disaggregated socioeconomic variable as a third independent regressor disrupt this consistent outcome in most cases.

Table D-36 Regression Results for Models of Annual Enplanements, Providence T. F. Greene Airport, 1990 to 2010, that Include both Aggregate and Disaggregated Socioeconomic Variables

Ind. Variable	Statistic	Regression					
		1	2	3	4	5	6
Constant	Coefficient	-33.10	-20.62	-2.58	-19.11	1.31	16.02
	t-Statistic	5.749 ***	5.749 ***	-0.200	-1.473	0.095	0.483
Oil Price	Coefficient	-0.319	13.282	-0.378	-0.354	-0.389	-0.368
	t-Statistic	-3.324 **	1.826 *	-4.076 ***	-4.405 ***	-3.973 ***	-2.764 **
Pop	Coefficient	7.105					
	t-Statistic	1.622					
Emp	Coefficient		-1.552				
	t-Statistic		3.148 **				
Total Earnings	Coefficient			1.953			
	t-Statistic			1.841 *			
Wages and Salaries	Coefficient				3.411		
	t-Statistic				3.106 **		
Gross Regional Product	Coefficient					1.574	
	t-Statistic					1.441	
Average HH Income	Coefficient						0.412
	t-Statistic						0.155
Disaggregated SE Variable: % HH with Income > $100K	Coefficient	2.126	0.139	1.348	0.111	1.671	2.948
	t-Statistic	2.909 **	0.006	1.280	0.108	1.505	1.686
	R Squared	0.886	0.930	0.895	0.919	0.878	0.849
	Adj R Squared	0.866	0.917	0.876	0.905	0.856	0.822

* Coefficient estimate significant at 0.05 level ** at 0.01 level *** at 0.001 level

To assess the forecasting performance of these models, we limit the analysis to consideration of a single pair of regressions, those that use the Providence MSA Gross Regional Product as the aggregate socioeconomic variable. As mentioned above, it is not possible to obtain the Woods and Poole projections that would have been available for the years 2011 through 2015, which would have been included as part of the Woods and Poole 2012 data release. To create a test projection of values for Providence MSA GRP for those years, we use the percentage annual growth terms from the Woods and Poole 2016 data release for the years 2015 through 2019. With those percentage values for annual GRP growth, notional projected values for Providence GRP can be calculated for 2011 through 2015. To create notional projected values for the disaggregated socioeconomic variable, the percentage of Providence a MSA households with 2009 dollar denominated income exceeding $100,000 (in 2009 dollars), a similar procedure was used.

Forecast values for the Oil Price dependent variable are also needed to calculate out of sample projections. These were obtained for the years 2011 to 2015 from the data reported in the FAA's 2012 Annual Aerospace Forecast document. These national values are used for each of the case study airports out of sample forecast exercises.

Four out of sample model forecasts were calculated for the baseline model using Providence MSA GRP and the model that also includes the disaggregated socioeconomic variable. The first three are based on the out of sample forecasts described above. For the first forecast, the projected values – based on the percentage increases projected for the years 2015 through 2019 – for GRP, and the percentage of Providence household incomes exceeding $100,000 in the case of the model using the disaggregated socioeconomic variable as input, are used as described above. For the second, the values for percentage annual increases over the projection period (2011 to 2015) are increased by 50 percent, indicating a more rapidly growing Providence economy. For the third PVD projection, the values for percentage annual increases are reduced by 50 percent, representing a projection for a more slowly growing economic environment in the Providence MSA. For the fourth projection, the actual values observed in 2011 through 2015 for GRP and the share of Providence MSA households with incomes exceeding $100,000 are used to drive the "forecast" of PVD annual O&D enplanements.

Using the projections published by Woods and Poole for 2015 and subsequent years for simulating a range of forecast scenarios based on inputs that would have been available to analysts in 2011 or 2012, we can compare the forecasting performance of the baseline regressions for PVD annual O&D enplanements using the annual oil price variable and the Providence MSA GRP, with those of the regression parameters based additionally on input from the disaggregated socioeconomic variable. The independent variable data and projections are shown in Table D-35 below. The projections for the Providence MSA variables reported in the Woods and Poole 2016 data have annual GRP growing at a rate of around two percent, and the percentage of Providence MSA households with incomes exceeding $100,000 growing at about 1.7 percent initially, growing to 1.9 percent after three years. The annual percentage changes used to calculate the Oil Price projected values for 2011 through 2015 are those implied by the Oil Price projections reported in the 2012 FAA Aerospace Forecast, which relies on projected values for the years following 2010. Thus we have four "forecast" scenarios for the two regression specifications for annual PVD O&D enplanements:

- Projected values for the socioeconomic independent variables based on the annual percentage changes in the variables, as reported by Woods and Poole for the five years beginning in 2015,
- Projected values based on annual percentage changes that are 1.5 times these values,
- Projected values based on annual percentage changes that are half these values, and]
- Projected values that equal the actual values for the independent variables that occurred in the years 2011 through 2015.

Table D-37 Inputs to Out-of-Sample PVD O&D Enplanement Projections

PVD Inputs to Out-of-Sample Enplanement Forecasts	PVD					
	Sources	2011	2012	2013	2014	2015
Oil Price Forecast	*2012 FAA Aero Forecast*	$84.32	$92.71	$97.16	$100.65	$104.64
GRP Forecast 1 (Medium)	*W&P 2016 %*	$75,364.7	$76,893.4	$78,411.0	$79,928.8	$81,451.2
GRP Forecast 2 (High)	*1.5 x F'cast 1*	$76,141.7	$78,458.5	$80,781.2	$83,126.8	$85,501.7
GRP Forecast 3 (Low)	*0.5 x F'cast 1*	$74,587.6	$75,344.1	$76,087.6	$76,824.0	$77,555.6
HHInc% Forecast 1 (Medium)	*W&P 2016 %*	21.1%	21.5%	21.9%	22.3%	22.7%
HHInc% Forecast 2 (High)	*1.5 x F'cast 1*	21.3%	21.9%	22.5%	23.1%	23.7%
HHInc% Forecast 3 (Low)	*0.5 x F'cast 1*	21.0%	21.1%	21.3%	21.5%	21.7%
	2010					
Actual Oil Price	$83.99	$101.87	$100.93	$100.49	$92.02	$48.39
Actual GRP	$73,810.5	$73,604.1	$74,491.0	$74,943.6	$77,186.9	$78,644.6
Actual HHInc%	20.8%	19.7%	19.6%	19.6%	19.6%	19.8%

The forecasting or projection results for the baseline regression equation for annual enplanements at PVD are shown in Figure D-35 below.

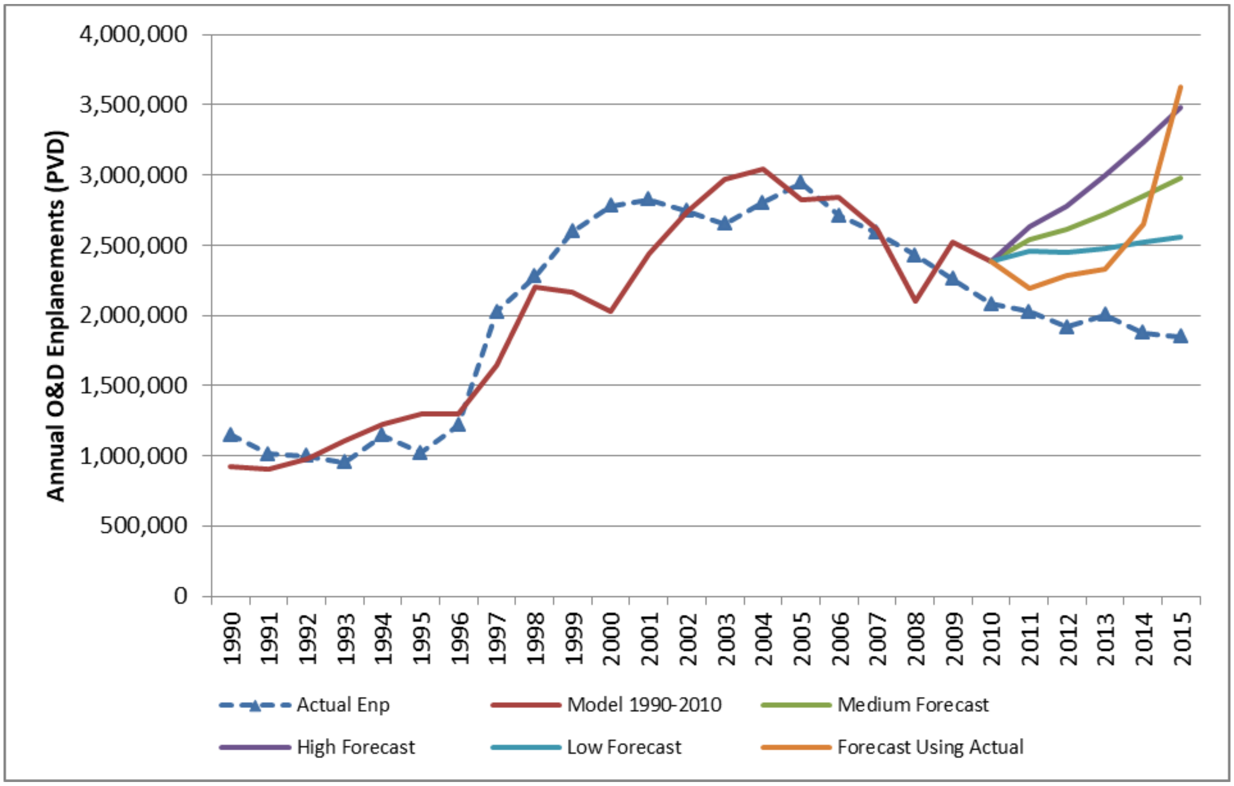

Figure D-35 Baseline Model Forecasts of PVD Annual Enplanements, 2011 to 2015, Using Providence MSA Gross Regional Product, 1990 to 2010

The results for the regression equation using the disaggregated socioeconomic variable along with Providence MSA GRP are shown in Figure D-36 below.

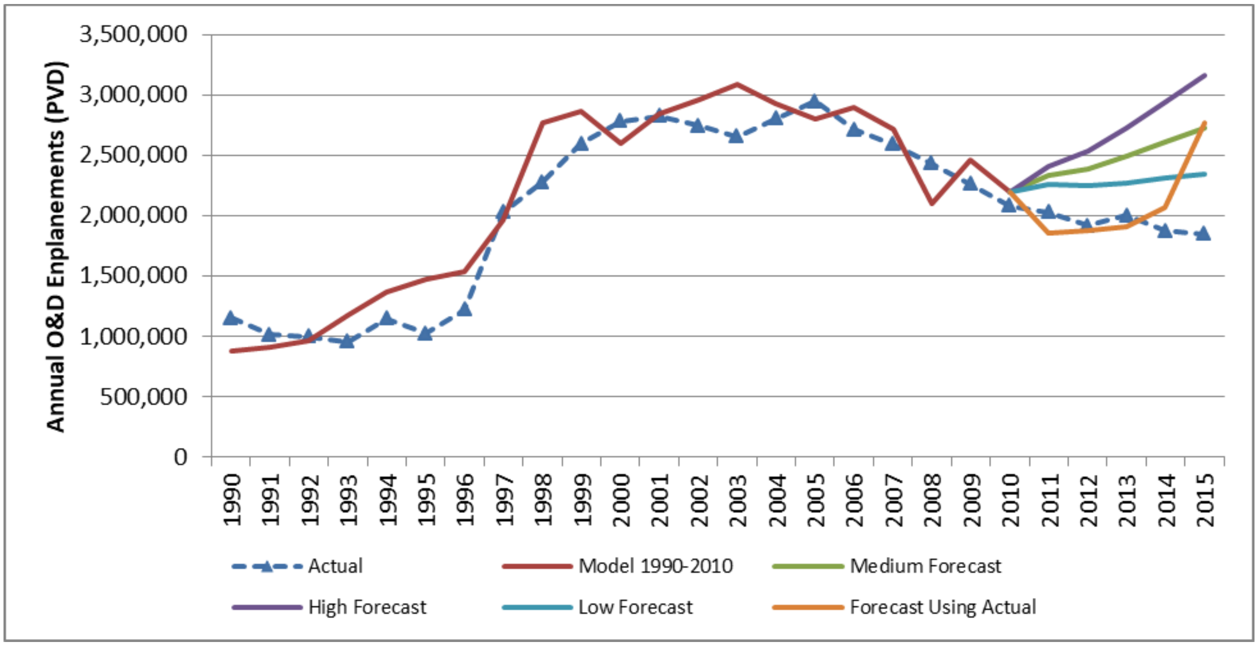

Figure D-36 Disaggregated SE Variable Model Forecasts of PVD Annual Enplanements, 2011 to 2015, Using Providence MSA GRP and Share of Providence Households with Incomes exceeding $100,000, 1990 to 2010

Enplanements at PVD declined over the forecast period, in contrast to the increases that were forecast under all eight forecast scenarios. The overforecasts are more pronounced or significant for the baseline specification, and for the alternative specification the forecast that uses the actual 2011 to 2015 values as independent variables in the model provides an accurate forecast for PVD enplanements through 2013. After 2013, however, the forecast increases sharply with the decline in oil prices from $92.02 in 2014 to $48.39 the following year, while actual PVD enplanements decline modestly.[9] The following tables D-38 through D-40 quantify this, reporting the percentage error for each of the forecasts and the Root Mean Squared Error (RMSE) for each forecast, a calculation which can be used to compare the accuracy of the baseline model forecasts and the forecasts calculated from the model also using the disaggregated socioeconomic variable.

[9] The baseline regression equation is ln(PVD_Enp) = -18.95 - 0.397ln(OilPrice) + 3.16ln(GRP),), with an adjusted R squared of 0.843, while the estimated equation that includes the disaggregated socioeconomic variable is ln(PVD_Enp) = 1.31 - 0.389ln(OilPrice) + 1.57ln(GRP) + 1.67ln(%HHInc>$100K) with an adjusted R squared of 0.856.

Table D-38 Percentage Forecast Errors, Baseline Model Projections of PVD Annual O&D Enplanements, 2011 to 2015

PVD using GRP	Percentage Forecast Error (Baseline Regression with Oil Price and GRP)			
	Medium	High	Low	Actual
2010	--	--	--	--
2011	25.5%	29.6%	21.4%	8.0%
2012	36.2%	45.1%	27.7%	19.1%
2013	36.1%	49.6%	23.8%	16.5%
2014	52.2%	72.3%	34.3%	41.3%
2015	61.3%	88.0%	38.2%	96.2%

The percentage errors in Tables D-38 and D-39 reflect the fact that the model-based projections of PVD enplanements consistently overestimated the number of annual passenger, even (at least for 2014 and 2015) for the "forecast" based on the values that actually occurred for the Oil Price and Providence MSA GRP. For the baseline model using Providence MSA GRP as the aggregate socioeconomic independent variable, the percentage differences between the forecasts and the actual PVD enplanement values were quite large, differences that narrowed modestly for the alternative model, results for which are shown in Table D-39.

Table D-39 Percentage Forecast Errors, Disaggregated SE Model Projections of PVD Annual O&D Enplanements, 2011 to 2015

PVD using GRP	Percentage Forecast Error (Specification including Disaggregated SE Variable)			
	Medium	High	Low	Actual
2010	--	--	--	--
2011	15.0%	18.6%	11.6%	-8.7%
2012	24.5%	32.3%	17.2%	-2.0%
2013	24.4%	36.1%	13.6%	-4.8%
2014	39.0%	56.7%	23.2%	10.2%
2015	47.5%	71.3%	26.9%	49.9%

The final table for the PVD analysis reports the Root Mean Squared Error (RMSE) of the different forecasts for the two model specifications. This statistic can be used to compare the accuracy of two or more of these forecasts, with a smaller RMSE value indicating a more accurate forecast. As can be seen, the inclusion of the additional disaggregated socioeconomic variable in the model specification provided significant improvement in the accuracy of the five year forecasts, compared to those based on the baseline specification, although neither specification resulted in good forecasting outcomes.

Table D-40 Root Mean Squared Error (RMSE) for PVD Baseline and Disaggregated SE Variable Model Forecasts of Annual PVD O&D Enplanements, 2011 to 2015

PVD (using GRP)	Root Mean Squared Error		% Difference BL v Alt
	Baseline	Disagg Var	
Medium Forecast	838,143	609,936	27%
High Forecast	1,146,777	884,463	23%
Low Forecast	567,269	367,406	35%
Actual Values	897,788	430,884	52%

g. Los Angeles International Airport (LAX)

The remaining case study airports and airport regions involve regions served by large hub airports. The first of these regions is the Los Angeles, California MSA, served by several passenger service airports including Los Angeles International Airport (LAX). The Los Angeles region has also experienced greater growth in more economic socioeconomic variables, such as gross regional product (GRP) or measures of regional income, than for more demographic socioeconomic variables, such as population or employment. Average household incomes grew during the sample period, while the proportion of MSA households with incomes exceeding $100,000 was flat to declining during the second half of the sample period. These growth patterns are shown in Figure D-37.

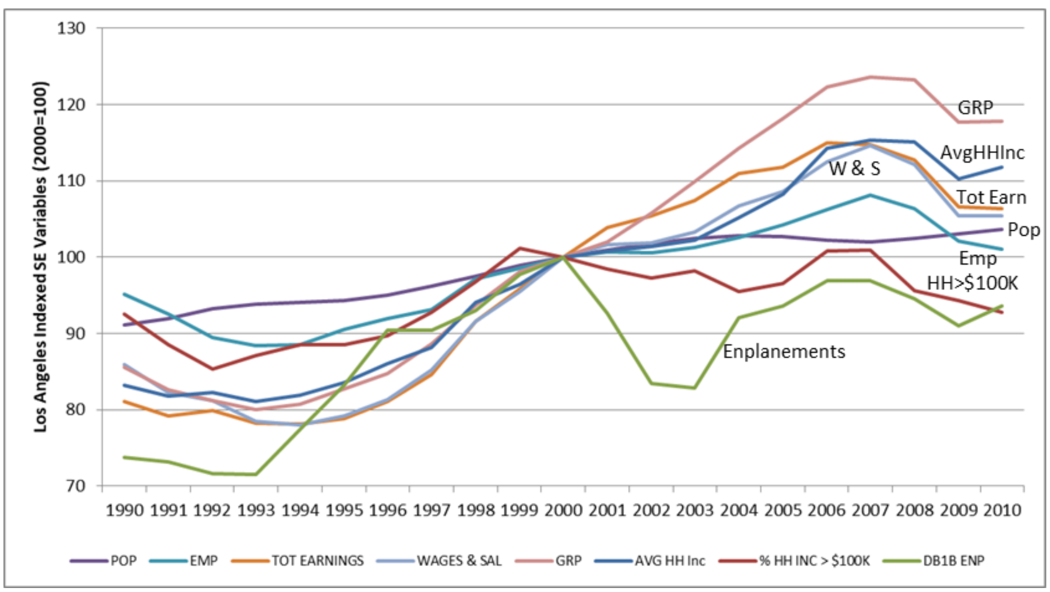

Figure D-37 Los Angeles, CA MSA Indexed Socioeconomic Variables, 1990 to 2010, 2000=100

Table D-41 reports the socioeconomic variables for the Los Angeles - Long Beach-Anaheim MSA for years 1990 to 2010.

Table D-41 Values of Socioeconomic Variables, Los Angeles, CA MSA (1990-2010)

	Year	Los Angeles International Airport (LAX)									
		Annual O&D Enplanements	Real Oil Price	Regional Population (000s)	Regional Employment (000s)	Regional Total Earnings ($M)	Regional W&S Earnings ($M)	Regional Product (GRP) ($M)	Regional Avg HH Income	% Regional HHs with Income > $100K
	1990	14,530,215	$40.16	11,297.1	6,881.7	$348,773.9	$262,676.1	$567,496.6	$105,118	25.5%
	1991	14,415,051	$33.20	11,400.8	6,700.5	$340,569.1	$251,402.2	$548,256.1	$103,235	24.6%
	1992	14,110,530	$31.33	11,552.4	6,471.1	$343,760.4	$248,143.0	$538,465.7	$103,860	23.8%
	1993	14,087,063	$26.96	11,632.8	6,397.4	$336,652.4	$239,670.8	$531,053.0	$102,345	24.3%
	1994	15,263,756	$25.31	11,663.2	6,402.5	$336,134.7	$238,338.3	$535,832.6	$103,429	24.8%
	1995	16,411,458	$27.16	11,692.7	6,550.1	$339,349.0	$241,952.5	$549,222.4	$105,482	24.9%
	1996	17,820,548	$31.73	11,771.0	6,656.4	$348,915.6	$248,687.2	$562,303.6	$108,574	25.2%
	1997	17,829,368	$27.80	11,915.8	6,736.3	$364,289.7	$260,329.1	$588,097.7	$111,313	26.1%
	1998	18,326,396	$17.87	12,086.8	7,027.2	$394,011.5	$279,901.9	$622,216.4	$118,752	27.2%
	1999	19,269,411	$25.03	12,253.2	7,135.0	$412,645.5	$291,965.9	$650,574.4	$121,789	28.4%
	2000	19,714,223	$38.86	12,392.7	7,236.8	$430,361.1	$305,626.4	$663,513.7	$126,253	28.2%
	2001	18,275,880	$29.99	12,511.5	7,286.0	$446,923.5	$310,755.8	$676,701.2	$127,316	27.6%
	2002	16,452,450	$31.83	12,614.2	7,279.7	$453,680.5	$311,248.0	$702,056.9	$127,970	27.4%
	2003	16,330,619	$36.38	12,696.5	7,327.0	$462,054.0	$315,850.9	$729,084.8	$129,066	27.8%
	2004	18,143,108	$45.87	12,735.0	7,426.7	$477,801.2	$326,266.8	$758,319.6	$132,752	27.1%
	2005	18,462,728	$60.44	12,726.4	7,542.9	$481,291.6	$331,955.8	$784,224.7	$136,670	27.5%
	2006	19,106,577	$70.72	12,670.2	7,687.1	$495,037.2	$343,978.4	$811,201.9	$144,308	28.7%
	2007	19,108,614	$78.22	12,632.0	7,822.7	$493,607.9	$350,285.9	$819,910.0	$145,645	28.7%
	2008	18,635,537	$103.82	12,692.7	7,694.3	$485,074.2	$342,911.3	$818,068.4	$145,391	27.2%
	2009	17,947,766	$66.42	12,774.6	7,389.4	$458,886.8	$322,351.5	$780,826.1	$139,252	26.9%
	2010	18,453,834	$83.99	12,845.3	7,310.5	$457,515.0	$322,306.0	$781,679.2	$141,191	26.4%
Used to assess out-of-sample forecasts	*2011*	*19,483,496*	*$101.87*	*12,954.5*	*7,465.9*	*$465,995.6*	*$325,666.3*	*$782,811.2*	*$141,939*	*25.7%*
	2012	*19,968,501*	*$100.93*	*13,064.8*	*7,749.1*	*$485,614.7*	*$336,217.2*	*$798,456.0*	*$148,612*	*25.7%*
	2013	*20,246,615*	*$100.49*	*13,175.8*	*7,989.3*	*$491,842.3*	*$337,394.2*	*$815,283.8*	*$146,032*	*25.8%*
	2014	*21,782,807*	*$92.02*	*13,262.2*	*8,158.4*	*$507,196.9*	*$350,290.2*	*$841,180.9*	*$149,844*	*25.8%*
	2015	*23,325,656*	*$48.39*	*13,347.4*	*8,288.9*	*$519,346.3*	*$359,651.7*	*$861,848.6*	*$151,674*	*26.1%*

Socioeconomic variables for the Los Angeles region were also strongly correlated during the sample period, as reported in Table D-42. However, reflecting the gradual decline in the proportion of Los Angeles MSA households with incomes exceeding $100,000 while other socioeconomic indexes were increasing, the correlation between the proportion of high income households and other socioeconomic variables is not as strong as between other variables. The correlations between the regional socioeconomic variables and either oil prices or annual O&D enplanements at LAX are also positive but less strong.

Table D-42 Correlations among Socioeconomic Variables, Los Angeles, CA MSA (1990 – 2010)

	Population	Emp	Total Earn	Wages & Salaries	GRP	Avg HH Inc	% > $100K	DB1B Enp	Real Oil Price
Population	1								
Employment	0.950	1							
Total Earn	0.979	0.976	1						
Wages & Salaries	0.978	0.984	0.998	1					
GRP	0.970	0.966	0.996	0.992	1				
Avg HH Inc	0.942	0.947	0.985	0.981	0.984	1			
% > $100K	0.940	0.956	0.978	0.981	0.978	0.989	1		
DB1B Enplanements	0.912	0.913	0.881	0.903	0.863	0.828	0.861	1	
Real Oil Price	0.559	0.650	0.692	0.675	0.705	0.763	0.712	0.364	1

The Los Angeles-Long Beach-Anaheim, California MSA saw modest changes in its population's age distribution over the case study period. Those 19 years old or younger made up about 30 percent of the population over the case study period, and there was a reduction in the share of those aged between 20 and 34, from 29 percent of the MSA's population to about 23 percent by 2010. There was a modest increase in the share of those aged between 50 and 64, from about 11 percent to around 17% by 2010, as shown in Figure D-38.

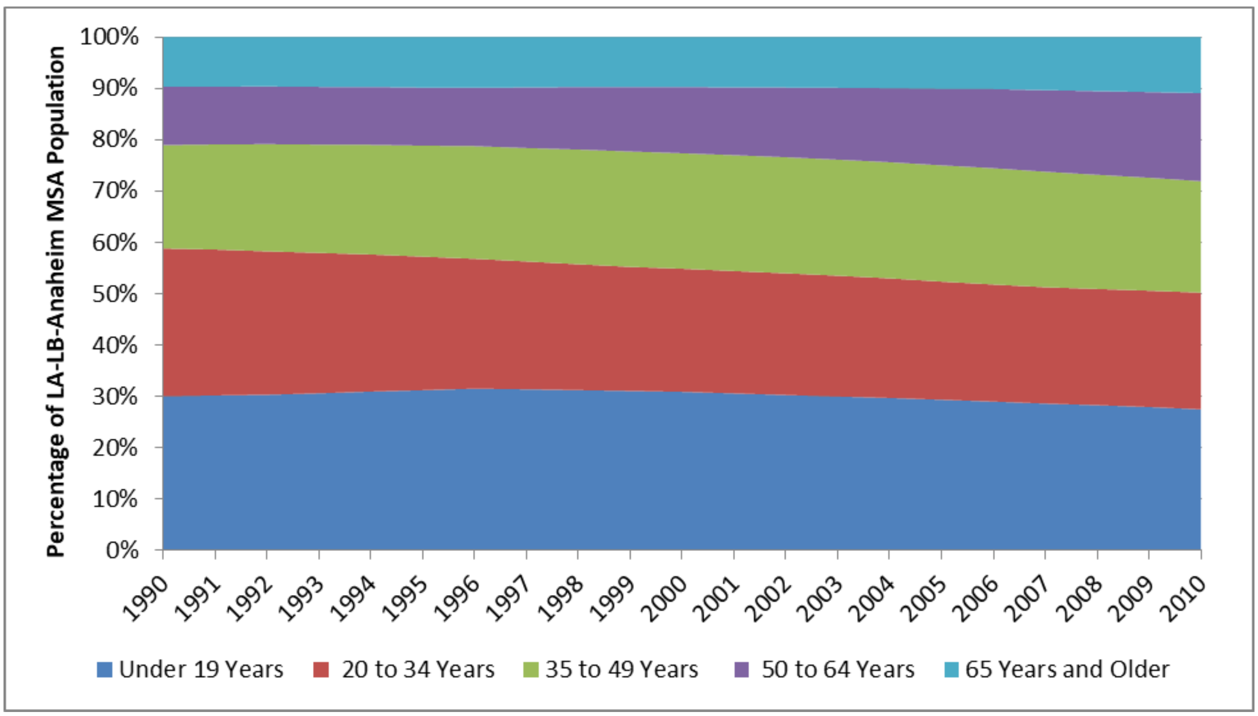

Figure D-38 Population Age Distribution, LA-Long Beach-Anaheim, 1990 to 2010

Figure D-39 shows the distribution of the Los Angeles area's adult population, those 20 and older. During the case study period Los Angeles area's adult population distribution shifted somewhat, with those between 20 and 34 years of age declining from around 41 percent of the adult population to around 31 percent in 2010, and the proportion aged 65 or older was stable at around 14 percent of the MSA's adult population. The share of the adult population between 50 and 64 years also grew, from about 16 percent of the adult population in 1985 to over 24 percent of the population 20 or older by 2010.

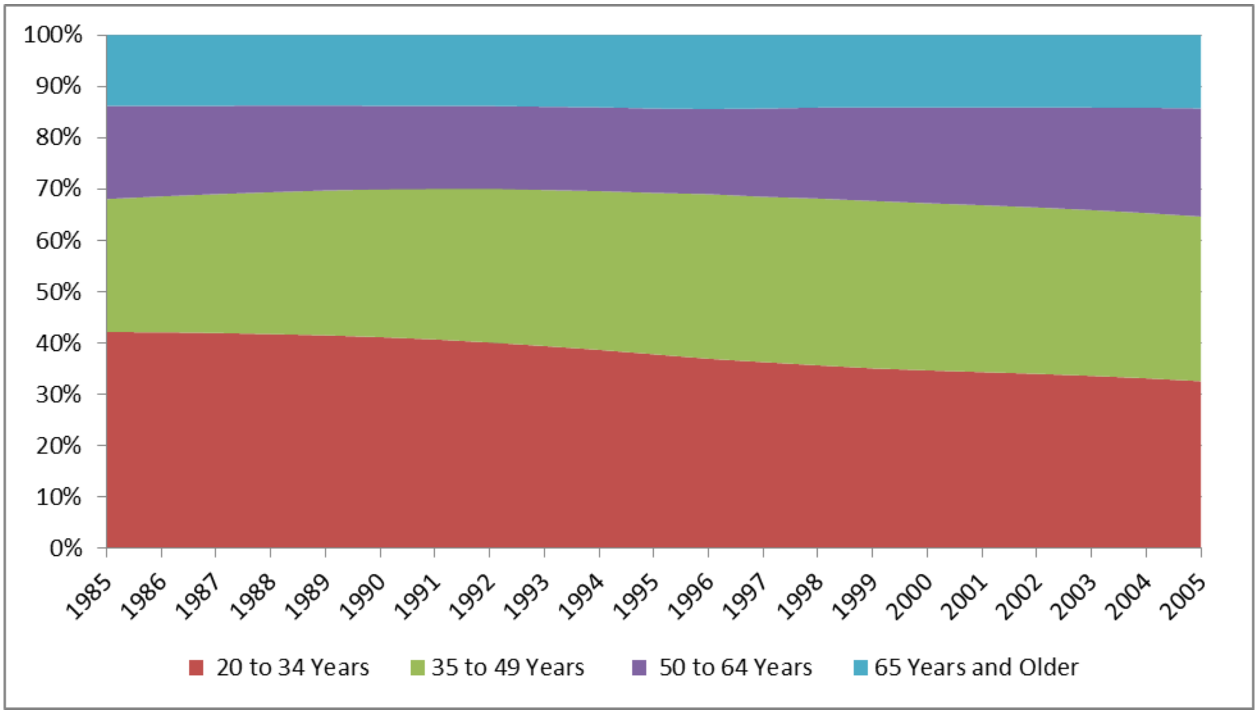

Figure D-39 Adult Population Age Distribution, LA-Long Beach-Anaheim, CA MSA, 1990 to 2010

Figure D-40 depicts household income distributions for the residents of the greater Los Angeles MSA over the case study sample period. The figure shows a steady distribution of household incomes over the case study sample period, with each of the four income groups identified made up by about a quarter of Los Angeles area households – around 23 percent – have incomes exceeding $100,000 over the sample period, a higher share that other communities considered so far in the case study analysis.

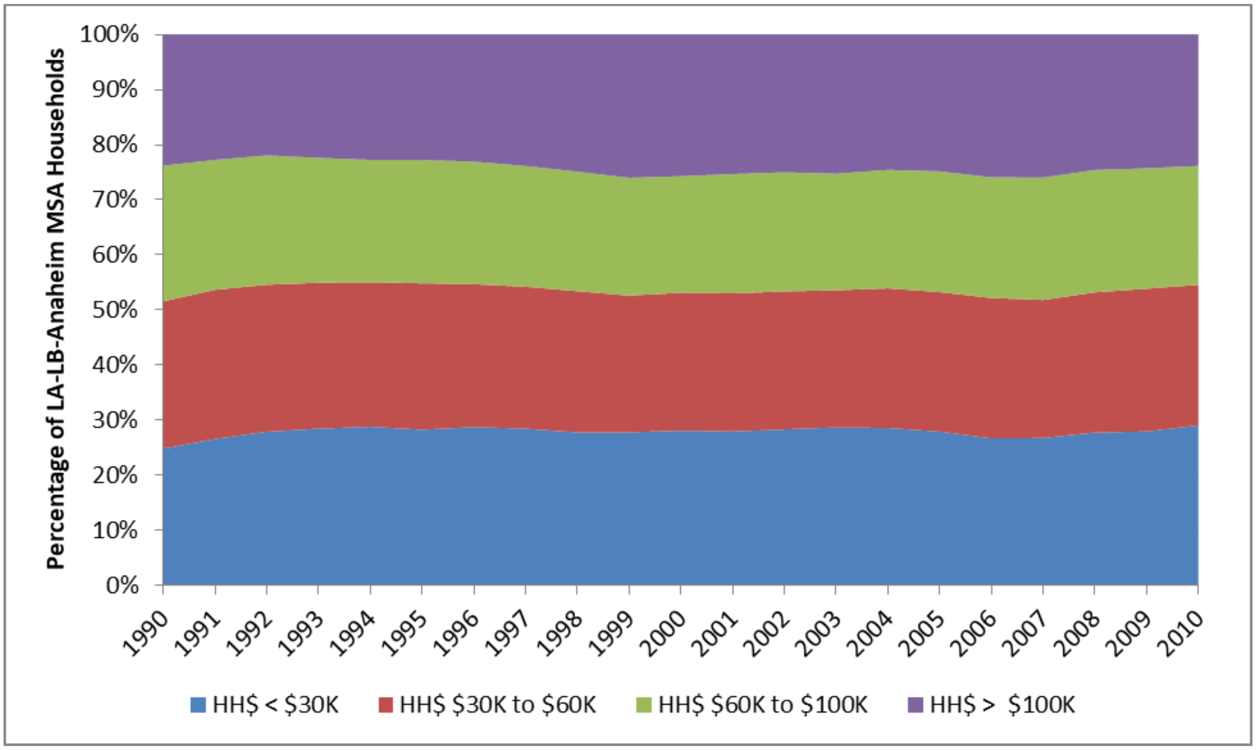

Figure D-40 Household Income Distribution, LA-Long Beach-Anaheim, CA MSA, 1990 to 2010

Figure D-41 reports specifically the evolution of the percentage of Los Angeles-Long Beach-Anaheim MSA households with annual incomes of $100,000 (in 2009 dollars), a group shown as the uppermost band in Figure D-40 above. This percentage hovered around 25 percent of Los Angeles area households over the case study sample period.

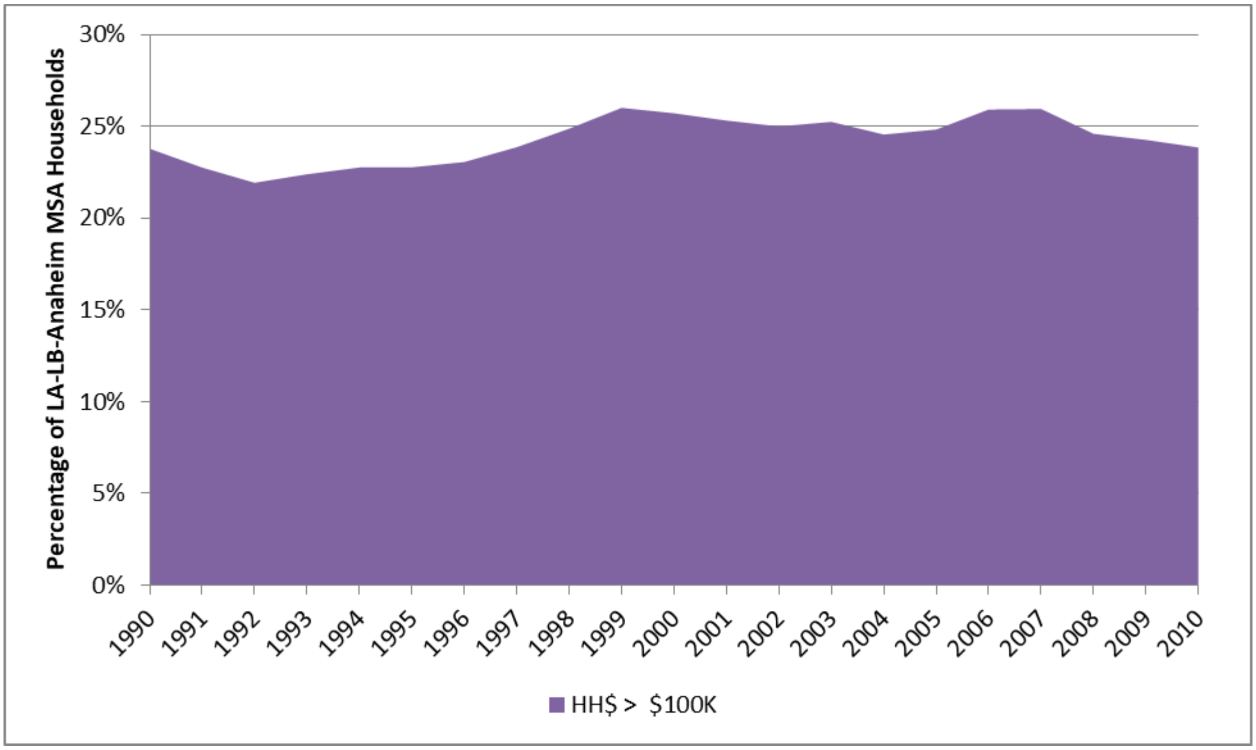

Figure D-41 Percentage of Households with Incomes Exceeding $100,000, Los Angeles-Long Beach-Anaheim, CA 1990 to 2010

Characteristics of the results reported in Table D-43 below for the six baseline regressions modeling annual passenger enplanements at Los Angeles International Airport (LAX) from 1990 to 2010 include:

- The model estimates gave similar degrees of goodness of fit to the actual LAX annual enplanements series, with adjusted R-squared statistics ranging from 0.473 to 0.618, indicating that the baseline models regressing annual enplanements on the annual real oil price and one of six aggregate socioeconomic variables for the Los Angeles-Long Beach-Anaheim (henceforth, Los Angeles) MSA accounted for no more than half of the observed variation in the annual enplanements series.
- The estimated regression constant terms varied across the baseline models, although the estimates for the models using Total Earnings in the Los Angeles MSA, Annual Wages and Salaries in the MSA, Gross Regional Product, and Average Real Household Income were roughly comparable, ranging from 5.88 to 8.68. All of these estimates at the 0.01 level (or better). Estimated constant terms for the models using MSA population and MSA annual employment were not statistically significant.
- The coefficients estimated in the baseline LAX models for the oil price variable, which in a log linear regression represents the elasticity of changes in annual LAX enplanements with respect to changes in the oil price, is relatively small (and negative) in all baseline regressions (ranging between -0.02 and -0.11) and statistically significant only in the two baseline regressions using either Los Angeles MSA GRP or Average Real Household Income as the aggregate socioeconomic explanatory variable (at the 0.05 level of

significance). This indicates that over the sample period there was on average little to no responsiveness of annual air passenger enplanements to changes in oil prices.

- The influence of the aggregate socioeconomic variables on annual enplanements at LAX is positive and strongly statistically significant across all of the baseline models. The range of coefficient (elasticity) estimates exhibits the same pattern as that described for the baseline regressions covering the Baltimore-Washington CSA, with the largest elasticity estimate (2.07, significant at the 0.001 level) coming for Los Angeles MSA population, a smaller value for the regression estimate using Los Angeles MSA total employment (1.63, also significant at the 0.001 level), and then estimates between 0.63 and 0.76 for Total Earnings, Wages and Salaries, and GRP (all also significant at the 0.001 level). The elasticity estimate for Average Real Household Income in the Los Angeles MSA was comparable to the estimates for these three, estimated at 0.95, and also significant at the 0.001 level.

The LAX regression equation substituting the annual series of the percentage of Los Angeles MSA households with incomes of $100,000 or more (a disaggregated socioeconomic variable) for the aggregate socioeconomic variables used in the six baseline regressions is very similar to the results from those baseline regressions. The estimated constant term of 18.67, is much larger that the constant term that resulted in the six baseline regressions, and it is significant at the 0.005 significance level. In addition to the difference in the constant term, the coefficient estimate on the Oil Price variable (+0.40) is positive rather than negative, statistically significant only at the 0.2 significance level. The coefficient estimate for the contribution of the disaggregated socioeconomic variable, 1.52, is, however, comparable to the estimates obtained for some of the aggregate socioeconomic variables in the baseline regressions. That estimate is statistically significant at the 0.001 significance level. The adjusted R-squared value is 0.656, slightly higher than the values for the six baseline regressions.

Table D-43 Baseline and Other Regression Results for Models of Annual Enplanements, Los Angeles International Airport, 1990 to 2010

Ind. Variable	Statistic	Regression 1	2	3	4	5	6	7
Constant	Coefficient	-2.76	2.44	8.68	7.39	7.33	5.88	18.67
	t-Statistic	-0.601	0.742	4.420 ***	3.282 **	3.468 **	3.006 **	37.169 ***
Oil Price	Coefficient	-0.02	-0.07	-0.05	-0.07	-0.10	-0.11	0.04
	t-Statistic	-0.516	-1.261	-0.883	-1.209	-1.755 *	-2.187 *	1.194
Pop	Coefficient	2.07						
	t-Statistic	4.163 ***						
Emp	Coefficient		1.63					
	t-Statistic		4.218 ***					
Total Earnings	Coefficient			0.63				
	t-Statistic			3.896 ***				
Wages and Salaries	Coefficient				0.76			
	t-Statistic				3.974 ***			
Gross Regional Product	Coefficient					0.72		
	t-Statistic					4.259 ***		
Average HH Income	Coefficient						0.95	
	t-Statistic						5.342 ***	
Disaggregated SE Variable: % HH with Income > $100K	Coefficient							1.52
	t-Statistic							4.841 ***
	R Squared	0.543	0.569	0.526	0.544	0.557	0.657	0.691
	Adj R Squared	0.492	0.521	0.473	0.494	0.508	0.618	0.656

* Coefficient estimate significant at 0.05 level ** at 0.01 level *** at 0.001 level

When the disaggregated socioeconomic variable is added as a new variable to the baseline regression specifications, many of the relationships that were estimated in the baseline equations no longer hold. These results are reported in Table D-44 below.

- For these regressions, the adjusted R-squared estimates are somewhat higher than those reported for the baseline regression equations, ranging from 0.636 to 0.662, suggesting that inclusion of the disaggregated socioeconomic variable in the models provided some additional information for modeling the O&D passenger enplanements at LAX over the period from 1990 to 2010.
- The equation constant estimates are larger by an order of magnitude, no doubt a reflection of the much smaller magnitude of the values that make up the disaggregated socioeconomic variable. With the exception of the equation using the Los Angeles MSA population series, the constant estimates are statistically significant at the 0.01 level or better.
- With the exception of the specification using Los Angeles employment as the aggregate socioeconomic variable, the coefficient estimates for the Oil Price variable are small in magnitude, sometimes positive rather than negative, and not statistically significant.
- The coefficient estimates for the aggregate socioeconomic variables are, with the exceptions of the specifications using the Los Angeles MSA annual employment series (negative coefficient, significant at the 0.1 level) and the Los Angeles MSA Average Real Household Income (positive coefficient, as expected, and also significant at the 0.1 significance level), not statistically significant and are sometimes negative rather than positive.
- Finally, the coefficient estimates for the disaggregated socioeconomic variable are positive as expected, of reasonable magnitude, and statistically significant at the 0.05 significance level or better, with the exception of the regression specification using the Average Real Household Income, for which the coefficient estimate is usually significant at the 0.01 level.

Table D-44 Regression Results for Models of Annual Enplanements, Los Angeles International Airport, 1990 to 2010, that Include both Aggregate and Disaggregated Socioeconomic Variables

Ind. Variable	Statistic	Regression					
		1	2	3	4	5	6
Constant	Coefficient	12.82	29.96	21.12	24.51	18.73	13.24
	t-Statistic	3.039 **	3.039 **	4.727 ***	4.189 ***	3.840 ***	2.540 **
Oil Price	Coefficient	0.002	0.091	0.043	0.077	0.022	-0.038
	t-Statistic	0.048	1.315	0.830	1.191	0.316	-0.576
Pop	Coefficient	0.580					
	t-Statistic	0.856					
Emp	Coefficient		-1.174				
	t-Statistic		-1.153				
Total Earnings	Coefficient			-0.170			
	t-Statistic			-0.568			
Wages and Salaries	Coefficient				-0.417		
	t-Statistic				-1.013		
Gross Regional Product	Coefficient					-0.009	
	t-Statistic					-0.028	
Average HH Income	Coefficient						0.410
	t-Statistic						1.031
Disaggregated SE Variable: % HH with Income > $100K	Coefficient	1.229	2.437	1.831	2.184	1.536	0.943
	t-Statistic	2.819 **	2.908 **	2.992 **	3.085 **	2.527 *	1.515
	R Squared	0.698	0.713	0.698	0.708	0.691	0.703
	Adj R Squared	0.644	0.662	0.645	0.657	0.636	0.651

* Coefficient estimate significant at 0.05 level ** at 0.01 level *** at 0.001 level

To assess the forecasting performance of these model estimates, we limit the analysis to consideration of a single pair of regressions, those that use the Los Angeles MSA Gross Regional Product as the aggregate socioeconomic variable. As mentioned above, it is not possible to obtain the Woods and Poole projections that would have been available for the years 2011 through 2015, which would have been included as part of the Woods and Poole 2012 data release. To create a test projection of values for Los Angeles MSA GRP for those years, we use the percentage annual growth terms from the Woods and Poole 2016 data release for the years 2015 through 2019. With those percentage values for annual GRP growth, notional projected values for Los Angeles GRP can be calculated for 2011 through 2015. To create notional projected values for the disaggregated socioeconomic variable, the percentage of Los Angeles MSA households with 2009 dollar denominated income exceeding $100,000 (in 2009 dollars), a similar procedure was used.

Forecast values for the Oil Price dependent variable are also needed to calculate out of sample projections. These were obtained for the years 2011 to 2015 from the data reported in the FAA's 2012 Annual Aerospace Forecast document. These national values are used for each of the case study airports out of sample forecast exercises.

Four out of sample model forecasts were calculated for the baseline model using Los Angeles MSA GRP and the model that also includes the disaggregated socioeconomic variable. The first three are based on the out of sample forecasts described above. For the first forecast, the projected values – based on the percentage increases projected for the years 2015 through 2019 – for GRP, and the percentage of Los Angeles household incomes exceeding $100,000 in the case of the model using the disaggregated socioeconomic variable as input, are used as described above. For the second, the values for percentage annual increases over the projection period (2011 to 2015) are increased by 50 percent, indicating a more rapidly growing Los Angeles area economy. For the third LAX projection, the values for percentage annual increases are reduced by 50 percent, representing a projection for a more slowly growing economic environment in the Los Angeles MSA. For the fourth projection, the actual values observed in 2011 through 2015 for GRP and the share of Los Angeles MSA households with incomes exceeding $100,000 are used to drive the "forecast" of LAX annual O&D enplanements.

Using the projections published by Woods and Poole for 2015 and subsequent years for simulating a range of forecast scenarios based on inputs that would have been available to analysts in 2011 or 2012, we can compare the forecasting performance of the baseline regressions for LAX annual O&D enplanements using the annual oil price variable and the Los Angeles MSA GRP, with those of the regression parameters based additionally on input from the disaggregated socioeconomic variable. The independent variable data and projections are shown in Table D-45 below. The projections for the Los Angeles MSA variables reported in the Woods and Poole 2016 data have annual GRP growing at a rate of around 2.1 to 2.4 percent, and the percentage of Los Angeles MSA households with incomes exceeding $100,000 growing at about 1.5 percent over the forecast period. The annual percentage changes used to calculate the Oil Price projected values for 2011 through 2015 are those implied by the Oil Price projections reported in the 2012 FAA Aerospace Forecast, which relies on projected values for the years following 2010. Thus we have four "forecast" scenarios for the two regression specifications for annual LAX O&D enplanements:

- Projected values for the socioeconomic independent variables based on the annual percentage changes in the variables, as reported by Woods and Poole for the five years beginning in 2015,
- Projected values based on annual percentage changes that are 1.5 times these values,
- Projected values based on annual percentage changes that are half these values, and]
- Projected values that equal the actual values for the independent variables that occurred in the years 2011 through 2015.

Table D-45 Inputs to Out-of-Sample LAX O&D Enplanement Projections

LAX Inputs to Out-of-Sample Enplanement Forecasts	LAX					
	Sources	2011	2012	2013	2014	2015
Oil Price Forecast	*2012 FAA Aero Forecast*	$84.32	$92.71	$97.16	$100.65	$104.64
GRP Forecast 1 (Medium)	*W&P 2016 %*	$800,079.2	$818,401.0	$836,759.8	$855,254.0	$873,920.4
GRP Forecast 2 (High)	*1.5 x F'cast 1*	$809,279.3	$837,077.9	$865,244.5	$893,930.3	$923,196.1
GRP Forecast 3 (Low)	*0.5 x F'cast 1*	$790,879.2	$799,934.7	$808,907.0	$817,846.3	$826,771.3
HHInc% Forecast 1 (Medium)	*W&P 2016 %*	26.8%	27.2%	27.7%	28.1%	28.5%
HHInc% Forecast 2 (High)	*1.5 x F'cast 1*	27.0%	27.6%	28.3%	29.0%	29.6%
HHInc% Forecast 3 (Low)	*0.5 x F'cast 1*	26.6%	26.8%	27.0%	27.2%	27.5%
	2010					
Actual Oil Price	$83.99	$101.87	$100.93	$100.49	$92.02	$48.39
Actual GRP	$781,679.2	$782,811.2	$798,456.0	$815,283.8	$841,180.9	$861,848.6
Actual HHInc%	26.4%	25.7%	25.7%	25.8%	25.8%	26.1%

The forecasting or projection results for the baseline regression equation for annual enplanements at LAX are shown in Figure D-42 below.

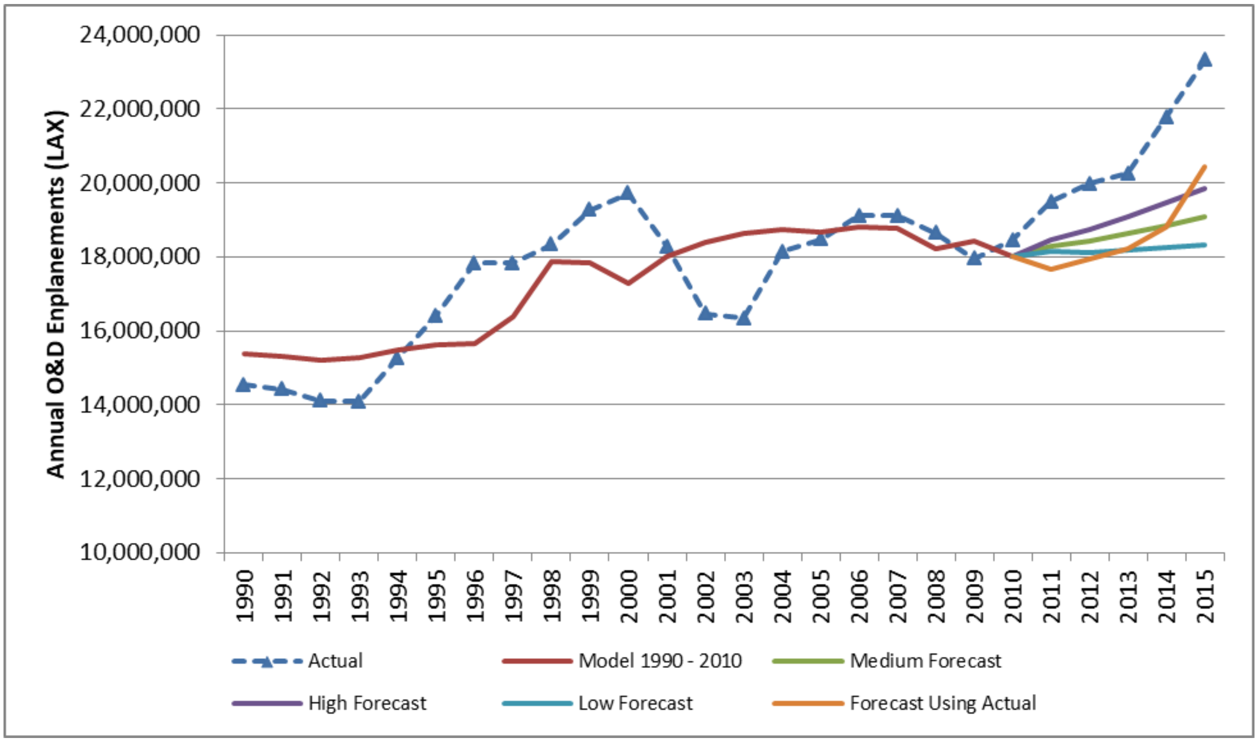

Figure D-42 Baseline Model Forecasts of LAX Annual Enplanements, 2011 to 2015, Using Los Angeles MSA Gross Regional Product, 1990 to 2010

The results for the regression equation using the disaggregated socioeconomic variable along with Los Angeles MSA GRP are shown in Figure D-43 below.

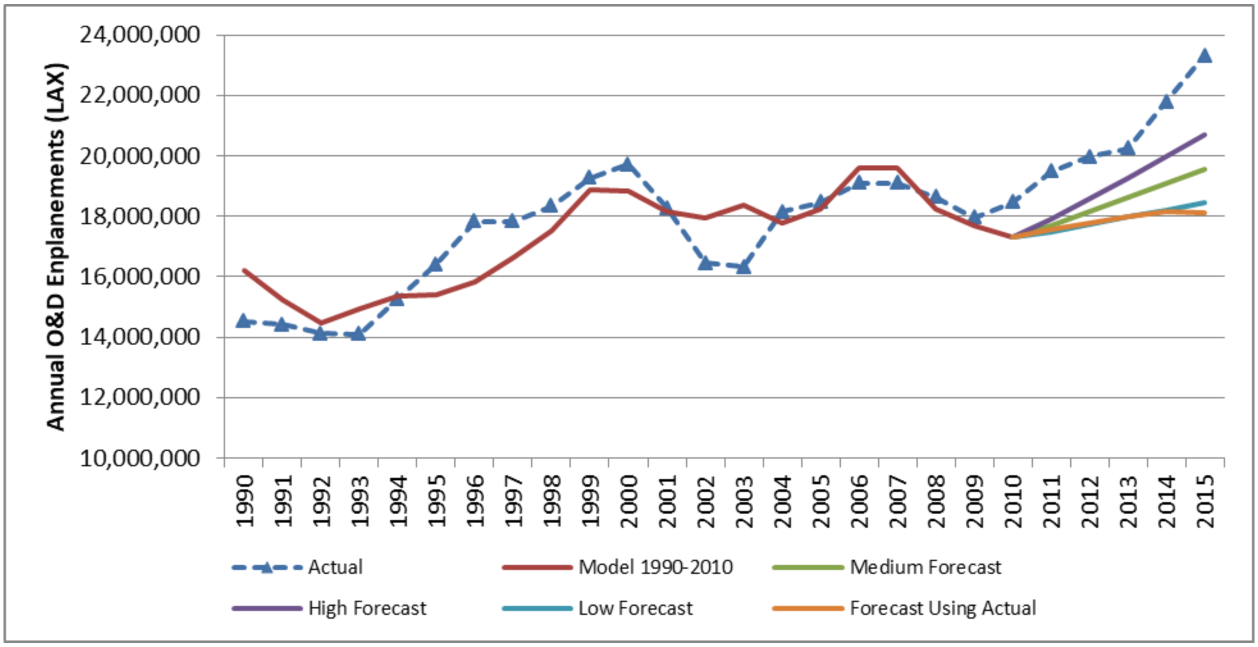

Figure D-43 Disaggregated SE Variable Model Forecasts of LAX Annual Enplanements, 2011 to 2015, Using Los Angeles MSA GRP and Share of Los Angeles Households with Incomes exceeding $100,000, 1990 to 2010

Enplanements at LAX increased more strongly than was forecast by either model and under any of the forecast scenarios for the independent variables. The underforecasting by the baseline forecasting scenarios is more severe for the baseline model specification, which does not include the disaggregated socioeconomic variable.[10] The accuracy of the projections driven by the actual values taken by the Oil Price, Los Angeles MSA GRP, and the Share of Los Angeles Households with Income exceeding $100,000 in the years 2011 through 2015 is also modest, with the projected enplanement counts undershooting the actual history over those years. The following tables D-46 to D-48 quantify this, reporting the percentage error for each of the forecasts and the Root Mean Squared Error (RMSE) for each forecast, a calculation which can be used to compare the accuracy of the baseline model forecasts and the forecasts calculated from the model also using the disaggregated socioeconomic variable.

[10] The baseline regression equation is ln(LAX_Enp) = 7.33 – 0.10ln(OilPrice) + 0.72ln(GRP), with an adjusted R squared of 0.508,while the estimated equation that includes the disaggregated socioeconomic variable is ln(LAX_Enp) = 18.73 + 0.02ln(OilPrice) - 0.01ln(GRP) + 1.54ln(%HHInc>$100K) with an adjusted R squared of 0.636.

Table D-46 Percentage Forecast Errors, Baseline Model Projections of LAX Annual O&D Enplanements, 2011 to 2015

LAX using GRP	Percentage Forecast Error (Baseline Regression with Oil Price and GRP)			
	Medium	High	Low	Actual
2010	--	--	--	--
2011	-6.1%	-5.3%	-6.9%	-9.3%
2012	-7.8%	-6.2%	-9.3%	-10.2%
2013	-8.0%	-5.7%	-10.2%	-10.0%
2014	-13.4%	-10.6%	-16.2%	-13.7%
2015	-18.2%	-14.9%	-21.4%	-12.4%

The percentage errors in the tables reflect the fact that the model-based projections of SAX enplanements consistently underestimated the number of annual O&D passengers at LAX, even for the “forecast” based on the actual 2011 to 2015 values for the Oil Price and Los Angeles MSA GRP and the proportion of higher income households. For the Medium, High, and Low forecast scenarios, the alternative model performs better than the baseline model (although neither performs well).

Table D-47 Percentage Forecast Errors, Disaggregated SE Model Projections of LAX Annual O&D Enplanements, 2011 to 2015

LAX using GRP	Percentage Forecast Error (Specification including Disaggregated SE Variable)			
	Medium	High	Low	Actual
2010	--	--	--	--
2011	-9.2%	-8.1%	-10.2%	-9.8%
2012	-9.1%	-6.9%	-11.2%	-11.0%
2013	-8.1%	-4.8%	-11.2%	-11.2%
2014	-12.4%	-8.3%	-16.4%	-16.6%
2015	-16.2%	-11.2%	-21.0%	-22.4%

The final table presenting the LAX regression analysis reports the Root Mean Squared Error (RMSE) of the different forecasts for the two model specifications. This statistic can be used to compare the accuracy of two or more of these forecasts, with a smaller RMSE value indicating a more accurate forecast. As can be seen, the inclusion of the additional disaggregated socioeconomic variable in the model specification provided only very modest improvement in the accuracy of the five year forecasts, compared to those based on the baseline specification.

Table D-48 Root Mean Squared Error (RMSE) for LAX Baseline and Disaggregated SE Variable Model Forecasts of Annual LAX O&D Enplanements, 2011 to 2015

LAX (using GRP)	Root Mean Squared Error		% Difference BL v Alt
	Baseline	Disagg Var	
Medium Forecast	2,569,496	2,480,409	3%
High Forecast	2,067,589	1,757,616	15%
Low Forecast	3,062,846	3,188,493	-4%
Actual Values	2,400,050	3,282,685	-37%

h. Phoenix Sky Harbor International Airport (PHX)

Indexed annual socioeconomic data from 1990 to 2010 are shown for the Phoenix, Arizona MSA in Figure D-44. This region is served by Phoenix Sky Harbor International Airport (PHX), a large hub. The data for the Phoenix MSA exhibit stronger growth patterns than those occurring in the other case study MSAs. Demographic and economic variables grew strongly until the Great Recession, which began in 2007. In the wake of that recession, economic variables declined, although population in the region continued to grow. For the Phoenix area, for example, gross regional product, indexed to 100 for 2000, grew to an indexed level of over 130 by 2007, having begun the sample period at a much lower indexed level. Phoenix GRP then declined in the wake of the Financial Crisis. Population growth in the Phoenix MSA has been robust, starting from an indexed level of less than 70 in 1990, reaching the index baseline level of 100 in 2000, and reaching an indexed value of nearly 130 by 2010.

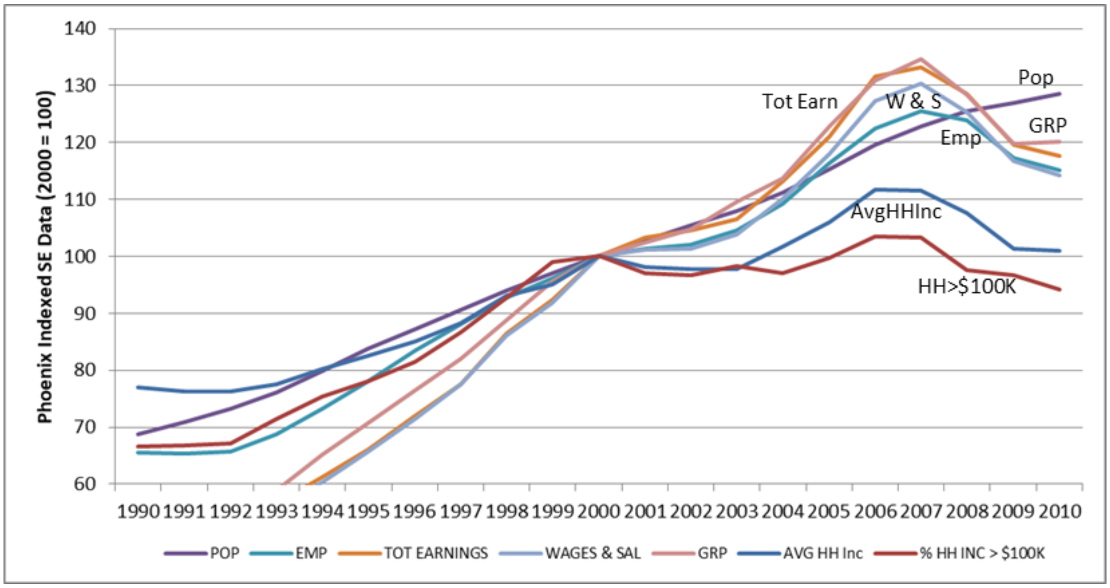

Figure D-44 Values of Socioeconomic Variables, Phoenix, AZ MSA (1990-2010)

Table D-49 reports values for the socioeconomic and other variables for the Phoenix MSA over the years from 1990 to 2010.

Table D-49 Values of Socioeconomic Variables, Phoenix, AZ MSA (1990-2010)

Year		Phoenix Sky Harbor International Airport (PHX)								
		Annual O&D Enplanements	Real Oil Price	Regional Population (000s)	Regional Employment (000s)	Regional Total Earnings ($M)	Regional W&S Earnings ($M)	Regional Product (GRP) ($M)	Regional Avg HH Income	% Regional HHs with Income > $100K
	1990	2,511,285	$40.16	2,249.1	1,137.9	$50,352.1	$39,324.5	$81,882.4	$79,300	14.7%
	1991	2,514,645	$33.20	2,319.2	1,172.3	$51,483.8	$39,130.4	$81,818.2	$78,568	14.8%
	1992	2,721,233	$31.33	2,398.8	1,211.2	$53,916.3	$40,430.1	$89,956.8	$78,709	14.8%
	1993	2,986,400	$26.96	2,491.8	1,256.3	$56,726.0	$42,227.3	$94,986.6	$79,965	15.8%
	1994	3,857,175	$25.31	2,613.5	1,315.2	$61,419.0	$45,705.8	$105,035.0	$82,720	16.6%
	1995	4,438,203	$27.16	2,744.0	1,378.1	$66,282.2	$49,833.0	$114,138.8	$85,025	17.2%
	1996	5,068,697	$31.73	2,855.7	1,431.6	$72,156.8	$54,123.0	$123,377.4	$87,744	18.0%
	1997	5,088,458	$27.80	2,963.7	1,483.0	$77,683.8	$58,795.4	$132,181.6	$91,028	19.1%
	1998	5,360,061	$17.87	3,074.5	1,536.0	$86,736.6	$65,221.9	$143,306.2	$95,969	20.5%
	1999	5,665,832	$25.03	3,178.3	1,585.2	$92,571.1	$69,649.2	$154,144.8	$97,997	21.8%
	2000	5,935,671	$38.86	3,273.5	1,630.5	$100,199.4	$75,775.3	$161,367.1	$103,099	22.1%
	2001	5,628,830	$29.99	3,363.7	1,676.1	$103,553.7	$76,667.8	$165,398.3	$101,099	21.4%
	2002	5,439,588	$31.83	3,452.5	1,721.7	$104,687.2	$76,740.3	$168,968.0	$100,869	21.3%
	2003	5,437,761	$36.38	3,536.4	1,765.6	$106,702.4	$78,664.6	$176,739.1	$100,808	21.7%
	2004	5,630,069	$45.87	3,637.3	1,817.4	$113,414.1	$83,479.3	$183,332.0	$104,924	21.4%
	2005	6,007,806	$60.44	3,774.7	1,888.4	$121,371.7	$89,438.8	$198,175.3	$109,323	22.0%
	2006	6,196,638	$70.72	3,914.2	1,958.9	$131,791.2	$96,494.9	$211,326.6	$115,170	22.9%
	2007	6,343,397	$78.22	4,018.1	2,012.4	$133,516.7	$98,792.7	$217,160.7	$114,942	22.8%
	2008	6,232,328	$103.82	4,106.4	2,058.6	$128,884.7	$95,022.8	$207,253.9	$110,950	21.5%
	2009	5,779,652	$66.42	4,153.6	2,085.6	$119,915.6	$88,426.8	$193,384.5	$104,536	21.3%
	2010	5,789,690	$83.99	4,209.3	2,115.5	$117,848.8	$86,593.1	$193,892.6	$104,028	20.8%
Used to assess out-of-sample forecasts	*2011*	*5,967,308*	*$101.87*	*4,254.1*	*2,139.8*	*$121,216.9*	*$88,475.9*	*$196,521.1*	*$103,964*	*19.7%*
	2012	*6,144,758*	*$100.93*	*4,331.0*	*2,178.6*	*$125,183.8*	*$91,421.2*	*$202,798.3*	*$105,617*	*19.6%*
	2013	*6,401,976*	*$100.49*	*4,404.1*	*2,216.2*	*$126,708.4*	*$93,132.9*	*$205,867.0*	*$103,041*	*19.6%*
	2014	*6,836,676*	*$92.02*	*4,489.1*	*2,259.3*	*$131,392.3*	*$96,193.5*	*$213,481.5*	*$105,119*	*19.6%*
	2015	*7,302,152*	*$48.39*	*4,567.2*	*2,296.1*	*$135,583.2*	*$99,598.4*	*$220,708.0*	*$106,418*	*19.8%*

The correlation between Phoenix aggregate and disaggregated socioeconomic variables are strongly correlated over the sample period, as reported in Table D-50. With the exception of the correlation between the population series values and the percentage of Phoenix households with incomes exceeding $100,000 (correlation of 0.88), all of the pairs of variables are correlated at a 0.9 level or greater. The correlations between the Phoenix MSA socioeconomic variables and either oil prices or annual O&D enplanements at PHX are also positive and strong.

Table D-50 Correlations among Socioeconomic Variables, Phoenix, AZ MSA (1990 – 2010)

	Population	Emp	Total Earn	Wages & Salaries	GRP	Avg HH Inc	% > $100K	DB1B Enp	Real Oil Price
Population	1								
Employment	0.982	1							
Total Earn	0.975	0.997	1						
Wages & Salaries	0.973	0.997	1.000	1					
GRP	0.978	0.998	0.998	0.998	1				
Avg HH Inc	0.943	0.987	0.991	0.993	0.988	1			
% > $100K	0.880	0.929	0.933	0.938	0.938	0.952	1		
DB1B Enplanements	0.888	0.925	0.914	0.919	0.927	0.929	0.968	1	
Real Oil Price	0.771	0.740	0.725	0.723	0.717	0.685	0.464	0.504	1

The Phoenix, Arizona MSA saw only very slight changes in its population's age distribution over the case study period. Those 19 years old or younger made up about 29 percent of the area's population over the case study period, and there was a slight reduction in the share of those aged between 20 and 34, from 26 percent of the MSA's population to about 21 percent by 2010. Those aged between 35 and 49 remained steady at about 20 percent of the area's population between 1990 and 2010, and those aged between 50 and 64, and those aged 65 or over rose as a total from 25 percent to 30 percent by 2010, as shown in Figure D-45.

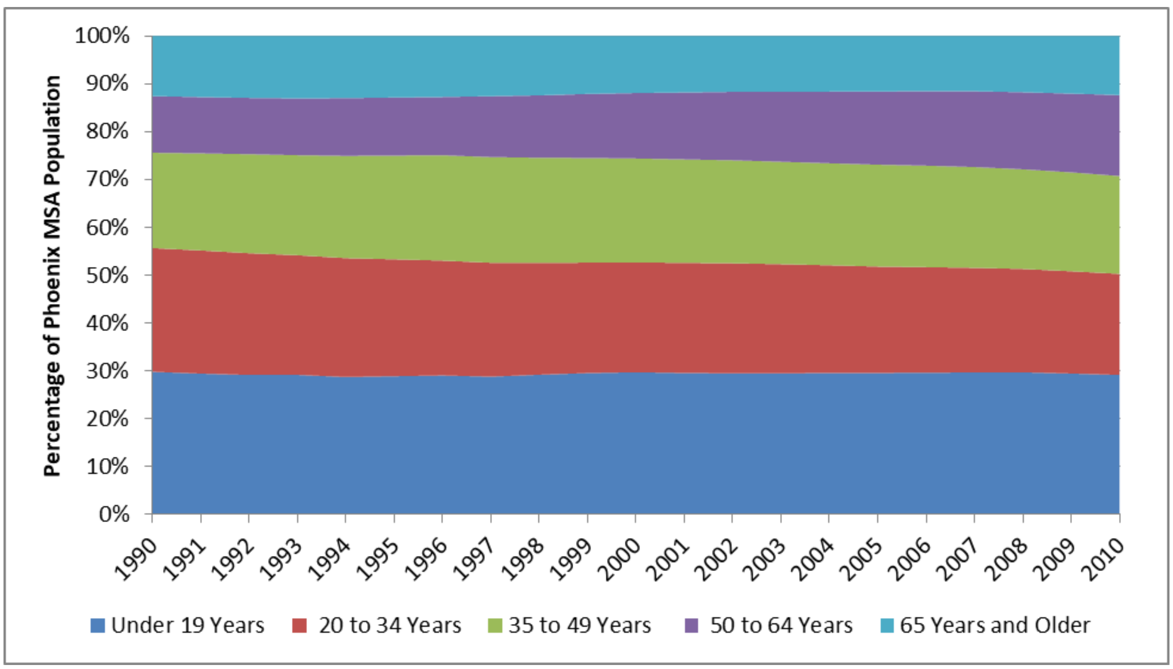

Figure D-45 Population Age Distribution, Phoenix, AZ MSA, 1990 to 2010

Figure D-46 shows the distribution of the Phoenix area's adult population, those 20 and older. During the case study period Phoenix area's adult population distribution shifted somewhat, with those between 20 and 34 years of age declining from around 37 percent of the adult population to around 30 percent in 2010, although the proportion aged 65 or older remaining stable at around 18 percent of the MSA's adult population. The share of the adult population between 50 and 64 years also grew, from about 17 percent of the adult population in 1985 to 24 percent of the population 20 or older by 2010.

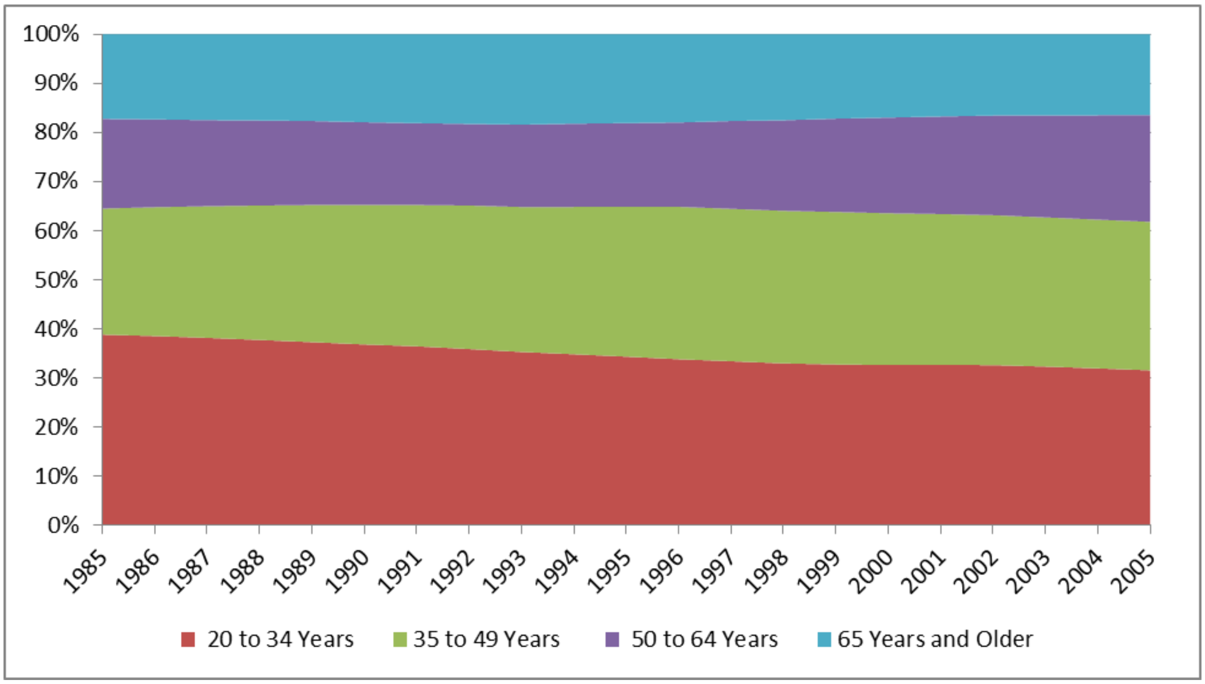

Figure D-46 Adult Population Age Distribution, Phoenix, AZ MSA, 1990 to 2010

Figure D-47 depicts household income distributions for the residents of the Phoenix MSA over the case study sample period. Outside of the growth shown by the share of Phoenix area households shown by the data, other income strata also remained steady over the case study time period, with households earning $30,000 or less per year declining modestly from around 30 percent of Phoenix households to around 25%. As reported in the analysis of aggregate socioeconomic trends over the case study sample period for the Phoenix area (and other case study locations), average household incomes also increased over the period.

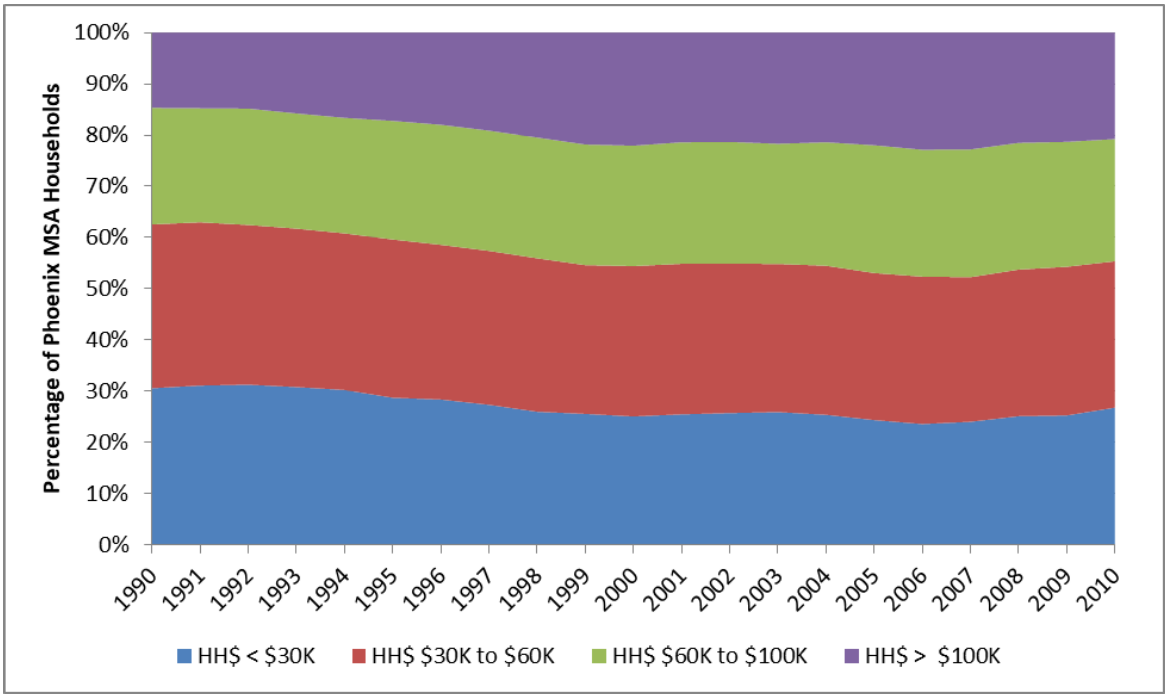

Figure D-47 Household Income Distribution, Phoenix, AZ MSA, 1990 to 2010

Figure D-48 reports specifically the evolution of the percentage of Phoenix MSA households with annual incomes of $100,000 (in 2009 dollars), a group shown as the uppermost band in Figure D-47 above. This percentage hovered around 22 percent of Phoenix area households since 1999, after a period of growth in the share of these high income households, from around 14 percent.

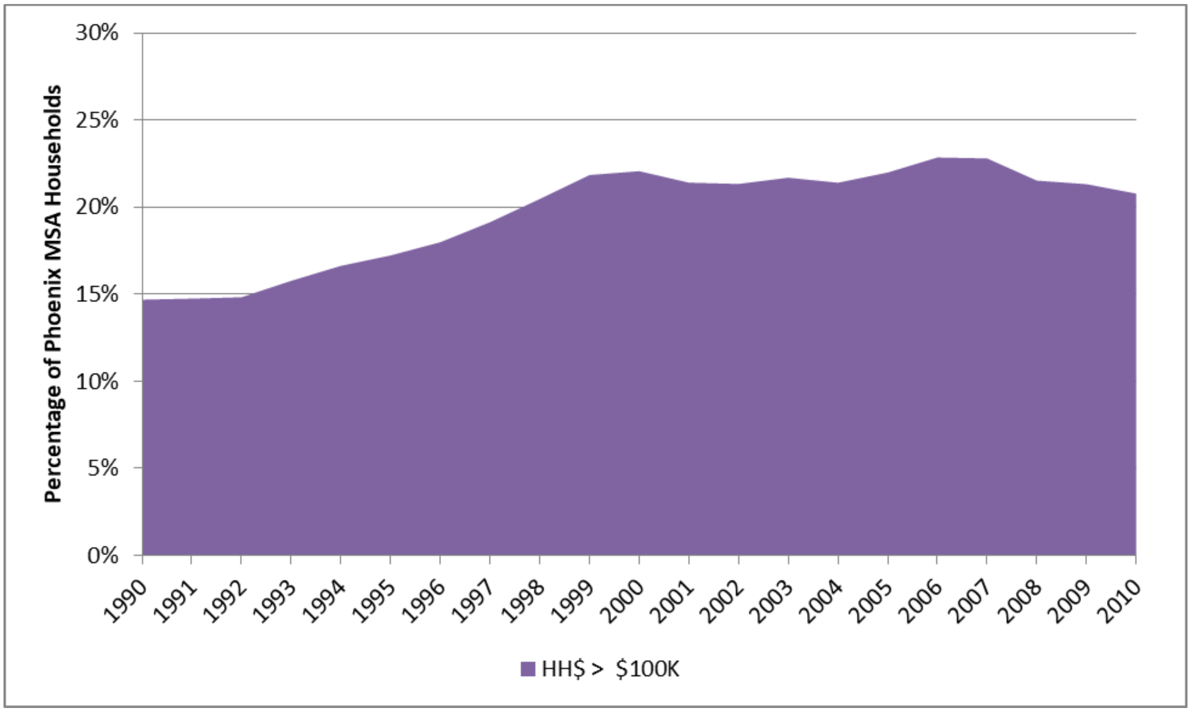

Figure D-48 Percentage of Households with Incomes Exceeding $100,000, Phoenix, AZ MSA 1990 to 2010

Characteristics of the results for the six baseline regressions modeling annual passenger enplanements at Phoenix Sky Harbor International Airport (PHX) from 1990 to 2010 are reported in Table D-51 below. These include:

- The model estimates had similar strong goodness of fit to the actual PHX annual enplanements series, with adjusted R-squared statistics ranging from 0.824 to 0.906, indicating that the baseline models regressing annual enplanements on the annual real oil price and one of the six aggregate socioeconomic variables for the Phoenix MSA accounted for nearly all of the observed variation in the annual PHX enplanements series.
- The estimated regression constant terms varied across the baseline models, although the estimates for the models using Total Earnings in the Phoenix MSA, Annual Wages and Salaries in the MSA, and Gross Regional Product were roughly comparable, ranging from 3.01 to 3.94. All of these estimates were at the 0.005 significance level (or better). The estimated constant term for the model using MSA annual employment was estimated at 1.93, and was statistically significant at the 0.1 level. The constant estimate for the log linear model using Phoenix MSA population took a value of 0.93 and was not statistically significant, while the baseline model using Average Real Household Income took a negative value of -15.24, and was also statistically significant at the 0.005 significance level.
- The coefficients estimated in the baseline PHX models for the oil price variable, which in a log linear regression represents the elasticity of changes in annual PHX enplanements with respect to changes in the oil price, were a bit larger in magnitude than those seen for other case study airports (and also negative, as would be expected) in all baseline

regressions. The estimated coefficient values were within a close range of one another, ranging between -0.201 and -0.313, and all were statistically significant at the 0.005 level. This indicates that over the sample period there was on average a one percent reduction (increase) in the oil price would result in an increase(reduction) in annual PHX enplanements of between 0.2 percent and 0.3 percent, and that these levels of response are precisely estimated.

- The influence of the individual aggregate socioeconomic variables on annual enplanements at PHX is positive and strongly statistically significant across all of the baseline models. The range of coefficient (elasticity) estimates exhibits a pattern somewhat similar to that described for the baseline regressions covering the Phoenix CSA, with one of the larger elasticity estimates (1.93, significant at the 0.005 level) occurring for Phoenix MSA population, a similar value for the regression estimate using Phoenix MSA total employment (1.98, also significant at the 0.005 level), and then estimates between 1.08 and 1.12 for Total Earnings, Wages and Salaries, and GRP (all also significant at the 0.005 level). The elasticity estimate for Average Real Household Income in the Phoenix MSA was the highest of those for the baseline aggregate socioeconomic variables used, estimated at 2.76, and also significant at the 0.005 level.

The PHX regression equation substituting the annual series of the percentage of Phoenix MSA households with incomes of $100,000 or more (a disaggregated socioeconomic variable) for the aggregate socioeconomic variables used in the six baseline regressions is broadly similar to the results from those baseline regressions.

The first comparison of the effects of using disaggregated socioeconomic variables to model annual O&D enplanements at PHX is to replace the individual aggregated socioeconomic variables used individually in the baseline regressions with a disaggregated socioeconomic variable – the percentage in each year of Phoenix MSA households with real household incomes (in 2009 dollars) exceeding $100,000. This estimation resulted in a model constant value of 18.62, which was statistically significant at the 0.001 level of significance. Note that this value is considerably larger than the constant values estimated in the six baseline questions. The estimated coefficient on the Oil Price variable, -0.009, is smaller in magnitude than the coefficients estimated in the baseline equations, and the estimated coefficient value is not statistically significant. The coefficient estimated for the disaggregated socioeconomic variable is comparable to those estimated for each of the individual aggregate variables, and is also statistically significant, at the 0.001 level of significance. The goodness of fit of the overall model to the actual data is also high, with an adjusted R-squared statistic of 0.906, a value comparable to the adjusted R-squared scores for the baseline regressions.

Table D-51 Baseline and Other Regression Results for Models of Annual Enplanements, Phoenix Sky Harbor International Airport 1990 to 2010

Ind. Variable	Statistic	Regression						
		1	2	3	4	5	6	7
Constant	Coefficient	0.93	1.93	3.94	3.81	3.01	-15.24	18.62
	t-Statistic	0.781	1.695	4.590 ***	4.490 ***	3.888 ***	-5.373 ***	53.301 ***
Oil Price	Coefficient	-0.30	-0.31	-0.22	-0.22	-0.20	-0.21	-0.01
	t-Statistic	-4.346 ***	-4.356 ***	-3.658 ***	-3.770 ***	-4.132 ***	-3.099 **	-0.200
Pop	Coefficient	1.93						
	t-Statistic	11.444 ***						
Emp	Coefficient		1.98					
	t-Statistic		11.051 ***					
Total Earnings	Coefficient			1.08				
	t-Statistic			12.313 ***				
Wages and Salaries	Coefficient				1.12			
	t-Statistic				12.610 ***			
Gross Regional Product	Coefficient					1.11		
	t-Statistic					14.867 ***		
Average HH Income	Coefficient						2.76	
	t-Statistic						10.452 ***	
Disaggregated SE Variable: % HH with Income > $100K	Coefficient							1.95
	t-Statistic							13.443 ***
	R Squared	0.852	0.842	0.891	0.898	0.916	0.860	0.916
	Adj R Squared	0.836	0.824	0.878	0.886	0.906	0.844	0.906

* Coefficient estimate significant at 0.05 level ** at 0.01 level *** at 0.001 level

The second comparison of the contribution to model performance of the disaggregated socioeconomic variable – the percentage of Phoenix households with relatively high incomes – is made by adding the disaggregated variable as an additional independent regressor added to the baseline regressions. Results from these regressions are reported for all the case study airports in Table D-52. For PHX and the Phoenix MSA, there is remarkable similarity among the six regression specifications that include the disaggregated socioeconomic variable.

- For each, there is a slight improvement in the model goodness of fit to the actual data, with the adjusted R-squared scores rising slightly to values between 0.901 and 0.907.
- In four of the six regressions adding the disaggregated socioeconomic variable to the baseline regression specifications, the regression constant is much larger in magnitude than the constants estimated for the baseline regressions, and the estimates are statistically significant at the 0.05 significance level or better. In the other specification – that using the Phoenix MSA GRP as the aggregate socioeconomic variable, the constant estimate is smaller – 6.78 – and is significant at the 0.1 level.
- The coefficient estimates for the Oil Price variable are negative (in five of the six regressions) and smaller in magnitude than those estimated in the baseline regressions for PHX. Three of the estimates are significant at the 0.1 level of statistical significance or better – those in regressions using Phoenix MSA Population, Employment, or Gross Regional Product. The remaining coefficient estimates are not statistically significant and are much smaller in magnitude (or positive).
- The estimated coefficients for the aggregate socioeconomic variables used in each equation are much smaller in magnitude compared to the estimates seen in the baseline regressions. The estimates are not statistically significant for three of the specifications, those using Phoenix MSA Total Earnings, Wages and Salaries, and Average Real Household Income as aggregate socioeconomic variables. In the case of the regression using Phoenix MSA Average Real Household Income, the parameter estimate is also negative (-1.61) rather than positive as expected. For the remaining three specifications – those using Population, Employment, or Phoenix MSA Gross Regional Product – the estimates are smaller than the baseline coefficient estimates and statistically significant at the 0.1 level of significance or better.
- Finally, for five of the six specifications the coefficient estimates for the disaggregated socioeconomic variable introduced into these regressions are positive as expected and statistically significant at the 0.1 level of significance or better. In these equations, the coefficient estimates range from 1.28 to 3.05, with the relatively high 3.05 value occurring in the specification that includes the aggregate socioeconomic variable Average Real Household Income, which, as noted above, had an unanticipated negative coefficient. In the case of the specification using Phoenix MSA Gross Regional Product as the aggregate socioeconomic variable, the coefficient estimate on the disaggregated socioeconomic variable is relatively small (0.48) and is not statistically significant.

For the PHX and Phoenix MSA data, the addition of the disaggregated socioeconomic variable does not improve overall model fit very much. While the Oil Price variable and the individual aggregate socioeconomic variables are consistently significant in the baseline regressions, and replacing the aggregate socioeconomic variables with the disaggregated one gives similar results, adding the disaggregated socioeconomic variable as a third independent regressor muddles this consistent baseline model outcome in most cases.

Table D-52 Regression Results for Models of Annual Enplanements, Phoenix Sky Harbor International Airport, 1990 to 2010, that Include both Aggregate and Disaggregated Socioeconomic Variables

Ind. Variable	Statistic	Regression					
		1	2	3	4	5	6
Constant	Coefficient	12.16	12.88	16.26	16.55	6.78	38.33
	t-Statistic	3.791 ***	3.791 ***	2.290 *	1.896 *	1.134	2.458 *
Oil Price	Coefficient	-0.124	-0.124	-0.044	-0.039	-0.157	0.117
	t-Statistic	-1.573	-1.546	-0.383	-0.290	-1.835 *	1.072
Pop	Coefficient	0.716					
	t-Statistic	1.736 *					
Emp	Coefficient		0.694				
	t-Statistic		1.697				
Total Earnings	Coefficient			0.175			
	t-Statistic			0.332			
Wages and Salaries	Coefficient				0.158		
	t-Statistic				0.237		
Gross Regional Product	Coefficient					0.844	
	t-Statistic					1.984 *	
Average HH Income	Coefficient						-1.615
	t-Statistic						-1.264
Disaggregated SE Variable: % HH with Income > $100K	Coefficient	1.281	1.325	1.645	1.684	0.480	3.051
	t-Statistic	3.114 **	3.346 **	1.746 *	1.465	0.636	3.471 **
	R Squared	0.923	0.923	0.915	0.916	0.921	0.919
	Adj R Squared	0.909	0.909	0.901	0.901	0.907	0.905

* Coefficient estimate significant at 0.05 level ** at 0.01 level *** at 0.001 level

To assess the forecasting performance of these models, we limit the analysis to consideration of a single pair of regressions, those that use the Phoenix MSA Gross Regional Product as the aggregate socioeconomic variable. As mentioned above, it is not possible to obtain the Woods and Poole projections that would have been available for the years 2011 through 2015, which would have been included as part of the Woods and Poole 2012 data release. To create a test projection of values for Phoenix MSA GRP for those years, we use the percentage annual growth terms from the Woods and Poole 2016 data release for the years 2015 through 2019. With those percentage values for annual GRP growth, notional projected values for Phoenix GRP can be calculated for 2011 through 2015. To create notional projected values for the disaggregated socioeconomic variable, the percentage of Phoenix MSA households with 2009 dollar denominated income exceeding $100,000 (in 2009 dollars), a similar procedure was used.

Forecast values for the Oil Price dependent variable are also needed to calculate out of sample projections. These were obtained for the years 2011 to 2015 from the data reported in the FAA's 2012 Annual Aerospace Forecast document. These national values are used for each of the case study airports out of sample forecast exercises.

Four out of sample model forecasts were calculated for the baseline model using Phoenix MSA GRP and the model that also includes the disaggregated socioeconomic variable. The first three are based on the out of sample forecasts described above. For the first forecast, the projected values – based on the percentage increases projected for the years 2015 through 2019 – for GRP, and the percentage of Phoenix household incomes exceeding $100,000 in the case of the model using the disaggregated socioeconomic variable as input, are used as described above. For the second, the values for percentage annual increases over the projection period (2011 to 2015) are increased by 50 percent, indicating a more rapidly growing Phoenix economy. For the third PHX projection, the values for percentage annual increases are reduced by 50 percent, representing a projection for a more slowly growing economic environment in the Phoenix MSA. For the fourth projection, the actual values observed in 2011 through 2015 for GRP and the share of Phoenix MSA households with incomes exceeding $100,000 are used to drive the "forecast" of PHX annual O&D enplanements.

Using the projections published by Woods and Poole for 2015 and subsequent years for simulating a range of forecast scenarios based on inputs that would have been available to analysts in 2011 or 2012, we can compare the forecasting performance of the baseline regressions for PHX annual O&D enplanements using the annual oil price variable and the Phoenix MSA GRP, with those of the regression parameters based additionally on input from the disaggregated socioeconomic variable. The independent variable data and projections are shown in Table D-53 below. The projections for the Phoenix MSA variables reported in the Woods and Poole 2016 data have annual GRP growing at a rate of around 3.2 percent, and the percentage of Phoenix MSA households with incomes exceeding $100,000 growing at about 1.7 percent initially, remaining at about that level over the forecast period. The annual percentage changes used to calculate the Oil Price projected values for 2011 through 2015 are those implied by the Oil Price projections reported in the 2012 FAA Aerospace Forecast, which relies on projected values for the years following 2010. Thus we have four "forecast" scenarios for the two regression specifications for annual PHX O&D enplanements:

- Projected values for the socioeconomic independent variables based on the annual percentage changes in the variables, as reported by Woods and Poole for the five years beginning in 2015,
- Projected values based on annual percentage changes that are 1.5 times these values,
- Projected values based on annual percentage changes that are half these values, and]
- Projected values that equal the actual values for the independent variables that occurred in the years 2011 through 2015.

Table D-53 Inputs to Out-of-Sample PHX O&D Enplanement Projections

PHX Inputs to Out-of-Sample Enplanement Forecasts	PHX					
	Sources	2011	2012	2013	2014	2015
Oil Price Forecast	*2012 FAA Aero Forecast*	$84.32	$92.71	$97.16	$100.65	$104.64
GRP Forecast 1 (Medium)	*W&P 2016 %*	$200,223.7	$206,613.7	$213,101.2	$219,719.2	$226,480.1
GRP Forecast 2 (High)	*1.5 x F'cast 1*	$203,389.2	$213,125.7	$223,163.8	$233,559.5	$244,339.6
GRP Forecast 3 (Low)	*0.5 x F'cast 1*	$197,058.1	$200,202.6	$203,345.8	$206,503.3	$209,680.4
HHInc% Forecast 1 (Medium)	*W&P 2016 %*	21.1%	21.5%	21.9%	22.3%	22.7%
HHInc% Forecast 2 (High)	*1.5 x F'cast 1*	21.3%	21.9%	22.5%	23.1%	23.7%
HHInc% Forecast 3 (Low)	*0.5 x F'cast 1*	21.0%	21.1%	21.3%	21.5%	21.7%
	2010					
Actual Oil Price	$83.99	$101.87	$100.93	$100.49	$92.02	$48.39
Actual GRP	$193,892.6	$196,521.1	$202,798.3	$205,867.0	$213,481.5	$220,708.0
Actual HHInc%	20.8%	19.7%	19.6%	19.6%	19.6%	19.8%

The forecasting or projection results for the baseline regression equation for annual enplanements at PHX are shown in Figure D-49 below.

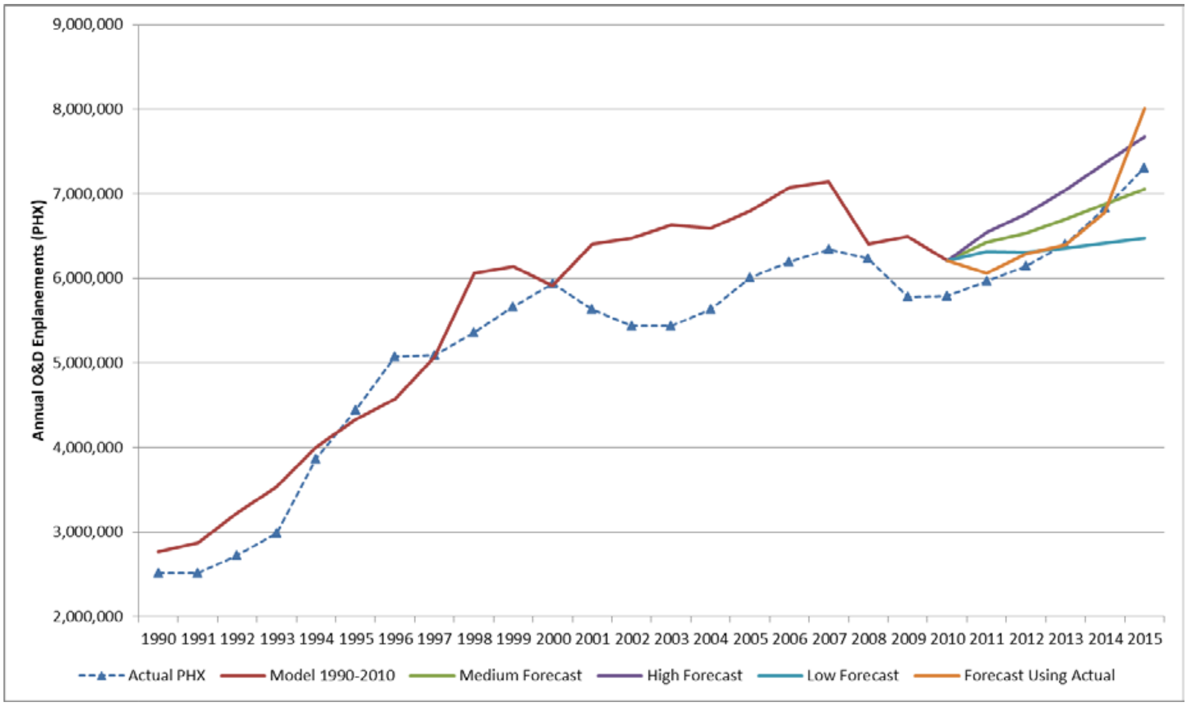

Figure D-49 Baseline Model Forecasts of PHX Annual Enplanements, 2011 to 2015, Using Phoenix MSA Gross Regional Product, 1990 to 2010

The results for the regression equation using the disaggregated socioeconomic variable along with Phoenix MSA GRP are shown in Figure D-50 below.

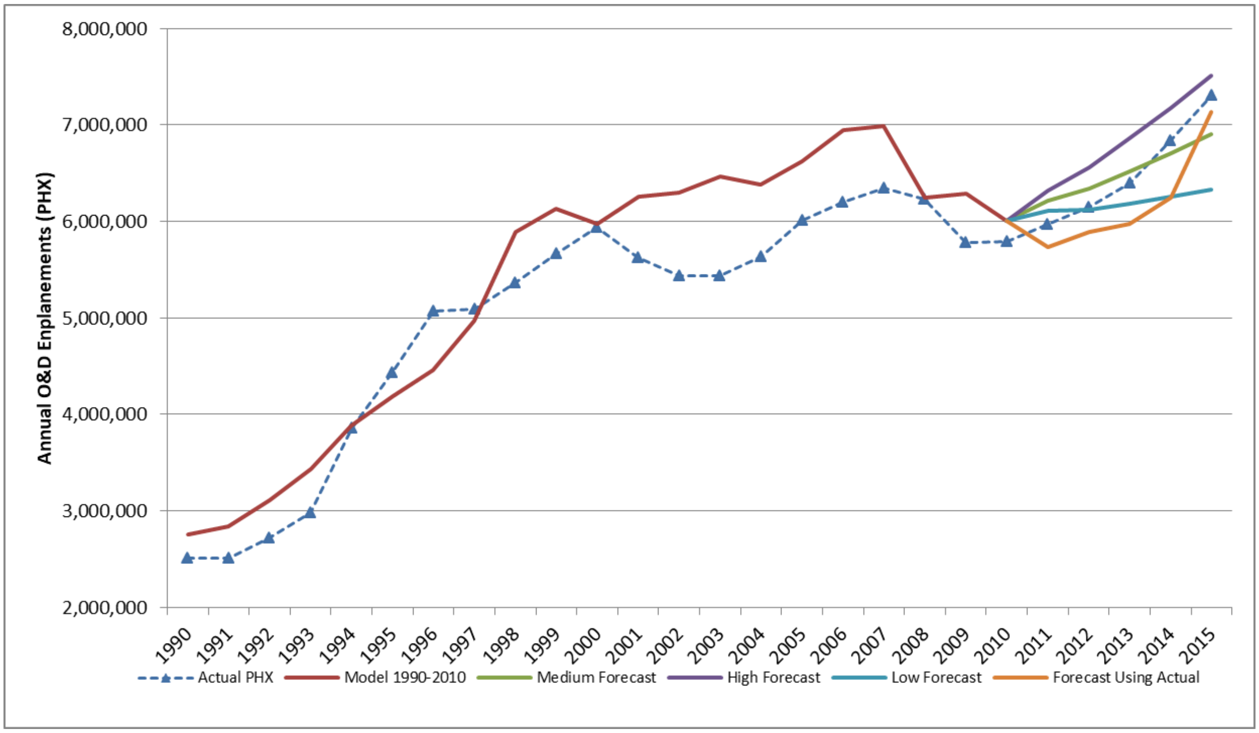

Figure D-50 Disaggregated SE Variable Model Forecasts of PHX Annual Enplanements, 2011 to 2015, Using Phoenix MSA GRP and Share of Phoenix Households with Incomes exceeding $100,000, 1990 to 2010

Enplanements at PHX from 2011 to 2015 were fairly well forecast by the four forecasting scenarios using both the baseline and alternative model specifications. In fact, the two forecast charts (for the baseline and alternative models) look remarkably similar, suggesting that even though the regression model estimates differ in some respects between the two specifications, they are using similar information from the available data.[11] This is also suggested by the high R-squared scores for both model specifications, without and with the disaggregated socioeconomic variable. The following tables D-54 to D-56 quantify this, reporting the percentage error for each of the forecasts and the Root Mean Squared Error (RMSE) for each forecast, a calculation which can be used to compare the accuracy of the baseline model forecasts and the forecasts calculated from the model also using the disaggregated socioeconomic variable.

[11] The baseline regression equation is ln(PHX_Enp) = 3.01 - 0.123ln(OilPrice) + 1.11ln(GRP), with adjusted R squared of 0.906, while the estimated equation that includes the disaggregated socioeconomic variable is ln(PHX_Enp) = 6.78 - 0.157ln(OilPrice) + 0.84ln(GRP) + 0.48ln(%HHInc>$100K), with adjusted R squared of 0.907.

Table D-54 Percentage Forecast Errors, Baseline Model Projections of PHX Annual O&D Enplanements, 2011 to 2015

PHX (using GRP)	Percentage Forecast Error (Baseline Regression with Oil Price and GRP)			
	Medium	High	Low	Actual
2010	--	--	--	--
2011	7.7%	9.6%	D-8%	1.6%
2012	6.3%	10.0%	2.6%	2.3%
2013	4.6%	10.1%	-0.7%	0.0%
2014	0.6%	7.6%	-6.1%	-0.8%
2015	-3.4%	5.1%	-11.3%	9.7%

The mix of negative and positive percentage errors in Tables D-54 and D-55 reflect the fact that the model-based projections of PHX enplanements create a range for PHX enplanements that contain the actual enplanement results for those years. As can be seen in the figures and the tables of percentage errors, the model forecasts using the actual independent variable values as forecasting inputs do not differ from the forecasts based on the projected values for the independent variables, in spite of the significant drop in Oil Prices in 2015.

Table D-55 Percentage Forecast Errors, Disaggregated SE Model Projections of PHX Annual O&D Enplanements, 2011 to 2015

PHX (using GRP)	Percentage Forecast Error (Specification including Disaggregated SE Variable)			
	Medium	High	Low	Actual
2010	--	--	--	--
2011	4.2%	6.0%	2.4%	-3.8%
2012	3.2%	6.8%	-0.3%	-4.2%
2013	1.8%	7.1%	-3.4%	-6.7%
2014	-1.9%	5.0%	-8.4%	-8.7%
2015	-5.5%	2.9%	-13.3%	-2.3%

The final table showing forecasting performance of the estimated PHX regression models reports the Root Mean Squared Error (RMSE) of the different forecasts for the two model specifications. This statistic can be used to compare the accuracy of two or more of these forecasts, with a smaller RMSE value indicating a more accurate forecast. As can be seen, the inclusion of the additional disaggregated socioeconomic variable in the model specification provided some improvement in the accuracy of two of the five year forecasting scenarios, but for the two other forecast scenarios, the baseline model specification that did not include the disaggregated socioeconomic independent variable provided more accurate forecasts.

Table D-56 Root Mean Squared Error (RMSE) for PHX Baseline and Disaggregated SE Variable Model Forecasts of Annual PHX O&D Enplanements, 2011 to 2015

PHX (using GRP)	Root Mean Squared Error		% Difference BL v Alt
	Baseline	Alternative	
Medium Forecast	318,590	241,766	24%
High Forecast	554,205	367,045	34%
Low Forecast	447,829	518,924	-16%
Actual Values	326,203	369,242	-13%

i. Baltimore-Washington Regional Airport System (BWI, DCA, IAD)

The eighth and final case study airport region is the region served by the three Washington DC area airports, Ronald Reagan Washington National Airport (DCA), Washington Dulles International Airport (IAD), and Baltimore-Washington International Thurgood Marshall Airport (BWI). The region served by the three airport system is made up by the Baltimore-Washington Combined Statistical Area (CSA). Figure D-51 depicts the indexed series for the region's socioeconomic variables of interest.

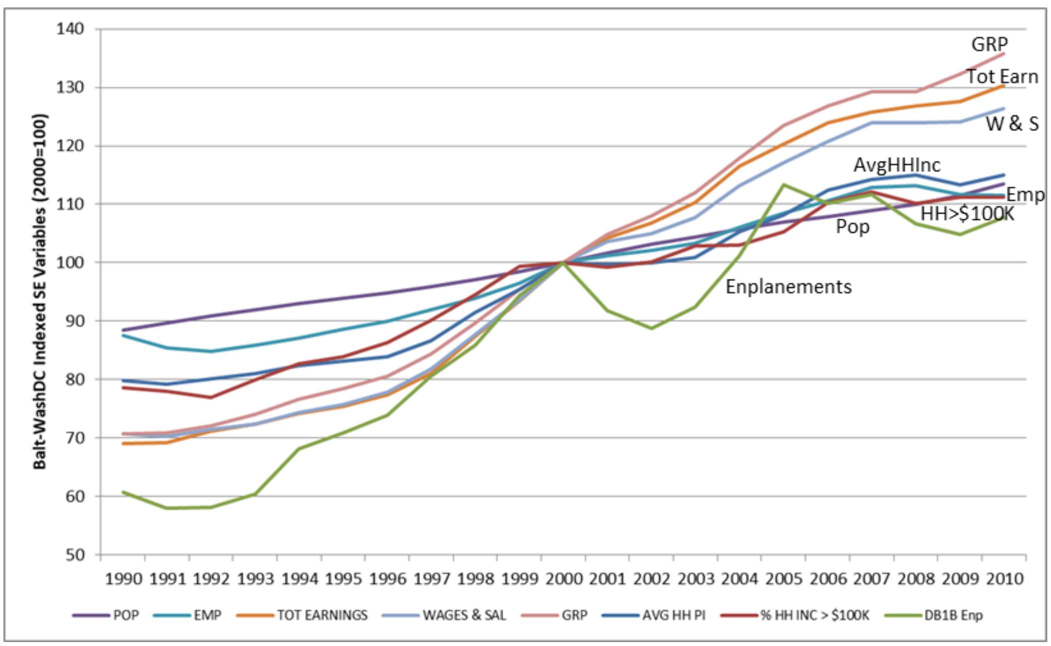

Figure D-51 Baltimore-Washington CSA Indexed Socioeconomic Variables, 1990 to 2010, 2000=100

Table D-57 reports socioeconomic and other variables for the Baltimore-Washington CSA for the years from 1990 to 2010.

Table D-57 Values of Socioeconomic Variables, Baltimore-Washington CSA (1990-2010)

	Year	Baltimore Washington Airport System (BWI/DCA/IAD)								
		Annual O&D Enplanements	Real Oil Price	Regional Population (000s)	Regional Employment (000s)	Regional Total Earnings ($M)	Regional W&S Earnings ($M)	Regional Product (GRP) ($M)	Regional Avg HH Income	% Regional HHs with Income > $100K
	1990	12,476,709	$40.16	7,089.4	4,613.6	$206,643.1	$160,076.5	$314,863.6	$96,246	25.4%
	1991	11,931,371	$33.20	7,188.4	4,500.1	$206,989.6	$158,813.3	$315,204.4	$95,613	25.2%
	1992	11,939,424	$31.33	7,282.6	4,469.9	$212,875.7	$161,645.9	$320,616.0	$96,646	24.8%
	1993	12,428,273	$26.96	7,367.3	4,526.8	$216,588.6	$163,769.7	$329,685.1	$97,788	25.8%
	1994	13,991,513	$25.31	7,451.2	4,592.2	$221,765.0	$168,083.9	$341,174.6	$99,414	26.7%
	1995	14,570,808	$27.16	7,525.9	4,675.0	$225,701.4	$171,242.8	$348,991.4	$100,271	27.1%
	1996	15,175,041	$31.73	7,603.7	4,743.2	$231,394.4	$176,072.8	$358,699.5	$101,256	27.9%
	1997	16,555,886	$27.80	7,688.3	4,847.4	$242,319.0	$185,063.6	$375,598.3	$104,444	29.1%
	1998	17,653,146	$17.87	7,777.7	4,953.9	$261,351.1	$198,416.2	$399,432.0	$110,273	30.5%
	1999	19,388,282	$25.03	7,889.9	5,084.4	$279,473.5	$211,715.9	$424,675.9	$115,091	32.1%
	2000	20,548,395	$38.86	8,014.1	5,271.0	$298,961.0	$226,206.6	$445,151.1	$120,610	32.3%
	2001	18,876,837	$29.99	8,146.5	5,331.6	$311,551.9	$234,389.8	$466,875.4	$120,150	32.1%
	2002	18,242,847	$31.83	8,264.0	5,385.1	$319,367.1	$237,419.6	$480,942.7	$120,573	32.4%
	2003	18,992,432	$36.38	8,363.3	5,449.6	$329,732.6	$243,806.4	$498,337.6	$121,715	33.2%
	2004	20,782,115	$45.87	8,470.4	5,591.1	$348,150.5	$256,172.0	$525,031.4	$127,008	33.3%
	2005	23,277,051	$60.44	8,573.5	5,718.1	$359,596.0	$264,885.0	$549,959.9	$130,421	34.0%
	2006	22,648,490	$70.72	8,648.8	5,828.7	$370,509.8	$273,145.3	$564,850.5	$135,548	35.6%
	2007	22,948,023	$78.22	8,724.5	5,947.3	$376,153.6	$280,284.0	$575,105.6	$137,838	36.2%
	2008	21,911,075	$103.82	8,817.7	5,964.6	$379,162.3	$280,308.0	$575,263.7	$138,747	35.6%
	2009	21,542,546	$66.42	8,942.9	5,882.5	$381,480.4	$280,785.3	$588,932.3	$136,704	35.9%
	2010	22,140,825	$83.99	9,088.9	5,879.3	$389,567.9	$285,829.1	$604,278.6	$138,616	35.9%
Used to assess out-of-sample forecasts	*2011*	*22,105,283*	*$101.87*	*9,226.9*	*5,991.1*	*$396,420.8*	*$288,989.3*	*$607,780.9*	*$138,993*	*35.8%*
	2012	*21,981,813*	*$100.93*	*9,348.1*	*6,073.4*	*$400,512.1*	*$291,866.5*	*$611,636.2*	*$139,908*	*36.2%*
	2013	*21,919,338*	*$100.49*	*9,463.6*	*6,161.2*	*$399,431.7*	*$290,675.0*	*$613,427.5*	*$136,063*	*36.4%*
	2014	*22,402,769*	*$92.02*	*9,546.6*	*6,219.8*	*$406,639.7*	*$294,978.0*	*$624,734.9*	*$137,963*	*36.4%*
	2015	*23,731,344*	*$48.39*	*9,662.3*	*6,328.4*	*$416,861.6*	*$303,177.7*	*$640,644.3*	*$139,440*	*36.7%*

Table D-58 reports the correlations between these socioeconomic variables over the sample period. As with the other seven case study regions, the evolution of the socioeconomic variables over the sample period is strongly and positively correlated. This includes the correlations between the Baltimore-Washington CSA socioeconomic variables and either oil prices or annual O&D enplanements at the three airports making up the Baltimore-Washington system, which are also positive and strong.

Table D-58 Correlations among Socioeconomic Variables, Baltimore-Washington CSA (1990 – 2010)

	Population	Emp	Total Earn	Wages & Salaries	GRP	Avg HH Inc	% > $100K	DB1B Enp	Real Oil Price
Population	1								
Employment	0.981	1							
Total Earn	0.990	0.995	1						
Wages & Salaries	0.988	0.997	1.000	1					
GRP	0.994	0.994	0.999	0.998	1				
Avg HH Inc	0.984	0.996	0.995	0.997	0.994	1			
% > $100K	0.973	0.982	0.978	0.982	0.977	0.984	1		
DB1B Enplanements	0.938	0.958	0.951	0.956	0.949	0.961	0.979	1	
Real Oil Price	0.780	0.805	0.787	0.786	0.797	0.804	0.716	0.685	1

The Baltimore-Washington CSA saw modest changes in its population's age distribution over the case study period. Those 19 years old or younger consistently made up about 27 percent of the population over the case study period, and there was a reduction in the share of those aged between 20 and 34, from 27 percent of the CSA's population to about 21 percent by 2010. There was a modest increase in the share of those aged between 50 and 64, from about 13 percent to around 19 percent by 2010, as shown in Figure D-52. Residents aged 65 or more remained about 10 percent of the region's population throughout the case study sample period.

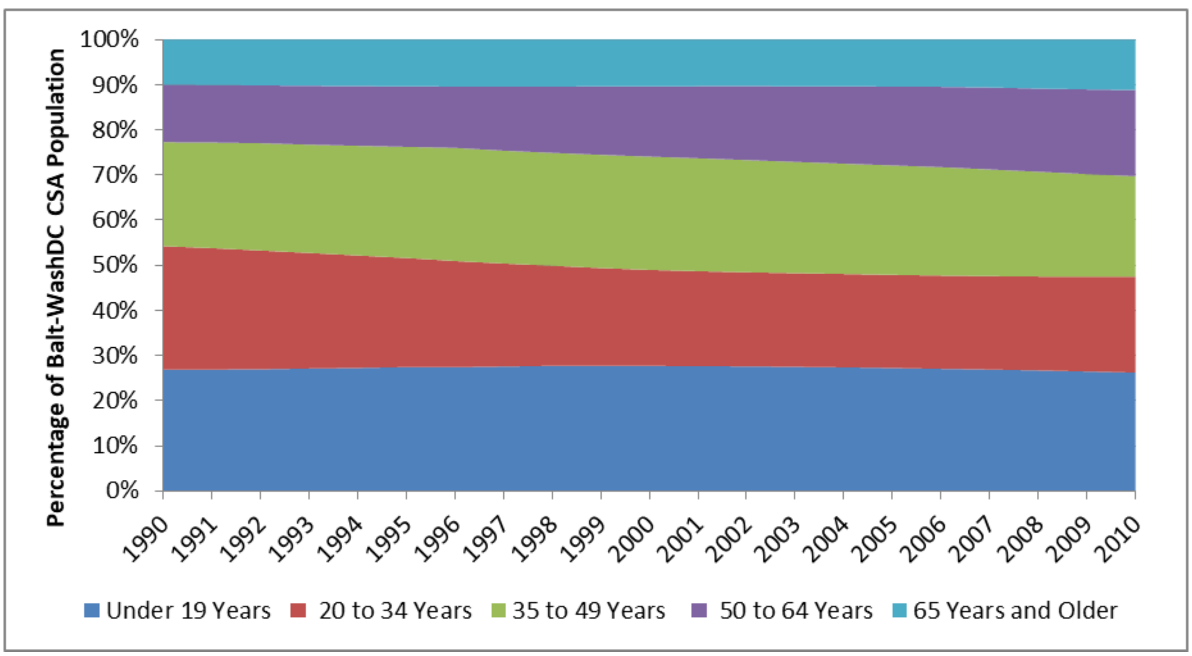

Figure D-52 Population Age Distribution, Baltimore-Washington CSA, 1990 to 2010

Figure D-53 shows the distribution of the Baltimore-Washington area's adult population, those 20 and older. During the case study period the Baltimore-Washington area's adult population distribution shifted somewhat, with those between 20 and 34 years of age declining from around 37 percent of the adult population to around 29 percent in 2010, although the proportion aged 65 or older was stable at around 15 percent of the CSA's adult population. The share of the adult population between 50 and 64 years also grew, from about 17 percent of the adult population in 1985 to over 26 percent of the Baltimore-Washington CSA's population 20 or older by 2010.

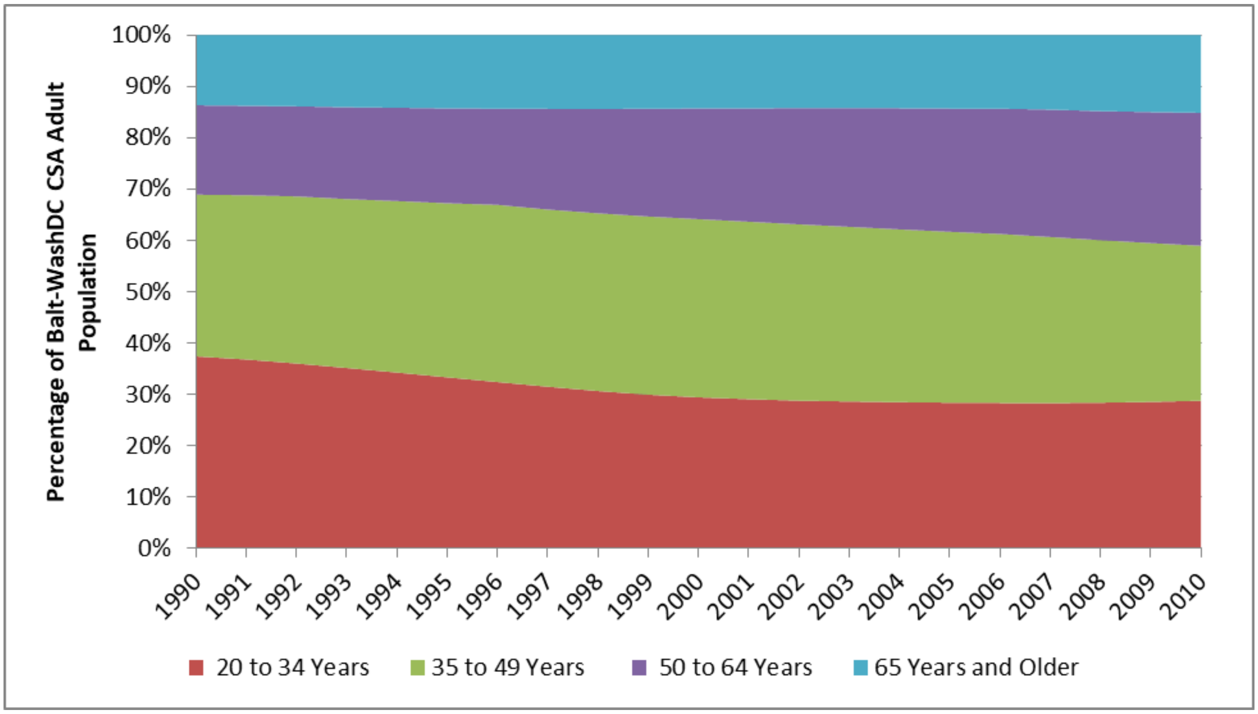

Figure D-53 Adult Population Age Distribution, Baltimore-Washington CSA, 1990 to 2010

Figure D-54 depicts household income distributions for the residents of the Baltimore-Washington CSA over the case study sample period. The figure shows a large and growing share of Baltimore-Washington area households – from around 53 percent in 1990 to over 30 percent in 2010 – had incomes exceeding $100,000, a higher share than the other communities included in the case study analysis. The other lower income strata each declined modestly over the study period. As reported in the analysis of aggregate socioeconomic trends over the case study sample period for the combined statistical area (and other case study locations), average household incomes also increased over the period.

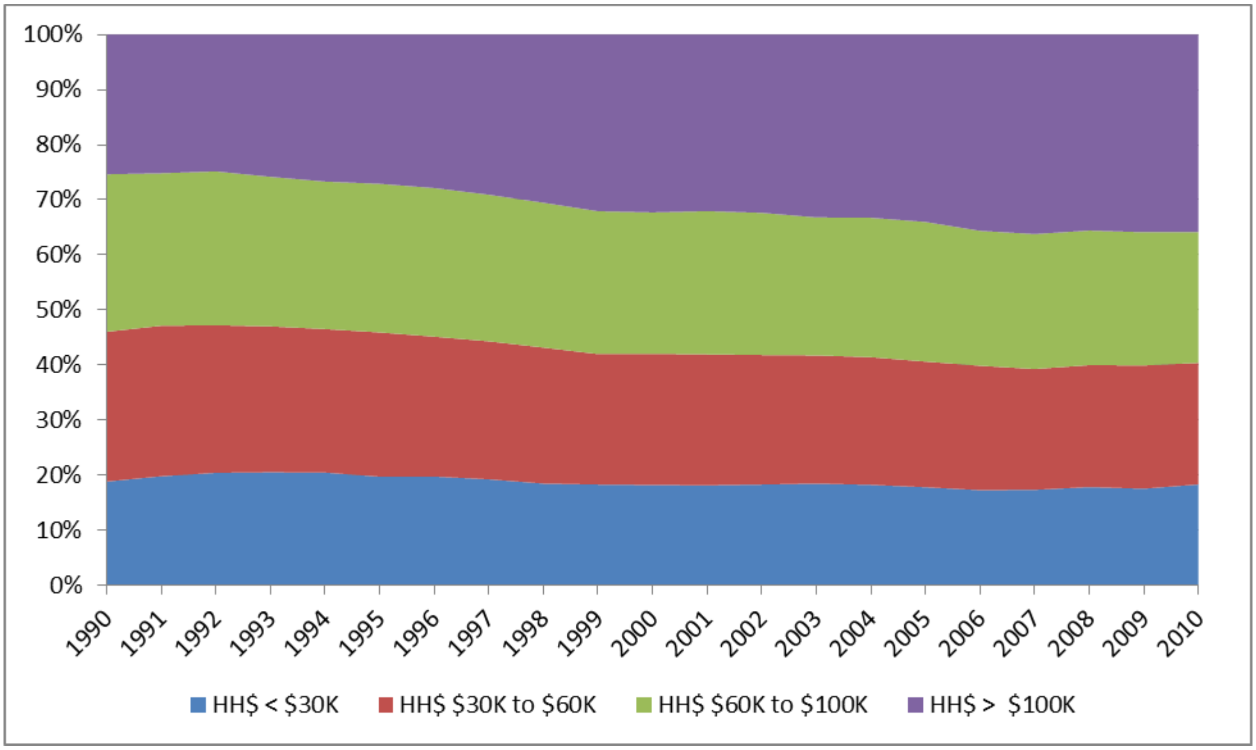

Figure D-54 Household Income Distribution, Baltimore-Washington CSA, 1990 to 2010

Figure D-55 reports specifically the evolution of the percentage of Baltimore-Washington CSA households with annual incomes of $100,000 (in 2009 dollars). This percentage has slowly risen from around 25 percent of the CSA's households in 1990 to over 35 percent of households by 2010.

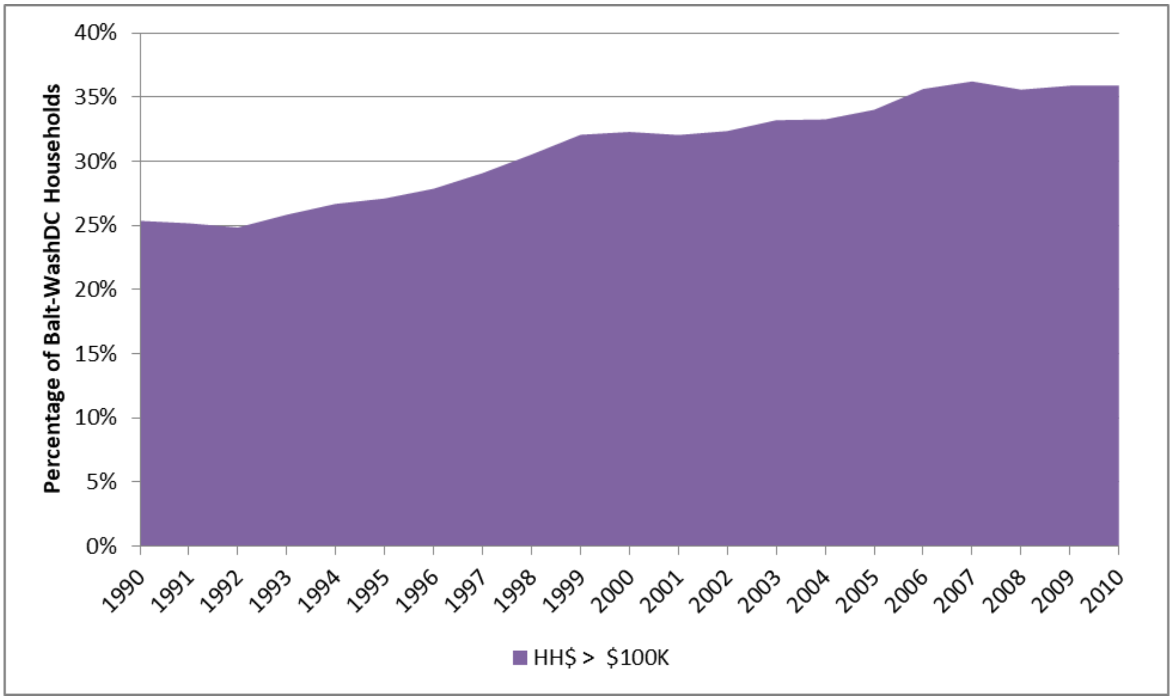

Figure D-55 Percentage of Households with Incomes Exceeding $100,000, Baltimore-Washington DC CSA, 1990 to 2010

Results for the baseline regressions modeling annual passenger origin and destination (O&D) enplanements at the three airports of the Washington, DC airport system (Reagan National Airport (DCA), Dulles International Airport (IAD), and Baltimore-Washington International Airport (BWI), henceforth WDC) from 1990 to 2010 are shown in the first six columns of Table D-59. These results are summarized below.

- The baseline model estimates based on the individual regional aggregate socioeconomic variables had very good results for goodness of fit to the actual WDC annual O&D enplanements series, with adjusted R-squared statistics ranging from 0.847 to 0.932, indicating that the baseline models regressing annual enplanements on the annual real oil price and one of the six aggregate socioeconomic variables for the Baltimore-Washington CSA accounted for a large majority of the observed variation in the annual WDC enplanements series.
- The estimated regression constant terms varied across the baseline models, which was also true for some of the other case study airports. Across the six baseline regressions using an individual aggregate socioeconomic variable, the constant term estimates for the WDC annual enplanements series ranged between -12.45 and 3.21, and all of these estimates were significant at the 0.05 significance level, or better.
- While the coefficients estimated in the baseline WDC models for the oil price variable, which in a log linear regression represents the elasticity of changes in annual WDC enplanements with respect to changes in the oil price, were relatively small in all baseline regressions, the estimated coefficients in each equation were statistically significant at the 0.05 significance level or better (with the exception of the regression using the Baltimore-

Washington CSA population as the aggregate socioeconomic variable, which was significant at the 0.1 significance level). The coefficient estimates for the six equations ranged from -0.87 to -0.160.

- The influence of the individual aggregate socioeconomic variables on annual enplanements at Baltimore-Washington airports was positive and strongly statistically significant across all of the baseline models. The range of coefficient (elasticity) estimates exhibits a pattern that occurs at other case study airport baseline regression, with the largest elasticity estimate (3.28, significant at the 0.005 level) occurring in the regression specification using Baltimore-Washington CSA population, a slightly smaller value for the regression estimate using Baltimore-Washington CSA total employment (2.74, also significant at the 0.005 level), and then parameter estimates between 1.10 and 1.18 for Baltimore-Washington CSA Total Earnings, Wages and Salaries, and GRP (all also significant at the 0.005 level). The elasticity estimate for the baseline regression using Average Real Household Income in the Baltimore-Washington CSA was slightly greater than the estimates for these three, estimated at 1.99, and is also significant at the 0.005 level.

The first comparison of the effects of using disaggregated socioeconomic variables to model annual O&D enplanements at the WDC airports is to replace the individual aggregated socioeconomic variables used individually in the six baseline regressions with a disaggregated socioeconomic variable – the percentage in each year of Baltimore-Washington CSA households with real household incomes (in 2009 dollars) exceeding $100,000. Coefficient estimates for this specification are shown as Regression 7 (Disagg) in Table D-57, with the disaggregated socioeconomic variable labeled "% HH with Income > $100K." This estimation resulted in a model constant value of 18.91, which was highly statistically significant. Note that this value is considerably larger than the constant values estimated in the six baseline questions, reflecting the smaller magnitude of the values taken by this independent variable. The estimated coefficient on the Oil Price variable, -0.029, is smaller in magnitude than the coefficients estimated in the baseline equations, and the estimated coefficient value is not statistically significant. The coefficient estimated for the disaggregated socioeconomic variable is comparable to those estimated for the aggregate variables, and is highly significant, at the 0.005 level of significance. The goodness of fit of the overall model to the actual data is also high, with an adjusted R-squared statistic of 0.949, a value modestly greater than the adjusted R-squared scores for the baseline regressions.

Table D-59 Baseline and Other Regression Results for Models of Annual Enplanements, Baltimore-Washington Airport System, 1990 to 2010

Ind. Variable	Statistic	Regression						
		1	2	3	4	5	6	7
Constant	Coefficient	-12.45	-6.20	3.21	2.52	2.47	-6.08	18.91
	t-Statistic	-3.766 ***	-3.624 ***	2.749 **	2.202 *	2.066 *	-3.602 ***	87.867 ***
Oil Price	Coefficient	-0.09	-0.16	-0.10	-0.10	-0.11	-0.13	-0.03
	t-Statistic	-1.460	-3.521 ***	-1.942 *	-2.146 *	-2.202 *	-3.094 **	-1.004
Pop	Coefficient	3.28						
	t-Statistic	8.474 ***						
Emp	Coefficient		2.74					
	t-Statistic		12.718 ***					
Total Earnings	Coefficient			1.10				
	t-Statistic			10.609 ***				
Wages and Salaries	Coefficient				1.18			
	t-Statistic				11.433 ***			
Gross Regional Product	Coefficient					1.12		
	t-Statistic					10.998 ***		
Average HH Income	Coefficient						1.99	
	t-Statistic						12.823 ***	
Disaggregated SE Variable: % HH with Income > $100K	Coefficient							1.80
	t-Statistic							16.678 ***
	R Squared	0.863	0.930	0.914	0.923	0.914	0.939	0.954
	Adj R Squared	0.847	0.922	0.905	0.915	0.905	0.932	0.949

* Coefficient estimate significant at 0.05 level ** at 0.01 level *** at 0.001 level

The second approach to comparing the contribution to model performance of the disaggregated socioeconomic variable – the percentage of Baltimore-Washington households with relatively high incomes – to the baseline results is made by adding the disaggregated variable as an additional independent regressor variable to individual aggregate socioeconomic variables used in the baseline regressions. Results from these regressions are reported for all the case study airports in Table D-60. For the WDC airports and the Baltimore-Washington CSA, there is remarkable similarity among the six regression specifications that include the disaggregated socioeconomic variable.

- For each, there is a slight improvement in the model goodness of fit to the actual data, with the R-squared scores rising slightly to values between 0.946 and 0.952.
- In each of the six regressions adding the disaggregated socioeconomic variable to the baseline regression specifications, the regression constant is much larger in magnitude than the constants estimated for the baseline regressions, and the estimates are statistically significant at the 0.01 significance level or better.
- The coefficient estimates for both the Oil Price variable and the aggregate socioeconomic variables used in each equation are much smaller in magnitude compared to the baseline regressions, and the estimates are not statistically significant. In these equations, the influence of these variables on the estimates of the annual enplanement values is modest compared to the baseline equations.
- Finally, the coefficient estimates for the disaggregated socioeconomic variable introduced into these regressions fall within a very tight range for all six specifications (2.12 to 2.24), and the estimates are significant at the 0.005 significance level.

For the Baltimore-Washington CSA data, the addition of the disaggregated socioeconomic variable does not improve the already high baseline model goodness of fit meaningfully. While the Oil Price variable and the individual aggregate socioeconomic variables are consistently significant in the baseline regressions, the addition of the disaggregated socioeconomic variable reverses this statistical significance, which is probably to some extent a reflection of the high correlation between the disaggregated socioeconomic variable and the aggregate variables over the sample period.

Table D-60 Regression Results for Models of Annual Enplanements, Baltimore-Washington Airport System, 1990 to 2010, that Include both Aggregate and Disaggregated Socioeconomic Variables

Ind. Variable	Statistic	Regression					
		1	2	3	4	5	6
Constant	Coefficient	27.40	23.53	22.43	22.76	23.02	23.87
	t-Statistic	2.646 **	2.646 **	5.525 ***	4.717 ***	5.043 ***	2.623 **
Oil Price	Coefficient	-0.001	-0.002	-0.008	-0.008	-0.004	-0.006
	t-Statistic	-0.030	-0.041	-0.222	-0.202	-0.100	-0.107
Pop	Coefficient	-0.898					
	t-Statistic	-1.373					
Emp	Coefficient		-0.507				
	t-Statistic		-0.519				
Total Earnings	Coefficient			-0.250			
	t-Statistic			-0.868			
Wages and Salaries	Coefficient				-0.281		
	t-Statistic				-0.798		
Gross Regional Product	Coefficient					-0.284	
	t-Statistic					-0.900	
Average HH Income	Coefficient						-0.397
	t-Statistic						-0.545
Disaggregated SE Variable: % HH with Income > $100K	Coefficient	2.236	2.124	2.185	2.208	2.231	2.150
	t-Statistic	6.729 ***	3.384 **	4.824 ***	4.257 ***	4.578 ***	3.329 **
	R Squared	0.959	0.954	0.954	0.954	0.955	0.954
	Adj R Squared	0.952	0.946	0.946	0.946	0.947	0.946

* Coefficient estimate significant at 0.05 level ** at 0.01 level *** at 0.001 level

To assess the forecasting performance of these models, we limit the analysis to consideration of a single pair of regressions, those that use the Baltimore-Washington CSA Gross Regional Product as the aggregate socioeconomic variable. As mentioned above, it is not possible to obtain the Woods and Poole projections that would have been available for the years 2011 through 2015, which would have been included as part of the Woods and Poole 2012 data release. To create a test projection of values for Baltimore-Washington CSA GRP for those years, we use the percentage annual growth terms from the Woods and Poole 2016 data release for the years 2015 through 2019. With those percentage values for annual GRP growth, notional projected values for the region's GRP can be calculated for 2011 through 2015. To create notional projected values for the disaggregated socioeconomic variable, the percentage of Baltimore-Washington CSA households with 2009 dollar denominated income exceeding $100,000 (in 2009 dollars), a similar procedure was used.

Forecast values for the Oil Price dependent variable are also needed to calculate out of sample projections. These were obtained for the years 2011 to 2015 from the data reported in the FAA's 2012 Annual Aerospace Forecast document. These national values are used for each of the case study airports out of sample forecast exercises.

Four out of sample model forecasts were calculated for the baseline model using Baltimore-Washington CSA GRP and the model that also includes the disaggregated socioeconomic variable. The first three are based on the out of sample forecasts described above. For the first forecast, the projected values – based on the percentage increases projected for the years 2015 through 2019 – for GRP, and the percentage of Baltimore-Washington household incomes exceeding $100,000 in the case of the model using the disaggregated socioeconomic variable as input, are used as described above. For the second, the values for percentage annual increases over the projection period (2011 to 2015) are increased by 50 percent, indicating a more rapidly growing Baltimore-Washington economy. For the third MSO projection, the values for percentage annual increases are reduced by 50 percent, representing a projection for a more slowly growing economic environment in the Baltimore-Washington CSA. For the fourth projection, the actual values observed in 2011 through 2015 for GRP and the share of Baltimore-Washington CSA households with incomes exceeding $100,000 are used to drive the "forecast" of the Washington DC airport system's annual O&D enplanements.

Using the projections published by Woods and Poole for 2015 and subsequent years for simulating a range of forecast scenarios based on inputs that would have been available to analysts in 2011 or 2012, we can compare the forecasting performance of the baseline regressions for MSO annual O&D enplanements using the annual oil price variable and the Baltimore-Washington CSA GRP, with those of the regression parameters based additionally on input from the disaggregated socioeconomic variable. The independent variable data and projections are shown in Table D-61 below. The projections for the Baltimore-Washington CSA variables reported in the Woods and Poole 2016 data have annual GRP growing at a rate of around 2.5 percent, and the percentage of Baltimore-Washington CSA households with incomes exceeding $100,000 growing at about 1.3 percent over the forecast period. The annual percentage changes used to calculate the Oil Price projected values for 2011 through 2015 are those implied by the Oil Price projections reported in the 2012 FAA Aerospace Forecast, which relies on projected values for the years following 2010. Thus we have four “forecast” scenarios for the two regression specifications for annual MSO O&D enplanements:

- Projected values for the socioeconomic independent variables based on the annual percentage changes in the variables, as reported by Woods and Poole for the five years beginning in 2015,
- Projected values based on annual percentage changes that are 1.5 times these values,
- Projected values based on annual percentage changes that are half these values, and]
- Projected values that equal the actual values for the independent variables that occurred in the years 2011 through 2015.

Table D-61 Inputs to Out-of-Sample Baltimore-Washington O&D Enplanement Projections

Balt-Wash Inputs to Out-of-Sample Enplanement Forecasts	Sources	2011	2012	2013	2014	2015
Oil Price Forecast	*2012 FAA Aero Forecast*	$84.32	$92.71	$97.16	$100.65	$104.64
GRP Forecast 1 (Medium)	*W&P 2016 %*	$619,279.9	$634,354.8	$649,597.3	$665,066.8	$680,784.8
GRP Forecast 2 (High)	*1.5 x F'cast 1*	$626,780.6	$649,666.8	$673,082.5	$697,125.6	$721,839.2
GRP Forecast 3 (Low)	*0.5 x F'cast 1*	$611,779.2	$619,225.4	$626,664.9	$634,126.6	$641,620.0
HHInc% Forecast 1 (Medium)	*W&P 2016 %*	36.4%	36.8%	37.3%	37.8%	38.3%
HHInc% Forecast 2 (High)	*1.5 x F'cast 1*	36.6%	37.3%	38.1%	38.8%	39.6%
HHInc% Forecast 3 (Low)	*0.5 x F'cast 1*	36.1%	36.4%	36.6%	36.9%	37.1%
	2010					
Actual Oil Price	$83.99	$101.87	$100.93	$100.49	$92.02	$48.39
Actual GRP	$604,278.6	$607,780.9	$611,636.2	$613,427.5	$624,734.9	$640,644.3
Actual HHInc%	35.9%	35.8%	36.2%	36.4%	36.4%	36.7%

The forecasting or projection results for the baseline regression equation for annual enplanements at the Baltimore-Washington airport system are shown in Figure D-56 below.

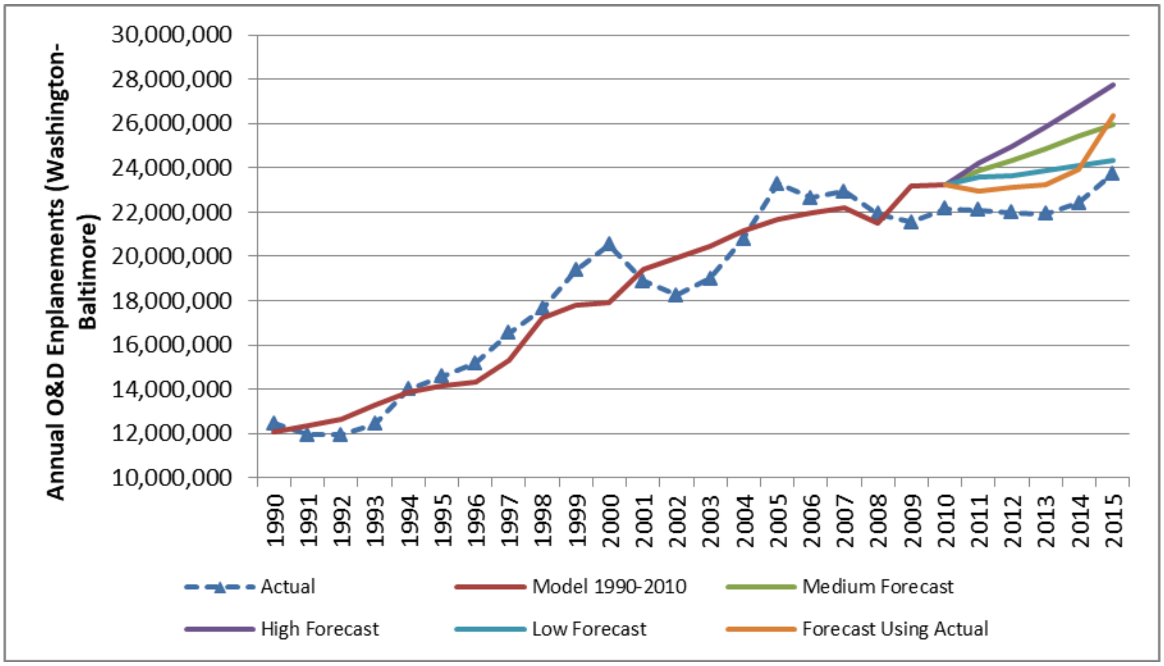

Figure D-56 Baseline Model Forecasts of Baltimore-Washington Annual Enplanements, 2011 to 2015, Using CSA Gross Regional Product, 1990 to 2010

The results for the regression equation using the disaggregated socioeconomic variable along with Baltimore-Washington CSA GRP are shown in Figure D-57 below.

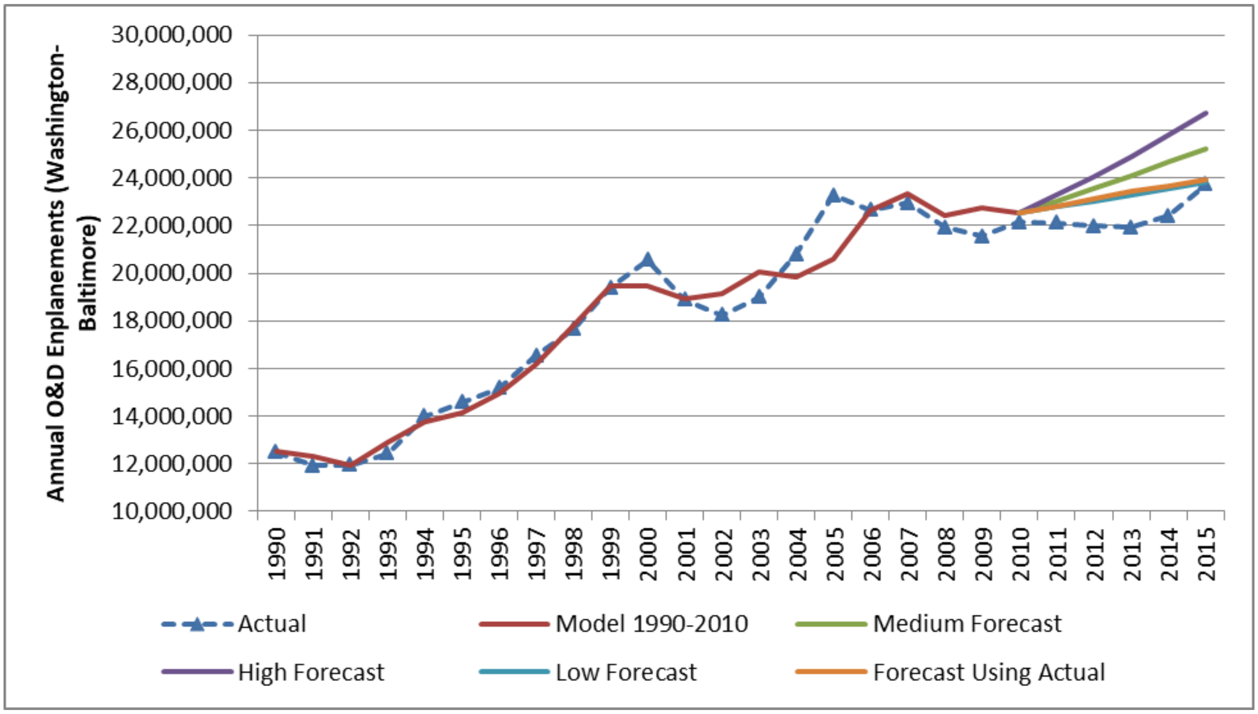

Figure D-57 Disaggregated SE Variable Model Forecasts of Baltimore-Washington Annual Enplanements, 2011 to 2015, Using CSA GRP and Share of CSA Households with Incomes exceeding $100,000, 1990 to 2010

Baltimore-Washington airport system enplanements were overforecasted in all of the forecasting scenarios, as can be seen in Figure D-57, with slightly better performance by the alternative model specification that included both the Baltimore-Washington CSA GRP as an aggregated socioeconomic variable and the percentage of CSA households with annual incomes exceeding $100,000 (in 2009 dollars) as a disaggregated socioeconomic variable.[12] The accuracy of the projections driven by the actual values taken by the Oil Price, Baltimore-Washington CSA GRP, and the Baltimore-Washington CSA Share of Households with Income exceeding $100,000 (for the years 2011 through 2015) is similarly modest, with the projected enplanement counts overshooting the actual history over those years. The following tables D-62 to D-64 quantify this, reporting the percentage error for each of the forecasts and the Root Mean Squared Error (RMSE) for each forecast, a calculation which can be used to compare the accuracy of the baseline model forecasts and the forecasts calculated from the model also using the disaggregated socioeconomic variable.

[12] The baseline regression equation is ln(Balt-Wash_Enp) = 2.47 – 0.106ln(OilPrice) + 1.12ln(GRP), with an adjusted R-squared of 0.905, while the estimated equation that includes the disaggregated socioeconomic variable is ln(Balt-Wash_Enp) = 23.02 - 0.004ln(OilPrice) - 0.28ln(GRP) + 2.23ln(%HHInc>$100K) with an adjusted R-squared of 0.947.

Table D-62 Percentage Forecast Errors, Baseline Model Projections of Baltimore-Washington Annual O&D Enplanements, 2011 to 2015

Balt-Wash using GRP	Percentage Forecast Error (Baseline Regression with Oil Price and GRP)			
	Medium	High	Low	Actual
2010	--	--	--	--
2011	8.1%	9.6%	6.7%	3.8%
2012	10.6%	13.6%	7.6%	5.2%
2013	13.3%	18.0%	8.9%	5.9%
2014	13.4%	19.6%	7.5%	6.8%
2015	9.5%	16.9%	2.4%	11.0%

The percentage errors in the tables reflect the fact that the model-based projections of system enplanements consistently underestimated the number of annual passenger, even for the "forecast" based on the values that actually occurred for the Oil Price and Baltimore-Washington CSA GRP. As can be seen in the figures and the tables of percentage errors, the model forecasts using these actual values do not differ from those based on the projected values for the independent variables, in spite of the significant drop in Oil Prices in 2015. This is because the estimated parameter for the Oil Price independent variable is very small in magnitude, therefore exerting little influence on the forecast of Washington DC airport system enplanements.

Table D-63 Percentage Forecast Errors, Disaggregated SE Model Projections of Baltimore-Washington Annual O&D Enplanements, 2011 to 2015

Balt-Wash using GRP	Percentage Forecast Error (Specification including Disaggregated SE Variable)			
	Medium	High	Low	Actual
2010	--	--	--	--
2011	4.2%	5.4%	3.1%	3.2%
2012	7.1%	9.5%	4.8%	5.1%
2013	9.9%	13.6%	6.3%	6.9%
2014	10.1%	15.1%	5.2%	5.7%
2015	6.3%	12.5%	0.5%	0.8%

The final table reports the Root Mean Squared Error (RMSE) of the different forecasts for the two model specifications. This statistic can be used to compare the accuracy of two or more of these forecasts, with a smaller RMSE value indicating a more accurate forecast. As can be seen, the inclusion of the additional disaggregated socioeconomic variable in the model specification provided consistent improvement in the accuracy of the five year forecasts, compared to those based on the baseline specification, although all eight forecasting scenarios for the Baltimore-Washington airport system resulted in forecasts that exceeded the actual annual enplanements over the forecast period of 2011 to 2015.

Table D-64 Root Mean Squared Error (RMSE) for Baltimore-Washington Baseline and Disaggregated SE Variable Model Forecasts of Annual System O&D Enplanements, 2011 to 2015

Balt-Wash (using GRP)	Root Mean Squared Error		% Difference BL v Alt
	Baseline	Disagg Var	
Medium Forecast	2,503,856	1,753,329	30%
High Forecast	3,588,934	2,644,458	26%
Low Forecast	1,544,198	984,074	36%
Actual Values	1,596,912	1,070,290	33%

D.3 Detailed Analysis of the Baltimore-Washington Region

The initial case study analysis described in Section D.2 developed a consistent set of fairly simple air travel demand models for each of the case study regions or airports. In order to explore how a more detailed demand analysis may shed additional light on the potential contribution of disaggregated socioeconomic data to air passenger demand analysis, further analysis was undertaken of the Baltimore-Washington region. This region was chosen because it is served by three major commercial service airports and air traffic data has been assembled for all three airports, allowing an analysis of regional demand and thereby avoiding distortions from changes in the regional share of specific airports. In addition, air passenger surveys have been performed at each airport by the Metropolitan Washington Council of Governments on a regular basis at the same time using a consistent survey methodology and questionnaires.

a. Analysis of Household Income Distribution

The initial model estimation regressions used the percent of households with personal incomes of $100,000 or more (in constant 2009 dollars) as the disaggregated measure of income. The potential difficulty with this measure is that it is not independent of the average household income since as the average household income has increased over time, the percent of households with incomes above any given threshold (e.g., $100,000) will also have increased as a growing percentage of households move into that income range, even if the relative distribution of incomes does not change. This is reflected in a strong correlation between average household income and the percent of households with incomes of $100,000 or more (0.984 for the Baltimore-Washington region for 1990 to 2010). As a result if this correlation, adding the disaggregated income variable to the model specification typically results in the average household income variable (or other aggregate economic variables that are strongly correlated with household income) becoming statistically insignificant, leaving the disaggregated income variable as the only statistically significant economic variable in the model.

In order to explore alternative disaggregated measures of household income that are independent of the average household income, an analysis was undertaken of the household income distributions for the Baltimore-Washington region, using data from Woods & Poole Economics. The change in the household income distribution from 1990 to 2010 is shown in Figure D-58. The highest income category in the Woods & Poole data is $200,000 or more (in constant 2009 dollars), so the data provide no information about the shape of the distribution above $200,000. In order to estimate the general shape of the distribution for higher income households with incomes above $200,000, an analysis was performed of the household income distribution reported by respondents to the 2010 Consumer Expenditure Survey (CES), who reported their actual household income in various categories. The result of this analysis is shown in Figure D-59.

CES respondents reported their income over the past year in each of up to four quarterly surveys (not all respondents completed four interviews in successive quarters). In some cases, respondents did not report their income and their incomes were inferred based on other household characteristics.

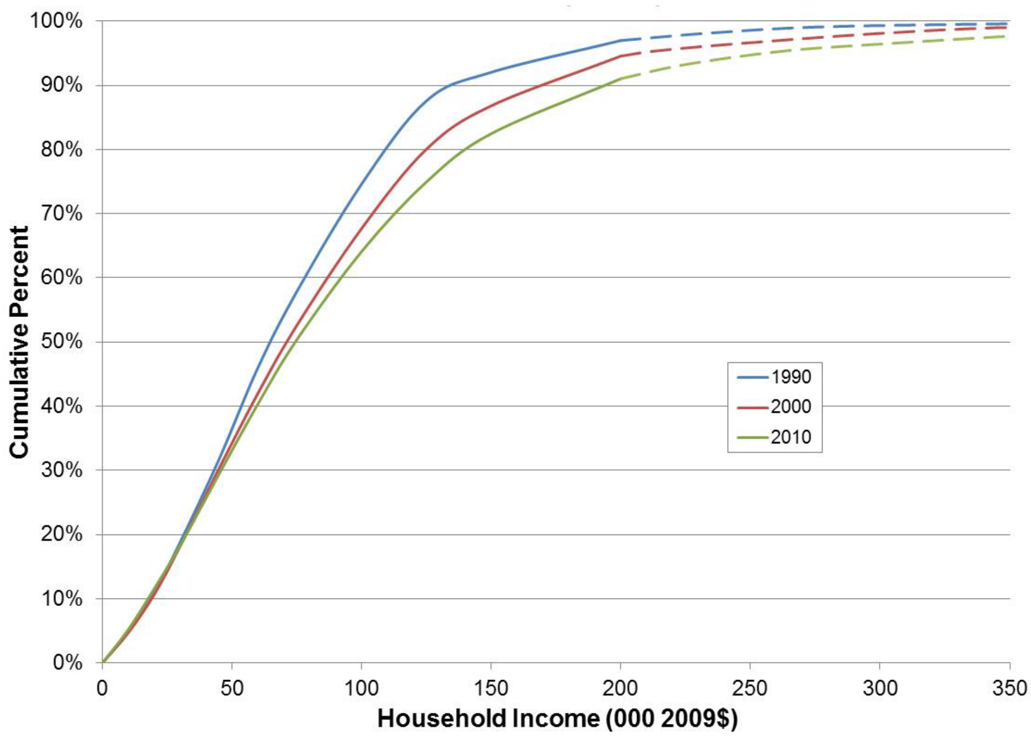

Figure D-58 Baltimore-Washington Region Household Income Distribution

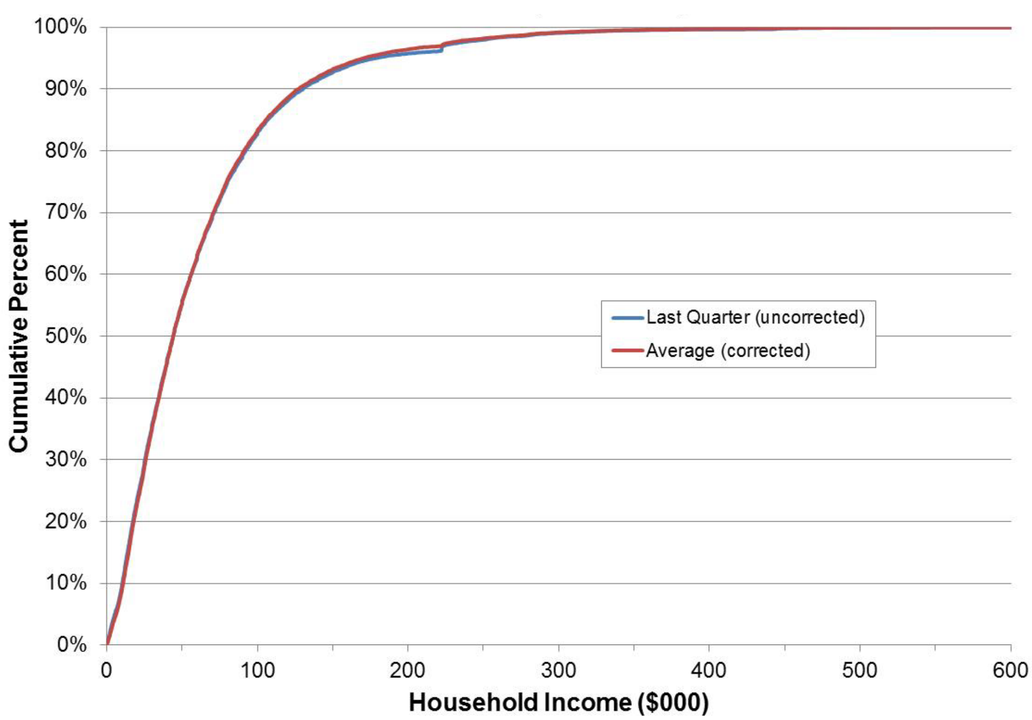

Figure D-59 Household Income Distribution – 2010 Consumer Expenditure Survey

Figure D-59 shows the distribution of the reported or inferred income for the last quarter in which the respondents participated in the survey. This ranged from the first quarter of 2010 to the first quarter of 2011. Examination of the data showed that a few values appeared in the data an unusually large number of times and were often very different from the values for the other quarters. It was assumed that these were an artifact of the income inference process and were most likely incorrect and were omitted from the analysis. Where incomes were reported in more than one quarter, these were averaged. This gave a smoother distribution curve but was not otherwise significantly different from the unadjusted distribution. It can be seen from Figure D-58 and Figure D-59 that for a given cumulative percentage, the household income for the Baltimore-Washington region in 2010 was significantly higher than given by the 2010 CES data, although this is not unexpected, since the CES data is for the U.S. overall and the average household income in the Baltimore-Washington region is significantly higher than the national average.

The extrapolation of the cumulative distributions shown in Figure D-58 above $200,000 were obtained by taking the distribution from the 2010 CES data and proportionally increasing the income for a given cumulative percent to correspond to the cumulative percentage at $200,000 shown in Figure D-58. It can be seen from Figure D-58 that the cumulative distribution curves for the Baltimore-Washington region flatten out for incomes between about $125,000 and $200,000 compared to the national distribution shown in Figure D-59. It is unclear whether this reflects a particular characteristic of household incomes in the region or is a consequence of the way that Woods & Poole adjusted the proportions of households in each income range to convert them to a consistent set of income ranges in constant dollars. It should be noted that the Woods & Poole data had a single income range from $150,000 to $200,000, so this may also have partly been reflected in the curve fitting logic in Excel.

The household income distributions shown in Figure D-58 show that there was very little increase in real incomes between 1990 and 2010 for households with incomes below the 30th percentile. For households with incomes between the 30th percentile and the 80th percentile, the increase in real income from 2000 to 2010 was less than from 1990 to 2000 and increased with income until by the 80th percentile the increase in real income from 2000 to 2010 was approximately the same as from 1990 to 2000. The largest increase in real incomes from 1990 to 2000 appears to have occurred around the 90th percentile, although the increase in real incomes from 2000 to 2010 for households between the 80th and 90th percentiles appears to have remained relatively constant. Above the 90th percentile the situation is unclear due to the absence of data points above a household income of $200,000 in the Woods & Poole data and the need to make assumptions to estimate the shape of the distribution above this income level and potential uncertainties in the data on which the calculation of the distribution was based.

The real average household income increased by about 25% from 1990 to 2000 and by about 15% from 2000 to 2010. It appears from the cumulative income distributions shown in Figure D-58 that the increase in average household income from 2000 to 2010 occurred disproportionately for households in the upper income ranges, in particular for households with incomes between the 70th and 90th percentile. This range corresponds to an average household income in 2010 between about $115,000 and $200,000 (in 2009 dollars). Findings from air passenger and household travel surveys show that households in this income range have higher air travel propensity that households in lower income ranges, not surprisingly. Therefore it appears likely that the growth in air travel from 1990 to 2010 was due not just to an increase in the real average household income, but to the disproportionate increase in income for households with incomes above the 70th percentile.

b. An Alternative Disaggregated Income Measure

As an alternative to the percent of households with a personal income of $100,000 or more, the percent of total personal income for all households that was received by the top 10% of households by income is a measure that is independent of changes in the average household income. This measure only reflects changes in the relative distribution of household income after adjusting for changes in average household income.

In order to calculate this measure from the Woods & Poole data, it is necessary to make two calculations:

1. Determine the household income that corresponds to the 90th percentile of the cumulative distribution
2. Estimate the average household income for all households below the 90th percentile of the income distribution

Calculating the average household income for households below the 90th percentile, rather than directly calculating the total income for the 10% of households above the 90th percentile, was necessary due to the uncertainly about the shape of the income distribution above $200,000. The percent of total personal income received by households in the top 10% by income can then be easily calculated from the average income for all households and the average income for households below the 90th percentile.

The household income corresponding to the 90th percentile was calculated by assuming that the distribution curve between the two data points on the cumulative income distribution on either side of the 90th percentile was approximated by the logistic function:

$$P = 1 / (1 + e^{aY^n})$$

where *P* is the cumulative percentage, *Y* is the household income, and *a* and *n* are parameters that are fitted to the data.

The average income for households below the 90th percentile was calculated by integrating the cumulative income distribution curve below the 90th percentile, assuming that the curve between each pair of data points was approximated by a quadratic function. The parameters of the quadratic function between each pair of data points were determined from the two data points and the subsequent data point in the distribution.

The resulting changes in the values of the percent of total personal income received by households in the top 10% by income from 1990 to 2010, together with the corresponding percent of households with a personal income of $100,000 or more and the average household income, is shown in Figure D-60.

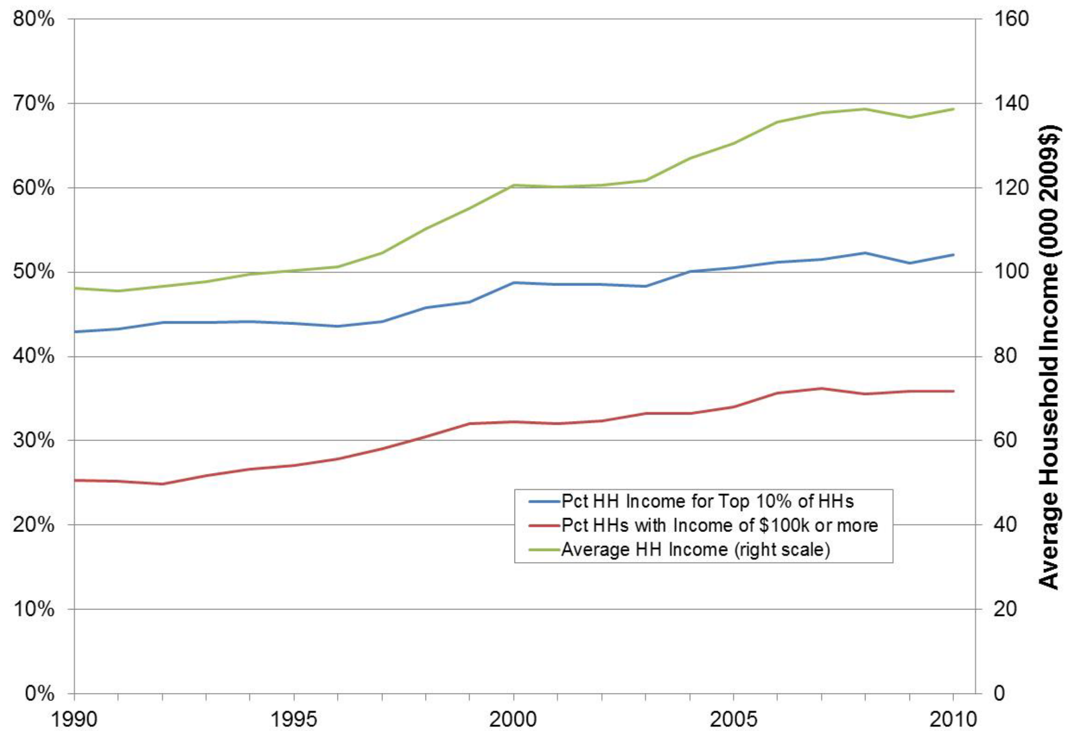

Figure D-60 Changes in Disaggregated Household Income Metrics over Time

It can be seen that the percent of total personal income received by households in the top 10% by income increased at a slower rate than percent of households with a personal income of $100,000 or more from 1990 to 2010, as would be expected given the increase in average household income over the period, but the difference in the rate of increase became less after 2000. In fact, from 1999 to 2010 the increase in the two measures was essentially identical. However, it can be seen that the increase in the two measures differed considerably from year to year, with the changes in the percent of total personal income received by households in the top 10% by income from year to year corresponding more closely to the changes in the average household income (the correlation coefficient for the two measures was 0.99).

Since the percent of total personal income received by the top 10% of households is independent of the average household income, the combination of an increasing average household income and an increasing share of total income received by households in the top 10% means that the average household income of the top 10% of households increased significantly over the period from 1990 to 2010, in fact by 75%. In contrast, the average household income of other households only increased by 21%. During the same period, the regional enplaned origin and destination (O&D) passengers increased by a little over 77%.

However, while it is clear that the much greater increase in average household income in the top 10% of households than for the other 90% of households would have contributed disproportionately to the growth in enplaned O&D passengers, deciding how best to reflect this in an air passenger demand model is less obvious. Simply adding a variable for the percent of total personal income received by the top 10% of households to a log-linear model of air passenger demand that incorporates a measure of total household income, such as a model of enplaned passengers per capita with a variable for the average household income or per capita gross regional product, could prove unsatisfactory for two reasons. The first is that the correlation between the two income variables will result in multicollinearity that makes it difficult to obtain reliable estimates of the model coefficients. The second is that if the coefficient for the average income variable becomes statistically insignificant and is omitted, the projected passenger demand depends only on the percent of total personal income received by the top 10% of households, which ignores the effect of the overall changes in real income.

A trial regression adding a variable for the percent of total personal income received by the top 10% of households to a log-linear model of O&D enplanements per capita for the Baltimore-Washington region for the period 1990 to 2010 with variables for the average household income and oil price gave statistically significant coefficients for all variables. In contrast, the same model with the percent of households with incomes of $100,000 in place of the percent of total personal income received by the top 10% of households resulted in a statistically significant coefficient for the percent of households with incomes of $100,000 but the coefficients for the average household income and oil price became statistically insignificant and the coefficient for average household income had a counterintuitive sign.

However, adding the variable for the percent of total personal income received by the top 10% of households gave an elasticity for that variable of -2.74 (which was initially viewed as a counterintuitive sign) and increased the elasticity for the average household income variable from 1.40 (which seems plausible) to 2.71, which seems high. Furthermore, estimating the model separately for the periods 1990 to 2000 and 2000 to 2010 resulted in the oil price coefficient becoming statistically insignificant for the period 1990 to 2010 and the elasticities for the other two variables increasing in magnitude (i.e., even less plausible), and the coefficients for all three variables for the period 2000 to 2010 becoming statistically insignificant. This suggests that the estimated coefficients obtained for the full period with the percent of total personal income received by the top 10% of households included in the model were distorted by the correlation between the variables and, far from improving the model, present a misleading result.

Clearly, there is more going on that is captured by the alternative disaggregated income variable, at least for the Baltimore-Washington region, and until those other issues are resolved, simply adding the variable to the model does not really result in a better understanding of the factors influencing air passenger demand in the region.

Ideally, a demand model would contain separate terms that predict the enplaned O&D passengers generated by different income strata that would be summed to give the total enplaned O&D passengers. However, there are two difficulties in developing such a model. The first is that it cannot be estimated as a single equation using standard regression techniques, although it may be possible to estimate such a model using non-linear estimation techniques, such as maximum likelihood. The second is that estimating any such model from total annual passenger data would require far more data points than are provided by annual data for a 20-year period, due to the number of model coefficients for each income strata.

c. Exploratory Analysis of Air Passenger Demand

Therefore a more detailed exploratory analysis of air passenger demand in the Baltimore-Washington region was undertaken, based on the data on enplaned origin and destination (O&D) passengers at the three primary commercial service airports serving the region, socioeconomic data for the Washington-Baltimore-Arlington Combined Statistical Area (CSA), and national data for average airline yield, and oil price (measured as U.S. crude oil composite acquisition cost by refiners) assembled to support the initial development of models of air passenger demand as part of the case study analysis being performed as part of the project. With one exception, discussed below, all financial data are expressed in constant 2015 dollars.

In the following analysis, the enplaned O&D passenger data and aggregate socioeconomic data, such as employment, have been expressed on a per person basis in order to account for the effect of population growth on demand while avoiding problems from the correlation that exists between growth in population and growth in other socioeconomic factors. In addition, the model development was based on the use of a multiplicative (log-linear) demand model. This model specification ensures that the marginal change in the dependent variable for a given change in one of the independent variables varies with the overall level of the dependent variable, which seems intuitively more reasonable than assuming that it is constant. In addition, the coefficient estimates for log-linear models give the demand elasticity, which is helpful in interpreting the reasonableness of the model estimates. However, there are some potential drawbacks with this model specification that are discussed further below.

One consequence of expressing the enplaned passenger data on a per person basis in a log-linear model is that this implies that the elasticity of demand with respect to population is unity, which was felt to be a reasonable assumption. If the population increases by x% and all other factors remain unchanged, one would expect x% more trips, since there are x% more households with the same distribution of household characteristics facing the same set of circumstances and therefore presumably making the same average number of trips each.

The change in the enplaned O&D passengers per person and selected socioeconomic variables, as well as the U.S. average airline yield and the oil price measure adopted, over the period 1990 to 2010 is shown in Figure D-61, expressed as an index relative to the value of each data series in 1990. This allows data expressed in very different units to be shown on the same chart and provides a direct comparison of the relative change in each data series over time. It can be clearly seen that there were considerable differences in the relative change of the different data over the 21-year period.

The enplaned passengers per capita show an overall growth of about 32% from 1990 to 2010 with considerable fluctuation. In particular, enplaned passengers per capita declined during each of the three recessions, shown by the grey bands in Figure D-61, as well as a much greater decline from 2000 to 2001 than in the other two recessions, which largely reflects the effect of the terrorist attacks of September 11, 2001 (“9/11”) and subsequent changes in the aviation system. Enplaned passengers per capita increased fairly steadily from 1993 to 2000, following the end of the 1990 to 1991 recession. After dropping steeply in 2001 and 2002, enplaned passengers per capita grew again fairly strongly after 2002, reaching a peak in 2005 then declined slightly in 2006, remaining fairly constant in 2007 before declining in 2008 and 2009, and recovering slightly in 2010.

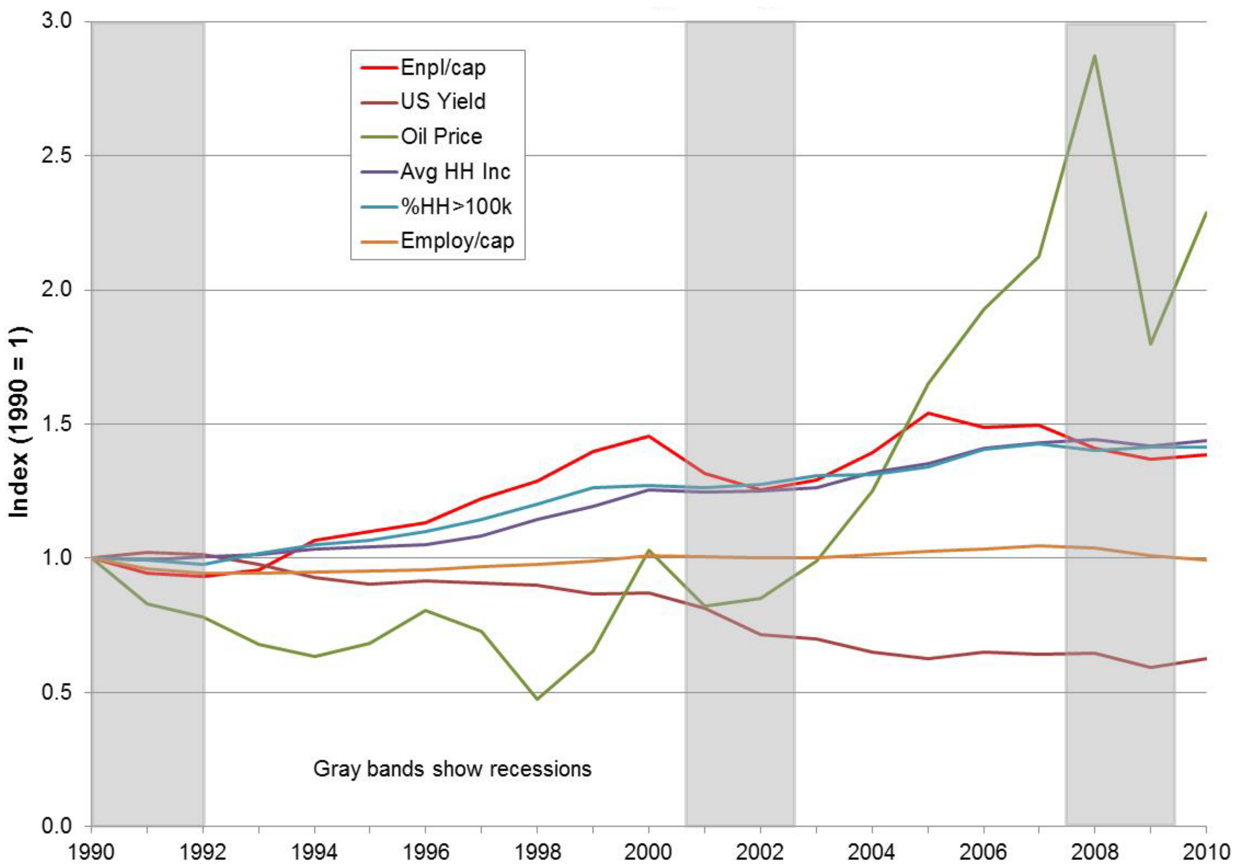

Figure D-61 Changes in Enplaned O&D Passengers and Socioeconomic and Other Data for Baltimore-Washington Region – 1990 to 2010

It should be noted that the periods of the three recessions shown in Figure D-61 are only approximate. The data series shown in Figure D-61 are on a calendar year basis but the official start and end dates of each recession did not coincide with the start or end of a calendar year. In addition, the effects of a recession often linger past the official end date of the recession, as can be seen from the employment per capita data in Figure D-61, which continued to decline from 1991 to 1992 and from 2009 to 2010.

The average household income increased by about 41% from 1990 to 2010, or somewhat more than the corresponding increase in enplaned passengers per capita. However, the growth from year to year was much less volatile than the changes in enplaned passengers per capita, with only modest declines during each recession. The strongest growth in average household income occurred from 1996 to 2000, at an average rate of 4.8% per year, followed by the period from 2003 to 2008, at an average rate of 2.8% per year.

Figure D-61 also shows the disaggregated household income measure of the percent of households with an income over $100,000 in constant 2009 dollars. This generally followed the changes in average household income, although the strong growth in the late 1990's started earlier, in 1994, and ended a year earlier in 1999.

The employment per capita showed very little overall change from 1990 to 2010, perhaps not surprisingly, although it declined slightly in the 1990 to 1991 and 2008 to 2009 recessions and showed a slow recovery after each recession. Somewhat unexpectedly, the employment per capita changed much less during the 2001 recession than during the other two recessions, declining by only 1% from 2000 to 2003. This suggests that the steep decline in enplaned passengers per capita in 2001 and 2002 was due much more to the effects of 9/11 than the 2001 recession.

In contrast to the overall growth in the socioeconomic variables real average airline yield declined fairly steadily over the period, with some year to year fluctuation. After increasing by about 2.3% during the 1990 to 1991 recession, average yield declined fairly steadily to 1995 and then more slowly to 2000, by which point it was 13% below its 1990 level. It then declined more steeply again in 2001 and 2002, as the airlines attempted to cope with the decline in demand following 9/11, and then continued to decline more slowly to 2005 before recovering slightly in 2006 and remaining relatively unchanged until 2008. It declined again in 2009 by about 8.6% and partly recovered in 2010, ending 37% below its level in 1990.

In contrast to the fairly steady trends in the socioeconomic variables and average airline yield, oil price showed much greater volatility over the period 1990 to 2010 as well as much greater overall change. The peak price in 2008 was approximately six times higher than the lowest price in 1998. Real oil price declined fairly steeply from 1990 to 1994, then increased to 1996 before declining steeply again to its lowest value in 1998, when it was approximately half the price in 1990. It then increased sharply to 2000, when the price was approximately the same as it had been in 1990, declined again in 2001, increased slightly in 2002 then increased rapidly to its peak value in 2008. It declined sharply in 2009 to about 63% of its peak value before recovering in 2009 to end at about 20% below its peak value or 130% above the price in 1990.

Surprisingly, there appears to be very little relationship between the average airline yield and oil price. Although both average yield and oil price declined from 1990 to 1998, average yield continued to decline to 2005 despite oil price rising sharply and remained relatively unchanged to 2008 while oil price continued to rise to its peak value. Thereafter average yield changed in the same direction as oil price in 2009 and 2010. Although changes in oil price obviously would have affected airline costs, these changes do not appear to have been passed through to air travelers through changes in fare. Of course, changes in refiners acquisition cost is not the same thing as changes in the price that airlines paid for jet fuel, particularly if they engaged in fuel price hedging. However, what matters for air travel demand is the level of airfares that are available, not whether or not the airlines are making or losing money.

The other effect that oil price could have on air travel demand is from changes in consumer spending patterns. If households have to spend a higher proportion of their income on gasoline and home heating fuel, and possibly other purchases that are disproportionately affected by changes in oil price, then this is likely to reduce their spending on discretionary purchases, such as air travel, even if their total income is unchanged.

Modeling Approach and Model Development

It seems clear from the data shown in Figure D-61 that the level of enplaned passengers per capita after 2000 was influenced by factors other than the ongoing trends in socioeconomic factors and airfares as represented by average airline yield, most notably changes in the airline industry following 9/11. In order to avoid the effect of these factors distorting the estimates of the effect of the socioeconomic, airline yield and oil price variables on the level of enplaned passengers and to provide a way to quantify the magnitude of the effect of any additional factors that occurred after 2000, demand models were first estimated using data for the period 1990 to 2000. These models were then used to project enplaned passengers per capita for the period 2001 to 2010 and the resulting projected traffic compared to the actual traffic. This gave the ratio by which the projected traffic exceeded the actual traffic. Based on the resulting pattern of this ratio, year-specific dummy variables were defined to account for the level of over-prediction in each year. The models were then re-estimated for the full period including the dummy variables.

Of course, the inclusion of year-specific dummy variables forces the model to fit the data for those years. But their inclusion in the model provides two important benefits:

1. It allows the estimation of the coefficients of the continuous variables to be based on the full 21 years of data without being distorted by year-specific effects (such as the effect of 9/11 in 2001 and 2002)
2. It provides an estimate of the magnitude of any year-specific effects during the period from 2001 to 2010 that reduce the projected enplaned passengers per capita below the level attributable to the effects of the continuous variables

Naturally, the inclusion of the dummy variables provides no information on what factors may have caused the effect measured by each dummy variable. It merely measures the magnitude of the effect in each year. Interpreting the likely or potential cause of these effects requires additional analysis or thought. However, knowing how the magnitude of the effect changes from year to year may lead to the addition of new continuous variables or changes in the definition of the continuous variables (e.g., using the average fare for the airports in question from the U.S. Department of Transportation 10% O&D Survey in place of the national average yield) that account for these effects without the need for dummy variables.

Even without such additional analysis, separating out the year-specific effects from the effects of the continuous variables included in the model has value from the perspective of the use of the models for forecasting. If the year-specific effects are considered unlikely to occur in the future or to recur under assumed conditions (e.g., future recessions), the effects measured by the dummy variables can be omitted from the forecasts or used to define future scenarios in which these effects continue but at a different level of frequency. In contrast, developing models that rely on a limited set of continuous variables without considering any year-specific effects runs the risk of the resulting models accounting for these effects by distorting the estimated coefficients of the continuous variables. Use of such models to prepare forecasts can result in a situation where it is unclear whether and to what extent the forecasts implicitly assume a continuation of the year-specific effects that occurred during the model estimation period.

However, one limitation of the three-step approach described above is the limited number of data points that are available in the period 1990 to 2000 to estimate the models to be used to assess the pattern of over- or under-prediction during the period from 2001 to 2010. Of course, including dummy variables in the resulting model estimated on the full period reduces the degrees of freedom in the model estimation, which needs to be carefully considered in deciding how many dummy variables to include. Therefore the first step in the process started with very simple models and progressively added variables, retaining them if they improved the model fit and had statistically significant coefficients of the expected sign and the values seemed plausible.

The resulting evolution of the model specification is shown in Table D-65. The model evolution sequence shown in D-65 began with a model specification using average household income and oil price as explanatory variables (Model 1a). This is not quite the same specification as used in the initial model estimation that explored a range of different socioeconomic variables together with oil price, since the dependent variable is different (enplaned O&D passengers per capita). The oil price coefficient was found to be not significantly significant. Replacing this variable with the average airline yield (Model 1b) improved the model fit, as measured by the adjusted R squared, and the yield coefficient has the expected sign, a reasonable value, and is statistically significant at the 1% level. It is also intuitively reasonable that airfares (as measured by average airline yield) would influence air travel demand.

Table D-65 Model Estimation Results – 1990-2000 (Model 1)

Variable		Model 1a	Model 1b	Model 1c	Model 1d
Intercept	Coefficient	-21.55	-10.72	-4.93	-4.63
	t-statistic	*(-10.01)*	*(-3.21)*	*(-1.474)*	*(-1.326)*
Average Household Income	Coefficient	1.937	1.261	0.869	0.848
	t-statistic	*(10.52)*	*(5.68)*	*(3.87)*	*(3.63)*
Oil Price	Coefficient	-0.0383			-0.0300
	t-statistic	*(-0.579)*			*(-0.714)*
US Average Airline Yield	Coefficient		-1.038	-1.282	-1.241
	t-statistic		*(-3.61)*	*(-5.42)*	*(-4.93)*
Employment/person	Coefficient			1.122	1.318
	t-statistic			*(2.65)*	*(2.54)*
Adjusted R Squared		0.917	0.967	0.981	0.980

Model 1c added the variable employment per person. This reflects employment levels on a per-person basis and is intended to act as a surrogate measure for business trips. It can be expected that an increase in employment levels on a per-person basis would lead to more business travel, even if average household incomes do not change and thus the number of personal trips does not change. This further improved the fit of the model and the estimated coefficient for employment per person has the expected sign and is statistically significant at the 5% level. The coefficient value appears somewhat higher than expected, implying that a 1% increase in employment per person would lead to about a 1.1% increase in total enplaned passengers. It would seem reasonable that a given increase in regional employment would lead to a proportional increase in business travel, if all other factors remain unchanged. However, business travel only accounts for about half of all air travel, so *a priori* one would expect an elasticity of demand with respect to employment of around 0.5. Adding the employment variable reduced the estimated coefficient for average household income (as would be expected, since in the absence of the employment variable the change in average household income is accounting for the growth in total enplaned passengers, including both personal and business travel) and increased the estimated coefficient for airline yield. Since the employment and yield variables have opposite signs, it is possible that an overestimation of the magnitude of the coefficient of the employment variable resulted in the absolute magnitude of the coefficient of the yield variable also being overstated, as indicated by the subsequent estimation results that included a disaggregated household income variable discussed below

Model 1d includes both the oil price and airline yield variables, in order to see whether including the yield and employment variables would result in a statistically significant effect of oil price. The t-statistic for the oil price coefficient improved slightly, but the coefficient was still not statistically significant at even the 50% level and there was no improvement in model fit.

The next set of models added disaggregated household income variables to Model 1c, as shown in Table D-66.

Table D-66 Model Estimation Results – 1990-2000 (Models 2 and 3)

Variable		Model 2a	Model 2b	Model 2c	Model 3
Intercept	Coefficient	3.62	4.64	4.49	-22.47
	t-statistic	*(0.978)*	*(8.65)*	*(7.58)*	*(-3.69)*
Average Household Income	Coefficient	0.086			2.111
	t-statistic	*(0.279)*			*(4.93)*
Oil Price	Coefficient			0.0204	
	t-statistic			*(0.705)*	
US Average Airline Yield	Coefficient	-0.729	-0.704	-0.685	-0.782
	t-statistic	*(-2.94)*	*(-3.28)*	*(-3.05)*	*(-3.46)*
Employment/person	Coefficient	0.953	0.966	0.797	0.539
	t-statistic	*(3.19)*	*(3.52)*	*(2.14)*	*(1.58)*
Pct of HH with Income >$100k	Coefficient	0.989	1.069	1.115	
	t-statistic	*(2.95)*	*(6.81)*	*(6.34)*	
Pct of HH Income by Top 10%	Coefficient				-1.751
	t-statistic				*(-3.10)*
Adjusted R Squared		0.991	0.992	0.992	0.992

Adding the variable for the percent of households with income over $100,000 in 2009 constant dollars (Model 2a) improved the fit of the model over Model 1c. However, the coefficient for average household income became statistically insignificant. Dropping the average household income from the model (Model 2b) gave a slight improvement in model fit and the statistical significance of the remaining estimated coefficients improved, with all being statistically significant at the 5% level or better. Adding the oil price variable to the model (Model 2c) did not improve the model fit, the estimated coefficient for oil price has a counterintuitive sign and is not statistically significant, and the statistical significance of the estimated coefficients of the other variables decreased. Therefore Model 2b is considered the best of the three models.

Replacing the disaggregated household income variable in Model 2a with the alternative disaggregated household income variable of the percent of total household income received by the top 10% of households by income (Model 3) resulted in the estimated coefficient for average household income becoming statistically significant while the model fit remained essentially the same as for Model 2b. All the estimated coefficients are statistically significant at the 5% level or better, with the exception of the coefficient for employment per person, which was statistically significant at only about the 18% level. However, the estimated coefficient value (0.54) is closer to the expected value, as discussed above. Given the statistically significant value for the estimated coefficient for average household income and the apparently reasonable value for the employment coefficient estimate, it was felt that Model 3 gave the best representation of enplaned passenger demand for the period 1990 to 2000.

In all the models that included a disaggregated household income variable, the estimated yield coefficient varied between -0.69 and -0.78, which seems intuitively reasonable. Since airfares are only one component of the costs of making an air trip, it seems reasonable that the demand elasticity of airfare (yield) would be less than unity.

It may seem counterintuitive that the coefficient for the alternative disaggregated income variable would be negative, but in fact this is what would be expected. If the top 10% of households have a higher share of total income, this implies that the rest have a lower share, particularly middle income households, as noted in the discussion of the alternative disaggregated income variable in the previous section. As their income drops relative to the average income, their air travel propensity would tend to drop as well, and there are far more of them than those in the top 10% of households, whose air travel propensity in any case is probably not as greatly affected by changes in their income as lower income households.

Extending the Model to the Full Period

Model 3 was then used to project the enplaned O&D passengers for the period from 2001 to 2010, assuming that the relationships between enplaned passengers and the socioeconomic and airline yield variables estimated for 1990 to 2000 continued to 2010, and the ratio of the projected passengers to the actual enplaned O&D passengers was calculated. This comparison gave the result shown in Figure D-62. As expected, the projected passengers exceeded the actual passengers in each year from 2001 to 2010.

The extent of the overestimation increased significantly from 2001 to 2002, but then grew more slowly to 2010, with two significant deviations. The first occurred in 2005, when the actual traffic peaked by more than the model predicted that it should have, based on the values of the socioeconomic and yield variables, causing a drop in the level of overestimation. The second occurred in 2009, when the actual traffic declined, whereas the model predicted that it should have grown, based on the values of the socioeconomic and yield variables.

It seems plausible that the 2009 deviation is in some way a consequence of the 2008 to 2009 recession that is not fully accounted for by the changes in household income and employment levels. Moreover, it seems plausible that the reason for the 2005 deviation was at least partly a consequence of the sad tale of Independence Air. In July 2004 Atlantic Coast Airlines, which had previously been a United Express regional airline, commenced service as a new low-fare airline based at Dulles International Airport named Independence Air. The airline expanded rapidly in late 2004 and into 2005 and ended up in a fare war with the legacy carriers (particularly United, for which Dulles is a major hub). By the end of 2005 the airline was in deep financial trouble and ceased operation in early January 2006. Although the US airline yield data shows a slight dip in yield in 2005, it seems quite likely that fares at Dulles (and indeed the other two regional airports) dropped by more than that in 2005. Using the actual average fares at the three airports rather than the average U.S. yield would have increased the projected passenger traffic in 2005 and potentially eliminated the deviation in the model overestimation. Of course, the only way to find out if this is correct would be to assemble the fare data and redo the analysis. In the meantime, the analysis used the value of the 2005 dummy variable as a surrogate for the effect of Independence Air on demand in the region.

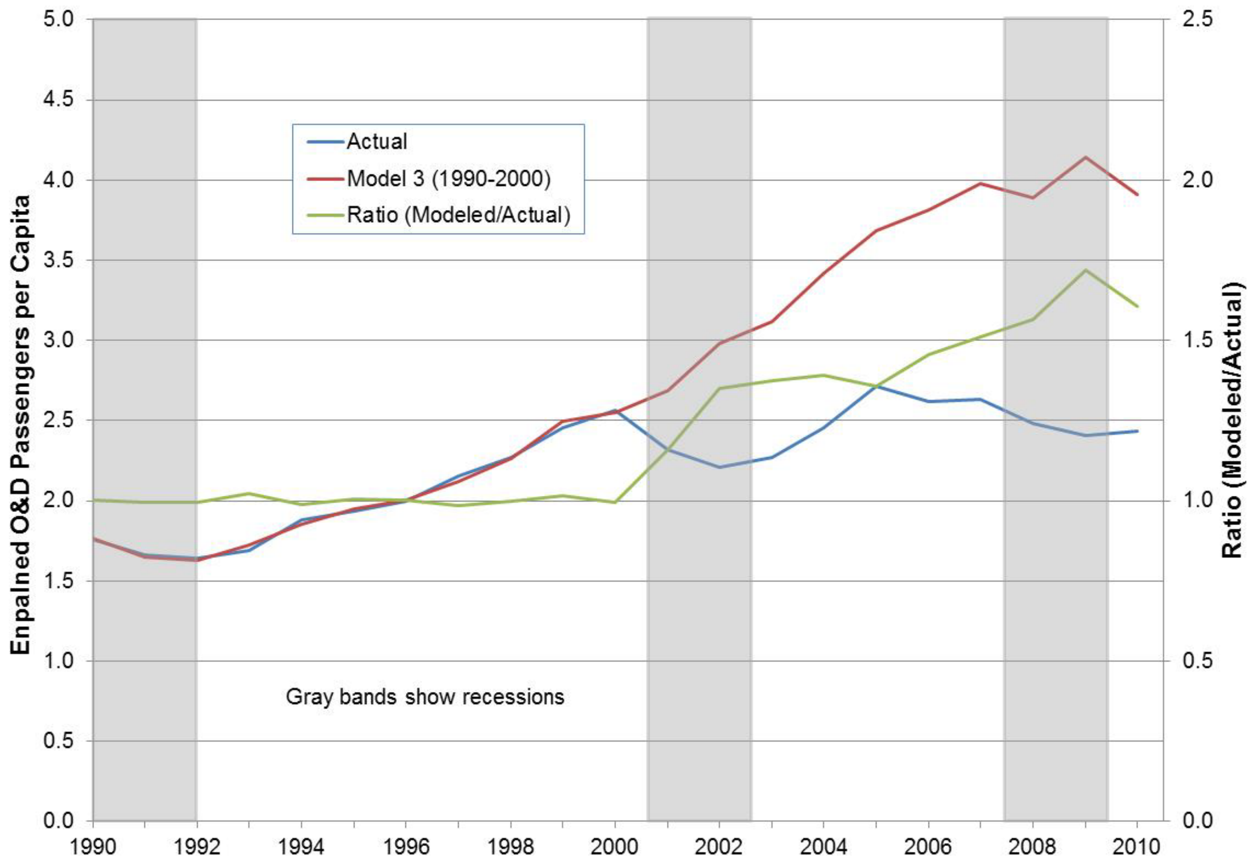

Figure D-62 Comparison of Projected and Actual Enplaned O&D Passengers for the Baltimore-Washington Region (Model 3) – 1990 to 2010

The extent of the overestimation in 2001 and 2002 is not unexpected, although perhaps higher than one might have expected (a 36% overestimation in 2002). Obviously 9/11 had a dramatic effect on demand in 2001, although only for the last four months of the year. 2002 was the first full year in the post-9/11 regime, so it is not surprising that the overestimation was a little over twice that of 2001 (which also experienced the effects of the 2001 recession, which may well have continued into 2002).

However, the subsequent increase in the overestimation in subsequent years was unexpected. The initial expectation was that the reduction in demand in 2002 would gradually dissipate over time as people got used to the new security arrangements, but this does not appear to have happened. Of course, there may well have been other factors causing this that are not reflected in the socioeconomic variables included in the model.

In order to account for the factors causing this overestimation, a set of year-specific dummy variables were defined, including a dummy variable for 2001, a dummy variable for 2002 on that was applied to each year from 2002 to 2010, and a dummy variable for 2003 on that took the value (Y-2002) for year Y. Thus the effect of the last two variables together gives a linearly increasing coefficient from 2002 to 2010. This implies that from 2002 on, projected enplaned passengers showed a small but continuous decline from the drop in 2002 (the first full year after 9/11) after accounting for the effect of the other variables.

A dummy variable for 2005 was included to reflect the effect of the fare war generated by Independence Air from late 2004 to its collapse in January 2006. Finally a dummy variable for 2009 was included to reflect the effect of the 2008 to 2009 recession over and above the effects of the decline in employment and household income.

This expanded model (Model 4) was then estimated on the full period from 1990 to 2010. The results of the model estimation are shown in Table D-67 and compared to the final model (Model 3) estimated on the period from 1990 to 2000. It can be seen that the coefficients of the continuous variables are similar in both models although the statistical significance is much higher (not only are the t-statistics higher in each case, but the model has five more degrees of freedom since there are ten more data points but only five more variables. All the coefficient estimates are statistically significant at the 5% level or better and all but that for the employment variable are statistically significant at the 1% level or better.

Two additional models were estimated to examine the effect of including oil price in the model (Model 5a) or using the original disaggregated income variable (Models 5b and 5c), the results for which are also shown in Table D-67.

Adding the oil price variable to Model 4 produces a small decrease in model fit and the coefficient for oil price is statistically insignificant, has a counterintuitive sign, and in any case has a very small value. Replacing the disaggregated household income variable in Model 4 with the percent of households with income over $100,000 (Model 5b) also results in a small decrease in model fit. As with Model 2a, the coefficient for average household income becomes statistically insignificant and the statistical significance of the estimated coefficients for airline yield and the dummy variables are lower than in Model 4. The statistical significance of the estimated coefficient for employment per person is greater than for Model 4 and the estimated value is closer to 1. This appears to reflect the weaker role of average household income in the model.

Dropping the variable for average household income (Model 5c) does not significantly change the model fit but the statistical significance of the estimated coefficients improves, although with the exception of the coefficients for employment per person and the dummy variable for 2005 the statistical significance is still lower than for Model 4 and the difference for the 2005 dummy variable is very small. The estimated coefficient for the employment per person is even higher, almost 1.02, or about twice the expected value, suggesting that this variable is partly accounting for the absence of average household income in the model.

Table D-67 Model Estimation Results – 1990-2000 (Models 3, 4 and 5)

Variable		Model 3	Model 4	Model 5a	Model 5b	Model 5c
Intercept	Coefficient	-22.47	-21.50	-21.81	2.80	4.99
	t-statistic	*(-3.69)*	*(-5.68)*	*(-4.82)*	*(0.917)*	*(13.32)*
Average Household Income	Coefficient	2.111	2.047	2.069	0.175	
	t-statistic	*(4.93)*	*(7.62)*	*(6.45)*	*(0.721)*	
Oil Price	Coefficient			0.00373		
	t-statistic			*(0.143)*		
US Average Airline Yield	Coefficient	-0.782	-0.849	-0.847	-0.861	-0.851
	t-statistic	*(-3.46)*	*(-5.93)*	*(-5.63)*	*(-5.51)*	*(-5.58)*
Employment/person	Coefficient	0.539	0.520	0.485	0.952	1.016
	t-statistic	*(1.58)*	*(2.40)*	*(1.454)*	*(4.34)*	*(5.18)*
Pct of HH with Income >$100k	Coefficient				0.844	0.978
	t-statistic				*(3.87)*	*(8.71)*
Pct of HH Income by Top 10%	Coefficient	-1.751	-1.697	-1.725		
	t-statistic	*(-3.10)*	*(-4.44)*	*(-3.85)*		
Dummy Variable 2001	Coefficient		-0.1514	-0.1512	-0.1456	-0.1391
	t-statistic		*(-8.31)*	*(-7.91)*	*(-7.00)*	*(-7.58)*
Dummy Variable 2002 on	Coefficient		-0.3059	-0.3059	-0.3030	-0.2960
	t-statistic		*(-9.71)*	*(-9.26)*	*(-8.57)*	*(-8.88)*
Dummy Variable 2003 on	Coefficient		-0.02347	-0.02408	-0.01632	-0.01467
	t-statistic		*(-10.86)*	*(-5.00)*	*(-5.04)*	*(-6.53)*
Dummy Variable 2005	Coefficient		0.0539	0.0538	0.0577	0.0601
	t-statistic		*(3.20)*	*(3.05)*	*(3.06)*	*(3.31)*
Dummy Variable 2009	Coefficient		-0.0901	-0.0892	-0.0856	-0.0873
	t-statistic		*(-5.19)*	*(-4.63)*	*(-4.51)*	*(-4.73)*
Adjusted R Squared		0.992	0.994	0.993	0.993	0.993

On balance, none of the three model variations offers an improvement over Model 4 from a statistical perspective and each has significant conceptual limitations.

The overall fit of the enplaned O&D passenger projections using Model 4 to the actual data is shown in Figure D-63. The overall model fit is very close. Of course, one would expect a near-perfect fit from 2001 on, since this is forced by the choice of dummy variables. But the fit from 1990 to 2000 (which did not include the effect of any dummy variables) is essentially unchanged from that shown in Figure D-62 in spite of the minor changes in the estimated coefficients of the continuous variables resulting from estimating the model on the full period from 1990 to 2010.

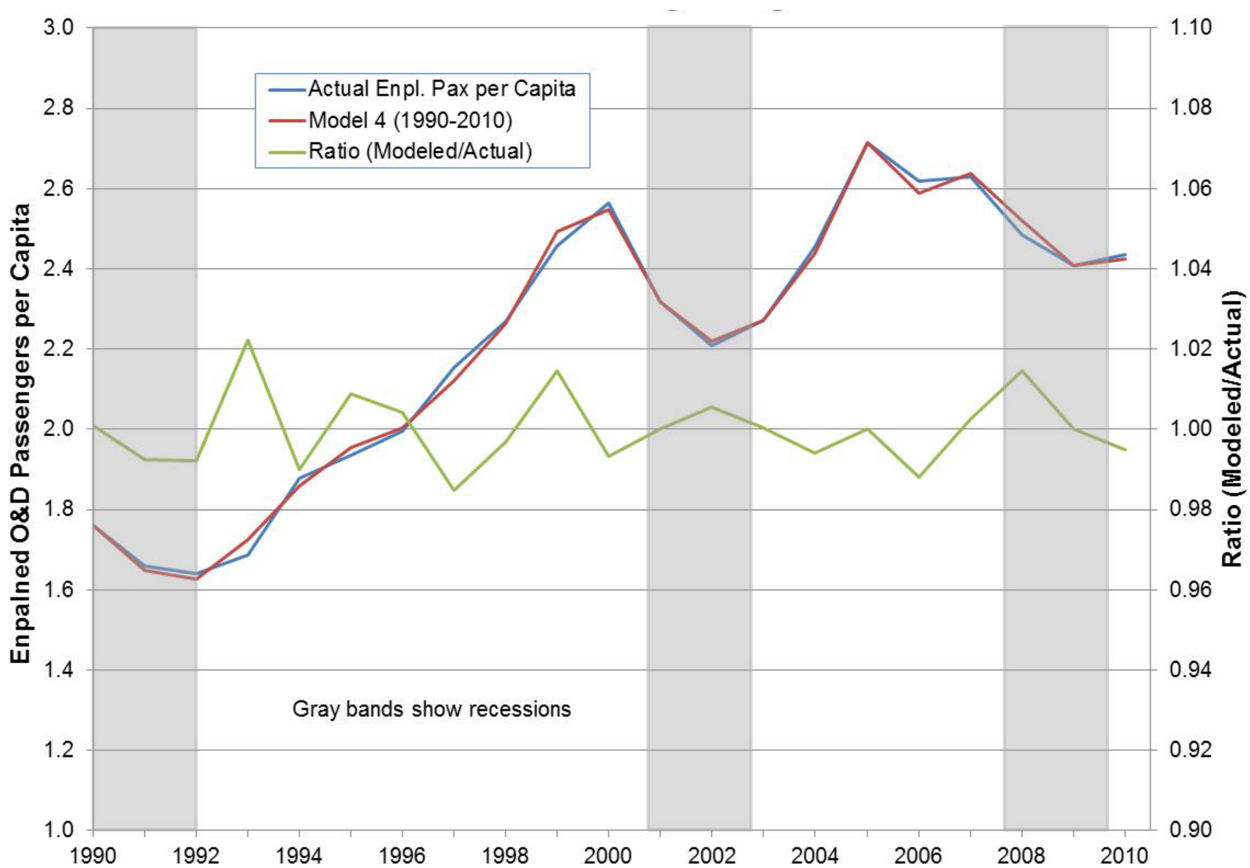

Figure D-63 Comparison of Projected and Actual Enplaned O&D Passengers for the Baltimore-Washington Region (Model 4) – 1990 to 2010

Contribution of the Disaggregated Household Income Variable

In order to quantify the contribution to Model 4 of the disaggregated household income variable, the model was re-estimated without the disaggregated income variable (Model 6). The results for this model estimation are shown in Table D-68, together with the results for Model 4.

Model 6 provides a poorer fit to the data, with an adjusted R squared of 0.985 compared to 0.994 for Model 4, in spite of having one more degree of freedom. The estimated coefficients have lower statistical significance with the exception of the coefficients for airline yield and employment. However, the increase in statistical significant for those two variables is fairly small and the estimated values of both coefficients are considerably larger than for Model 4 and in fact appear too high. It would be surprising if the elasticity of demand with respect to airfare was really higher than unity, since airfares are only one aspect of the cost of making an air trip and the other costs either are unaffected by changes in airfares or tend to move in the opposite direction (a decline in air travel due to an increase in airfares may result in lower hotel rates as hotels attempt to attract travelers in a declining market). The estimated elasticity of employment per person in Model 6 of 0.89 also appears too high for the reasons discussed above.

Table D-68 Model Estimation Results – 1990-2000 (Models 4 and 6)

Variable		Model 4	Model 6
Intercept	Coefficient	-21.50	-6.54
	t-statistic	*(-5.68)*	*(-2.37)*
Average Household Income	Coefficient	2.047	0.973
	t-statistic	*(7.62)*	*(5.19)*
US Average Airline Yield	Coefficient	-0.849	-1.185
	t-statistic	*(-5.93)*	*(-6.10)*
Employment/person	Coefficient	0.520	0.885
	t-statistic	*(2.40)*	*(2.75)*
Pct of HH Income by Top 10%	Coefficient	-1.697	
	t-statistic	*(-4.44)*	
Dummy Variable 2001	Coefficient	-0.1514	-0.1897
	t-statistic	*(-8.31)*	*(-7.39)*
Dummy Variable 2002 on	Coefficient	-0.3059	-0.3870
	t-statistic	*(-9.71)*	*(-9.41)*
Dummy Variable 2003 on	Coefficient	-0.02347	-0.02513
	t-statistic	*(-10.86)*	*(-7.38)*
Dummy Variable 2005	Coefficient	0.0539	0.0257
	t-statistic	*(3.20)*	*(1.029)*
Dummy Variable 2009	Coefficient	-0.0901	-0.0941
	t-statistic	*(-5.19)*	*(-3.39)*
Adjusted R Squared		0.994	0.985

The overall fit of Model 6 to the actual traffic is shown in Figure D-64. It is clear that the fit is not as close as Model 4, particularly for the period from 1990 to 2000, although it would probably be considered a perfectly acceptable fit to the data for most air passenger demand studies. Even for the period from 2001 to 2010, when the fit to the actual data is largely forced by the dummy variables, the fit in years that do not have a dummy variable for that specific year is not as good as Model 4.

However, the other important difference between the two models is the difference in the estimated coefficients for average household income, average airline yield, and employment per person, which have significantly different implications for the effect of future assumed changes in these variables. In particular, assuming that the relative household income distribution (measured by the percent of total income received by the top 10% of households) remains constant, the elasticity of demand with respect to average household income changes from 2.05 in Model 4 to 0.97 in Model 6. This difference points out that failing to take account of changes in income distribution in air passenger demand models can result in significant biases in estimated model coefficients.

Discussion

Although Model 4 fits the actual enplaned O&D passenger data closely, there are a number of important caveats. The first is that the model is based on the socioeconomic characteristics of the Baltimore-Washington region, but the enplaned O&D passenger data includes both air trips by residents of the region and by visitors to the region. The model thus implicitly assumes that the proportion of resident to visitor air passengers is constant, so that relationships based on the characteristics of the residents of the region also predict changes in visitor trips. The second is that the model uses U.S. average airline yield as a surrogate measure for the average airfares that were available at the airports serving the region. To the extent that these average airfares differed from the national average yield, this could have introduced biases or errors in the model. Indeed, the value of the coefficient for the dummy variable for 2005 suggests that the pattern of average regional airfares differed from that of the national average airline yield in at least one year.

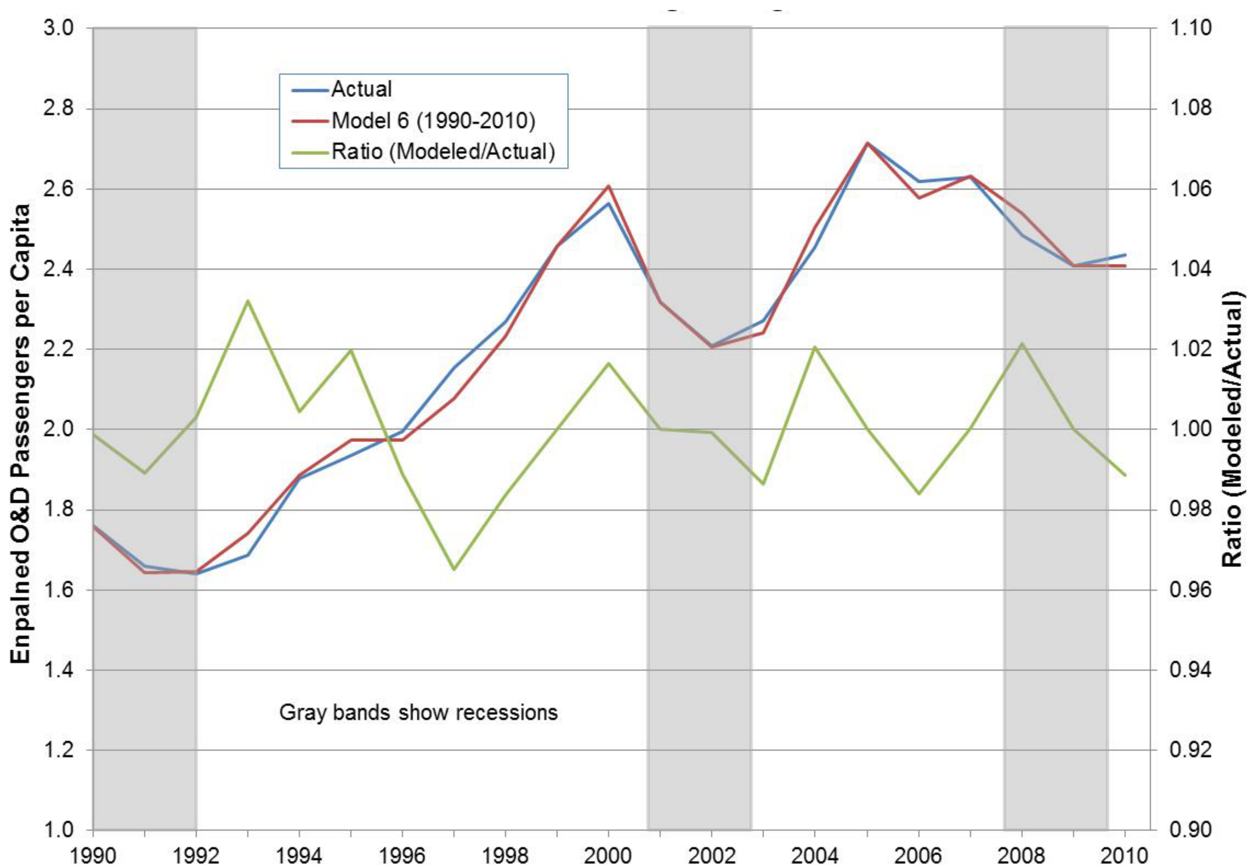

Figure D-64 Comparison of Projected and Actual Enplaned O&D Passengers for the Baltimore-Washington Region (Model 6) – 1990 to 2010

It is significant that when the model was estimated on the full 1990 to 2010 data with the dummy variables, the coefficients for the continuous socioeconomic variables did not change very much in value with the added 10 years of data (although the t-statistics improved substantially). The airline yield coefficient increased in magnitude by about 9% and the other three variables declined in value by between 3% and 4%. This suggests that the underlying relationships between demand and the socioeconomic variables are fairly stable and the reduction in demand reflected by the dummy variables really occurred and is not simply an artifact of a change in the causal relationships between the socioeconomic variables and air travel demand over time

Even so, the change in the estimated coefficients for the continuous variables from those in Model 3 to those in Model 4 could have resulted from a change in the causal relationships over time rather than simply being a consequence of having more data on which to estimate the coefficients for Model 4 although there is no statistical basis for rejecting the hypothesis that the true values of the coefficients in Model 3 are the same as the corresponding coefficient values in Model 4.

The increase in the effect of the dummy variables each year from 2002 to 2010 raises some unexpected and interesting issues. It seems unlikely that potential air travelers would be finding the hassle or inconvenience of the post-9/11 security measures increasingly onerous eight years after they were introduced, so the dummy variables may be reflecting other factors that were changing progressively over the period from 2002 to 2010 but are not included in the model. Either way, the dummy variables reflect the effect of these factors (whatever they are) on demand.

It is also possible that the elasticity of demand with respect to the socioeconomic variables is not constant but has been changing over time. There is in fact no reason to believe that the elasticity of demand with respect to any given variable is constant. This is a difficult issue to address empirically because what could appear as a change in the elasticity over time (whether because the elasticity varies with the levels of the independent variables and thus appears to vary over time or because it does actually vary over time for any given levels of the independent variables) could be due to the effect of omitted variables.

Whatever the cause of the increasing reduction in enplaned O&D passengers from the level projected by the socioeconomic variables alone that is reflected by the dummy variables, an important question is whether to assume that this trend will continue into the future. Some indication of the answer to this question may be gained from the *ex-post* forecasts discussed below.

It should also be noted that the apparent reduction in demand reflected by the dummy variables could be the result of changes in the composition of the O&D air travel market in the Baltimore-Washington region. The factors affecting travel by visitors to the region are likely to differ from those affecting travel by residents of the region, while the mix of personal and business travel is also likely to have changed over time. Some insights into changes in market composition over time may be gained from the air passenger surveys that have been conducted every two to three years at the three primary airports by the Metropolitan Washington Council of Governments. These surveys have been conducted at a much greater frequency than at many other airports and have used a consistent methodology and survey questions over time.

Clearly there is much more research that could be undertaken to obtain a better understanding of the factors that determine O&D passenger demand in the Baltimore-Washington region. However, the current preferred model (Model 4) appears to provide a solid basis for a more detailed case study analysis. It contains all the principal socioeconomic variables that are believed to exert a strong influence of air travel demand, namely population, household income, changes in employment levels, and average airfares, the estimated coefficients have plausible values and the expected sign, and the model fits the enplaned O&D passenger data very closely. It also includes a variable that reflects changes in the distribution of household income, the inclusion of which resulted in significant changes in the estimated elasticity of demand with respect to average household income, average airline yield, and employment per person.

d. Further Analysis of Air Passenger Demand

In order to understand how well the models of enplaned O&D passengers for the Baltimore-Washington region developed in the previous section would predict air passenger demand in subsequent years, two additional analyses were undertaken.

The first analysis used Model 4 to prepare an *ex-post* forecast projection of enplaned O&D passengers from 2010 to 2015. Obviously one issue that needs to addressed in applying Model 4 is whether to assume that the dummy variable for 2003 on, which was progressively increasing from 2003 to 2010, continues to do so at the same rate and for how long. The model was run with two different assumptions. The first assumption had the dummy variable continue to increase linearly (Model 4a). This resulted in a progressively increasing under-projection of the enplaned passengers per capita (relative to the actual enplaned passengers per capita), as shown in Figure D-65, reaching a 9% under-projection by 2015.

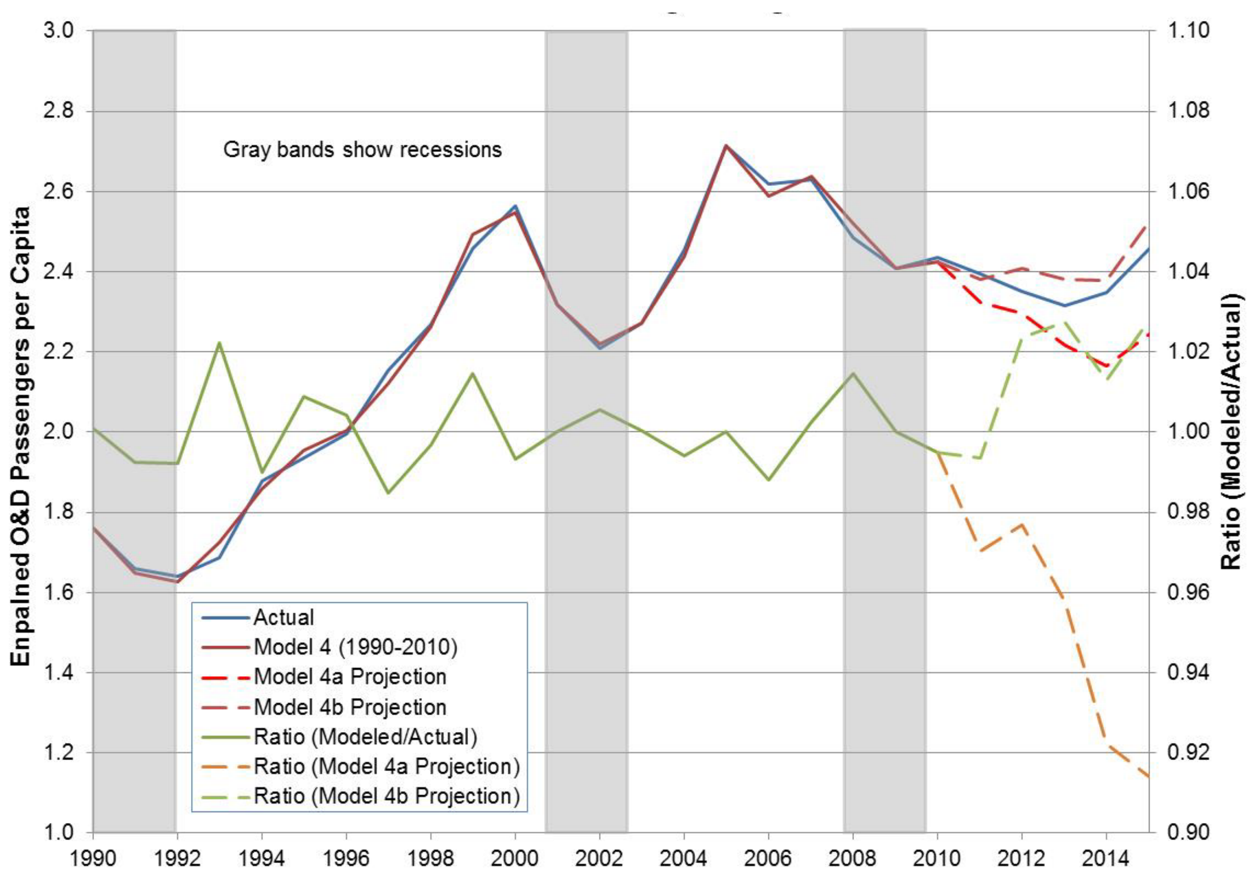

Figure D-65 Projection of Enplaned O&D Passengers per Capita for the Baltimore-Washington Region 2010 to 2015 – Model 4 Variants

The second assumption had the dummy variable remain constant at the 2010 value thereafter (Model 4b). This gave a very small under-projection for 2011 (less than 1%) and a small over-projection thereafter varying between 1% and 3%, with what appears to be a very slowly increasing trend, although this could simply reflect fluctuations in the data and results over the five-year period.

If instead of keeping the dummy variable constant at the 2010 value, it was kept constant at a slightly higher value, suggesting that whatever factors were accounted for by the dummy variable continued to increase slightly in 2011 and then remained constant, the projected enplaned passengers per capita would have fitted the actual data quite closely. The average over-prediction from 2011 to 2015 was 1.7%, which would have required the dummy variable value to have been held constant at 8.7 rather than 8 (the value in 2010) to offset. Of course, there is no reason to expect that any changes to the factors affecting the level of enplaned O&D passengers would have occurred exactly at the end of a calendar year, or that they would have occurred at a single point in time, so an increase in the dummy variable value after 2010 is as plausible as no increase.

All that can really be concluded is that there appear to be unexplained factors influencing the number of enplaned O&D passengers that are not accounted for by the continuous variables in the model (and that are accounted for by the dummy variables), and that while these factors appear to have had an increasing effect from 2002 to 2010, their effect appears to have remained relatively constant after 2010.

The second analysis explored whether the factor behind the progressively increasing dummy variable for 2003 on was the significant increase in oil price from 2003 to 2008, which would have reduced discretionary income for any given household income (which was part of the rationale for including oil price in the original models). The previous analysis tested a model specification (Model 5a) that included oil price and found its coefficient statistically insignificant but that model also included the dummy variable for 2003 on, which may have overpowered (and potentially double-counted) any oil price effect. So Model 5a was re-estimated without the dummy variable for 2003 on (Model 7), which gave the results shown in Table D-69. The oil price coefficient became statistically significant with the expected sign (negative). However, the overall model fit was not as good as Model 4, particularly for the period 1990 to 2000, as shown in Table D-69. Of course, from 2001 on the dummy variables force a perfect fit for four of the ten years. Even so, the fit for the other years from 2003 to 2010 is not as good as with Model 4. In addition, the statistical significance of the coefficient for the disaggregated household income variable (percent of total household income received by the top 10% of households) declined to less than the 20% level and the statistical significance of the coefficients for average household income and average airline yield became weaker (although still significant at the 5% level).

The statistical significance for the coefficient for employment per person become stronger but the estimated value increased to 1.64, which seems implausibly high based on the reasoning in the previous section. The coefficients for average household income and the disaggregated household income variable dropped in absolute value, implying a much weaker income distribution effect.

Table D-69 Model Estimation Results – 1990-2010 (Models 4 and 7)

Variable		Model 4	Model 7
Intercept	Coefficient	-21.50	-12.63
	t-statistic	*(-5.68)*	*(-1.714)*
Average Household Income	Coefficient	2.047	1.409
	t-statistic	*(7.62)*	*(2.70)*
Oil Price	Coefficient		-0.1113
	t-statistic		*(-5.09)*
US Average Airline Yield	Coefficient	-0.849	-0.866
	t-statistic	*(-5.93)*	*(-3.23)*
Employment/person	Coefficient	0.520	1.641
	t-statistic	*(2.40)*	*(3.82)*
Pct of HH Income by Top 10%	Coefficient	-1.697	-1.000
	t-statistic	*(-4.44)*	*(-1.324)*
Dummy Variable 2001	Coefficient	-0.1514	-0.1445
	t-statistic	*(-8.31)*	*(-4.25)*
Dummy Variable 2002 on	Coefficient	-0.3059	-0.3014
	t-statistic	*(-9.71)*	*(-5.12)*
Dummy Variable 2003 on	Coefficient	-0.02347	
	t-statistic	*(-10.86)*	
Dummy Variable 2005	Coefficient	0.0539	0.0620
	t-statistic	*(3.20)*	*(1.978)*
Dummy Variable 2009	Coefficient	-0.0901	-0.1296
	t-statistic	*(-5.19)*	*(-4.15)*
Adjusted R Squared		0.994	0.979

Notwithstanding these concerns, Model 7 was used to prepare projections of enplaned O&D passengers per capita to 2015. The model under-predicted the enplaned passengers by about 1% in 2011, then over-predicted the enplaned passengers by between 1% and 2% from 2010 to 2014 (when oil prices were fairly high) but by 11% in 2015 when oil price dropped to less than half the level from 2011 to 2014, as shown in Figure D-66.

Although the *ex-post* forecast projection gave a similar fit to the actual enplaned passengers from 2011 to 2014 as Model 4b (slightly better for 2012 and 2013 and slightly worse for 2011 and 2014), the fit for 2015 was significantly worse. This was a direct result of the decline in oil price in 2015, which reduced the effect of the oil price term, and would presumably continue as long as oil prices remain low.

Therefore, on balance it does not appear that oil price fully accounts for whatever factors the dummy variable for 2003 on is accounting for, or that adding oil price in place of the dummy variable gives a better model. The effect on the model fit for the period 1990 to 2000 and for the years from 2003 to 2010 that do not have year-specific dummy variables, as well as the over-prediction of demand for 2015, suggests that the effect of oil price on enplaned passenger demand indicated by the estimated oil price coefficient at best only partly explains the decline in demand after 2002 accounted for by the dummy variable for 2003 on and may even be coincidental.

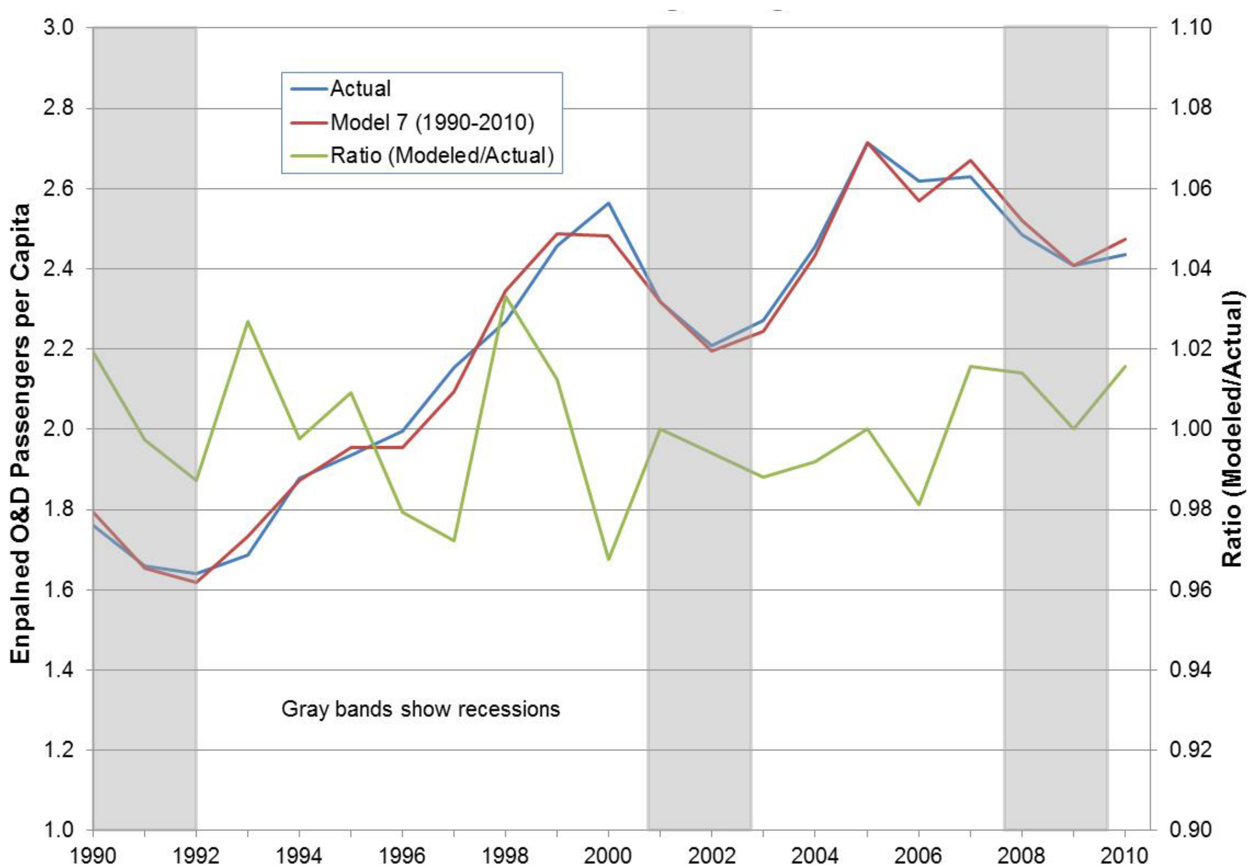

Figure D-66 Comparison of Projected and Actual Enplaned O&D Passengers for the Baltimore-Washington Region (Model 7) – 1990 to 2010

Discussion

The estimated coefficient for oil price in Model 7 implies an enplaned passenger demand elasticity of about 0.11. Thus if this elasticity is correct the three-fold increase in oil price from 2003 to 2008 (291%) would have reduced demand by about 32%. The estimated coefficient for average household income in Model 7 implies an enplaned passenger demand elasticity of 1.41. Thus the increase in average household income required to offset the effect of this increase in oil price over this period would have been about 23%. If the effect of the increase in oil price was to reduce discretionary household income available for air travel, this implies that on average households spent about 12% of their household income (0.23/1.91) on gasoline and other oil products in 2003, increasing to about 28% of their household income (2.91*0.12/1.23) in 2008, both of which seem far too high.

A recent paper by Gelman *et al.* (2017) on the response of consumer spending to changes in gasoline prices indicated that average household spending on gasoline in 2014 (when oil price was about 90% of its peak in 2008) was about 5% of average total household spending. While household spending is not the same thing as household income, and averages across all households can be misleading, this supports the argument that the above implied percentages are far too high.

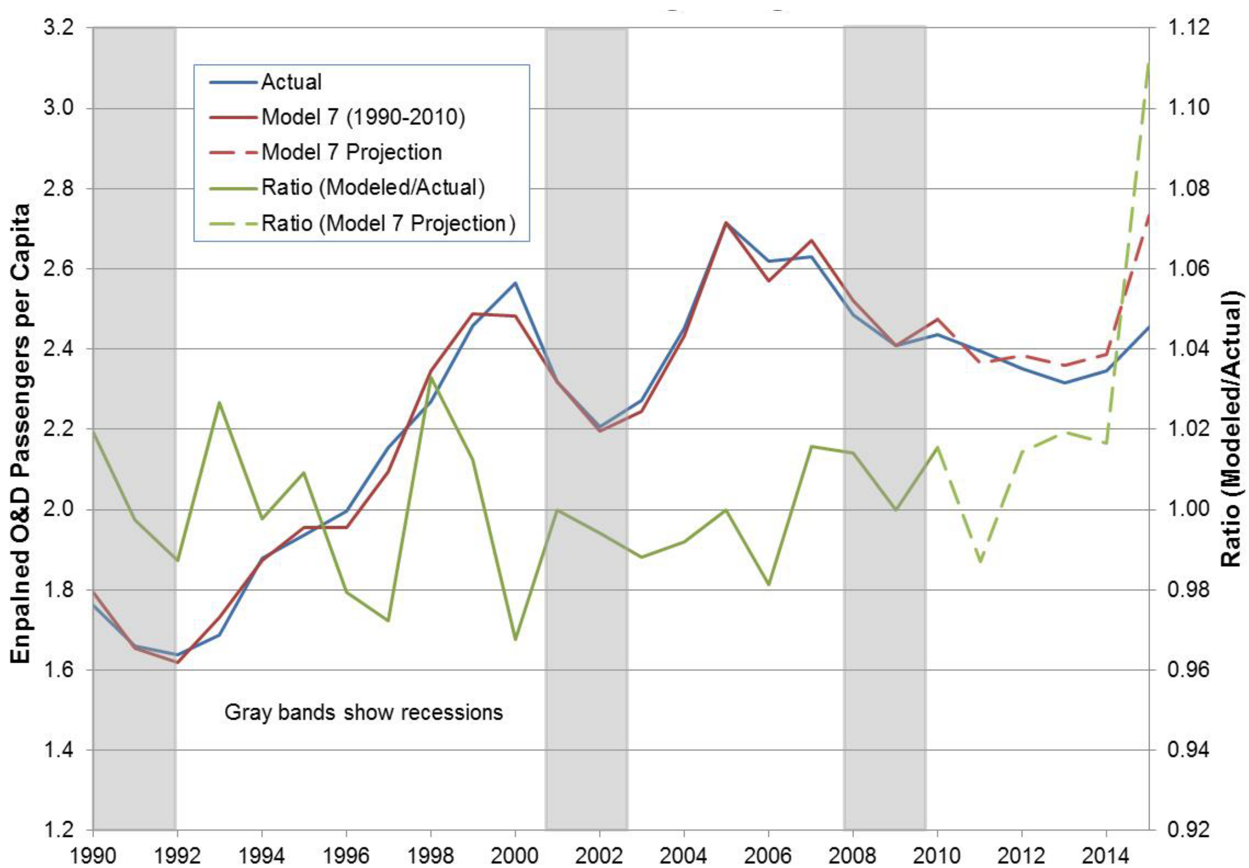

Figure D-67 Projection of Enplaned O&D Passengers per Capita for the Baltimore-Washington Region 2010 to 2015 – Model 7

It should also be noted that refiners acquisition cost and gasoline price are also not the same thing and the change in retail gasoline price over time may well differ from the change in refiners acquisition cost, as well as differ from region to region. This points out the importance of ensuring that the variables used in modeling are carefully chosen to reflect the factors that they are assumed to be reflecting to the greatest extent possible.

A similar concern arises with the use of national average airline yield, the changes in which from year to year may differ considerably from the changes in average airfares in a given region or at a given airport.

Contribution of the Disaggregated Income Variable to Model Performance

A third additional analysis examined how much the inclusion of the disaggregated income variable (percent of total household income received by the top 10% of households by income) in Model 4 improved the *ex-post* forecast projection of enplaned O&D passengers from 2010 to 2015.

This analysis estimated a new model (Model 8) that did not include the disaggregated income variable but retained the dummy variables used in Model 4. The estimated model coefficients for both models are shown in Table D-70.

Table D-70 Model Estimation Results – 1990-2010 (Models 4 and 8)

Variable		Model 4	Model 8
Intercept	Coefficient	-21.50	-6.54
	t-statistic	*(-5.68)*	*(-2.37)*
Average Household Income	Coefficient	2.047	0.973
	t-statistic	*(7.62)*	*(5.19)*
US Average Airline Yield	Coefficient	-0.849	-1.185
	t-statistic	*(-5.93)*	*(-6.10)*
Employment/person	Coefficient	0.520	0.885
	t-statistic	*(2.40)*	*(2.75)*
Pct of HH Income by Top 10%	Coefficient	-1.697	
	t-statistic	*(-4.44)*	
Dummy Variable 2001	Coefficient	-0.1514	-0.1897
	t-statistic	*(-8.31)*	*(-7.39)*
Dummy Variable 2002 on	Coefficient	-0.3059	-0.3870
	t-statistic	*(-9.71)*	*(-9.41)*
Dummy Variable 2003 on	Coefficient	-0.02347	-0.02513
	t-statistic	*(-10.86)*	*(-7.38)*
Dummy Variable 2005	Coefficient	0.0539	0.0257
	t-statistic	*(3.20)*	*(1.029)*
Dummy Variable 2009	Coefficient	-0.0901	-0.0941
	t-statistic	*(-5.19)*	*(-3.39)*
Adjusted R Squared		0.994	0.985

The adjusted R squared for Model 8 is slightly lower than for Model 4 and the statistical significance of the estimated coefficients of the dummy variables is somewhat lower although the coefficients for all the dummy variables except the year-specific dummy variable for 2005 are still statistically significant at better than the 1% level. The estimated coefficient for average household income dropped to less than half its value in Model 4, as would be expected without the disaggregated income variable based on the previous model development, and its statistical significance decreased slightly, although it is still statistically significant at better than the 1% level. The statistical significance of the other two continuous variables increased slightly and their estimated values also increased.

Model 8 was then used to project enplaned O&D passengers from 2010 to 2015, assuming that the factor behind the progressively increasing dummy variable for 2003 on remained constant from 2010 (Model 4b). The fit for both models for the period 1990 to 2010 and the comparative projections to 2015 are shown in Figure D-68.

It can be seen that the fit of Model 8 to the actual data for the period 1990 to 2010 is not as good as for Model 4, although the difference is not large, as would be expected from the model estimation results. The projected enplaned O&D passengers per person for 2011 to 2015 are significantly lower than given by Model 4 and lower than the actual enplaned O&D passengers per person. Whereas Model 4 tended to slightly over-predict the passengers, Model 8 under-predicted the passengers by an average of about 2.2% over the period from 2011 to 2015. If the assumptions regarding the value of the dummy variable for 2003 on used for the projections were adjusted as discussed above to reduce the over-prediction by Model 4 (a small increase in the value of the dummy variable in 2011 with the value remaining constant thereafter), the under-prediction by Model 8 would increase.

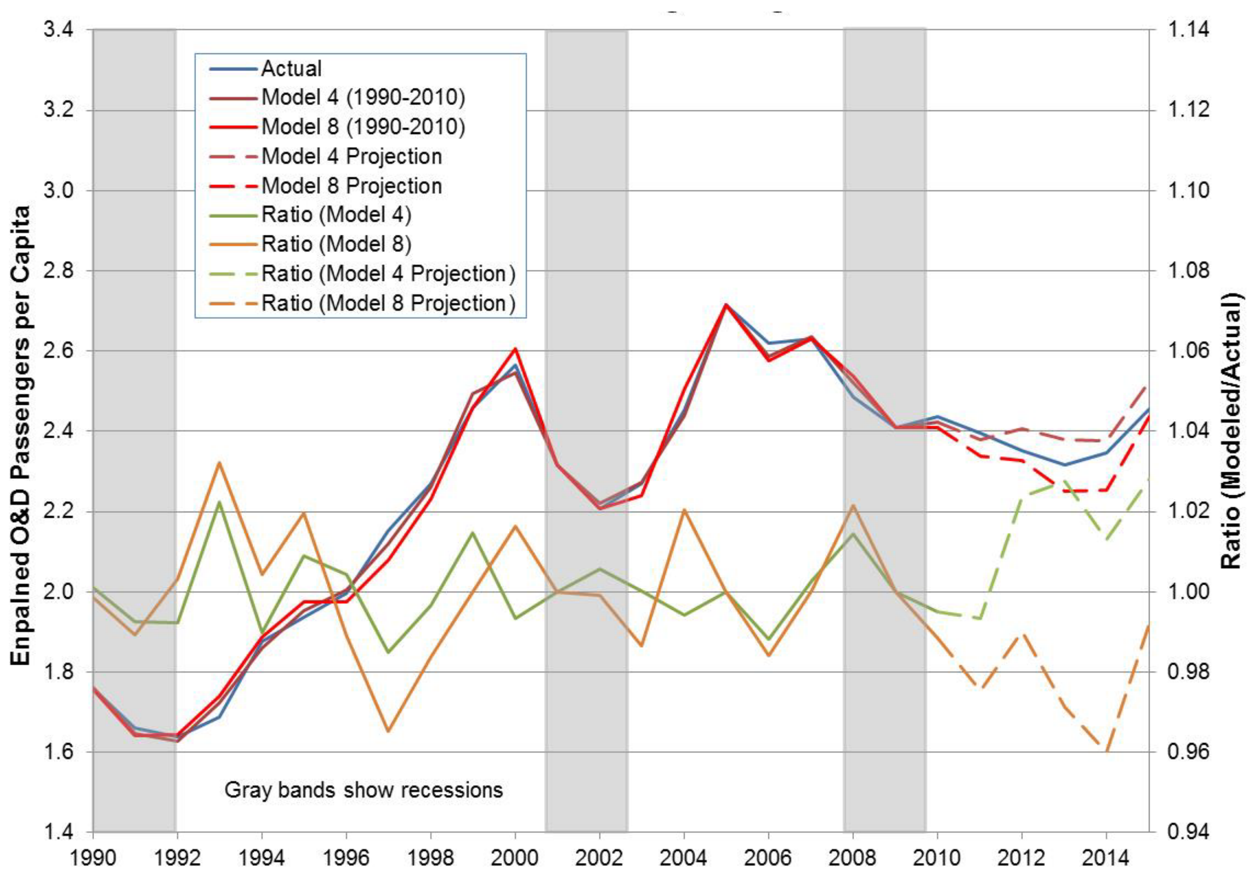

Figure D-68 Projection of Enplaned O&D Passengers per Capita for the Baltimore-Washington Region 2010 to 2015 – Models 4 and 8

Therefore it appears that the inclusion of the disaggregated income variable not only improved the model fit for the period 1990 to 2010, but also improved the projections of enplaned O&D passengers for the period from 2010 to 2015, although the extent of any improvement in the projections is very dependent on the assumptions regarding the value of the dummy variable for 2003 on after 2010 and future changes in household income distribution reflected in the disaggregated income variable.

Changing the Assumptions about the Decline in Air Passenger Demand after 2002

A fourth additional analysis examined what happens to the model coefficients and fit if the dummy variables are revised to force the decline in enplaned passengers accounted for by the dummy variables for 2002 on and 2003 on to progressively reduce from 2002 to 2010, rather than increase as assumed in Model 4, to give a progressively decreasing reduction of the projected enplaned passengers below the level given by the continuous variables and the other year-specific dummy variables until this effect disappeared by 2010. This was achieved by linearly reducing the value of the dummy variable for 2002 on from 1 in 2002 to 0 in 2010 and eliminating the dummy variable for 2003 on. This revised model (Model 9) was estimated on the period 1990 to 2010 and the estimated coefficients are shown in Table D-71, together with the previous results for Model 4.

Table D-71 Model Estimation Results – 1990-2010 (Models 4 and 9)

Variable		Model 4	Model 7
Intercept	Coefficient	-21.50	-56.29
	t-statistic	*(-5.68)*	*(-5.07)*
Average Household Income	Coefficient	2.047	4.388
	t-statistic	*(7.62)*	*(5.33)*
US Average Airline Yield	Coefficient	-0.849	0.757
	t-statistic	*(-5.93)*	*(2.38)*
Employment/person	Coefficient	0.520	-0.497
	t-statistic	*(2.40)*	*(-0.609)*
Pct of HH Income by Top 10%	Coefficient	-1.697	-4.688
	t-statistic	*(-4.44)*	*(-3.76)*
Dummy Variable 2001	Coefficient	-0.1514	0.0095
	t-statistic	*(-8.31)*	*(0.172)*
Dummy Variable 2002 on	Coefficient	-0.3059	0.0690
	t-statistic	*(-9.71)*	*(1.157)*
Dummy Variable 2003 on	Coefficient	-0.02347	
	t-statistic	*(-10.86)*	
Dummy Variable 2005	Coefficient	0.0539	0.1640
	t-statistic	*(3.20)*	*(2.78)*
Dummy Variable 2009	Coefficient	-0.0901	-0.0395
	t-statistic	*(-5.19)*	*(-0.581)*
Adjusted R Squared		0.994	0.903

The model fit as given by the adjusted R squared is not as good as for Model 4 and several of the estimated coefficients are statistically insignificant, including the coefficient for employment per person and all the dummy variables except the year-specific dummy variable for 2005. The estimated coefficients for the dummy variables for 2001 and 2002 on are not only insignificant, but positive, which is contrary to the expected effect. What appears to have happened is that because the declining value of the dummy variable for 2003 on is opposite to the effect found when a model estimated on the continuous variables for 1990 to 2000 is compared to the actual traffic for 2001 to 2010, the estimated coefficients for the income variables become much larger in order to overpower the effect of the dummy variable, while the estimated coefficient for average airline yield becomes positive, which is the opposite of the expected sign. The estimated coefficient for the employment per person becomes negative, which is also the opposite of the expected sign, although this value is statistically insignificant. Figure D-69 shows the fit of Model 9 to the actual data compared to that of Model 4.

Clearly, Model 9 does not fit the actual data nearly as well as Model 4, except for the years from 2001 on with year-specific dummy variables (2001, 2005 and 2009) where the dummy variables force a perfect fit. The distortions in the estimated coefficients for the continuous variables resulting from the declining effect of the dummy variable for 2002 on can be seen in the much worse fit of Model 9 to the actual data compared to Model 4 during the period 1990 to 2000, which has no dummy variables so the projected enplaned passengers depend entirely on the continuous variables. Of course the estimated coefficients for the continuous variables in Model 9 have obvious problems, aside from how well the model fits the actual data, as discussed above.

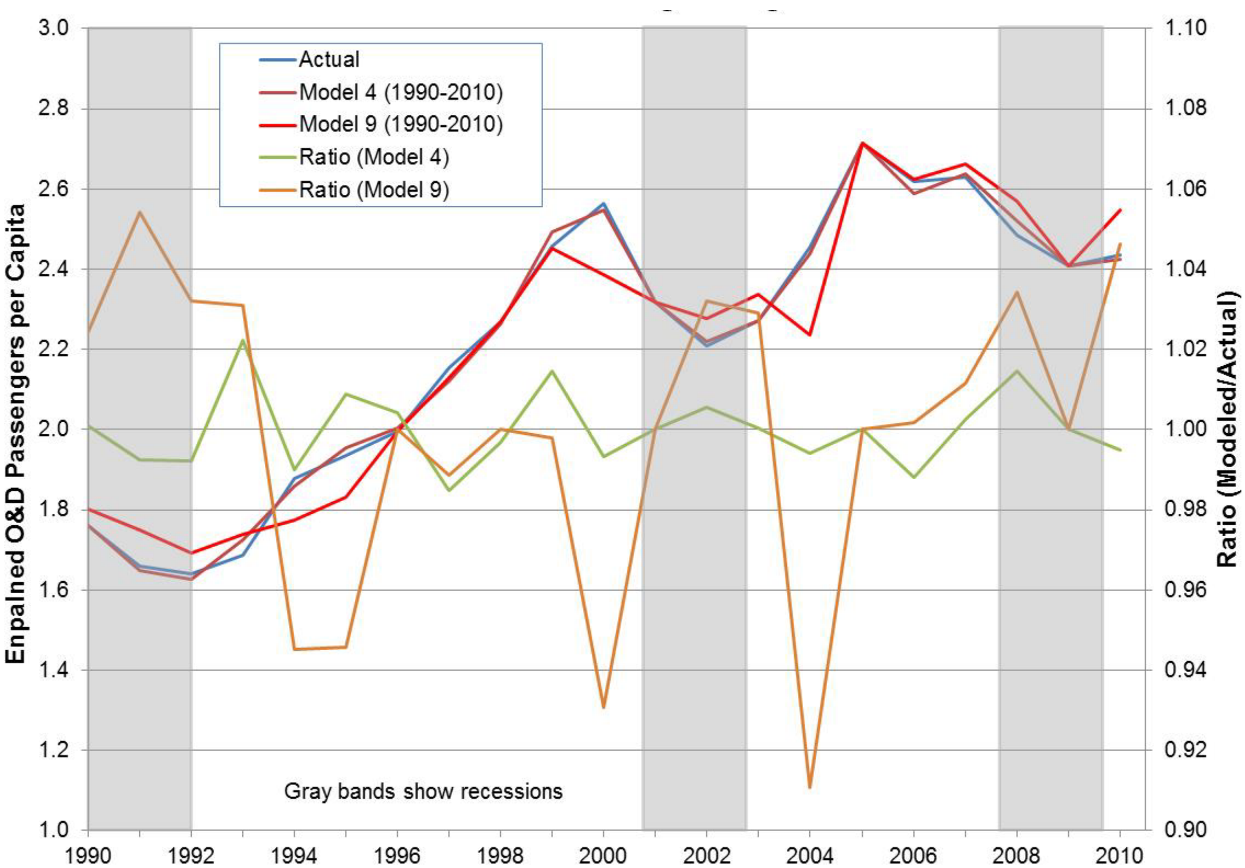

Figure D-69 Comparison of Projected and Actual Enplaned O&D Passengers for the Baltimore-Washington Region (Models 4 and 9) – 1990 to 2010

In summary, it appears that the progressive reduction in enplaned O&D passengers after 2002 relative to the level projected by the continuous variables that is accounted for by the dummy variable for 2003 on in Model 4 is not simply an artifact of the way that the dummy variables were defined and that replacing the dummy variables for 2002 on and 2003 on with one that declines in value from 1 in 2002 to zero in 2010 results in a much worse model on multiple counts.

D.4 References

Gelman, Michael, Yuriy Gorodnichenko, Shachar Kariv, Dmitri Koustas, Matthew D. Shapiro, Dan Silverman, and Steven Tadelis (2017). *The Response of Consumer Spending to Changes in Gasoline Prices*. Cambridge, MA: National Bureau of Economic Research, March (NBER Working Paper No. 22969).

Kennedy, Peter (1985). *A Guide to Econometrics (2nd Edition).* Cambridge, MA: The MIT Press.