URBAN SCIENCE

Determining Automotive Demand: Demographic or Registration Data?

21 - 25 May 2011

James G. Wendelberger, Ph.D.
Director of Statistical Analysis

We are Urban Science.

We are retail performance experts that help our clients increase market share and improve profitability.

Conquering the toughest business challenges through our unique combination of:

- Cutting edge analytics
- Proprietary software
- Real-world experience

One of the critical 4 P's in marketing a product is placement. The focus here has to do with determining in which channels one distributes and services a product. In other words this is retail Network planning and management. Before performing this analysis, the issue to tackle is what data should one use? The data required in the automotive market for this type of analysis needs to be an indicator of consumer automotive potential. Two possible sources were considered, demographics (population, households, etc.) and vehicle registrations.

Registration Data or Demographic/Other Data?

Registration Data

- By State or Other Government Level
- New Vehicle Registration

Demographic (or other) Data

- Demographic
- Economic
- Other

Why?

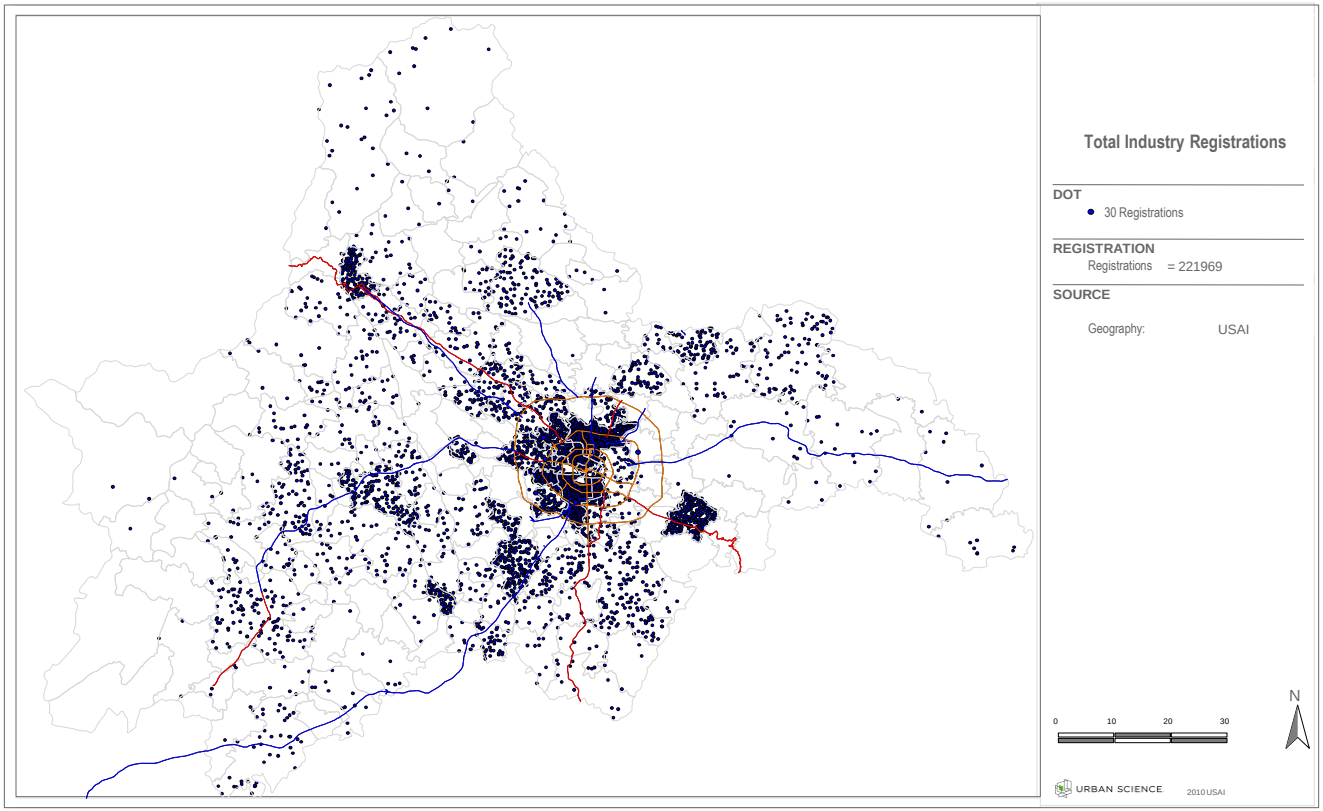
- Network Planning

Study

- Future Demand

Registration Data

- Who bought what from whom?
- Address provides location.



Other Data

Demographics

1. Population
2. Households
3. Etc.

Economic

1. Wealth
2. GDP
3. Inflation
4. Stock Market

Geographic

1. Barriers – Mountains, Rivers, Lakes, etc.
2. Satellite Data

Analysis Level

1. 2,832 Postal Districts

Registration Data

- Total 234 variables
- Three Years of Data (2005 – 2007)
- For Each of Nine Segments, Sixteen Vehicle Models and Total
  - Competitive Market
  - Brand Registrations
  - Expected

Demographic, Economic and Other Data (Selected)

1,"Age 0 to 15"
2,"Age 16 to 24"
3,"Age 25 to 44"
4,"Age 45 to 64"
5,"Age 65 and over"
6,"Total for Age Structure"
7,"Households with no cars"
8,"Households with one car"
9,"Households with two cars"
10,"Households with three cars"
11,"Households with four or more cars"
12,"All cars or vans in the area"
13,"Total for Cars"
14,"Economically Active - Employees - Part time"
15,"Economically Active - Employees - Full time"
16,"Economically Active - Self employed"
17,"Economically Active - Unemployed"
18,"Economically Active - Full time student"
19,"Economically Inactive - Retired"
20,"Economically Inactive - Student"
21,"Economically Inactive - Looking after home/family"
22,"Economically Inactive - Permanently sick/disabled"
23,"Economically Inactive - Other"
24,"Number of unemployed - Aged 16 to 24"
25,"Number of unemployed - Aged 50 and over"
26,"Number of unemployed - Who have never worked"
27,"Number of unemployed - Who are long term unemployed"
28,"Total for Economic Activity"
65,"One person households - pensioner"
66,"One person households - other"
67,"One family only households - all pensioners"
68,"One Family Married households - no children"
69,"One Family Married households - with dependent children"
70,"One Family Married households - all children non-dependent"
71,"One Family Cohabiting households - no children"
72,"One Family Cohabiting households - with dependent children"
73,"One Family Cohabiting households - all children non dependent"
74,"One Family Lone parent households - with dependent children"
75,"One Family Lone parent households - all children non-dependent"
76,"Other households - with dependent children"
77,"Other households - all student"
78,"Other households - all pensioner"
79,"Other households - other"
80,"Total Household Composition"
81,"Total - Males"
82,"Married or re-married - Males"
83,"Single (never married), separated, divorced or widowed - Males"
84,"Total - Females"
85,"Married or re-married - Females"
86,"Single (never married), separated, divorced or widowed - Females"
87,"Total for Marital Status"
88,"Male population"
89,"Female population"
90,"Total for Population"
91,"People aged 16 to 74 with highest qualification attained level 1"
92,"People aged 16 to 74 with highest qualification attained level 2"
93,"People aged 16 to 74 with highest qualification attained level 3"
94,"People aged 16 to 74 with highest qualification attained level 4/5"
95,"People aged 16 to 74 with no or other qualifications"
96,"Total number of full-time students and school children aged 16 to 17"
97,"Total number of full-time students and school children aged 18 to 74"
98,"Full-time students aged 18 to 74 Economically active - In employment"
99,"Full-time students aged 18 to 74 Economically active - Unemployed"
100,"Full-time students aged 18 to 74 Economically inactive"
101,"Total for Qualifications & Students"
102,"AB - Higher and intermediate managerial/administrative/professional"
103,"C1 - Supervisory, clerical, junior managerial/administrative/professional"
104,"C2 - Skilled manual workers"
105,"D - Semi-skilled and unskilled manual workers"
106,"E - On state benefit, unemployed, lowest grade workers"
107,"Total for Social Grade"
108,"Total for Households"

Why?

Network Planning Needs

1. Dealership Location
2. Customer Convenience

Tradeoff Between Data Cost and Information Added

Concentrate on Information in Data Types

Demographic and registration data from prior years are used to predict the current year of registration data. The contribution to R-squared of the demographic data and registration data is measured. The prior year registration data is shown to be a superior predictor than the demographic data. The model with prior year registration data has a higher R-squared value than the model with demographic data when predicting automotive demand.

Study

Compare Different Data Types

Demographic versus Registration

Procedure

Compare Registration Only versus Demographics Only GainSmarts Models

Analytical Results

Highest R-Squared Values Occur for the Registration Data

Conclusion

Registration Data Implicitly Includes the Demographic Data Information

Compare Different Data Types

Demographic Data versus Registration Data

Compare Registration versus Demographics Models Using GainSmarts Models

GainSmarts Software has won Knowledge and Data Discover Awards – see

http://urbanscience.com/services/customer.html

Background: An analysis is conducted that utilized both demographics and prior year registrations to predict automotive demand in an effort to determine which is a better indicator or predictor of future demand. The results of the analysis of data from the United Kingdom is utilized.

Analytical Procedure

- The automotive registrations by census tract/postal code for the current year are predicted.

- Independent variables are demographics and prior year automotive registrations.

- The GainSmarts software is used to determine the best model to predict the dependent variables from the independent variables.

- The analysis proceeds by including and excluding groups of independent variables.

- GainSmarts is used to select the best model using the available data.

- The R-squared is used to evaluate the quality of the resulting models relative to one another.

Analytical Results

Dependent Variable	Actual R-Squared	Demographics	Prior All Type Registrations	Prior Year Only Actual Registrations	All Prior Years Only Actual Registrations	Missing	Demographics
2007 Actual Registrations	80.06%	X	X	X	X	Removed	10,24,31
2007 Actual Registrations	79.91%		X	X	X	Removed	
2007 Actual Registrations	59.20%	X				Removed	8,10,11,21,23,24,36,37,38,45,48,65,68,69,70,72,85,93,102,104,105
2007 Actual Registrations	81.60%	X	X	X	X	Set to Zero	11,45,54,67,71,93,102
2007 Actual Registrations	81.74%		X	X	X	Set to Zero	
2007 Actual Registrations	62.71%	X				Set to Zero	11,19,22,23,31,42,79,92,93,100,102,104
2007 Actual Registrations	75.69%			X		Set to Zero	
2007 Actual Registrations	79.54%			X	X	Set to Zero	
2006 Actual Registrations	68.00%	X				Set to Zero	13,23,32,36,43,72,77,79,92,93,99,102,104
2006 Actual Registrations	85.36%	X	X	X	X	Set to Zero	14,20,30,34,42,77,79,102
2006 Actual Registrations	84.91%		X	X		Set to Zero	
2005 Actual Registrations	65.45%	X				Set to Zero	15,22,30,34,36,37,39,43,79,100,102,104,108

Detailed Demographic Data R-Squared

- Detailed Demographic Data Does Not Add Practically to the overall R-squared value – 80.06% - 79.91% = .15%

- Detailed Demographic Data Alone Explains Only A Fraction of the Overall R-Squared Without demographic Data 59.20% versus 79.91% (or .5920/.7991 = 74.08%)

Conclusion

When available then prior year registrations are a better predictor of automotive demand than are demographic data.

1. Prior registrations are a better predictor of current registrations than demographic information alone – this is determined by the r-squared when using the Prior year registration data being larger than the r-squared when using the demographic data alone.
2. Prior year registrations can give an r-squared around 80% or more when predicting current year registrations. This is a very large value of r-squared in absolute terms.

Demographic data:

3. The contribution to r-squared of the demographic data deteriorates in time since the census. The r-squared value for 2007 is the least in the data tested here.

In the demographic data only models many of the variables entering the final model are different from year to year. The demographic data is an inconsistent predictor.

4. If there is no registration data available then demographic data may be used for prediction.

The registration data intrinsically accounts for all demographic and any other data that may be useful in forecasting automotive demand. For this reason it makes intuitive sense that adding in demographic information will not substantially help in predicting automotive demand.

Overall Conclusion

Demographic Data are of Limited Additional Value,
over Automotive Registrations,
in Explaining Performance or Registration Demand

James Wendelberger
Urban Science Applications, Inc.
620 Los Pueblos
Los Alamos, NM 87544/2615

JGWendelberger@UrbanScience.com

1-505-662-2838

Thank You!

Dr. James G. Wendelberger

Director of Statistical Analysis