University of Central Florida
College of Business

QMB 6911
Capstone Project in Business Analytics

Solutions: Problem Set #7

1 Data Description

This analysis follows the script `assignment7.R` to produce a more accurate model for used Home prices with the data from `HomeSales.dat` in the `Data` folder. The dataset includes the following variables.

Variable	Definition
`year_built`	the year in which the house was constructed
`num_beds`	the number of bedrooms in the house
`num_baths`	the number of bathrooms in the house
`floor_space`	the area of floor space in the house, in square feet
`lot_size`	the area of lot on which the house was built, in square feet
`has_garage`	an indicator for whether the house has a garage
`has_encl_patio`	an indicator for whether the house has an enclosed patio
`has_security_gate`	an indicator for whether the property is accessed through a security gate
`has_pool`	an indicator for whether the property includes a pool
`transit_score`	an integer to represent the convenience of transportation options
`school_score`	an integer to represent the quality of the schools in the county
`type_of_buyer`	a categorical variable to indicate the type of buyer, either "Owner-Occupied" or "Rental"
`price`	the price at which the home was sold

Two new variables were created:

Variable	Definition
`log_price`	logarithm of the house price
`age`	the age of the house, calculated based on the year built

I will first estimate a model with nonlinear functional forms, and then consider exclusions of insignificant variables from the full model, with nonlinearity taken into account. This approach allows for inclusion of possibly irrelevant variables and avoids excluding any relevant variables.

2 Example with a Categorical Numerical Variable

2.1 School Score

Decided to try the variable school score as a factor, differently than what it was used on Problem Set #6. Table 1 shows odel 1 being the full model, model 2 the best reduced model found in Problem Set #6 and model 3 the full model using school score as a factor. We see there is no difference in R squared for any of these models, however school schore goes from being insiginifcant is model 1 to having each different factor being significant in model 3. Floor space continues to be statistically insiginificant in both models.

	Model 1	Model 2	Model 3
(Intercept)	12.22***	12.28***	12.16***
	(0.06)	(0.04)	(0.07)
as.factor(num_beds)2	0.08**	0.09**	0.07*
	(0.03)	(0.03)	(0.03)
as.factor(num_beds)3	0.20***	0.23***	0.19***
	(0.04)	(0.04)	(0.04)
as.factor(num_beds)4	0.42***	0.46***	0.41***
	(0.05)	(0.04)	(0.05)
as.factor(num_beds)6	0.48***	0.54***	0.47***
	(0.08)	(0.07)	(0.08)
as.factor(num_beds)8	0.83***	0.92***	0.83***
	(0.10)	(0.07)	(0.10)
as.factor(num_baths)2	0.07***	0.08***	0.07***
	(0.02)	(0.02)	(0.02)
as.factor(num_baths)3	0.09*	0.10**	0.09**
	(0.03)	(0.03)	(0.03)
floor_space	0.00		0.00
	(0.00)		(0.00)
lot_size	0.00***	0.00***	0.00***
	(0.00)	(0.00)	(0.00)
has_garage	0.28***	0.31***	0.28***
	(0.03)	(0.02)	(0.03)
has_encl_patio	0.05**	0.05**	0.05**
	(0.01)	(0.01)	(0.01)
has_security_gate	0.23***	0.23***	0.23***
	(0.02)	(0.02)	(0.02)
has_pool	0.10***	0.10***	0.10***
	(0.03)	(0.03)	(0.03)
transit_score	0.07***	0.07***	0.07***
	(0.00)	(0.00)	(0.00)
school_score	0.00		
	(0.00)		
type_of_buyerRental	0.11***	0.11***	0.11***
	(0.02)	(0.02)	(0.02)
age	−0.01***	−0.01***	−0.01***
	(0.00)	(0.00)	(0.00)
as.factor(school_score)2			0.10**
			(0.03)
as.factor(school_score)3			0.10**
			(0.03)
as.factor(school_score)4			0.09**
			(0.03)
as.factor(school_score)5			0.09**
			(0.03)
as.factor(school_score)6			0.10**
			(0.03)
as.factor(school_score)7			0.10**
			(0.03)
as.factor(school_score)8			0.10**
			(0.03)
as.factor(school_score)9			0.09*
			(0.04)
as.factor(school_score)10			0.10*

	Model 1	Model 2	Model 3
			(0.04)
R^2	0.65	0.65	0.65
Adj. R^2	0.65	0.65	0.65
Num. obs.	1862	1862	1862

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 1: Dollar Value of Home Prices

2.2 Transit Score

Separated variable for transit score in 4 different categories to analyze significance in the model. Table 2 shows model 1 being the full model with school score as a factor and model 2 has the adtional separate categories for transit score. None of the new catecories are statistically significant while the original transit score variable is. There is also no diffrence on R squared and adjusted R squared New categories will not be added to the model.

	Model 1	Model 2
(Intercept)	12.16***	12.19***
	(0.07)	(0.07)
as.factor(num_beds)2	0.07*	0.07*
	(0.03)	(0.03)
as.factor(num_beds)3	0.19***	0.19***
	(0.04)	(0.04)
as.factor(num_beds)4	0.41***	0.40***
	(0.05)	(0.05)
as.factor(num_beds)6	0.47***	0.46***
	(0.08)	(0.08)
as.factor(num_beds)8	0.83***	0.81***
	(0.10)	(0.10)
as.factor(num_baths)2	0.07***	0.08***
	(0.02)	(0.02)
as.factor(num_baths)3	0.09**	0.09**
	(0.03)	(0.03)
floor_space	0.00	0.00
	(0.00)	(0.00)
lot_size	0.00***	0.00***
	(0.00)	(0.00)
has_garage	0.28***	0.27***
	(0.03)	(0.03)
has_encl_patio	0.05**	0.05**
	(0.01)	(0.01)
has_security_gate	0.23***	0.23***
	(0.02)	(0.02)
has_pool	0.10***	0.10***
	(0.03)	(0.03)
transit_score	0.07***	0.05***
	(0.00)	(0.01)
as.factor(school_score)2	0.10**	0.10**
	(0.03)	(0.03)
as.factor(school_score)3	0.10**	0.10**
	(0.03)	(0.03)
as.factor(school_score)4	0.09**	0.09**

	Model 1	Model 2
	(0.03)	(0.03)
as.factor(school_score)5	0.09^{**}	0.09^{**}
	(0.03)	(0.03)
as.factor(school_score)6	0.10^{**}	0.10^{**}
	(0.03)	(0.03)
as.factor(school_score)7	0.10^{**}	0.10^{**}
	(0.03)	(0.03)
as.factor(school_score)8	0.10^{**}	0.10^{**}
	(0.03)	(0.03)
as.factor(school_score)9	0.09^{*}	0.09^{*}
	(0.04)	(0.04)
as.factor(school_score)10	0.10^{*}	0.10^{*}
	(0.04)	(0.04)
type_of_buyerRental	0.11^{***}	0.11^{***}
	(0.02)	(0.02)
age	-0.01^{***}	-0.01^{***}
	(0.00)	(0.00)
ts_cat3-6		0.03
		(0.04)
ts_cat6-8		0.05
		(0.06)
ts_cat8+		0.11
		(0.07)
R^2	0.65	0.65
Adj. R^2	0.65	0.65
Num. obs.	1862	1862

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 2: Dollar Value of Home Prices

2.3 Lot Size

Transformed lot size into 5 categories. On table 3 model 1 is our current best model (full model with school score as factor) and model 2 is the full model with the new lot size categories. We can see that category 5050-6650 is statistically significant while the others are not, this sugestsa non linear relationship between lot size and the home price. There is no difference on the R squared or adjusted R squared for these models, but since lot size is a continuous variable, we can do better.

3 Nonlinear Model Specifications

3.1 Quadratic Specification for lot size

4 Example with a Categorical Numerical Variable

I will revisit the recommended linear model from Problem Set #6, augmented with a quadratic specification for lot size. This allowed for an increasing relationship between price and lot size, for Homes with low lot size, but a decreasing relationship for the tractors with high lot size. In doing so, I will further investigate nonlinear relationships by incorporating another nonlinear but parametric specification for the value of lot size. This parametric analysis will be performed using the Box-Tidwell framework to investigate whether the value of these characteristics are best described with parametric nonlinear forms.

	Model 1	Model 2	Model 3	Model 4	Model 5
(Intercept)	12.22***	12.28***	12.16***	12.19***	12.15***
	(0.06)	(0.04)	(0.07)	(0.07)	(0.08)
as.factor(num_beds)2	0.08**	0.09**	0.07*	0.07*	0.07*
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
as.factor(num_beds)3	0.20***	0.23***	0.19***	0.19***	0.21***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)
as.factor(num_beds)4	0.42***	0.46***	0.41***	0.40***	0.45***
	(0.05)	(0.04)	(0.05)	(0.05)	(0.05)
as.factor(num_beds)6	0.48***	0.54***	0.47***	0.46***	0.53***
	(0.08)	(0.07)	(0.08)	(0.08)	(0.09)
as.factor(num_beds)8	0.83***	0.92***	0.83***	0.81***	0.89***
	(0.10)	(0.07)	(0.10)	(0.10)	(0.11)
as.factor(num_baths)2	0.07***	0.08***	0.07***	0.08***	0.08***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
as.factor(num_baths)3	0.09*	0.10**	0.09**	0.09**	0.09*
	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)
floor_space	0.00		0.00	0.00	0.00
	(0.00)		(0.00)	(0.00)	(0.00)
lot_size	0.00***	0.00***	0.00***	0.00***	0.00**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
has_garage	0.28***	0.31***	0.28***	0.27***	0.28***
	(0.03)	(0.02)	(0.03)	(0.03)	(0.04)
has_encl_patio	0.05**	0.05**	0.05**	0.05**	0.03
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
has_security_gate	0.23***	0.23***	0.23***	0.23***	0.23***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
has_pool	0.10***	0.10***	0.10***	0.10***	0.11***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
transit_score	0.07***	0.07***	0.07***	0.05***	0.07***
	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)
school_score	0.00				
	(0.00)				
type_of_buyerRental	0.11***	0.11***	0.11***	0.11***	0.13***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
age	−0.01***	−0.01***	−0.01***	−0.01***	−0.01***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
as.factor(school_score)2			0.10**	0.10**	0.11**
			(0.03)	(0.03)	(0.04)
as.factor(school_score)3			0.10**	0.10**	0.08*
			(0.03)	(0.03)	(0.03)
as.factor(school_score)4			0.09**	0.09**	0.10**
			(0.03)	(0.03)	(0.03)
as.factor(school_score)5			0.09**	0.09**	0.08*
			(0.03)	(0.03)	(0.03)
as.factor(school_score)6			0.10**	0.10**	0.08*
			(0.03)	(0.03)	(0.04)
as.factor(school_score)7			0.10**	0.10**	0.08*
			(0.03)	(0.03)	(0.04)
as.factor(school_score)8			0.10**	0.10**	0.09*
			(0.03)	(0.03)	(0.04)
as.factor(school_score)9			0.09*	0.09*	0.09*
			(0.04)	(0.04)	(0.04)
as.factor(school_score)10			0.10*	0.10*	0.08
			(0.04)	(0.04)	(0.05)
ts_cat3-6				0.03	
				(0.04)	

5 Linear Regression Model

A natural staring point is the recommended linear model from Problem Set #6, augmented with the quadratic specification for lot size.

5.1 Quadratic Specification for Lot Size

In the demo for Problem Set #6, we considered the advice of a used tractor dealer who reported that overpowered used tractors are hard to sell, since they consume more fuel. This implies that tractor prices often increase with lot size, up to a point, but beyond that they decrease. To incorporate this advice, I created and included a variable for squared lot size. A decreasing relationship for high values of lot size is characterized by a positive coefficient on the lot size variable and a negative coefficient on the squared lot size variable.

The results of this regression specification are shown in Table 4. The squared lot size variable has a coefficient of $-2.081e-05$, which is nearly ten times as large as the standard error of $2.199e-06$, which is very strong evidence against the null hypothesis of a positive or zero coefficient. I conclude that the log of the sale price does decline for large values of lot size.

With the squared lot size variable, the $\bar{R}^2$ is 0.764, indicating that it is a much stronger model than the others we considered. The F-statistic is large, indicating that it is a better candidate than the simple average log sale price. The new squared lot size variable is statistically significant and the theory behind it is sound, since above a certain point, added lot size may not improve performance but will cost more to operate. This new model is much improved over the previous models with a linear specification for lot size. Next, I will attempt to improve on this specification, as we did for Problem Set #8.

6 Nonlinear Specifications

6.1 The Box–Tidwell Transformation

The Box–Tidwell function tests for non-linear relationships to the mean of the dependent variable. The nonlinearity is in the form of an exponential transformation in the form of the Box-Cox transformation, except that the transformation is taken on the explanatory variables.

6.1.1 Transformation of lot size

Performing the transformation on the lot size variable produces a modified form of the linear model. This specification allows a single exponential transformation on lot size, rather than a quadratic form.

```
 MLE of lambda Score Statistic (z) Pr(>|z|)
       0.13681             -2.0662  0.03881 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

iterations =  7
```

The R output is the statistics for a test of nonlinearity: that the exponent λ in the Box–Tidwell transformation is zero. The "`MLE of lambda`" statistic is the optimal exponent on lot size. Similar to the Box-Cox transformation, with Box-Tidwell, the exponents are on the explanatory variables and are all called lambda, in contrast to the parameter τ in our class notes. The exponent is significantly different from 0, although it is a small positive value, which suggests an increasing relationship for the value of lot size with a slope that is sharply declining. Next I consider the possibility of a changing relationship for the next continuous variable.

6.1.2 Transformation of Age

```
 MLE of lambda Score Statistic (z) Pr(>|z|)
       0.69941              2.3896  0.01687 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

	Model 1	Model 2	Model 3	Model 4	Model 5
(Intercept)	12.21521***	12.16200***	12.06040***	12.11585***	12.22161***
	(0.06433)	(0.06661)	(0.08494)	(0.06955)	(0.04377)
as.factor(num_beds)2	0.08297**	0.06977*	0.06338*	0.07027*	0.07708*
	(0.03183)	(0.03213)	(0.03228)	(0.03171)	(0.03154)
as.factor(num_beds)3	0.20487***	0.19005***	0.18141***	0.19585***	0.20529***
	(0.04011)	(0.04060)	(0.04081)	(0.03879)	(0.03852)
as.factor(num_beds)4	0.41779***	0.41201***	0.40457***	0.42722***	0.43580***
	(0.04733)	(0.04742)	(0.04754)	(0.04317)	(0.04298)
as.factor(num_beds)6	0.47769***	0.47122***	0.48108***	0.52910***	0.52126***
	(0.07916)	(0.07923)	(0.07934)	(0.06717)	(0.06710)
as.factor(num_beds)8	0.83073***	0.82682***	0.85696***	0.92787***	0.90046***
	(0.09758)	(0.09756)	(0.09873)	(0.07654)	(0.07530)
as.factor(num_baths)2	0.07490***	0.07468***	0.07439***	0.07970***	0.08024***
	(0.01662)	(0.01664)	(0.01663)	(0.01596)	(0.01597)
as.factor(num_baths)3	0.08553*	0.08607**	0.08890**	0.09920**	0.09679**
	(0.03330)	(0.03336)	(0.03337)	(0.03212)	(0.03212)
floor_space	0.00007	0.00006	0.00006		
	(0.00005)	(0.00005)	(0.00005)		
lot_size	0.00003***	0.00003***	0.00007***	0.00007***	0.00003***
	(0.00000)	(0.00000)	(0.00002)	(0.00002)	(0.00000)
has_garage	0.27672***	0.27667***	0.26832***	0.29584***	0.30557***
	(0.03296)	(0.03296)	(0.03322)	(0.02275)	(0.02222)
has_encl_patio	0.04608**	0.04636**	0.04609**	0.04625**	0.04653**
	(0.01436)	(0.01439)	(0.01438)	(0.01438)	(0.01439)
has_security_gate	0.22997***	0.23140***	0.23056***	0.22970***	0.23051***
	(0.02056)	(0.02082)	(0.02081)	(0.02080)	(0.02081)
has_pool	0.09967***	0.09999***	0.10085***	0.10102***	0.10016***
	(0.02560)	(0.02558)	(0.02557)	(0.02557)	(0.02559)
transit_score	0.06665***	0.06650***	0.06676***	0.06687***	0.06661***
	(0.00346)	(0.00346)	(0.00346)	(0.00346)	(0.00346)
school_score	0.00405				
	(0.00327)				
type_of_buyerRental	0.11080***	0.11148***	0.11205***	0.11083***	0.11019***
	(0.01879)	(0.01884)	(0.01882)	(0.01879)	(0.01881)
age	−0.00974***	−0.00968***	−0.00967***	−0.00965***	−0.00967***
	(0.00048)	(0.00048)	(0.00048)	(0.00048)	(0.00048)
as.factor(school_score)2		0.10288**	0.10070**	0.10145**	0.10371**
		(0.03325)	(0.03325)	(0.03325)	(0.03325)
as.factor(school_score)3		0.10013**	0.09819**	0.09895**	0.10095**
		(0.03231)	(0.03230)	(0.03230)	(0.03231)
as.factor(school_score)4		0.09229**	0.08833**	0.08920**	0.09327**
		(0.03177)	(0.03181)	(0.03181)	(0.03176)
as.factor(school_score)5		0.08607**	0.08086*	0.08133*	0.08665**
		(0.03187)	(0.03196)	(0.03196)	(0.03187)
as.factor(school_score)6		0.10210**	0.09705**	0.09803**	0.10320**
		(0.03364)	(0.03372)	(0.03371)	(0.03363)
as.factor(school_score)7		0.09600**	0.09074**	0.09105**	0.09641**
		(0.03266)	(0.03275)	(0.03275)	(0.03266)
as.factor(school_score)8		0.09904**	0.09437**	0.09534**	0.10013**
		(0.03277)	(0.03283)	(0.03282)	(0.03276)
as.factor(school_score)9		0.08716*	0.08297*	0.08407*	0.08837*
		(0.03503)	(0.03507)	(0.03506)	(0.03502)
as.factor(school_score)10		0.09569*	0.10108*	0.10230*	0.09688*
		(0.04156)	(0.04162)	(0.04161)	(0.04155)
sq_lot_size			−0.00000	−0.00000	
			(0.00000)	(0.00000)	

```
iterations =  3
```

This coefficient is effectively 1, which is more evidence of a purely linear relationship between `log_saleprice` and age: the percentage depreciation rate is constant. Next, I will consider the possibility of nonlinearity in depreciation from hours of use.

6.1.3 Transformation of All Three Continuous Variables

```
         MLE of lambda Score Statistic (z) Pr(>|z|)
lot_size      0.082757             -2.0269  0.04267 *
age           0.697634              2.3932  0.01670 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

iterations =  8
```

The performance is similar to the other models with forms of nonlinearity for the value of lot size. Now consider the full set of such models in a table for a final comparison.

7 Final Comparison of Candidate Models

I created a variable `lot size_bt` by raising lot size to the optimal exponent $\hat{\lambda} = 0.1143693$. Then, I included this variable in the place of the lot size variables a the linear regression model. Table 5 collects the results of the set of models from the three forms of nonlinearity. Model 1 is the linear regression model with a quadratic form for lot size. Model 2 has the same specification as the other one, except that the lot size variable is transformed using the optimal exponent for the Box-Tidwell transformation. The last model has the highest R-squared among the ones we have estimated. Again, the differences are marginal, so the practical recommendation is the model with the quadratic relationship for lot size, which has a simpler interpretation. In either case, we conclude that John Deere tractors are worth approximately thirty percent more valuable than an equivalent tractor of another brand. Compare this with the lower premium of 17%, which was not even statistically significant, when we estimated a simpler, linear specification in which we ignored the nonlinearity in the model.

Table **??** compare previours model with variable age transformed using $\hat{\lambda} = 0.1143693$.

	Model 1	Model 2
(Intercept)	12.11585***	10.86297***
	(0.06955)	(0.21050)
as.factor(num_beds)2	0.07027*	0.07009*
	(0.03171)	(0.03159)
as.factor(num_beds)3	0.19585***	0.19584***
	(0.03879)	(0.03862)
as.factor(num_beds)4	0.42722***	0.42747***
	(0.04317)	(0.04297)
as.factor(num_beds)6	0.52910***	0.52439***
	(0.06717)	(0.06692)
as.factor(num_beds)8	0.92787***	0.91419***
	(0.07654)	(0.07482)
as.factor(num_baths)2	0.07970***	0.08001***
	(0.01596)	(0.01595)
as.factor(num_baths)3	0.09920**	0.09896**
	(0.03212)	(0.03206)
lot_size	0.00007***	
	(0.00002)	
sq_lot_size	−0.00000	
	(0.00000)	
has_garage	0.29584***	0.29816***
	(0.02275)	(0.02246)
has_encl_patio	0.04625**	0.04614**
	(0.01438)	(0.01438)
has_security_gate	0.22970***	0.22968***
	(0.02080)	(0.02077)
has_pool	0.10102***	0.10156***
	(0.02557)	(0.02557)
transit_score	0.06687***	0.06684***
	(0.00346)	(0.00346)
as.factor(school_score)2	0.10145**	0.10112**
	(0.03325)	(0.03322)
as.factor(school_score)3	0.09895**	0.09932**
	(0.03230)	(0.03228)
as.factor(school_score)4	0.08920**	0.08966**
	(0.03181)	(0.03176)
as.factor(school_score)5	0.08133*	0.08141*
	(0.03196)	(0.03188)
as.factor(school_score)6	0.09803**	0.09800**
	(0.03371)	(0.03365)
as.factor(school_score)7	0.09105**	0.09175**
	(0.03275)	(0.03269)
as.factor(school_score)8	0.09534**	0.09602**
	(0.03282)	(0.03278)
as.factor(school_score)9	0.08407*	0.08538*
	(0.03506)	(0.03501)
as.factor(school_score)10	0.10230*	0.10047*
	(0.04161)	(0.04125)
type_of_buyerRental	0.11083***	0.11068***
	(0.01879)	(0.01879)
age	−0.00965***	−0.00964***
	(0.00048)	(0.00048)
lot_bt		0.47925***
		(0.06601)
R^2	0.65314	0.65315
Adj. R^2	0.64842	0.64862

	Model 1	Model 2
(Intercept)	10.86297^{***}	10.90045^{***}
	(0.21050)	(0.20893)
lot_bt	0.47925^{***}	0.47974^{***}
	(0.06601)	(0.06549)
as.factor(num_beds)2	0.07009^{*}	0.07393^{*}
	(0.03159)	(0.03135)
as.factor(num_beds)3	0.19584^{***}	0.20211^{***}
	(0.03862)	(0.03832)
as.factor(num_beds)4	0.42747^{***}	0.43336^{***}
	(0.04297)	(0.04263)
as.factor(num_beds)6	0.52439^{***}	0.53277^{***}
	(0.06692)	(0.06638)
as.factor(num_beds)8	0.91419^{***}	0.92310^{***}
	(0.07482)	(0.07425)
as.factor(num_baths)2	0.08001^{***}	0.07966^{***}
	(0.01595)	(0.01582)
as.factor(num_baths)3	0.09896^{**}	0.09720^{**}
	(0.03206)	(0.03181)
has_garage	0.29816^{***}	0.29609^{***}
	(0.02246)	(0.02229)
has_encl_patio	0.04614^{**}	0.04641^{**}
	(0.01438)	(0.01427)
has_security_gate	0.22968^{***}	0.22896^{***}
	(0.02077)	(0.02061)
has_pool	0.10156^{***}	0.09923^{***}
	(0.02557)	(0.02537)
transit_score	0.06684^{***}	0.06675^{***}
	(0.00346)	(0.00343)
as.factor(school_score)2	0.10112^{**}	0.10255^{**}
	(0.03322)	(0.03296)
as.factor(school_score)3	0.09932^{**}	0.09679^{**}
	(0.03228)	(0.03203)
as.factor(school_score)4	0.08966^{**}	0.08604^{**}
	(0.03176)	(0.03152)
as.factor(school_score)5	0.08141^{*}	0.07768^{*}
	(0.03188)	(0.03163)
as.factor(school_score)6	0.09800^{**}	0.09447^{**}
	(0.03365)	(0.03338)
as.factor(school_score)7	0.09175^{**}	0.09305^{**}
	(0.03269)	(0.03243)
as.factor(school_score)8	0.09602^{**}	0.09403^{**}
	(0.03278)	(0.03253)
as.factor(school_score)9	0.08538^{*}	0.08368^{*}
	(0.03501)	(0.03473)
as.factor(school_score)10	0.10047^{*}	0.10047^{*}
	(0.04125)	(0.04093)
type_of_buyerRental	0.11068^{***}	0.11198^{***}
	(0.01879)	(0.01864)
age	-0.00964^{***}	
	(0.00048)	
age_bt		-0.03120^{***}
		(0.00150)
R^2	0.65315	0.65856
Adj. R^2	0.64862	0.65409
Num. obs.	1862	1862

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$