This is a working paper based on unverifiable data. Someday, this paper will be reworked as an exercise in spatial hedonics.

Valuing the Bosque: A Hedonic Study of Open Space Amenities and Home Prices

David S. Dixon[1]

MARCH 2005

ABSTRACT

Albuquerque is bisected by woods – *bosque* in Spanish – that edge the Rio Grande along most of its course in New Mexico. Unique, natural open space, the bosque is also critical habitat and a fragile ecosystem. What is the value of this shared open space amenity offsetting the maintenance and protection costs? The present study sets out to determine if, after accounting for all the usual factors in home value – home features and neighborhood – there is a component related to distance from the bosque. Preliminary results appear to indicate that a fraction of home value is associated with the square of the distance, as much as half of the selling price for homes nearest the bosque. (JEL R52)

Keywords: Hedonic analysis, open space, urban forest, property values.

Introduction

Albuquerque, New Mexico, which was founded on the east bank of the Rio Grande, is now bisected by the river. The natural wooded areas along the banks of the

[1] University of New Mexico, Department of Economics. The author acknowledges the guidance of Kristine Grimsrud, Alok Bohara and Phil Ganderton, all of University of New Mexico Economics. The author also thanks Danelle Callan for the data and for initial set up of the model.

river, called *bosque* in Spanish, form 6300 acres of urban forest and open space. The bosque is a multiuse common amenity, providing open space, urban woods, recreation, and habitat for threatened and endangered species. Authority over the bosque is shared by the City of Albuquerque, the Mid Rio Grande Conservancy District, the State of New Mexico, and the U.S. Army Corps of Engineers. These agencies may share or divide responsibility and costs for maintenance, restoration and development of the bosque. Some of these activities, like reparation of fire damage and restoration of native flora, can be quite costly. Sometimes, securing budgetary commitments for these projects is difficult because of widely varying perceptions of the value of the bosque and the marginal value added by these projects. This was the motivation for the present study. It is hoped that this study will serve as a necessary first step in valuing the bosque for the purposes of policy-making and resource allocation.

Literature review

Hedonic analysis of land quality and land price dates back to 1922 (Colwell and Dilmore 1999) and the earliest study of environmental amenities based on housing prices was done in 1967 (Ridker and Henning 1967). An early effort to value open space amenities looked at farmland at the edge of urban housing areas (Beasley, Workman and Williams 1986). They found that household willingness to pay is related to risk of the open space being developed, and strongly affected by the kind development at risk: high or low density. Garrod and Willis (1992), looking specifically at urban woodland, found that value is either enhanced by proximity to deciduous woodland, but decreased by proximity to coniferous woodland. Tyrävinen and Miettinen (2000) studing the effects of urban forests on housing prices in Finland. They found a negative effect from distance to urban woodlands, but no effect related to either the size of the wooded area or distance from large forests.

Hedonic models of home prices

Taylor (2003) suggests two candidate property value models. The linear model is

$$P = \alpha_0 + \sum_{i=1}^{h} \beta_i H_i + \sum_{i=1}^{n} \gamma_i N_i + \sum_{i=1}^{l} \delta_i L_i + \varepsilon \qquad (1)$$

where

H_i = structural and property features 1 through h of the house
N_i = neighborhood characteristics 1 through n
L_i = location characteristics 1 through l

Parameters in (1) are interpreted as dollar amount change in the selling price per unit change in the regressor. The log-linear model is

$$\ln P = \alpha_0 + \sum_{i=1}^{h} \beta_i H_i + \sum_{i=1}^{n} \gamma_i N_i + \sum_{i=1}^{l} \delta_i L_i + \varepsilon \qquad (2)$$

for which the parameters are interpreted as the percent change in the selling price per change in the regressor. According to Taylor, the log-linear model is the most frequently encountered in hedonic property value studies.

The effect of distance may be linear or quadratic. The theoretical basis for a quadratic in distance is that the value of a shared amenity is proportional to the number of people sharing it, and that, with uniform density, the population within a certain distance of the amenity increases as the square of the distance.

This study will test three hypotheses. The first is that equation (2) is a suitable log-linear hedonic model of home prices in Albuquerque. The second hypothesis is that home prices will monotonically increase with the number of bathrooms and bedrooms and the size of the house. The third hypothesis is that home prices will be a decreasing function of the square of the distance from the bosque.

Data

Home sales data were taken from Multiple Listing Service (MLS) data for Albuquerque home sales in 2003 and 2004. MLS data included address, AGE_OF_HOUSE, SQ_FEET, SELLING_PRICE, BEDROOMS, BATHROOMS, and CLOSING_DATE. From the MLS area identifier, four dummies were created: BOSQUE (homes near or in the bosque), MOUNTAIN (homes near the eastern edge of the city), MIDDLE (homes between BOSQUE and MOUNTAIN), and WEST (homes west of the bosque). From the address, dummies were created for each of the high school districts in the city: AHS, CHS, DNHS, EHS, HHS, LCHS, MHS, RGHS, SHS, VHS, and WMHS. The address was used to look up location coordinates from the Albuquerque GIS (AGIS) system. DIST, the distance of each address from the bosque, was computed from AGIS data on the bosque.

The house features H_i in equations (1) and (2) are AGE_OF_HOUSE, BEDROOMS and BATHROOMS.

The high school dummies serve as proxies for the neighborhood characteristics N_i in equations (1) and (2).

The location characteristics L_i in equations (1) and (2) are either the area dummies, or DIST and its square, DIST_SQ. The theoretical basis for a quadratic in distance is that the value of a shared amenity is proportional to the number of people sharing it, and that, with uniform density, the population within a certain distance of the amenity increases as the square of the distance.

There were 283 points in the final data set. Their distribution with regard to the bosque is shown in Figure 1.

Empirical Results

The first OLS regression is the linear model in equation (1) using area dummies. To avoid the dummy variable trap, the high school dummy RGHS and the area dummy MOUNTAIN are dropped. These were chosen because they represent houses and neighborhoods closest to the bosque, which will be considered the baseline. Results are shown in Table 1. The signs of the fitted parameters are all as expected: SQ_FEET, BATHROOMS and BEDROOMS all add to the value, AGE_OF_HOUSE and areas other than BOSQUE detract from the value. According to the t-value, however, AGE_OF_HOUSE is not a significant regressor. Parameter values indicate that each square foot adds $32 to a home price, while a bedroom adds $5,326 and a bathroom $53,951. Neighborhood makes up between $111,000 and $169,000 of the selling price. Distance from the bosque detracts $137,203 in the MOUNTAIN area, $152,571 for the west side of the river (WEST area), and $160,659 between the river and the mountains (MIDDLE area).

Next, a Box-Cox test is used to decide between the linear model in equation (1) or the log-linear model in equation (2). Output from a STATA boxcox lhsonly command is shown in Table 2. The P-values indicate that the areas model is linear.

Table 3 shows the OLS regression for the linear model in equation (1) using quadratic distance. Again, AGE_OF_HOUSE is insignificant. Here, however, it is noted that R-squared is lower than the linear areas model and that the high school parameters have become increasingly insignificant. Table 4 shows a STATA boxcox lhsonly test on the quadratic distance model. The P-value for theta = 0 indicates use of the log-linear model in equation (2).

Loss of significance of the high school dummies can be understand in light of the linear area model regression (Table 1). Note that for the areas outside the bosque, the fitted parameters indicate a loss in value of $152,571 west of the river (WEST area) and $160,659 east of the river (MIDDLE area). Further still to the east, however, near the mountains (MOUNTAIN area) the loss is only $137,203. That is, the value of bosque is not monotonic with distance. One likely explanation for this is that the Sandia Mountains at the eastern edge of the city are an alternative open space amenity. Future study could isolate this.

Table 5 shows the OLS regression for the log-linear model in equation (2) using quadratic distance and AGE_OF_HOUSE excluded.

The OLS regression of the log-linear model using DIST and DIST_SQ is shown in Table 5. Note that R-squared has decreased, and the t-values for the high school parameters have all worsened as well.

The Akaike information criterion (AIC) values for the three models are shown in Table 6, where AREAS refers to the linear model with area dummies, DIST is the linear model with quadratic distance, and LN_DIST is the log-linear model with quadratic distances. DIST_SQ and JUST_DIST are log-linear models with just the square distance term and just the linear distance term, respectively, and NO_HS is the log-linear model with quadratic distances and no high school dummies. Note that, despite the poor t-values for the high school dummies in the log-linear quadratic distance model, AIC prefers it to the same model with no high school dummies. However, AIC also slightly prefers log-linear models with just the linear or the square distance term. Table 7 shows the log-linear model with just linear distance. This model indicates that the value of the bosque decreases at 0.009% per foot, or 48.7% per mile. This is consistent with the linear areas model. For example, the MIDDLE area averages 2.5 miles from the bosque, which would mean an increase of 120% for the average selling price of $142,343. That would be a loss of $173,425, compared with $160,659 indicated by the linear model.

Conclusion

This study shows that hedonic equations (1) and (2) are suitable to modeling the components of the selling prices of homes in Albuquerque. Home prices are positively affected by the size of the house, the number of bedrooms and the number of bathrooms. Home age is not significant. The bosque significantly influences the selling prices of homes in Albuquerque. This effect is linear when considering broad areas of the city, but log-linear in statute distance from the bosque. These results suggest that there is another spatial factor – perhaps the proximity of an open space amenity in the form of the Sandia Mountains to the east of the city. Further study will isolate the two open space amenities.

References

Colwell, Peter F. and Gene Dilmore. 1999. "Who was First? An Examination of an Early Hedonic Study". *Land Econ.* 75:4, pp. 620-626.

Garrod, G., and K.Willis. 1992. “The Environmental Economic Impact of Woodland: A Two-Stage Hedonic Price model of the Amenity Value of Forestry in Britain,” *Appl. Econ.* 24:7, pp. 715-28.

Ridker, Ronald G., and John A. Henning. 1967.” The Determinants of Residential Property Values with Special Reference to Air Pollution,” *Rev. Econ. Statist.* 49:2, pp. 246-257.

Taylor, Laura O. 2003. “The Hedonic Method,” in *A primer on nonmarket valuation*. Patricia A. Champ, Kevin J. Boyle, and Thomas C. Brown, eds. Dordrecht: Kluwer, pp. 352-5.

Tyrävinen, L., and A. Meittinen. 2000. “Property Prices and Urban Forest Amenities.” J. *Environ. Econ. Manage.* 39:2, pp. 205-23.

Tables and figures

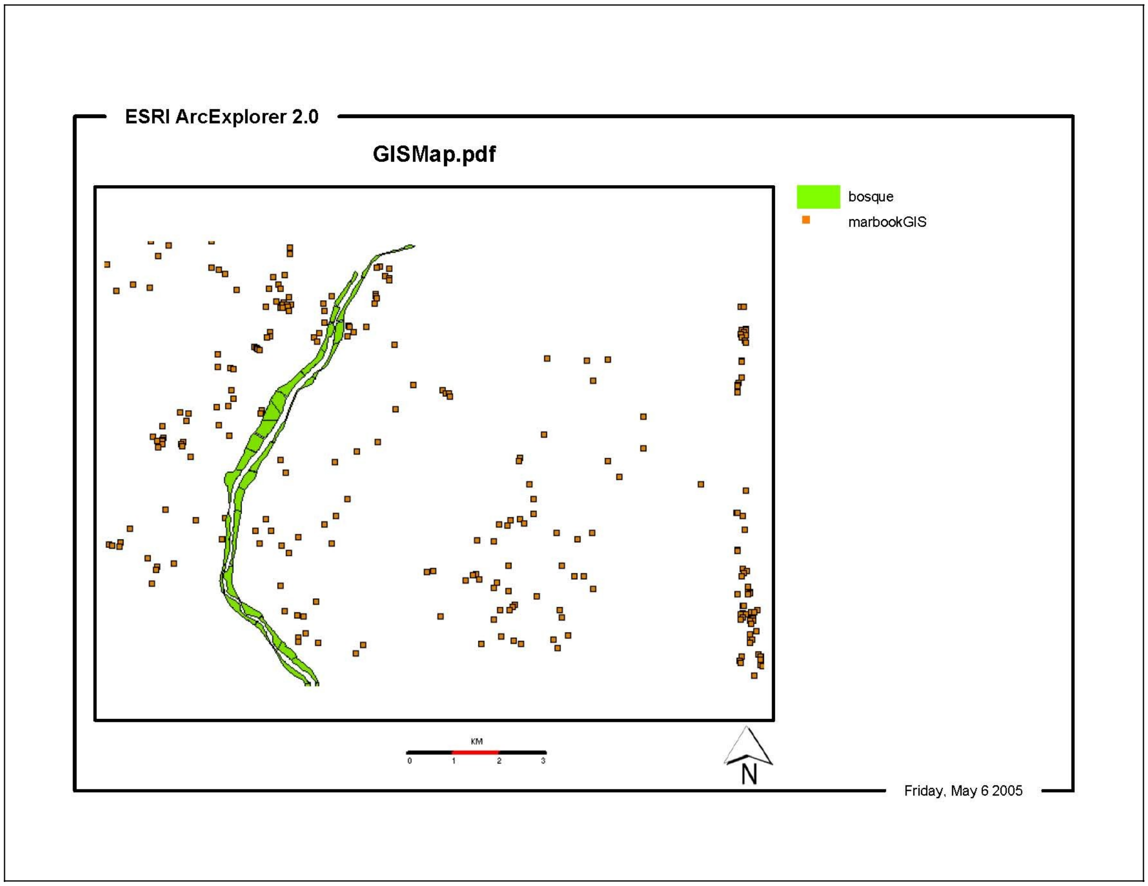

Figure 1

Table 1

Source	SS	df	MS		
Model	2.1028e+12	17	1.2369e+11	Number of obs =	283
Residual	8.3193e+11	265	3.1394e+09	F(17, 265) =	39.40
				Prob > F =	0.0000
Total	2.9347e+12	282	1.0407e+10	R-squared =	0.7165
				Adj R-squared =	0.6983
				Root MSE =	56030

selling_pr~e	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
age_of_house	2.928362	55.39491	0.05	0.958	-106.1418	111.9985
sq_feet	32.19336	5.003303	6.43	0.000	22.34208	42.04465
bedrooms	5325.769	6508.066	0.82	0.414	-7488.328	18139.87
bathrooms	53951.01	9131.61	5.91	0.000	35971.27	71930.76
west	-152571.3	16205.18	-9.41	0.000	-184478.6	-120664
middle	-160659.1	19219.85	-8.36	0.000	-198502.1	-122816
mountain	-137203.2	23632.91	-5.81	0.000	-183735.3	-90670.99
ahs	156217.9	60325.25	2.59	0.010	37440.14	274995.7
chs	121508.3	58156.94	2.09	0.038	6999.839	236016.8
dnhs	139172.6	60184.32	2.31	0.022	20672.34	257672.9
ehs	138096.2	61824.79	2.23	0.026	16365.84	259826.5
hhs	132009.3	60789.41	2.17	0.031	12317.58	251701
lchs	140643.7	60847.1	2.31	0.022	20838.4	260448.9
mhs	130759.7	62412.3	2.10	0.037	7872.601	253646.8
shs	110609.6	61849.26	1.79	0.075	-11168.87	232388.1
vhs	169023.5	57898.43	2.92	0.004	55024.05	283023
wmhs	137515	59879.34	2.30	0.022	19615.24	255414.8
_cons	-6352.1	59931.4	-0.11	0.916	-124354.4	111650.2

Table 2

```
Iteration 0:   log likelihood = -3664.8601
Iteration 1:   log likelihood = -3528.5547
Iteration 2:   log likelihood = -3527.4383
Iteration 3:   log likelihood = -3527.4331
Iteration 4:   log likelihood = -3527.4331

Fitting full model

Iteration 0:   log likelihood = -3486.4819
Iteration 1:   log likelihood = -3368.6114
Iteration 2:   log likelihood = -3368.6061
Iteration 3:   log likelihood = -3368.6061

                                                  Number of obs   =        283
                                                  LR chi2(17)     =     317.65
Log likelihood = -3368.6061                       Prob > chi2     =      0.000

------------------------------------------------------------------------------
selling_pr~e |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      /theta |    .148816   .0472296     3.15   0.002     .0562477    .2413843
------------------------------------------------------------------------------

Estimates of scale-variant parameters
----------------------------
             |      Coef.
-------------+--------------
Notrans      |
age_of_house |   .0001268
     sq_feet |   .0006639
    bedrooms |   .2496493
   bathrooms |   1.067963
        west |  -3.386891
      middle |  -3.494491
    mountain |  -2.536713
         ahs |   3.145867
         chs |   2.446933
        dnhs |   1.802921
         ehs |    2.51051
         hhs |   1.845107
        lchs |   2.652868
         mhs |   2.141384
         shs |   1.682832
         vhs |   3.255677
        wmhs |   2.493454
       _cons |   29.78742
-------------+--------------
      /sigma |   1.265039
----------------------------

---------------------------------------------------------
   Test         Restricted     LR statistic      P-Value
    H0:       log likelihood       chi2       Prob > chi2
---------------------------------------------------------
theta = -1      -3656.1101       575.01           0.000
theta =  0      -3373.7523        10.29           0.001
theta =  1      -3486.4819       235.75           0.000
---------------------------------------------------------
```

Table 3

Source	SS	df	MS
Model	1.7380e+12	16	1.0863e+11
Residual	1.1967e+12	266	4.4989e+09
Total	2.9347e+12	282	1.0407e+10

Number of obs =	283
F(16, 266) =	24.15
Prob > F =	0.0000
R-squared =	0.5922
Adj R-squared =	0.5677
Root MSE =	67074

selling_pr~e	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
age_of_house	-22.28858	66.26632	-0.34	0.737	-152.7618	108.1847
sq_feet	51.81821	6.143114	8.44	0.000	39.7229	63.91353
bedrooms	8395.092	7832.181	1.07	0.285	-7025.864	23816.05
bathrooms	72038.35	10917.41	6.60	0.000	50542.81	93533.89
dist	5.72084	20.4217	0.28	0.780	-34.48791	45.92959
dist_sq	-.0160517	.0098234	-1.63	0.103	-.0353932	.0032898
ahs	6034.984	68829.05	0.09	0.930	-129484.1	141554
chs	-11871.18	67599.17	-0.18	0.861	-144968.7	121226.3
dnhs	10024.64	69083.99	0.15	0.885	-125996.4	146045.7
ehs	5197.646	69053.99	0.08	0.940	-130764.3	141159.6
hhs	-352.771	68599.9	-0.01	0.996	-135420.6	134715.1
lchs	11739.82	68700.33	0.17	0.864	-123525.8	147005.4
mhs	-3316.542	69322.18	-0.05	0.962	-139806.5	133173.4
shs	-29045.22	70933.67	-0.41	0.683	-168708.1	110617.7
vhs	77733.09	68263.12	1.14	0.256	-56671.69	212137.9
wmhs	16266.77	69842.79	0.23	0.816	-121248.3	153781.8
_cons	-88467.61	71202.58	-1.24	0.215	-228659.9	51724.73

Table 4

```
Iteration 0:   log likelihood = -3664.8601
Iteration 1:   log likelihood = -3528.5547
Iteration 2:   log likelihood = -3527.4383
Iteration 3:   log likelihood = -3527.4331
Iteration 4:   log likelihood = -3527.4331

Fitting full model

Iteration 0:   log likelihood = -3537.9887
Iteration 1:   log likelihood = -3414.3205
Iteration 2:   log likelihood = -3414.3191
Iteration 3:   log likelihood = -3414.3191

                                                  Number of obs   =        283
                                                  LR chi2(15)     =     226.23
Log likelihood = -3414.3191                       Prob > chi2     =      0.000

------------------------------------------------------------------------------
selling_pr~e |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      /theta |   .0402465   .0495932     0.81   0.417    -.0569544    .1374474
------------------------------------------------------------------------------

Estimates of scale-variant parameters
----------------------------
             |      Coef.
-------------+--------------
Notrans      |
     sq_feet |   .0002719
    bedrooms |   .0936184
   bathrooms |   .3870022
        dist |   -.000087
     dist_sq |  -3.36e-08
         ahs |  -.0089555
         chs |  -.1287301
        dnhs |  -.2749581
         ehs |   .0308907
         hhs |  -.1705875
        lchs |   .0902875
         mhs |  -.0692594
         shs |  -.3498128
         vhs |   .3300921
        wmhs |  -.0602171
       _cons |   14.00878
-------------+--------------
      /sigma |    .402259
----------------------------

---------------------------------------------------------
   Test         Restricted     LR statistic      P-Value
    H0:       log likelihood       chi2       Prob > chi2
---------------------------------------------------------
theta = -1      -3657.6826       486.73           0.000
theta =  0      -3414.6537         0.67           0.413
theta =  1      -3537.9887       247.34           0.000
```

Table 5

Source	SS	df	MS		
Model	22.2370128	15	1.48246752	Number of obs =	283
Residual	17.4139677	267	.065220853	F(15, 267) =	22.73
Total	39.6509805	282	.140606314	Prob > F =	0.0000
				R-squared =	0.5608
				Adj R-squared =	0.5361
				Root MSE =	.25538

lnprice	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
sq_feet	.0001641	.0000234	7.02	0.000	.0001181	.0002102
bedrooms	.0580553	.0298148	1.95	0.053	-.0006468	.1167574
bathrooms	.2333537	.041568	5.61	0.000	.151511	.3151965
dist	-.000057	.0000777	-0.73	0.464	-.0002101	.000096
dist_sq	-1.84e-08	3.74e-08	-0.49	0.624	-9.20e-08	5.53e-08
ahs	-.0062741	.2619924	-0.02	0.981	-.522108	.5095598
chs	-.0796907	.2572692	-0.31	0.757	-.5862252	.4268437
dnhs	-.1765207	.2628767	-0.67	0.502	-.6940956	.3410542
ehs	.0188568	.2628223	0.07	0.943	-.498611	.5363246
hhs	-.1101595	.2610684	-0.42	0.673	-.624174	.403855
lchs	.0556155	.2614341	0.21	0.832	-.4591191	.57035
mhs	-.0431881	.2638337	-0.16	0.870	-.5626472	.4762711
shs	-.2172167	.2699665	-0.80	0.422	-.7487507	.3143173
vhs	.1963466	.2597975	0.76	0.450	-.3151656	.7078589
wmhs	-.0405107	.2657296	-0.15	0.879	-.5637028	.4826813
_cons	11.14208	.2707146	41.16	0.000	10.60907	11.67509

Table 6

Model	nobs	ll(null)	ll(model)	df	AIC	BIC
areas	283	-3664.86	-3486.482	18	7008.964	7074.582
dist	283	-3664.86	-3537.929	17	7109.857	7171.83
ln_dist	283	-123.4652	-7.032946	16	46.06589	104.393
dist_sq	283	-123.4652	-7.317814	15	44.63563	99.31733
just_dist	283	-123.4652	-7.160711	15	44.32142	99.00313
no_hs	283	-123.4652	-28.78687	6	69.57374	91.44642

Table 7

Source	SS	df	MS		
Model	22.221282	14	1.58723443	Number of obs =	283
Residual	17.4296985	268	.065036189	F(14, 268) =	24.41
Total	39.6509805	282	.140606314	Prob > F =	0.0000
				R-squared =	0.5604
				Adj R-squared =	0.5375
				Root MSE =	.25502

lnprice	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
sq_feet	.0001606	.0000222	7.23	0.000	.0001168	.0002043
bedrooms	.0596199	.0296021	2.01	0.045	.0013375	.1179022
bathrooms	.2371808	.0407732	5.82	0.000	.1569044	.3174573
dist	-.0000923	.0000299	-3.09	0.002	-.0001511	-.0000334
ahs	.0008886	.2612156	0.00	0.997	-.513407	.5151842
chs	-.0770748	.2568497	-0.30	0.764	-.5827746	.428625
dnhs	-.1709814	.2622625	-0.65	0.515	-.6873384	.3453755
ehs	.0253133	.2621214	0.10	0.923	-.4907658	.5413924
hhs	-.0994897	.2597942	-0.38	0.702	-.6109869	.4120075
lchs	.0657448	.26025	0.25	0.801	-.4466497	.5781394
mhs	-.0373824	.2631953	-0.14	0.887	-.5555758	.4808111
shs	-.2136444	.2694862	-0.79	0.429	-.7442237	.3169349
vhs	.2019228	.2591815	0.78	0.437	-.3083681	.7122137
wmhs	-.0404532	.2653532	-0.15	0.879	-.5628952	.4819887
_cons	11.13659	.2701002	41.23	0.000	10.6048	11.66837