Appendices for

Hedonic Markets and Explicit Demands: Bid-Function Envelopes for Public Services and Neighborhood Amenities

John Yinger

The Maxwell School

Syracuse University

May 2009

Appendix A: Data Collection Procedures and Descriptive Statistics

Appendix B: Endogenous Switching Regressions

Appendix A: Data Collection Procedures and Descriptive Statistics

Preliminaries. This appendix describes the procedures used to obtain the variables listed in Table 4 of the text. It also presents descriptive statistics for the variables in Tables 3 and 4.

In the following discussion, USGS stands for the U.S. Geological Survey. Lat stands for latitude. Long stands for longitude. Distance to a CBG means distance to the center of the CBG (using the latitude and longitude in the Brasington data set). All distances except commuting distances are translated into miles using the standard formula:

$$D_{Miles} = \sqrt{X^2 + Y^2}$$
$$X = 69.1\left(\mathrm{Lat}_2 - \mathrm{Lat}_1\right)$$
$$Y = 69.1\left(\mathrm{Long}_2 - \mathrm{Long}_1\right)\left(\cos\left\{\frac{\mathrm{Lat}_1}{57.3}\right\}\right)$$

Distances to worksites are calculated using the more accurate spherical law of cosines formula.

Lakefront; Distance to Lake. To obtain the lakefront variables, I divided the shoreline of Lake Erie into seven approximately linear segments. Then I used the USGS website to identify the latitude and longitude of a series of points along each segment. These points were then treated as data for an estimation of the line describing each segment. Each line was Lat = a + b(Long). These variables were expressed in decimals, not minutes and seconds.

Each CBG was linked to the shoreline segment directly to its north. The next step was to find the shortest distance from a CBG to its assigned line segment. Using the standard formula for the distance between two points and the expression for the shoreline with the estimated coefficients, I solved for the values of latitude and longitude on the shoreline that minimized the distance from the CBG to the shoreline. Plugging these values and the latitude and longitude of the CBG into the standard distance formula yields the desired distance measure. Let Y_i and X_i be latitude and longitude, respectively, for CBG i and let a and b be the estimated coefficients from the above regression. Then the formula for the minimum distance to the lake, $D_{\min}$, is:

$$D_{\min} = \sqrt{\left(\frac{X_i + bY_i - ab}{1+b^2} - X_i\right)^2 + \left(a + b\left(\frac{X_i + bY_i - ab}{1+b^2}\right) - Y_i\right)^2}$$

For some observations near one of the ends of a segment, the point on the "shoreline" identified by the above procedure was not actually on the shoreline but was instead a point on an extension of the line segment into the territory of another segment. This case is illustrated in Figure A-1. For a CBG located at point A, the above formula calculates the distance to point B, which is not correct because B is not on the shoreline. In these cases, the distance to Lake Erie was defined as the distance to the intersection between the two relevant line segments, which is point C in the figure. These cases are easy to identify because point B falls outside the region specified for line segment 1, which is identified here by the vertical line through point C.

In a few other cases, the linear approximation to the shoreline cuts off a piece of land that juts into the lake. These cases are easily identified as CBGs located to the north of their line segment. These CBGs were all assumed to be right on the lake, that is, to have a distance to the lake equal to zero.

Snowbelt 1; Snowbelt 2. Cleveland has an intense snowbelt in its eastern suburbs. The amount and location of the snowfall obviously varies from year to year, but in most years it is concentrated in an area about 10 miles from Lake Erie starting in the town of Pepper Pike and moving to the East. Because the shoreline of Lake Erie goes from Southwest to Northeast in this region, this snowbelt can also be described as being located about 10 miles Southeast of Lake Erie (and east of Pepper Pike). Cleveland snowfall maps are available from the National Weather Service at http://www.erh.noaa.gov/cle/climate/info/snowinfo.html.

I use the "distance to Lake Erie" variable to identify observations in the snowbelt. To be specific, the first snowbelt variable is distance to Lake Erie for all observations east of Pepper Pike (and only for observations within 20 miles of the lake). The second snowbelt variable is the square of the distance to Lake Erie for the same observations. If these variables are indeed capturing a snow belt effect, we would expect this quadratic specification to indicate a maximum effect about 10 miles from the lake. This effect could be positive or negative.

Ghetto; Near Ghetto. To define the ghetto variables, I plotted all the CBGs in which the population was at least 80 percent black. I then identified two sub-sets of these CBGs that were closely clustered together. See Figure B-3, which identifies CBGs at least 80 percent black and draws ovals around the clusters of these CBGs defined to be in one of the ghetto subsets. These two sub-sets were defined as the black ghetto. I then found the population weighted centroid of each sub-set and the distance from each CBG outside the ghetto to the nearest of these two centroids. CBGs were identified as near the ghetto if this distance was less than five miles. Note that this distance is to the center of the ghetto, not to its outer edge.

Near Airport; Airport Distance. The space occupied by the Cleveland Hopkins Airport is shaped approximately like a circle with a radius of two miles. I used the USGS site to find the (approximate) center of this circle and then measured the distance from each CBG to this center.

Near Public; Elementary Score. The data for elementary schools comes from two sources: the 2001 report card file posted by the Ohio Department of Education (at http://ilrc.ode.state.oh.us/Downloads.asp) and the latitudes and longitudes available at the U.S. Geological Survey Site (www.usgs.gov). The Ohio DOE file gives the name, district, and address of each elementary school that existed in 2001. (It lists other schools, too, but I did not collect any of that information.) In addition, this source gives the fourth-grade passing rates for each school for 1998-1999, 1999-2000, and 2000-2001 for each state test (citizenship, mathematics, reading, writing, and science). Following my treatment of the district level tests, I recorded the overall average passing rate averaged over the first two of the above years (except in a handful of schools for which only one of these years was available). I also recorded the average passing rate on the math, reading, and writing tests for these two years.

The Ohio DOE site identified 378 elementary schools in the school districts that appear in the Cleveland area subsample of the Brasington data. (None of these were charter schools, which are called community schools in Ohio. Ohio's charter school law passed in 1997, and there were few such schools in 2000. See the 2002-2003 *Annual Report on Community Schools*, http://education.ohio.gov/GD/Templates/Pages/ODE/ODEDetail.aspx?page=3&TopicRelationID=662&ContentID=42095&Content=61111.) Test scores are not available for 11.9 percent of these elementary schools, most of which were in small districts and none of which were in Cleveland, and missing scores were set to the district average. To obtain a location for each school, I first downloaded the list of schools in each county, complete with latitude and longitude, from the USGS site. This list is for 2009, so it includes some schools that did not exist in 2000 and does not include some schools that did exist. This source provided locations for about 70 percent of the 2001 schools. I searched for the addresses of the remaining schools on the 2000-2001 list on the USGS maps. Placing the pointer on each school location reveals its latitude and longitude.

Corrected School Assignments. When I first calculated the distance to the nearest elementary school in an observation's school district, I found that some of the distances were unreasonably large (almost 50 miles!). As a result, I plotted the latitude and longitude of each observation with a different color for each assigned school district. This approach revealed a few observations that were assigned to the wrong school district in the original Brasington data (as indicated by an observation with one color code surrounded by observations of a different color). On the basis of these plots I corrected the school-district assignments for about 120 observations. Ambiguities were resolved with http://tax.ohio.gov/online_services/thefinderschooldistrict.stm, which indicates the school district associated with any latitude and longitude.

Near Private. This variable comes from the Brasington data set.

Land Release. This variable comes from the Brasington data set.

Smog d; Smog Distance d. The Brasington data set indicates the total release (in pounds) of pollutants into the air from facilities within each CBG. This figure includes emissions through "confined air streams," such as stacks, vents , and pipes as well as "fugitive" emissions. These data come from the U.S. Environmental Protection Agency's Toxic Release Inventory. To obtain the Smog variables, I plotted the 34 CBGs with at least 25,000 pounds of air pollutants per year. This plot indicated that these CBGs all fell into one of 7 clusters.

I then looked at the total pollution from each of these clusters and selected 3 clusters with more than 800,000 pounds of air pollutants. (The other 4 clusters had air emissions below 250,000 pounds.) Then I found the effluent-weighted centroid of each of these 3 clusters. The Smog variables identify CBGs within 20 miles of one of these centroids. The Smog Distance variables indicate the distance from the CBG to the nearest of these centroids (given that the distance is less than 20). To account for wind patterns and perhaps other factors that influence the distribution of these emissions across space, the Smog and Smog Distance variables are interacted with direction indicators. Smog NE, for example, identifies CBGs located with 20 miles to the Northeast of one of the effluent concentrations.

Local Amenities; Freeway; Railroad; Shopping; Hospital; Small Airport; Big Park. To obtain these variables, I copied the latitude and longitude of each CBG into the search command on the website for Google Maps. All of these features are highlighted on the maps and are therefore easy to identify around each CBG location. Many features were verified on the satellite pictures also available through Google Maps.

- The local amenities include neighborhood parks, golf courses, rivers, and lakes.
- Freeways were defined as limited access highways.
- Railroads with a visible end were checked on the satellite pictures. Many of these segments had been removed and were therefore obviously not counted.
- Small streams were not counted as rivers; all rivers were verified on the satellite pictures.
- All lakes were verified on the satellite pictures; industrial ponds were not counted as lakes.
- Shopping includes malls, shopping centers, and business districts as identified on the maps; all of them were verified on the satellite pictures.
- Hospitals were defined to include medical centers, but not nursing homes, mental hospitals, rehabilitation centers, or the hospital for the Merchant Marines.
- Small Airports exclude Cleveland Hopkins Airport and small “air parks.”
- Big Parks were defined as county, metro, regional, state, and national parks, plus the Cleveland zoo.

For neighborhood amenities, freeways, and railroads, I determined whether the geographic feature was within one-quarter mile of the CBG. This distance was selected because one-quarter mile was thought to be a reasonable walking distance to neighborhood amenities and a reasonable indicator of the distance over which the noise and pollution from a freeway or railroad might be seen as a problem. Shopping, Hospital, Small Airport, and Big Park are switched on if the relevant geographic feature is within one mile of the CBG. One mile was selected as a reasonable indicator of convenient access or, in the case of airports, of noticeable air traffic. In all cases distance was measured as the crow flies, not along streets.

Historic District. Historic districts in the five-county Cleveland MSA are listed by the National Register of Historic Places (http://www.nationalregisterofhistoricplaces.com). This list indicates the name, address, size (in acres), and date of designation for the district; it also indicates whether the district contains any single family houses. I identified 38 registered historic districts containing single family houses and then looked up their latitudes and longitudes using either http://ohio.hometownlocator.com or the USGS site. I included 5 districts designated in 2001 and 2002 under the assumption that it takes several years to obtain a designation, so that these districts were well-known historic sites in 2000; only two sales, both in Geauga County, took place in one of these sites. To link CBGs and historic districts, I calculated the radius of a circle that would enclose an area the size of the district and then placed a CBG in the district if the distance between the center of the CBG and the center of the historic district was less than or equal to this radius.

Descriptive Statistics. Table A-1 provides descriptive statistics for all these variables, which are listed in Table 4 in the text. Table A-1 also provides descriptive statistics for the variables in Table 3 in the text; to ease interpretation, however, the first eight of these variables are presented in raw form instead of the form in which they enter the regressions. The county and worksite fixed effects are also included in this table.

Figure A-1: Corner Solutions for Distance to Lake Erie

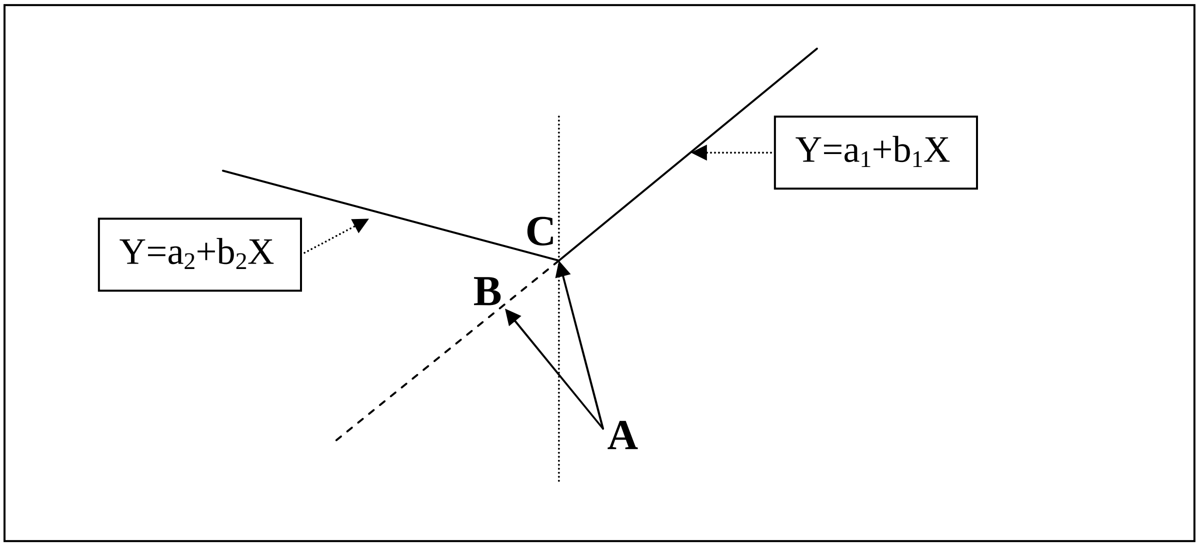

Table A-1. Descriptive Statistics

Variable	Mean	Std. Dev.	Minimum	Maximum
Bidprice	84099.11	23135.67	32165.38	342068.5
Elementary Score	1.03584	0.454023	0.039394	3.048485
9th Grade Score	0.548809	0.245568	0.172	0.9105
High School	11.96919	28.75614	0.198366	177.2222
Safety	31.33959	71.83275	3.853047	1034.333
Distance to Hazard	0.98421	0.749663	0.051717	4.910117
Non Black	2823.153	4502.245	0.01	10000
Non Hispanic	3513.967	4775.025	0.015806	10000
Commute 1	-0.09277	0.032664	-0.17694	-0.00562
Commute 2	0.91189	0.02982	0.837826	0.994394
Income Tax Rate	0.000147	0.001157	0	0.01
School Tax Rate	0.033632	0.008386	0.02	0.0653
City Tax Rate	0.054404	0.025672	0	0.1049
Tax Break Rate	0.030834	0.016738	0	0.080338
Not City	0.139339	0.346404	0	1
Lakefront	0.184384	0.387914	0	1
Distance to Lake	0.180218	0.465641	0	1.997163
Snowbelt 1	0.827995	3.065022	0	19.58192
Snowbelt 2	10.07429	46.86643	0	383.4517
Ghetto	0.104505	0.306006	0	1
Near Ghetto	0.27988	0.449075	0	1
Near Airport	0.320721	0.466894	0	1
Airport Distance	1.829285	2.902368	0	9.989283
Near Public	0.75976	0.427358	0	1
Near Private	0.18018	0.384453	0	1
Land Release	24.50486	682.5647	0	23917
Smog NE	0.085886	0.28028	0	1
Smog SE	0.023423	0.15129	0	1
Smog SW	0.082883	0.275788	0	1
Smog NW	0.004204	0.064723	0	1
Smog Direction NE	1.409507	4.801836	0	19.97808
Smog Direction SE	0.240389	1.819597	0	19.94106
Smog Direction SW	0.869267	3.168338	0	19.84942
Smog Direction NW	0.007141	0.116647	0	2.788374
Local Amenities	0.29009	0.485911	0	2
Freeway	0.135736	0.342611	0	1
Railroad	0.283484	0.450824	0	1
Shopping	0.250451	0.433403	0	1
Hospital	0.024024	0.15317	0	1

Small Airport	0.134535	0.341328	0	1
Big Park	0.207808	0.40586	0	1
Historic District	0.156156	0.363112	0	1
Worksite 2	0.244444	0.429886	0	1
Worksite 3	0.078679	0.269318	0	1
Worksite 4	0.03964	0.19517	0	1
Worksite 5	0.028829	0.167376	0	1
Geauga County	0.04024	0.196581	0	1
Lake County	0.085886	0.28028	0	1
Lorain County	0.13033	0.336768	0	1
Medina County	0.067868	0.251595	0	1

Appendix B: Endogenous Switching Regressions

The estimated coefficients in Table 5 and equation (15) can be used to calculate ψ for each observation. This slope is negative for some observations for every service/amenity. In three cases (9th Grade Score, Distance to Hazard, and Non Hispanic), the number of observations with negative ψ is small so these observations are simply dropped. The other cases are estimated with a selection model (Elementary Score) or an endogenous switching model (High School, Safety, and Non Black). The model to be estimated is (11).

Under normal circumstances, all the terms in (11) are assumed to be positive. But some term in the inverse demand function for *S* (as defined by (8) and incorporated into (11)) must be negative for ψ to be negative. With an endogenous switching regression, we do not have to identify ahead of time the variables that take on negative values for some CBGs. Instead, the model assumes that the sign of MB_S, and hence of ψ, in CBG *i* is determined by a latent variable I_i^*, which is realized as the indicator variable I_i :

$$I_i = 1 \ (\psi < 0) \text{ if } \quad I_i^* = \kappa_{00} + \sum_j \kappa_{0j} N_{ji} + \kappa_{0y} Y_i + \varepsilon_{0i} > 0$$
$$I_i = 0 \ (\psi \geq 0) \text{ if } \quad I_i^* = \kappa_{00} + \sum_j \kappa_{0j} N_{ji} + \kappa_{0y} Y_i + \varepsilon_{0i} \leq 0$$

where the *N*s are the variables in equation (8), the κs are coefficients to be estimated, and the εs are random errors. Because we observe whether the slope is negative, we can estimate this relationship using the realization of this equation.

Equation (11) is multiplicative and would normally be estimated in log form. When the dependent variable and at least one of the explanatory variables (or the implicit multiplicative error term) are negative, as is true for observations with negative ψ, it is not possible to simply take logs, but it is possible to multiply both sides of the equation by -1 so that all terms are positive. In this case the log transformation can be used and the equation can be estimated. This approach assumes, however, that the elasticities for the explanatory variables are equal regardless of whether ψ is positive or negative. An endogenous switching regression allows us to avoid this assumption and estimate separate regressions, and hence separate elasticities, for observations with positive ψ and observations with negative ψ.

More specifically, the endogenous switching model estimates the above equation simultaneously with the following two:

$$\text{Regime 1: } \log\{-\psi_i\} = \kappa_{10} + \sum_j \kappa_{1j} \log\{N_{ji}\} + \kappa_{1y} \log\{Y_i\} + \varepsilon_{1i} \text{ if } I_i = 1$$

$$\text{Regime 2: } \log\{\psi_i\} = \kappa_{20} + \sum_j \kappa_{2j} \log\{N_{ji}\} + \kappa_{2y} \log\{Y_i\} + \varepsilon_{2i} \text{ if } I_i = 0$$

These two equations and the selection equation are estimated using maximum likelihood methods assuming a tri-variate normal distribution for the error terms. See Lokshin and Sahaia (2004). More specifically, these three equations are estimated with the "movestay" command in STATA with robust standard errors.

The Brasington data includes an extensive set of household characteristics, including income, at the CBG level. I selected median family income as the variable most likely to correspond with homeowner income. In addition, I made use of a fairly comprehensive set of household characteristics in specifying the above three equations. The endogenous switching model requires one or more variables in the criterion function regression that are not in the slope regressions. I selected variables that might predict a negative bid that were not significant in the slope regressions.

As indicated in the text, Elementary Score requires a different approach because the negative segment of the bid envelope is very flat and the slope has little variation. So I estimated a selection model, which is equivalent to the endogenous switching model except that the determinants of the slope when it is negative are not estimated.

The empirical results for equation (11) for all seven services and amenities are presented in the following tables. The first table defines the variables and the next seven present STATA output. One table (for Elementary Score) is based on a selection model; two tables (for 9th Grade Scores, Distance to Hazard, and non Hispanic) are based on OLS; and three tables (for High School, Safety, and non Black) are based on endogenous switching regressions.

For the last category of tables, the output includes results for several ancillary parameters. The "lns1" and "lns2" parameters are the logs of "the square roots of the variances of the residuals" in the two slope equations and "r1" and "r1" are "transformations of the correlations between the errors of the two" slope equations and the first equation. See Lokshin and Sahaia (2004). This table also gives results for a test of the hypothesis that the two slope equations are independent. These models perform well. Most of the coefficients are significant with reasonable signs. In all three cases, many factors help identify the CBGs in which the slope of the bid function is negative (Panel 3), and many factors affect the slope of the bid function in a CBG (Panels 1 and 2). The results in Panel 4 indicate that we can reject the hypothesis of independent errors in every case except Non Black.

The results in these tables should be taken as representative of a broader set of regressions. Theory does not provide clear guidance on the set of variables to be included in the demand or selection equations, and the results vary somewhat depending on which variables are included.

Table B-1. Variable Names

Block Group Trait	Label in Estimation Results
log{ψ} for Elementary Score if $\psi > 0$	posslope0
log{ψ} for 9th Grade Score if $\psi > 0$	lslopes1
log{ψ} for High School if $\psi > 0$	lslopes2_0
log{-ψ} for High School $\psi < 0$	lslopes2_1
Dummy (=1 if ψ for High School < 0)	negs2
log{ψ} for Safety if $\psi > 0$	lslopes3_0
log{-ψ} for Safety if $\psi < 0$	lslopes3_1
Dummy (=1 if ψ for Safety < 0)	negs3
log{ψ} for Distance to Hazard if $\psi > 0$	lslopes4
log{ψ} for Non Black if $\psi > 0$	lslopes5_0
log{-ψ} for Non Black if $\psi < 0$	lslopes5_1
Dummy (=1 if ψ for Non Black < 0)	negs5
log{ψ} for Non Hispanic if $\psi > 0$	lslopes6
Median family income (log)	lfaminc
Percent on welfare	Pctwelfare~g
Poverty rate	Poverty_cbg
Percent married	Pctmar_cbg
Percent of households with kids	Pctkids_cbg
Percent of kids in private elementary	privel
Percent of kids in private high school	privhs
Percent with English as 2nd language	Pctenglish~g
Percent foreign born	Pctforeign~g
Percent Asian and Pacific Islander	pctapi_cbg
Percent minority (school district)	pctmin_sd
Percent on welfare (school districgt)	pctadc_sd
Percent Catholic (county)	catholic
Percent other Christian denomination	non_Catholic
Index of educational homogeneity	Leikedu_cbg
Index of racial homogeneity	Leikrace_cbg
Percent in house < 1 year (tract)	pctlt1_ct
Percent owner occupied	Ownerocc_cbg
Percent rural	Ruralhsg_cbg
Percent blue collar	Bluecoll_cbg
Unemployment rate	Unemp_cbg
Percent with high school degree only	pcths_cbg
Percent with some college	pctlesba_cbg
Percent with college degree	pctba_cb g
Percent with graduate degree	Pctgraddeg~g
Percent ages 5-17	Pct517_cbg
Percent elderly	Pct65pls_cbg

Table B-2. Selection Model for Elementary Score

Heckman selection model
(regression model with sample selection)

Number of obs = 1665
Censored obs = 1407
Uncensored obs = 258

Wald chi2(5) = 51.73
Prob > chi2 = 0.0000

Log pseudolikelihood = -936.1502

	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
posslope0						
lfaminc	-.8363613	.2407976	-3.47	0.001	-1.308316	-.3644068
Pct517_cbg	.034007	.0127502	2.67	0.008	.0090171	.058997
pctlt1_ct	-.0555479	.013278	-4.18	0.000	-.0815723	-.0295235
Leikeduc_cbg	3.705229	1.309254	2.83	0.005	1.139138	6.271319
Leikrace_cbg	.8819024	.2696652	3.27	0.001	.3533683	1.410437
_cons	6.409809	2.651532	2.42	0.016	1.212902	11.60672
select						
pctmin_sd	-.0039143	.0014369	-2.72	0.006	-.0067305	-.0010981
pctadc_sd	.0248778	.0026493	9.39	0.000	.0196853	.0300702
lfaminc	.1189556	.1474556	0.81	0.420	-.1700521	.4079633
Pct517_cbg	-.0305044	.0070974	-4.30	0.000	-.044415	-.0165938
Pctmar_cbg	.005747	.0027583	2.08	0.037	.0003408	.0111531
Pctenglish~g	-.0109289	.0032728	-3.34	0.001	-.0173434	-.0045143
Leikeduc_cbg	-2.134656	.7164364	-2.98	0.003	-3.538845	-.7304663
pctlt1_ct	.0038282	.0071609	0.53	0.593	-.0102068	.0178633
pcths_cbg	.0093422	.0041077	2.27	0.023	.0012912	.0173932
pctlesba_cbg	.0095965	.0042216	2.27	0.023	.0013224	.0178706
pctba_cbg	.0035491	.005	0.71	0.478	-.0062508	.0133489
Pctgraddeg~g	.0062402	.005388	1.16	0.247	-.0043202	.0168005
privel	.0058394	.0012862	4.54	0.000	.0033185	.0083602
Ruralhsg_cbg	-.0045808	.0014826	-3.09	0.002	-.0074867	-.0016749
catholic	-.0132374	.0068924	-1.92	0.055	-.0267463	.0002715
non_Catholic	-.0288937	.0108817	-2.66	0.008	-.0502215	-.0075659
_cons	-.2554976	1.650598	-0.15	0.877	-3.49061	2.979615
/athrho	-2.250431	.2767775	-8.13	0.000	-2.792905	-1.707957
/lnsigma	.5394176	.065625	8.22	0.000	.4107949	.6680402
rho	-.9780448	.01202			-.9925266	-.9363964
sigma	1.715008	.1125474			1.508016	1.950411
lambda	-1.677354	.1243319			-1.92104	-1.433668

Wald test of indep. eqns. (rho = 0): chi2(1) = 66.11 Prob > chi2 = 0.0000

Table B-3. Linear Regression for 9th Grade Score

Linear regression

Number of obs	= 1634
F(19, 1614)	= 261.88
Prob > F	= 0.0000
R-squared	= 0.7033
Root MSE	= .973

lslopes1	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
lfaminc	1.240655	.1539811	8.06	0.000	.9386315	1.542679
Pct65pls_cbg	.0343171	.0046112	7.44	0.000	.0252726	.0433616
Pctkids_cbg	.0114674	.0036005	3.18	0.001	.0044053	.0185296
Pctmar_cbg	.0195983	.0036083	5.43	0.000	.0125209	.0266757
Pctforeign~g	.0311619	.0102847	3.03	0.002	.0109893	.0513346
Pctenglish~g	-.0126463	.0060059	-2.11	0.035	-.0244265	-.0008661
pctapi_cbg	-.042579	.0134968	-3.15	0.002	-.0690521	-.016106
Unemp_cbg	-.0554094	.0075228	-7.37	0.000	-.070165	-.0406539
Bluecoll_cbg	-.0115635	.0038723	-2.99	0.003	-.0191587	-.0039684
Ownerocc_cbg	-.0084454	.0022408	-3.77	0.000	-.0128406	-.0040501
catholic	-.045059	.0043779	-10.29	0.000	-.0536461	-.036472
non_Catholic	-.0929897	.0078178	-11.89	0.000	-.1083238	-.0776555
pctlt1_ct	-.0087163	.0050898	-1.71	0.087	-.0186997	.0012671
privel	-.0126737	.0018365	-6.90	0.000	-.0162758	-.0090717
privhs	-.0162243	.0017172	-9.45	0.000	-.0195926	-.0128561
pcths_cbg	.035847	.0055246	6.49	0.000	.0250109	.0466831
pctlesba_cbg	.033242	.0054509	6.10	0.000	.0225504	.0439336
pctba_cbg	.0538031	.0059936	8.98	0.000	.0420469	.0655592
Pctgraddeg~g	.0169755	.0064816	2.62	0.009	.0042623	.0296887
_cons	-17.29602	1.709073	-10.12	0.000	-20.64826	-13.94378

Table B-4. Endogenous Switching Regression for High School

```
Endogenous switching regression model           Number of obs   =       1665
                                                Wald chi2(16)   =     502.40
Log pseudolikelihood = -2846.1234               Prob > chi2     =     0.0000
```

	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
lslopes2_1						
lfaminc	-1.121753	.1633665	-6.87	0.000	-1.441945	-.8015604
Pctkids_cbg	-.0015123	.0033219	-0.46	0.649	-.0080231	.0049985
Pctmar_cbg	-.0227395	.0041706	-5.45	0.000	-.0309137	-.0145653
Pctforeign~g	-.006065	.012154	-0.50	0.618	-.0298865	.0177565
Pctenglish~g	.0097096	.0055016	1.76	0.078	-.0010732	.0204924
pctapi_cbg	.0464976	.0142743	3.26	0.001	.0185204	.0744747
Bluecoll_cbg	.0055973	.0035329	1.58	0.113	-.001327	.0125216
Ownerocc_cbg	.0114569	.0024036	4.77	0.000	.0067458	.0161679
catholic	-.0117909	.0103067	-1.14	0.253	-.0319917	.0084099
non_Catholic	-.0562482	.0193014	-2.91	0.004	-.0940781	-.0184182
privel	.0067362	.0019644	3.43	0.001	.0028861	.0105863
privhs	.0126115	.001586	7.95	0.000	.009503	.01572
pcths_cbg	-.0064855	.0052612	-1.23	0.218	-.0167973	.0038263
pctlesba_cbg	.003706	.0056745	0.65	0.514	-.0074158	.0148279
pctba_cbg	-.0559247	.0088692	-6.31	0.000	-.0733079	-.0385414
Pctgraddeg~g	.0096664	.0130421	0.74	0.459	-.0158956	.0352284
_cons	10.71928	1.738593	6.17	0.000	7.311696	14.12685
lslopes2_0						
lfaminc	.622696	.2068991	3.01	0.003	.2171811	1.028211
Pctkids_cbg	.0031396	.0042943	0.73	0.465	-.0052771	.0115563
Pctmar_cbg	.0140266	.006165	2.28	0.023	.0019434	.0261099
Pctforeign~g	-.0188096	.0151405	-1.24	0.214	-.0484845	.0108653
Pctenglish~g	.0046831	.0120903	0.39	0.699	-.0190135	.0283796
pctapi_cbg	-.0039839	.0182517	-0.22	0.827	-.0397566	.0317889
Bluecoll_cbg	-.0473384	.0074827	-6.33	0.000	-.0620042	-.0326725
Ownerocc_cbg	-.0092604	.0030542	-3.03	0.002	-.0152464	-.0032744
catholic	.0195118	.0074953	2.60	0.009	.0048214	.0342023
non_Catholic	.087573	.0131714	6.65	0.000	.0617576	.1133884
privel	-.0105169	.0025875	-4.06	0.000	-.0155883	-.0054454
privhs	-.004484	.0027919	-1.61	0.108	-.0099559	.0009879
pcths_cbg	-.0237236	.0102956	-2.30	0.021	-.0439027	-.0035445
pctlesba_cbg	-.0200492	.0107079	-1.87	0.061	-.0410363	.0009379
pctba_cbg	-.0134364	.0096472	-1.39	0.164	-.0323446	.0054718
Pctgraddeg~g	-.0200437	.0105275	-1.90	0.057	-.0406771	.0005898
_cons	-9.213055	2.331883	-3.95	0.000	-13.78346	-4.642647
negs2						
Pctkids_cbg	-.0173752	.0064715	-2.68	0.007	-.0300591	-.0046914
Pctenglish~g	-.0255242	.0100148	-2.55	0.011	-.0451529	-.0058955
pctapi_cbg	-.0372519	.0213143	-1.75	0.081	-.0790272	.0045233
Ownerocc_cbg	.0033998	.0035675	0.95	0.341	-.0035924	.0103919
catholic	.0513722	.0097095	5.29	0.000	.032342	.0704024
non_Catholic	.047878	.0159898	2.99	0.003	.0165385	.0792174
privel	.0181957	.0026978	6.74	0.000	.0129082	.0234833
privhs	.0117428	.0025401	4.62	0.000	.0067644	.0167212
pcths_cbg	-.0137521	.0093744	-1.47	0.142	-.0321256	.0046215

pctlesba_cbg	-.0208429	.0099809	-2.09	0.037	-.0404052	-.0012806
Pctgraddeg~g	-.0657509	.0128959	-5.10	0.000	-.0910264	-.0404754
lfaminc	-1.476003	.3126844	-4.72	0.000	-2.088854	-.8631531
Pctmar_cbg	-.0072028	.0060062	-1.20	0.230	-.0189747	.0045691
Pctforeign~g	-.0517243	.0167118	-3.10	0.002	-.0844788	-.0189698
Bluecoll_cbg	.0054164	.006762	0.80	0.423	-.0078369	.0186696
pctba_cbg	-.0367199	.0110417	-3.33	0.001	-.0583612	-.0150786
Pct65pls_cbg	-.0190365	.0080275	-2.37	0.018	-.0347702	-.0033028
Unemp_cbg	.0541632	.0141479	3.83	0.000	.0264339	.0818926
pctlt1_ct	.0197749	.0091343	2.16	0.030	.0018721	.0376777
_cons	17.38806	3.401856	5.11	0.000	10.72054	24.05557
/lns1	-.1419104	.032031	-4.43	0.000	-.2046899	-.0791309
/lns2	.0350889	.0320982	1.09	0.274	-.0278223	.0980001
/r1	-.1856442	.1063755	-1.75	0.081	-.3941363	.022848
/r2	-.1287901	.0800474	-1.61	0.108	-.2856801	.0281
sigma_1	.867699	.0277932			.8149	.923919
sigma_2	1.035712	.0332444			.9725611	1.102963
rho_1	-.1835405	.102792			-.3749206	.022844
rho_2	-.1280827	.0787342			-.278154	.0280926

Wald test of indep. eqns. : chi2(1) = 5.52 Prob > chi2 = 0.0188

Table B-5. Endogenous Switching Model for Safety

Endogenous switching regression model — Number of obs = 1665; Wald chi2(12) = 67.52; Prob > chi2 = 0.0000

Log pseudolikelihood = -2487.5647

	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
lslopes3_1						
lfaminc	.080694	.0386007	2.09	0.037	.0050379	.15635
Pctkids_cbg	.0061213	.001449	4.22	0.000	.0032813	.0089614
Pctmar_cbg	.0033821	.0013437	2.52	0.012	.0007484	.0060157
Pctforeign~g	-.0040679	.0034957	-1.16	0.245	-.0109194	.0027836
Pctenglish~g	.0036669	.0010647	3.44	0.001	.0015801	.0057537
pctapi_cbg	-.0027265	.0049434	-0.55	0.581	-.0124154	.0069624
Bluecoll_cbg	-.001755	.0010592	-1.66	0.098	-.0038311	.0003211
Ownerocc_cbg	.0000583	.000579	0.10	0.920	-.0010767	.0011932
pcths_cbg	-.0044479	.0018345	-2.42	0.015	-.0080434	-.0008524
pctlesba_cbg	-.0086078	.002503	-3.44	0.001	-.0135135	-.0037021
pctba_cbg	.0054226	.0031331	1.73	0.083	-.0007181	.0115633
Pctgraddeg~g	.0135638	.0053734	2.52	0.012	.0030323	.0240954
_cons	.359315	.4292034	0.84	0.402	-.4819083	1.200538
lslopes3_0						
lfaminc	.8865481	.238963	3.71	0.000	.4181893	1.354907
Pctkids_cbg	-.0049618	.0046013	-1.08	0.281	-.0139802	.0040566
Pctmar_cbg	.0100831	.0045671	2.21	0.027	.0011317	.0190345
Pctforeign~g	-.0405614	.0142305	-2.85	0.004	-.0684527	-.0126702
Pctenglish~g	-.014966	.0071917	-2.08	0.037	-.0290614	-.0008706
pctapi_cbg	.0241734	.0232372	1.04	0.298	-.0213707	.0697175
Bluecoll_cbg	-.0006965	.0058064	-0.12	0.905	-.0120769	.0106838
Ownerocc_cbg	.002881	.0027436	1.05	0.294	-.0024964	.0082584
pcths_cbg	.0070114	.0077186	0.91	0.364	-.0081169	.0221396
pctlesba_cbg	-.0109707	.0079773	-1.38	0.169	-.0266059	.0046646
pctba_cbg	.0240075	.0090122	2.66	0.008	.0063438	.0416712
Pctgraddeg~g	.0143126	.0085303	1.68	0.093	-.0024064	.0310317
_cons	-3.719217	2.453304	-1.52	0.130	-8.527605	1.089172
negs3						
Pctforeign~g	.0117969	.0138305	0.85	0.394	-.0153104	.0389042
Pctenglish~g	.0059209	.0070123	0.84	0.398	-.0078228	.0196647
Bluecoll_cbg	-.0009846	.0059093	-0.17	0.868	-.0125666	.0105974
Ownerocc_cbg	.0199124	.0035674	5.58	0.000	.0129205	.0269043
pctlesba_cbg	-.0441972	.0091155	-4.85	0.000	-.0620633	-.0263311
pctba_cbg	-.0148349	.0100745	-1.47	0.141	-.0345806	.0049107
Pctgraddeg~g	-.0464891	.0129805	-3.58	0.000	-.0719304	-.0210477
lfaminc	-1.28429	.2570346	-5.00	0.000	-1.788068	-.7805113
Pctkids_cbg	-.0250793	.0066131	-3.79	0.000	-.0380408	-.0121178
Pctmar_cbg	-.0306322	.0055005	-5.57	0.000	-.0414129	-.0198515
pctapi_cbg	.0480246	.0232903	2.06	0.039	.0023765	.0936727
pcths_cbg	-.0379186	.0080915	-4.69	0.000	-.0537776	-.0220596
Pct65pls_cbg	-.0414523	.0079801	-5.19	0.000	-.057093	-.0258117
Unemp_cbg	.0445511	.012134	3.67	0.000	.0207688	.0683333
pctlt1_ct	.0071766	.0095694	0.75	0.453	-.0115791	.0259323
_cons	16.43827	2.727064	6.03	0.000	11.09332	21.78322

```
        /lns1 |  -1.587491   .5254123    -3.02   0.003     -2.61728   -.5577016
        /lns2 |   .1980925   .0314589     6.30   0.000     .1364342    .2597508
          /r1 |    .084207   .1906601     0.44   0.659    -.2894799    .4578938
          /r2 |   .1932835   .0729214     2.65   0.008     .0503601    .3362068
-------------+----------------------------------------------------------------
      sigma_1 |   .2044379   .1074142                      .0730011    .5725235
      sigma_2 |   1.219075   .0383507                      1.146179    1.296607
        rho_1 |   .0840085   .1893145                      -.281656    .4283661
        rho_2 |    .190912   .0702636                      .0503176    .3240868
------------------------------------------------------------------------------
Wald test of indep. eqns. :          chi2(1) =     7.12   Prob > chi2 = 0.0076
------------------------------------------------------------------------------
```

Table B-6. Linear Regression for Distance to Hazard

Linear regression

Number of obs	= 1658
F(19, 1638)	= 49.14
Prob > F	= 0.0000
R-squared	= 0.3349
Root MSE	= 1.3357

lslopes4	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
lfaminc	.4782995	.1829858	2.61	0.009	.1193888	.8372102
Pct65pls_cbg	.010302	.005763	1.79	0.074	-.0010017	.0216057
Pctkids_cbg	-.0011706	.0048515	-0.24	0.809	-.0106864	.0083453
Pctmar_cbg	.0244091	.0051097	4.78	0.000	.0143869	.0344312
Pctforeign~g	.004042	.0114778	0.35	0.725	-.0184708	.0265547
Pctenglish~g	-.0197158	.0056209	-3.51	0.000	-.0307407	-.0086909
pctapi_cbg	-.0403322	.0191062	-2.11	0.035	-.0778073	-.0028571
Unemp_cbg	-.0006826	.0086771	-0.08	0.937	-.0177019	.0163367
Bluecoll_cbg	-.0165719	.005164	-3.21	0.001	-.0267006	-.0064432
Ownerocc_cbg	-.0023606	.0028021	-0.84	0.400	-.0078566	.0031354
catholic	-.0355998	.0074734	-4.76	0.000	-.0502584	-.0209413
non_Catholic	.0174423	.0140033	1.25	0.213	-.010024	.0449086
pctlt1_ct	-.0255689	.0068016	-3.76	0.000	-.0389096	-.0122283
privel	-.0021771	.0019415	-1.12	0.262	-.0059852	.001631
privhs	-.0049928	.0020149	-2.48	0.013	-.008945	-.0010407
pcths_cbg	.0133039	.0061936	2.15	0.032	.0011557	.0254522
pctlesba_cbg	.0232486	.0069658	3.34	0.001	.0095857	.0369114
pctba_cbg	.0133721	.0073198	1.83	0.068	-.0009849	.0277292
Pctgraddeg~g	.0254572	.0076109	3.34	0.001	.010529	.0403854
_cons	-9.353588	1.994317	-4.69	0.000	-13.26527	-5.441908

Table B-7. Endogenous Switching Model for Non Black

```
Endogenous switching regression model           Number of obs   =       1665
                                                Wald chi2(14)   =      65.40
Log pseudolikelihood =  -5517.423               Prob > chi2     =     0.0000
```

	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
lslopes5_1						
lfaminc	-.0506707	.1408626	-0.36	0.719	-.3267564	.225415
Pctkids_cbg	-.0082459	.0047444	-1.74	0.082	-.0175448	.0010529
Pctmar_cbg	.0008474	.0036784	0.23	0.818	-.0063621	.0080569
Pctforeign~g	.0014628	.0222587	0.07	0.948	-.0421635	.0450891
Pctenglish~g	.0079274	.0138393	0.57	0.567	-.0191971	.0350519
pctapi_cbg	-.0374958	.0297414	-1.26	0.207	-.0957877	.0207962
Unemp_cbg	.0022544	.0053851	0.42	0.675	-.0083003	.012809
Bluecoll_cbg	.0010098	.0034635	0.29	0.771	-.0057785	.0077981
privel	.0017968	.0032262	0.56	0.578	-.0045265	.0081201
privhs	-.0002453	.0022956	-0.11	0.915	-.0047445	.0042539
pcths_cbg	-.001659	.0040005	-0.41	0.678	-.0094998	.0061819
pctlesba_cbg	.0043688	.0051123	0.85	0.393	-.005651	.0143887
pctba_cbg	-.0300602	.012035	-2.50	0.012	-.0536484	-.006472
Pctgraddeg~g	-.0347583	.0092857	-3.74	0.000	-.0529579	-.0165587
_cons	-14.91451	2.007322	-7.43	0.000	-18.84879	-10.98023
lslopes5_0						
lfaminc	4.6081	1.177103	3.91	0.000	2.30102	6.915181
Pctkids_cbg	-.1587467	.0271471	-5.85	0.000	-.211954	-.1055394
Pctmar_cbg	.2623286	.0286873	9.14	0.000	.2061025	.3185548
Pctforeign~g	-.2174386	.0734066	-2.96	0.003	-.3613128	-.0735644
Pctenglish~g	-.1334562	.0408167	-3.27	0.001	-.2134553	-.053457
pctapi_cbg	-.0918906	.108322	-0.85	0.396	-.3041978	.1204167
Unemp_cbg	-.1383367	.0685825	-2.02	0.044	-.2727559	-.0039175
Bluecoll_cbg	-.0172335	.0362555	-0.48	0.635	-.0882929	.0538259
privel	-.0009739	.0144643	-0.07	0.946	-.0293234	.0273756
privhs	-.0148305	.0139629	-1.06	0.288	-.0421973	.0125364
pcths_cbg	.1295372	.0484293	2.67	0.007	.0346176	.2244569
pctlesba_cbg	.05682	.0476963	1.19	0.234	-.036663	.150303
pctba_cbg	.1621187	.0529214	3.06	0.002	.0583946	.2658428
Pctgraddeg~g	-.2760793	.0556472	-4.96	0.000	-.3851459	-.1670127
_cons	-51.48258	11.97889	-4.30	0.000	-74.96077	-28.00439
negs5						
Pctmar_cbg	-.0533217	.0098145	-5.43	0.000	-.0725577	-.0340858
Pctforeign~g	-.046779	.0505295	-0.93	0.355	-.1458151	.0522571
Pctenglish~g	.2070663	.0335978	6.16	0.000	.1412159	.2729167
pctapi_cbg	.1973388	.0526772	3.75	0.000	.0940934	.3005843
Unemp_cbg	.083904	.018526	4.53	0.000	.0475937	.1202143
privel	-.0264212	.0064619	-4.09	0.000	-.0390863	-.0137562
privhs	.0012518	.0052441	0.24	0.811	-.0090265	.01153
pctlesba_cbg	.0251598	.0152875	1.65	0.100	-.0048031	.0551228
Pctgraddeg~g	-.0037998	.0190634	-0.20	0.842	-.0411634	.0335638
lfaminc	-1.042715	.4056351	-2.57	0.010	-1.837745	-.2476845
Pctkids_cbg	-.0084981	.0102609	-0.83	0.408	-.0286091	.0116129
Bluecoll_cbg	-.0607002	.0111983	-5.42	0.000	-.0826485	-.0387518
pcths_cbg	-.0316429	.0147697	-2.14	0.032	-.0605909	-.0026948

```
   pctba_cbg |  -.0981172   .0208449    -4.71   0.000    -.1389725   -.0572619
Pct65pls_cbg |  -.0129094   .0124819    -1.03   0.301    -.0373735    .0115547
Ownerocc_cbg |   .0151364    .006157     2.46   0.014     .0030688    .0272039
   pctlt1_ct |  -.0507908   .0200275    -2.54   0.011    -.0900439   -.0115377
       _cons |  -3.769285   4.889526    -0.77   0.441    -13.35258     5.81401
-------------+----------------------------------------------------------------
       /lns1 |  -.9018489   .2700505    -3.34   0.001    -1.431138   -.3725597
       /lns2 |    2.20451   .0013388  1646.65   0.000     2.201886    2.207134
         /r1 |  -.1656559   .1402384    -1.18   0.238    -.4405182    .1092064
         /r2 |   .0179451   .1328693     0.14   0.893    -.2424739    .2783641
-------------+----------------------------------------------------------------
     sigma_1 |   .4058187   .1095915                      .2390367    .6889685
     sigma_2 |   9.065808   .0121372                       9.04205    9.089627
       rho_1 |  -.1641571   .1364593                     -.4140739    .1087743
       rho_2 |   .0179432   .1328265                     -.2378311    .2713903
------------------------------------------------------------------------------
Wald test of indep. eqns. :          chi2(1) =     1.42   Prob > chi2 = 0.2339
------------------------------------------------------------------------------
```

Table B-8. Linear Regression for Non Hispanic

Linear regression

Number of obs	= 1650
F(19, 1630)	= 15.28
Prob > F	= 0.0000
R-squared	= 0.1440
Root MSE	= 9.2209

lslopes6	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
lfaminc	-.392665	1.241302	-0.32	0.752	-2.827381	2.042051
Pct65pls_cbg	.0958185	.043652	2.20	0.028	.0101985	.1814384
Pctkids_cbg	.0447367	.0345446	1.30	0.195	-.0230198	.1124933
Pctmar_cbg	-.0019403	.0323357	-0.06	0.952	-.0653642	.0614837
Pctforeign~g	.0478445	.1112355	0.43	0.667	-.1703351	.266024
Pctenglish~g	.3313235	.0738328	4.49	0.000	.1865064	.4761406
pctapi_cbg	.0989073	.1055382	0.94	0.349	-.1080975	.305912
Unemp_cbg	.0377606	.064039	0.59	0.556	-.0878468	.163368
Bluecoll_cbg	-.1065303	.0343992	-3.10	0.002	-.1740017	-.039059
Ownerocc_cbg	.0330984	.0199849	1.66	0.098	-.0061005	.0722972
catholic	.023357	.0482959	0.48	0.629	-.0713716	.1180856
non_Catholic	.1661785	.0910048	1.83	0.068	-.0123202	.3446772
pctlt1_ct	-.1766616	.0482908	-3.66	0.000	-.2713801	-.0819431
privel	-.00835	.0150667	-0.55	0.580	-.0379022	.0212022
privhs	-.021321	.0137686	-1.55	0.122	-.048327	.005685
pcths_cbg	-.0917815	.0485561	-1.89	0.059	-.1870205	.0034575
pctlesba_cbg	-.0917875	.0480249	-1.91	0.056	-.1859845	.0024095
pctba_cbg	-.0742896	.0537191	-1.38	0.167	-.1796553	.0310762
Pctgraddeg~g	-.0499533	.058884	-0.85	0.396	-.1654497	.065543
_cons	-22.8889	15.30265	-1.50	0.135	-52.90383	7.126031