DATA DOCUMENTATION FOR:
THE SPEED OF GASOLINE PRICE RESPONSE IN MARKETS WITH AND WITHOUT EDGEWORTH CYCLES

Matthew Lewis
The Ohio State University*

Michael Noel
University of California
San Diego†

This document describes the data used in the main article and provides additional details about the code that has been provided to estimate the models discussed in the article. The retail and wholesale gasoline price data used in the article were provided by the Oil Price Information Service (OPIS), and the terms of the data acquisition agreement prevent us from publicly providing the data here. However, we will describe how one can acquire this data from OPIS and how it can be used to replicate the analysis in the article.

Data

The retail prices used in the article are daily metropolitan average retail gasoline prices (before taxes) provided by OPIS for the period January 1, 2004 to August 28, 2005 for 90 cities. The specific cities used are listed in Table 1. For 85 of the 90 cities mean retail prices for the Metropolitan Statistical Area (MSA) are used. For 5 cities selected there was no appropriate corresponding MSA, so OPIS provided us mean prices for the geographic area corresponding to city as specified by the U.S. Census Bureau.

Wholesale prices provided by OPIS are the daily minimum rack prices for unbranded gasoline for the same time period at the terminals closest to each of the 90 selected cities. These terminal names are also provided in Table 1.

*Department of Economics, E-mail: mlewis@econ.ohio-state.edu.
†Department of Economics, E-mail: mdnoel@ucsd.edu

For the analysis in Section 5 of the article U.S. Census data was collected for the MSA surrounding each selected city. The number of gas stations in each MSA was collected using the 2002 Economic Census Establishments Data. We utilized the count of establishments in the 4471 NAICS code which is defined as "Gasoline Stations". We also collected total population and total land area for each MSA from the 2000 Decennial Census Summary File 3. Finally, to generate the brand market shares for each MSA we utilize the OPIS Rack to Retail Margins Report from October 10, 2005. For each MSA, this report lists the number of gasoline stations that are operated under each brand name. We use this to identify market shares for each brand and an HHI for the market. We categorize stations as *branded* or *independent*, where branded refers to stations that operate under a major refinery brand and independent refers to stations that are unbranded or operate under a convenience store brand not affiliated with a brand of gasoline from a major refinery. For our purposes, we define the major refinery brands as: Amoco, BP, 76, Chevron, Citgo, Conoco, Diamond Shamrock, Exxon, Gulf, Hess, Ashland, Mobil, Phillips, Shell, Sinclair, Sunoco, Tesoro, Texaco, and Valero.

Code and Estimation

Price Response in Non-cycling Markets

The estimation of price pass-through in non-cycling markets is done in Stata using the ECreg_noncycle.do program. This program inputs a Stata dataset (OPISCityData.dta) containing the following variables:

region: The MSA for which the average retail prices are observed

rack: The city where the wholesale distribution rack is located

rackstate: The state where the rack is located

c: The daily wholesale rack price

p: The daily average retail price (before taxes)

date: Stata's numerical date code for the relevant day

The program estimates the empirical model in Equations 3 & 4 and then calculates estimated cumulative response functions (CRFs). When complete, the program displays a matrix which contains the following columns:

Column 1 CRF for a positive cost change

Column 2 CRF for a negative cost change

Column 3 Upper bound of the 95% confidence CRF for the estimate of a positive cost change

Column 4 Lower bound of the 95% confidence CRF for the estimate of a positive cost change

Column 5 Upper bound of the 95% confidence CRF for the estimate of a negative cost change

Column 6 Upper bound of the 95% confidence CRF for the estimate of a negative cost change

Each column has 50 rows containing the corresponding price response estimates for the first 50 periods following a cost shock.

Price Response in Cycling Markets

Price response in cycling markets is modeled by a Markov switching regression in Gauss using the program Markov.txt. This file reads in a Gauss data file called pricedat.dat which contains the following 4 columns:

Column 1 Stata's numerical date code for the relevant day

Column 2 An MSA code from 1–90 as listed in Table 1

Column 3 The daily average retail price (before taxes)

Column 4 The daily wholesale rack price

All 90 cities are included in the original data file, so it contains 90*606 rows. At the top of the program file the *cities* vector is defined which determines which city numbers will be included as cycling cities. This is all cities that have a median daily price change of less than -.2. The program outputs two files:

x9g.fmt Contains the parameter vector

hnw9g.fmt Contains the variance-covariance matrix

The order of the parameters reported in these files is as follows:

Relenting Regime Price Equation:
1. constant
2. p_{t-1}
3. r_t
4. $r_t - r_{t-1}$
5. $r_{t-1} - r_{t-2}$
...
21. $r_{t-17} - r_{t-18}$

Undercutting Regime Price Equation:
22. constant
23. p_{t-1}
24. r_t
25. $r_t - r_{t-1}$
...
42. $r_{t-17} - r_{t-18}$

Switching Equation R - R
43 - 63. same as above

Switching Equation U - R
64 - 84. same as above

85. σ^R: variance on restoration phase price changes
86. σ^U: variance on undercutting phase price changes

87-118. city dummies in the restoration phase
119-150. city dummies in the undercutting phase

The x9g.fmt and hnw9g.fmt output files are read in by another Gauss program (Sim.txt) which simulates the cumulative response functions for a price increase and a price decrease and reports these in a table after the simulation is complete. Note that you can set the parameter that determines the size of the cost shock for which the CRF is simulated. This is the parameter dr which is set toward the beginning of the file. The output table contains 17 different estimates for each

of the 18 periods that are simulated.[1] The estimates from this table that are used in the paper are the 14th, 15th, and 16th estimate from every simulation period. These correspond to the estimates of the cumulative response to a cost increase (crf_up), the cumulative response to a cost decrease (crf_dn), and the difference between these two responses which reflects the response asymmetry (asym). The first two of these estimates are presented in Figures 3 & 4 in the article.

City Specific Price Response Estimates and the Speed of Price Response Regressions

Section 5 of the article contains a regression of the speed of price response in individual cities on various city specific characteristics. This requires city specific estimates of price response. For non-cycling cities this is performed using the program ECreg_noncycle_cityCRFs.do. This program uses the same input data as the ECreg_noncycle.do. It outputs a data file called noncycle_city_CRFs.dta which contains the estimated CRFs for each of the 39 non-cycling markets. The dataset has 78 rows: two for each city. The first row for each city contains the response to positive cost shock, the second contains the response for a negative cost shock. The dataset contains the following variables:

region: The MSA from which the CRFs in the current observation are estimated

up_down: A 1 if the CRF in the current observation is for a price increase, a 2 if it is for a price decrease

day1 - day50: 50 variables containing the percentage of passthrough on each of the 50 days following the cost shock

For cycling cities, CRFs can be simulated for each city using the Gauss program Sim_city.txt. This program operates similar to the Sim.txt program and reads in the same two files (x9g.fmt and hnw9g.fmt) as inputs. The two main difference from the Sim_city.txt program. First, this program simulates a CRF for a specific city as defined by the parameter *citynum* which is defined in the beginning of the file and can be set from 1 to 33. Note that these city numbers are different from the original city numbers ranging from 1 to 90 because the other cities have been dropped and only the 33 cycling cities remain. The second main difference is that Sim_city.txt only simulates the CRF for 8 periods after the cost change, so the table that is produced contains 17*8 rows containing the same 17 estimates per period that were produced by Sim.txt. This program is run 33 times, once for each *citynum*, to estimate CRFs for all the cycling cities.

[1] The table therefore has 17*18 rows. The 4 columns are: 1) the period of simulation, 2) the name of the estimated value, 3) the estimated value, and 4) the standard error of the estimate.

Once CRFs for the cycling cities are one can construct a dataset similar to that created for non-cycling cities. These datasets will be used as an inputs for estimating the speed of price response regressions, and should have the following format:

region: The MSA from which the CRFs in the current observation are estimated

up_down: A 1 if the CRF in the current observation is for a price increase, a 2 if it is for a price decrease

day1 - day8: 8 variables containing the percentage of passthrough on each of the 8 days following the cost shock

Note that since we estimate CRFs for cycling cities using two different sized cost shocks we need to construct two different data files with the variables above. We call these files Cycle_region_CRFs_2cents.dta and Cycle_region_CRFs_4cents.dta, and these become inputs for the final estimation program.

The CRF estimates for both cycling and non-cycling cities are used as inputs in the programs Response_Speed_2_69.do and Response_Speed_4_92.do. Since the response paths predicted for cycling markets depend on the size of the cost shock, these two programs analyze the differences in response across cities for two different sized cost shocks. The programs perform the regressions in Section 5 of the article relating response speeds to market characteristics for a 2.69 cent and a 4.92 cent cost shock respectively. The other input into these programs is a dataset of market characteristics named market_characteristics.dta which contains the following variables:

region: The Metropolitan Statistical Area

populationperstation: Population of the MSA divided by the number of gasoline stations in the MSA according to the 2002 Economic Census Establishments Data

indephhi: Herfindahl index within independent retail chains and stations

indepshr: Number of stations operating independently or under retail chains divided by the total number of observed stations in the market based on the OPIS Rack to Retail Margins Report from October 10, 2005.

Each program runs two of the 4 regressions in Table 2 of the article corresponding to the size of the cost shock considered in each program.

TABLE 1: METROPOLITAN AREAS INCLUDED IN DATA SET

City Number	Region Name	Rack Location
1	City - IL, Quincy	Palmyra, MO
2	City - IN, South Bend	Niles, MI
3	City - KY, Louisville	Louisville, KY
4	City - MI, Muskegon	Muskegon, MI
5	City - MI, Traverse City	Traverse City, MI
6	MSA-1000 Birmingham, Al	Birmingham, AL
7	MSA-1020 Bloomington, In	Switz City, IN
8	MSA-1040 Bloomington-Normal, Il	Heyworth, IL
9	MSA-1280 Buffalo-Niagara Falls, Ny	Buffalo, NY
10	MSA-1320 Canton-Massillon, Oh	Akron/Canton, OH
11	MSA-1360 Cedar Rapids, Ia	Iowa City, IA
12	MSA-1400 Champaign-Urbana, Il	Champaign, IL
13	MSA-1480 Charleston, Wv	Charleston, WV
14	MSA-1520 Charlotte-Gastonia-Rock Hill, (NC Only)	Charlotte, NC
15	MSA-1560 Chattanooga, Tn-Ga (TN Only)	Chattanooga, TN
16	MSA-160 Albany-Schenectady-Troy, Ny	Albany, NY
17	MSA-1600 Chicago, Il	Chicago, IL
18	MSA-1640 Cincinnati, Oh-Ky-In (OH Only)	Cincinnati, OH
19	MSA-1680 Cleveland-Lorain-Elyria, Oh	Cleveland, OH
20	MSA-1740 Columbia, Mo	Columbia, MO
21	MSA-1760 Columbia, Sc	Columbia, SC
22	MSA-1840 Columbus, Oh	Columbus, OH
23	MSA-1920 Dallas, Tx	Dallas/Ft. Worth, TX
24	MSA-1960 Davenport-Moline-Rock Island, (IA Only)	Bettendorf, IA
25	MSA-1960 Davenport-Moline-Rock Island, (IL Only)	Peoria, IL
26	MSA-2000 Dayton-Springfield, Oh	Dayton, OH
27	MSA-2040 Decatur, Il	Decatur/Forsythe, IL
28	MSA-2120 Des Moines, Ia	Des Moines, IA
29	MSA-2160 Detroit, Mi	Detroit, MI
30	MSA-2240 Duluth-Superior, Mn-Wi (MN Only)	Duluth, MN
31	MSA-2290 Eau Claire, Wi	Chippewa Falls, WI
32	MSA-240 Allentown-Bethlehem-Easton, Pa	Macungie, PA
33	MSA-2440 Evansville-Henderson, In-Ky (IN Only)	Evansville, IN
34	MSA-2520 Fargo-Moorhead, Nd-Mn (ND Only)	Fargo, ND
35	MSA-2580 Fayetteville-Springdale-Rogers	Rogers , AR
36	MSA-2640 Flint, Mi	Flint, MI
37	MSA-2720 Fort Smith, Ar-Ok (AR Only)	Ft. Smith, AR
38	MSA-2760 Fort Wayne, In	Huntington, IN
39	MSA-2960 Gary, In	Hammond, IN
40	MSA-3000 Grand Rapids-Muskegon-Holland,	Muskegon, MI
41	MSA-3080 Green Bay, Wi	Green Bay, WI
42	MSA-3120 Greensboro–Winston-Salem–Hig	Greensboro, NC
43	MSA-320 Amarillo, Tx	Amarillo, TX
44	MSA-3240 Harrisburg-Lebanon-Carlisle, P	Harrisburg , PA
45	MSA-3400 Huntington-Ashland, Wv-Ky-Oh (KY Only)	Ashland, KY

TABLE 1(CONTINUED): METROPOLITAN AREAS INCLUDED IN DATA SET

City Number	Region Name	Rack Location
46	MSA-3480 Indianapolis, In	Indianapolis, IN
47	MSA-3520 Jackson, Mi	Jackson, MI
48	MSA-3710 Joplin, Mo	Carthage, MO
49	MSA-3720 Kalamazoo-Battle Creek, Mi	Jackson, MI
50	MSA-3760 Kansas City, Mo-Ks (MO Only)	Kansas City, KS
51	MSA-3840 Knoxville, Tn	Knoxville, TN
52	MSA-4040 Lansing-East Lansing, Mi	Lansing, MI
53	MSA-4150 Lawrence, Ks	Topeka, KS
54	MSA-4280 Lexington, Ky	Lexington, KY
55	MSA-4360 Lincoln, Ne	Lincoln, NE
56	MSA-4400 Little Rock-North Little Rock,	Little Rock, AR
57	MSA-460 Appleton-Oshkosh-Neenah, Wi	Green Bay, WI
58	MSA-4720 Madison, Wi	Madison, WI
59	MSA-4920 Memphis, Tn-Ar-Ms (TN Only)	Memphis, TN
60	MSA-5080 Milwaukee-Waukesha, Wi	Milwaukee , WI
61	MSA-5120 Minneapolis-St. Paul, Mn-Wi (MN Only)	Minneapolis, MN
62	MSA-520 Atlanta, Ga	Atlanta, GA
63	MSA-5360 Nashville, Tn	Nashville, TN
64	MSA-5880 Oklahoma City, Ok	Oklahoma City, OK
65	MSA-5920 Omaha, Ne-Ia (NE Only)	Omaha, NE
66	MSA-5990 Owensboro, Ky	Owensboro, KY
67	MSA-6120 Peoria-Pekin, Il	Peoria, IL
68	MSA-6160 Philadelphia, Pa-Nj (PA Only)	Philadelphia, PA
69	MSA-6280 Pittsburgh, Pa	Pittsburgh, PA
70	MSA-6820 Rochester, Mn	Rochester, MN
71	MSA-6880 Rockford, Il	Rockford, IL
72	MSA-6960 Saginaw-Bay City-Midland, Mi	Bay City, MI
73	MSA-6980 St. Cloud, Mn	Sauk Center, MN
74	MSA-7000 St. Joseph, Mo	Kansas City, KS
75	MSA-7040 St. Louis, Mo-Il (MO Only)	St Louis, MO
76	MSA-7560 Scranton–Wilkes-Barre–Hazlet	Scranton, PA
77	MSA-7720 Sioux City, Ia-Ne (IA Only)	Sioux City, IA
78	MSA-7760 Sioux Falls, Sd	Sioux Falls, SD
79	MSA-7880 Springfield, Il	Decatur, IL
80	MSA-7920 Springfield, Mo	Springfield, MO
81	MSA-80 Akron, Oh	Akron/Canton, OH
82	MSA-8160 Syracuse, Ny	Syracuse, NY
83	MSA-8320 Terre Haute, In	Switz City , IN
84	MSA-8400 Toledo, Oh	Toledo, OH
85	MSA-8440 Topeka, Ks	Topeka, KS
86	MSA-8560 Tulsa, Ok	Tulsa, OK
87	MSA-870 Benton Harbor, Mi	Niles, MI
88	MSA-8920 Waterloo-Cedar Falls, Ia	Waterloo, IA
89	MSA-9040 Wichita, Ks	Wichita, KS
90	MSA-9320 Youngstown-Warren, Oh	Youngstown, OH