**EXERCISE 5. THE MARKET FOR GASOLINE**

(To be handed in on Nov. 30)

Several oil refinery accidents in California in the Spring of 1999 restricted production of gasoline for motor vehicles, and prices at the pumps soared. Some consumer groups argued that the price increases were not justified by the supply restrictions, and were in large part the result of price-fixing by the refiners. The oil companies responded that the price increases were simply the result of competitive supply and demand. To shed some light on the question of whether the increases in California gasoline prices were consistent with competitive pricing, this exercise asks you to estimate a simultaneous equations model for the demand and supply of gasoline, and test the hypothesis that the observed price changes in California are consistent with the demand elasticity for gasoline.

The (stylized) facts about the market for gasoline in California in 1999 are that refinery capacity was reduced by 20 percent for 3 months, and gasoline prices went from an average of $1.20 per gallon in January to $1.56 per gallon over the remainder of the year (up to the present). If the refineries are not restricting supply *more* than the reduction in capacity due to accidents, then these market changes should reflect the short-run elasticity of demand for gasoline. An exogenously driven 5 percent reduction in annual supply coupled with a 30 percent increase in price suggests the short-run price elasticity of gasoline demand should be about -0.167. The objective of this exercise is to test the hypothesis that the short-run demand elasticity for gasoline is this number.

The file gas.dat contains annual data from 1968 through 1999 for the U.S. on the variables listed below. The variables are delimited by spaces and can be read in free format.

| | |
|---|---|
| YEAR | 4-digit year of observation |
| PGAS | Real price of gasoline, U.S. retail average, in 1992 dollars per gallon, from EIA |
| QGAS | Consumption of gasoline by motor vehicles, in billions of gallons per year, from EIA |
| PCRUDE | Real price of crude oil, 1992 dollars per barrel, from EIA |
| POP | U.S. Population, millions, from ERP |
| GDP | Real Gross Domestic Product, in billions of 1992 dollars, from ERP |
| PGDP | Implicit Price Deflator for GDP, 1992 = 100, from ERP |
| MPG | Fuel efficiency of motor vehicles, in miles per gallon, from EIA |
| VEH | Number of registered vehicles, in millions, from FHWA |
| PVEH | Real average price of vehicles in 1992 dollars, from ERP, normalized so that (PVEH×VEH)/(IPC×POP) equals the expenditure share of consumers on motor vehicles in 1992 |
| IR | Nominal mortgage interest rate, from ERP |
| SCC | Service cost of motor vehicle capital, per dollar |
| IPC | Real income per capita, in 1992 dollars, from ERP |

With the exception of PGAS, PCRUDE, PGDP, and PVEH where preliminary estimates based on part-year data are available, the variables have missing data in 1999, coded as ".". Most statistical packages will read this as missing data. Also, the variables MPG and VEH are preliminary estimates in 1998.

*Model Specification*

The demand for gasoline has short-run and long-run elements. In the short run, the number of vehicles in operation and their fuel efficiency is fixed, as are many activities that require vehicle use. However, even in the short-run, consumers may show some response to price by rescheduling trips and modifying their driving habits. A plausible specification for the determination of short-run demand is

$$(VMT) = (PGAS)^{\beta}\cdot e^{\alpha+\varepsilon},$$

where VMT = QGAS·MPG/VEH is average vehicle-miles traveled per vehicle, $\alpha,\beta$ are parameters and $\varepsilon$ is a disturbance reflecting "unintended" travel. This is a constant elasticity form with QGAS proportional to VMT that facilitates answering the policy question initially posed. One can interpret $\beta$ as the short-run behavioral response of vehicle use to changes in operating cost. Because of lags in behavioral response, $\varepsilon$ is likely to exhibit serial correlation.

Estimation of the short-run demand function is the primary focus of this exercise. However, QGAS and PGAS are jointly determined in the market, and the demand function cannot be estimated in isolation without taking into account the operation of the overall market. Therefore, the exercise goes on to specify a long-run demand equation in which vehicles per capita is chosen by the consumer, and a pricing equation that reflects the technology and industrial structure of petroleum refining. For the immediate policy question, what is most important about specifying the other elements in the market is identification of proper instruments that can be used in 2SLS to estimate the short-run demand equation. For this purpose, it is not necessarily critical to get the exact functional forms of long-run demand and of supply correctly specified as long as one has correctly identified the variables upon which they depend. However, if these elements of the market are modeled well, there is the possibility of improving estimator efficiency by using full system estimation methods (e.g., 3SLS), and of carrying out tests of specification and over-identification that may increase confidence in the parameter estimates obtained for the short-run demand equation by 2SLS.

*Long-run demand*

In the long run, there will be behavioral response through changes in vehicle ownership.[1] Consumer theory provides a link between the short-run and long-run behavioral responses. The

[1] There is also some behavioral adjustment in vehicle fuel efficiency through choice of vehicle size. However, econometric analysis of this dimension would require additional data on vehicle sizes and on the determinants of fuel efficiency by size, including technological innovation and EPA mandates. For this exercise, assume as a reasonable approximation that fuel efficiency is exogenously determined.

short-run per capita demand for gas should be tied to the representative consumer's indirect utility function u = U(y,PGAS,PPV,MPG) of per capita income y (net of vehicle purchase cost), the price PGAS, the number of people per vehicle PPV = POP/VEH, and MPG: By Roy's identity,

$$-\left(\frac{\partial U/\partial PGAS}{\partial U/\partial y}\right) = (QGAS)/(POP),$$

or

$$\left(\frac{dy}{dPGAS}\right)_{u,MPG,PPV=const} = e^{\alpha+\varepsilon}\cdot(PGAS)^{\beta}/(PPV\cdot MPG).$$

Treating this as a differential equation and solving yields

$$y = e^{\alpha+\varepsilon}\cdot(PGAS)^{1+\beta}/((1+\beta)\cdot PPV\cdot MPG) + \psi(u,PPV,MPG),$$

where $\psi$ is an increasing function of u. Over the relevant range, we expect utility to decrease as persons per vehicle rises, so that $\psi$ should be increasing in PPV. Because consumers prefer larger and more powerful vehicles, other things equal, and MPG is negatively correlated with size and power, $\Psi$ should be increasing in MPG. However, because of technical progress, the MPG of vehicles of a given size and power has risen over time, so that the disutility of sizing down to a specified MPG is declining over time. Then, the rate of increase of $\psi$ with MPG should be declining over time.

The net per capita income y will equal IPC less the annual service cost of the motor vehicles owned per capita. The service cost of a dollar's worth of vehicle, neglecting taxes, is

$$SCC = (RIR + \mu - \Delta PVEH/PVEH),$$

where RIR is the real rate of interest to consumers (equal to the nominal rate, IR, less the rate of inflation, $\Delta$PGDP/PGDP), $\mu$ is a maintenance and depreciation rate, and $\Delta$PVEH/PVEH is the net rate of real capital gain for vehicles. For this exercise, assume that $\mu$ is 8 percent. This is a commonly used rate for vehicles, and corresponds to a service life of about 10 years. The implicit rental rate per motor vehicle is then RVEH = SCC·PVEH, and the indirect utility per capita satisfies

$$IPC - RVEH/PPV - e^{\alpha+\varepsilon}\cdot(PGAS)^{1+\beta}/((1+\beta)\cdot PPV\cdot MPG) = \psi(u,PPV,MPG).$$

Putting these elements together, the utility per capita is

$$u = f(IPC - RVEH/PPV - e^{\alpha+\varepsilon}\cdot(PGAS)^{1+\beta}/((1+\beta)\cdot PPV\cdot MPG),PPV,MPG),$$

where f is a function that is increasing in its first argument and decreasing in its second and third arguments. Maximizing this in PPV yields an "instantaneous" desired level for people per vehicle; approximate this solution by a log linear form,

$$PPV^{*} = \lambda\cdot IPC^{\gamma}\cdot[RVEH + e^{\alpha+\varepsilon}\cdot(PGAS)^{1+\beta}/((1+\beta)\cdot MPG)]^{\delta}$$
$$= \lambda\cdot IPC^{\gamma}\cdot[RVEH + CPV/(1+\beta)]^{\delta},$$

with the second form obtained by noting that $e^{\alpha+\varepsilon}\cdot(PGAS)^{1+\beta}/MPG = VMT\cdot PGAS/MPG = CPV$, the operating cost per vehicle. In this expression, $\lambda$,$\gamma$, and $\delta$ are parameters. This is the level that would maximize utility in a static environment in which current income, prices, and technology do not change in the future. If consumers have relatively naive proportional *adaptive* expectations, then the observed adjustment in people per vehicle will satisfy

$$PPV_{+1} =(PPV^{*})^{\theta}\cdot(PPV)^{(1-\theta)}\cdot e^{\nu} = \lambda^{\theta}\cdot IPC^{\gamma\theta}\cdot[RVEH + CPV/(1+\beta)]^{\delta\theta}\cdot(PPV)^{(1-\theta)}\cdot e^{\nu},$$

where $PPV_{+1}$ denotes the variable one year ahead and $\nu$ is a disturbance in the adaptive adjustment.

Summarizing, the demand side of the market is given in the short-run by the equation

$$\log(VMT) = \alpha + \beta\cdot\log(PGAS) + \varepsilon.$$

If serial correlation in $\varepsilon$ can be described by an AR1 process, then this equation becomes

$$\log(VMT) = \alpha\cdot(1-\rho) + \rho\cdot\log(VMT_{-1}) + \beta\cdot\log(PGAS) - \rho\beta\cdot\log(PGAS_{-1}) + \eta$$

with $\eta$ serially independent. In the long run, the demand side of the market is determined by

$$\log(PPV_{+1}) = \theta\cdot\log(\lambda) + \gamma\theta\cdot\log(IPC) + \delta\theta\cdot\log[RVEH + CPV/(1+\beta)] + (1-\theta)\cdot\log(PPV) + \nu.$$

*Supply Side Pricing*

On the supply side of the gasoline market, petroleum refining is a process industry with constant or mildly increasing returns to scale. Then, the total cost function of a refiner is an increasing function of crude oil price, and an increasing mildly concave function of output. For example, a plausible specification for total cost of a refiner with gasoline output q is

$$TC = FC + (m\cdot PCRUDE + w)\cdot e^{\zeta}\cdot q^{1-\tau}/(1-\tau),$$

where FC is fixed cost, m = 0.04524 is a conversion factor giving the barrels of crude necessary to produce a gallon of gasoline, w summarizes other variable unit costs, N is the number of refiners,

$\tau$ is a scale factor which should be small positive or zero, and $\zeta$ is a disturbance. This specification assumes that all the refiners have the same cost function. Assume the refiner knows the short-run gasoline demand function, which can be inverted to give

$$(PGAS) = (MPG\cdot((N-1)q_0+q)/VEH)^{1/\beta}e^{-(\alpha+\varepsilon)/\beta},$$

where $q_0$ is the average output of all other refiners and q is this refiner's output. Refiners have historically derived about 48 percent of their revenue from the sale of gasoline, and the remainder from other petroleum products. Refiners have some control over the mix of their products, but as an approximation assume they are produced in fixed proportions. Make the strong assumption that the prices of all the products move together. (A useful exercise in economic theory is to consider the circumstance under which this could be true.) The refiner's revenue as a function of its output q is then 2.09 times the revenue from gasoline, making its profit the function

$$2.09q\cdot(MPG\cdot((N-1)q_0+q)/VEH)^{1/\beta}e^{-(\alpha+\varepsilon)/\beta} - FC - (m\cdot PCRUDE + w)\cdot e^{\zeta}\cdot q^{1-\tau}/(1-\tau).$$

and this is maximized at a value q, equal by symmetry to $q_0$, that results in the price

$$PGAS = (VEH/MPG)^{-\tau/(1+\tau\beta)}\cdot(2.09\cdot(1+1/N\beta)/N^{\tau})^{-1/(1+\tau\beta)}\cdot(w+m\cdot PCRUDE)^{1/(1+\tau\beta)}\cdot e^{(\zeta-\varepsilon\tau)/(1+\tau\beta)}.$$

Treat w as a parameter. Estimation of this equation yields estimates of $\tau$, $\beta$, and w. Note that $\varepsilon$ enters the determination of PGAS when $\tau \neq 0$, inducing a correlation of the explanatory variable PGAS and the disturbance in the short-run demand equation. The presence of $\varepsilon$ will also induce some serial correlation, in addition to any lags there may be in refiner price response due to inventories and other factors.

*The System*

The discussion above has led to specification of a three-equation simultaneous equations system explaining *short-run demand* associated with dependent variable VMT, *long-run demand*, determined by the dependent variable PPV, and *supply price*, with dependent variable PGAS. The short-run demand function is log linear in variables, and linear in parameters, except for the non-linearities introduced by partial differencing to handle AR1 serial correlation. The remaining two equations are nonlinear in variables and parameters. Variables that can be identified as proper instruments, in addition to constants, are RVEH, PCRUDE, IPC, MPG, and POP, and in the case of serial correlation, lagged values of these variables.

There are some problematic elements in the specification of the long-run demand curve and in the supply price function. The representative consumer model may not hold, and heterogeneity in income, tastes, and vehicle holdings may be important. There may be feedback from vehicle holdings to vehicle use. The simple adaptive expectations model may be inappropriate, although economists rarely go wrong by assuming consumers are short-sighted. On the supply side, the Cournot assumption for refiner conduct, the assumed homogeneity in technologies, refiner myopia that ignores the long-run impact of price on vehicle ownership, treatment of non-gasoline petroleum products, and the absence of a theory of entry, are all possible limitations. This suggests caution in use of full system estimators that take the functional specifications for long run demand and for supply to be exactly right.

*Estimation Tasks*

a. Estimate the short run demand equation by OLS and by instrumental variables (2SLS). Test for AR1 serial correlation and for a unit root, and if necessary transform the equation before applying instrumental variables.

b. For the most satisfactory specification you estimate in part a, test if the short run demand elasticity is consistent with the price response to supply restrictions in 1999 in California described at the start of this exercise.

c. Conditioned on the estimate of $\beta$ obtained in part a, estimate the long-run demand and supply equations using nonlinear least squares. Then do NL2SLS estimation of each of these equations. Review Greene and the TSP manual to be sure you understand what the TSP command LSQ does in the case of instrumental variables and a nonlinear equation.

d. Do NL3SLS estimation of the full system. What can you conclude from the results about the validity of the specification of the system? About the policy conclusion from part b?