Nick Mastronardi

Spring 08

Heterogeneous Demand Patterns

Motivation

There are many possible normative questions that might require analysis such as that performed in this exercise. For example, a policy question about optimal commodity taxation would require an accurate modeling of price elasticities. Or, firm profit maximization would like to have a good model of price and income elasticities. Also, marketing organizations might like to have a total model of demand behavior accounting for heterogeneity.

Assumptions and Imposed Structure

A few basic assumptions were imposed on consumer unit behavior to simplify analysis. It was assumed that consumer units are rational Utility maximizers. Further, it was assumed consumer units two-stage budget. Given this assumption, I did not need to analyze demand substitution behavior between goods in different time periods. There may be reason to doubt this imposed structure in the future, but for the analysis provided it was assumed that this was the structure. It greatly reduces the size of our elasticity matrix and adds some fidelity to our data as it may not have been the case that the consumer units in different periods are the same consumer unit. Lastly, it was assumed that agents share a common utility form specification with differing subsistence parameter values, and that those specific parameters depend on demographics. In particular, the optimization problem of a consumer unit was assumed to be

$$\max_{\{x_{1k}, x_{2k}, \ldots x_{nk}\}} U(x_{1k}, x_{2k}, \ldots) = \prod_{j=1}^{n} (x_{jk} - \gamma_j)^{\alpha_j}$$

$$s.t. \quad \sum_{j=1}^{n} p_{jk} x_{jk} \leq (=) \; + expend_{kt}$$

n	Variable representation	description
1	Energy	Energy expenditure
2	Food	Food expenditure
3	Cg	Consumer goods
4	Hous	Housing services
5	Dur	Durable services
6	Cs	Consumer services

The focs imply structural behavioral demand patterns

$$p_{1k} x_{1k} = p_{1k} \gamma_1 + \alpha_1 (m_k - p_{1k} \gamma_1 - p_{2k} \gamma_2 - \dots p_{n-1,k} \gamma_{n-1})$$

$$p_{2k} x_{2k} = p_{2k} \gamma_2 + \alpha_2 (m_k - p_{1k} \gamma_1 - p_{2k} \gamma_2 - \dots p_{n-1,k} \gamma_{n-1})$$

$$\vdots$$

$$p_{5k} x_{5k} = p_{5k} \gamma_5 + \alpha_5 (m_k - p_{1k} \gamma_1 - p_{2k} \gamma_2 - \dots p_{5k} \gamma_5)$$

Through normalization of one of our gammas (the subsistence parameter), one of our prices, and assuming summability across equations, I was able to dismiss the demand for good 6 (consumer services) from the estimation. The demand for that good can be calculated from the other estimated prarameters given the assumptions above.

Econometric Methodology

The econometric methodology for estimating parameters of the above FOCs (a "linear expenditure system") is slightly complicated due to the fact that the equations are not linear in parameters. The system of FOCs implied by our structural assumptions is nonlinear and simultaneous. A key attribute however allows the econometrician to still proceed. Since the demand random variables (x's) only directly depend on parameter values, I was able to estimate the parameters that ultimately specify the demand data through the method of "seemingly unrelated regressions." I construct an objective function consisting of the sum of the LHS (data) – RHS (data and parameters) and set forth to minimize the sum of the square of the terms. This first stage of the methodology returns consistent parameter estimates. Using these estimates and comparing implied residual variances for each foc eqn, I reweighted the terms in the objective function such that terms with smaller variance were weighted more so, as they represent more precise data. The minimizers of the reweighted objective function are now consistent, and more efficient. The first stage corresponds to Nonlinear least squares estimation (nlls) , and the second to Generalized nonlinear least squares (gnls).

There is reason to believe that the error terms may be normally distributed. If we believe our structural assumptions, then the size of the data and the central limit theorem suggest that I would be justified to assume that our errors are normally distributed around 0. Also, after calculating estimates with and without the assumption of normality on the error terms, and recognizing the small difference, I was further confident that imposing this condition on the error terms was not problematic. Normally I would prefer to assume as few moments

on the error terms as necessary, however given the above justifications along with the fact that Stata has an easy to access MLE estimator and therefore likelihood ratio statistics would be simple and easy to generate, I imposed the normality restriction and estimated parameters by maximum likelihood estimate technique through Stata.

As a last initial comment about the econometric methodology, in order to analyze nonlinear effects of age on demand, I separated the age regressor into differing age bands of 10 years and augmented each band with its own dummy variable coefficient. I was then able to recognize a unique linear relationship between age and demand within each age band.

Results and Findings

Data. There exists some idiosyncracies within the data (negative prices, aberrant values and so forth), however, given the immense volume of the data set, overall I felt confident using it without too much treatment and trust many of the calculations. As for specific treatment, I did eliminate a category (rented vehicles) from the model and therefore needed to remove the expenditure on that good from the total expenditure data value for each consumer unit.

Parameters. I generated estimates of parameters for a homogenous model (alphas=commodity exponents of utility, and gammas=subsistence levels of commodities). The alphas were all less than one as expected by the utility functional form. Also they were all near each other in value indicating a degree of complementarity between commodities. The gammas varied more so, and were all positive save the subsistence estimate on consumer goods. Economic interpretation suggests that this result implies that consumer units generally are over-endowed with respect goods (ie apparel) and to enjoy bare subsistence levels could afford to sell off some goods and apparel. Also noteworthy was the high value of gamma 6, the subsistence level on consumer services. Interpretation of this parameter suggests high dependence by consumer units on the service market within our economy (a far cry from self reliance or autonomy).

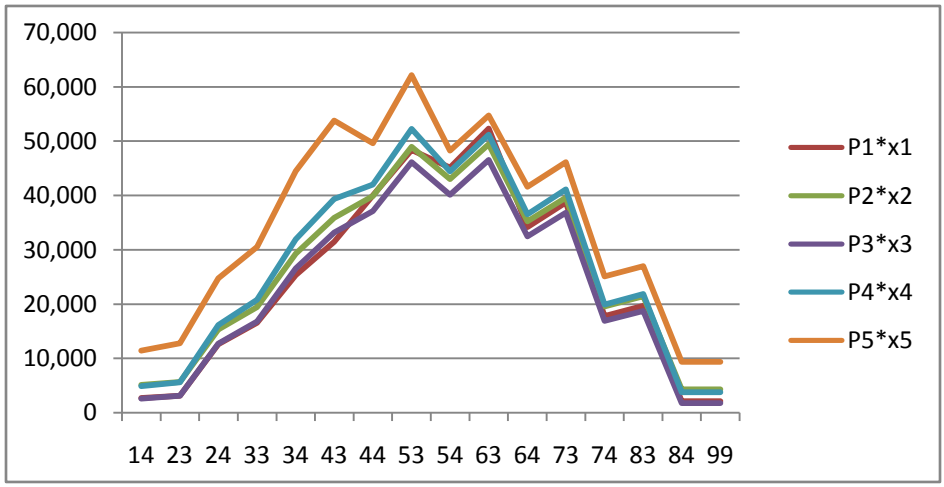
Next, inserting linear heterogeneity in demographics, I reestimated the model to yield alphas and now deltas (coefficients for demographic characteristics on subsistence levels). The alpha estimates did not change greatly. The biggest result of note was the improved "likelihood" estimate of the mle model by including heterogeneity. The LR test between homogeneity (all deltas=0 aside from the constant) and heterogeneity was highly significant (~45k>4.6=chi^s w 15 dof) implying the subsistence levels vary greatly across consumer units. Specifically, I chose to include as demographic characteristics consumer unit size, age of cu head, and urban rural residence type of the cu. As opposed to running LR tests for every possible combination of these characteristics, in the interest of time, I took the extremely high

LR stat to imply justification of inclusion of all of these characteristics and then I post-analyzed the demand dependence. In conclusion, as displayed in the graphs of demand relative to each of these demographics, age was the most significant source of heterogeneity, followed by size and then residence type, although the latter two were much smaller.

The estimates generated a consistent story for my price elasticity matrix. All elements were less than zero (representing that proportional increase in price result in proportional decreases in demand) except in the category of consumer goods. Consumer goods own price elasticity was negative implying the usual interpretation, but the cross price elasticities were positive. If prices for other goods increase, thus making a consumer unit less well off, they may will shift their expenditures to more necessities and goods that may provide more imputed income. I caution those looking at the price elasticity matrix to recognize that the values are calculated at mean levels because officially those elasticites are functions of the demographics and those demographics may have nonlinear effects (ie age). Therefore, that matrix is only one calculation that is subject to change for consumer units located away from mean demographic characteristics.

I present the same caution for the total expenditure elasticity vector. All values calculated at mean demographic levels. Consumer goods enjoy the largest proportional gain in expenditures with an increase in budget while energy was the smallest, but still positive. As a point of note, this is an analytical result that would not have been possible with more restrictive functional form of utility such as homothetic preferences.

The results for breaking age into different age bands reinforced my decision to do so. It was found that the resulting piecewise linear function of demand(age) for each commodity was highly nonlinear, and in fact exhibited a bell shape. Demand expenditures on all goods were higher when the age of the cu was in their working prime as opposed to at the age tails. The fact that all commodities exhibited this behavior indicates a lack of smoothing on behalf of the cu and that maybe 2-stage budgeting is not a bad assumption to impose on cu behavior.



Subsistence expenditures in general struck me as high when calculated at mean demographic values. And in particular, except for age, did not depend greatly on the demographic characteristics. Size and residence type effected most commodity expenditures positively and in the same fashion (not much effect). The graphs do not intend to be misleading and the vertical axis values are such that the effect can be seen. Consumer goods and services expenditures both decreased slightly with cu size for a cu at mean demographics. Also of note, expenditures on energy and consumer goods move together with respect to urban-rural residence type for a family at mean demographics, while food, housing, durables, and consumer services move together and opposite to energy and consumer goods.

Conclusion

The assumptions of the Linear Expenditure System and 2-stage budgeting contain strengths and weaknesses, and in general seemed to perform fairly well in the estimates of demand patterns with heterogeneity. The strengths include easy to use functional forms and estimable parameters. Weaknesses include restrictive assumptions on intertemporal substitutions and somewhat impractical notions of subsistence expenditures. Overall however, the model captured many of the salient characteristics of the data including the effects of heterogeneity along exogenous demographics.

The empirical methods of “seemingly unrelated regressions”, nonlinear least squares, generalized least squares in 2 stages, and maximum likelihood are powerful tools and being able to understand them, justify their use different settings, and interpret their results were integral to the analysis. The breadth of the data set and horsepower of stata combined to produce powerful estimates of the LES model. An econometrician could feel comfortable bringing the results of our “demand patterns with heterogeneity” analysis to those needing to make normative decisions.