

Problem Set 1

ECON 777-1: Empirical Industrial Organization

All homework submissions should be typed in LaTeX or a similarly suitable word processor. You may work in groups on this assignment, but you should complete your own write-up. You are welcome to use the programming language (e.g., R, Python, Matlab, Julia, Stata) of your choice to complete this assignment. Please include your computer code in your submission.

Background and Data Description

The ACA health insurance exchanges were established in every U.S. state in 2014 with the primary aims of (1) reducing the number of people without health insurance in the U.S. and (2) making health insurance more affordable for those ineligible for employer-provided health insurance or publicly-provided health insurance (e.g., Medicare and Medicaid). Exchange consumers can select a plan from one of the four actuarial value (AV) or “metal” tiers: (1) bronze (which has a 60% AV); (2) silver (70% AV); (3) gold (80% AV); and (4) platinum (90% AV). The actuarial value of a plan refers to the expected percentage of health care expenditures that the insurer will pay. In addition to plan AV, observable plan characteristics include whether the plan is a Health Maintenance Organization (HMO) and the firm who sells the plan. Eligible consumers can receive subsidies for purchasing a plan in the insurance exchanges. Consumers without insurance may be subject to a penalty. The penalty amount was phased in between 2014 and 2016. Starting in 2019, the penalty was eliminated.

You will be provided with a random sample of 2014-2019 enrollment data from the California ACA health insurance exchange (called Covered California). The California exchange is divided into 19 distinct rating areas which we will consider to be markets. Four large insurers – Anthem Blue Cross, Blue Shield, Health Net, and Kaiser – participate in the California exchange. There are also many small insurers, which have been combined for this analysis. In California (but not in most other states), all insurers are required to offer a plan in each metal tier with predetermined cost sharing levels (e.g., all silver plans have the same deductible, copays, coinsurance, etc.). Insurers can differentiate plans based on provider network type (e.g., HMO or PPO) and which providers are included in their networks.

Problem 1 - Model

- (a) Specify the consumer’s (indirect) utility function. Define every parameter and variable in your utility function. If your function includes vectors of variables, please define each element of the vector.
- (b) What is the outside option? What is the utility of the outside option?
- (c) Suppose policymakers are interested in how price sensitivity varies with age. Modify your utility function in (a) to allow price sensitivity to vary with age.
- (d) What variation is there in the data that can identify the price parameter?
- (e) Why might the premium be endogenous? What would be ideal instruments in this setting? Given the available data, what instruments can you construct?

- (f) Your utility equations do not capture cost sharing subsidies, which have the effect of increasing the actuarial value of plans in the silver tier. For the lowest-income exchange consumers the increase in AV is very substantial. How do you think the omission of cost sharing subsidies will bias your estimates of the premium and AV parameters?

Problem 2 - Estimation

With the exception of optional Part D, we will now estimate the model at the market-year level. For what follows, please compute market shares and the average premium paid **per person** for each plan in every market and year.

- (a) Estimate the parameters of the utility function in 1(a) as a logit and nested logit using the inversion approach of Berry (1994).
- (b) Suppose that the coefficients β_{ik} of the product characteristics AV and HMO are normally distributed with mean $\bar{\beta}_k$ and variance σ_k^2 . Estimate the utility and random coefficient parameters using the approach of BLP (1995). Compare your estimates to the logit and nested logit estimates in 2(a).
- (c) Use either the logit or nested logit to estimate the parameters of the utility function in 1(c). You can also try estimating separate models by age group. Discuss your results.
- (d) (Optional) Estimate the utility function as a logit at the household level using maximum likelihood. How do the estimated parameters compare to your estimates in 2(a)? (Note: there are packages in R or Stata to estimate a logit at the household level).

Problem 3 - Results

- (a) Compute the mean own-price elasticity of demand and mean cross-price elasticity of demand using the estimated parameters from each of the 3 models estimated in 2(a). How do the estimated elasticities vary across the 3 models?
- (b) Compute the mean own-price elasticity of demand and mean cross-price elasticity of demand by age group using the estimated parameters in 2(c). Do your estimated elasticities make sense?
- (c) Congress set the penalty for not having insurance to 0 starting in 2019. Compare enrollment before and after the policy change using your estimated model in 2(c). Discuss the effects on total enrollment in the insurance exchange, enrollment across the 4 metal tiers, enrollment by age group, and enrollment by firm. You can assume that insurers do not adjust premiums in response to the policy change.