Industrial Organization and Data Science

Instructor: Jacob LaRiviere, Director of Economics and Data Science at Amazon

Emails: [jlarivi1@uw.edu](mailto:jlarivi1@uw.edu), [lghhager@uw.edu](mailto:lghhager@uw.edu) (TA)

# Course Assignments & Reading

Course assignments should be printed and turned in at the start of class unless otherwise noted. Feel free to work in groups but everyone is required to turn in their own work with answers written in your own words. In both calculations and complex ideas, write down each step of logic used in reaching your conclusion. Keep in mind that in most cases a good answer is one precise sentence; quality is heavily favored over quantity. This will be graded on a full credit, half credit and no credit basis. All work must be typed

**Due November 15**

**Assignment to be turned in.** Please turn in typed out answers. Math with pen/pencil OK.

1. Consider a market in which the seller of a product knows there are two types of consumers, a high and a low preference type, which are indistinguishable from one another. The firm can produce along a product quality spectrum, such as with cars.
   1. If the firm decides to offer a low quality good, what risk do they run?
   2. What two options do they have to mitigate this risk?
   3. How much would the firm be willing to pay to identify each type of consumer and price discriminate accordingly?
2. Coding up Double ML.
   1. Use the Double ML algorithm to estimate the own price elasticity for Tropicana, Minute Maid and Dominick’s using all available lagged features in the ML models including feat as a predictor. You’ll estimate three separate regressions one for each brand for the final stage OLS residuals regression.
      1. I suggest using a random forest or XGBoost for P and Q.
      2. I also suggest using store and week fixed effects in your model.
      3. I also suggest using the interaction of important lagged price and quantity variables interacted with sociodemographic characteristics.
   2. Now do the same thing but estimate the full 3x3 elasticity matrix. Remember that for the OLS regression you’ll have residual log sales on the left and residual log price dom, residual log price MM and residual log price trop on the RHS and you’ll estimate three separate regressions one for each brand.