

ECON 6910
Applied Econometrics
Philip Shaw
Problem Set 7

Chapter 2 Exercises (Li and Racine (2007)):
2.4, 2.6

Chapter 12 Exercises (Li and Racine (2007)):
12.1, 12.2

Additional Problem:

For this problem consider a model that relates the number of cigarettes smoked per day ($cigs$) to the price of cigarettes ($cigpric$), age (age), and education in years ($educ$). The population model under consideration is given by $cigs = g(cigpric, age, educ) + u$. The dataset for this problem can be found in `smoke.csv`.

- a. Suppose a researcher specifies a linear model such that $cigs = \beta_0 + \beta_1 cigpric + \beta_2 age + \beta_3 educ + u$. Estimate the linear model and comment on the signs for each of the coefficients and statistical significance of each variable.
- b. Now suppose the researcher allows $g(\cdot)$ to be unspecified. Calculate the optimal bandwidths using least-squares-cross-validation. What do the bandwidths suggest for the relationship between each of the independent variables and the dependent variable?
- c. Now plot the relationship between number of cigarettes smoked per day and each of the independent variables. Include the asymptotic confidence bands when plotting the relationships. Do the functional relationships make sense?
- d. Using the wild bootstrap and the nonparametric consistent model specification test to test whether the linear model is correctly specified. What do the results of the specification test imply for the results from the linear model?