ECONOMICS 6002 CLASS 14
ROBUST AND NONPARAMETRIC ESTIMATION

1. Parametric vs. nonparametric estimation
   a. Parametric: complete specification of the estimation equation and the probability distribution of the random variables – e.g., maximum likelihood estimation
   b. Nonparametric estimation: no parameters to estimate - relationship entirely data-driven – e.g., kernel estimation
   c. Semi-parametric estimation: some parameters unspecified, for example the distribution of the disturbances – e.g., OLS and robust estimation
   d. Parametric estimators are more precise, and the conclusions that can be derived are more definitive, because the parametric specification provides additional information to sharpen the estimation process
   e. But they are less robust in the event of misspecification of the parametric structure

2. Kernel estimation
   a. Estimates a bivariate relationship $y = f(x) + \varepsilon$, where the functional relationship f( ) is completely unspecified.
   b. Kernel estimation is dependent on the specification of a **kernel function** and a **bandwidth size**.
   c. For any value of $x=x^*$, the value of $f(x^*)$ is established as a weighted average of the values of y in the bandwidth around x^*, where the weights depend on the distance of x from x^* in accordance with the kernel function.
   d. There are many kernel functions in use, and the results are not very sensitive to choice of function. The normal is a common choice.
   e. Choice of bandwidth is important.
      i. Too small a bandwidth leads to overfitting, which obscures the essential character of the relationship.
      ii. Too wide a bandwidth leads to oversmoothing, which obscures the nonlinearities in the relationship.
   f. Under certain conditions kernel estimation is consistent and asymptotically normal, but biased in small samples. The wider the bandwidth, the greater the bias but the smaller the variance.

3. Robust estimation
   a. OLS is BLUE if the disturbances are IID. But there may be non-linear estimators that are better.
   b. Because OLS minimizes the sum of **squares** of the residuals, a heavy penalty is imposed on large residuals. In small samples, outliers can have a disproportionate effect on the estimates.
   c. If the disturbances are normally distributed, OLS is best unbiased. But in economic applications, disturbances are often **leptokurtotic** - the distribution has greater mass in the tails (i.e., more outliers) than is predicted by the normal distribution.
   d. 'Robust' estimators do not weight outliers as heavily as OLS does, and so are more robust in small samples.
   e. There are many robust estimators. The simplest is that which minimizes the sum of the **absolute value** of the residuals (LAD - Least Absolute Deviation).
   f. Robust estimators are non-linear, and computation is typically complex.