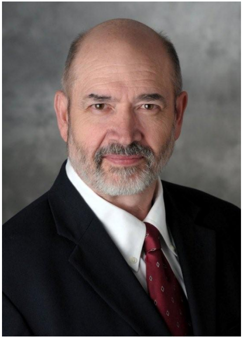
Transformations and Misspecification of Econometric Models



Dedicated to the Memory of Hal White (1950-2012)

Misspecification of Functional Form

One of Hal's earliest papers:

> *White, Halbert, (1980) "Using Least Squares to Approximate Unknown Regression Functions," International Economic Review, 21, 149-170.*

deals with misspecification of the functional form of econometric models. Let's begin a consideration of this topic with the following simple example. Suppose

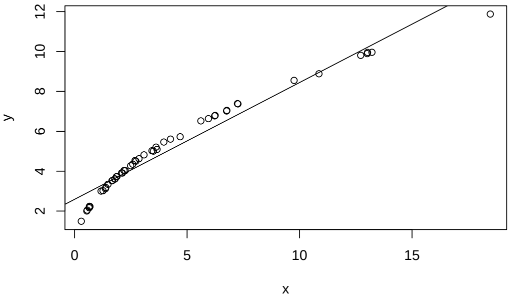
$$\log y_i = \alpha + \beta \log x_i + u_i$$

but unaware of this convenient formulation we instead estimate

$$y_i = a + b x_i + v_i.$$

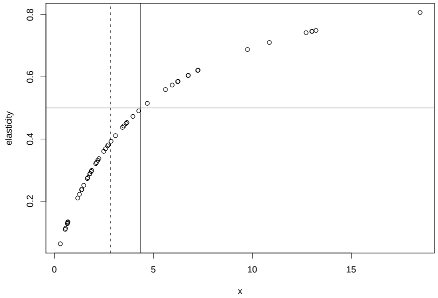
What relationship does $(\hat{a}, \hat{b})$ bear to (α, β) in the original model and can we hope to say anything reasonable having made this initial specification error?

To Log or Not to Log? That is the Question



A linear fit to a log-linear model: The figure illustrates 50 observations from a log-linear model and a superimposed least-squares linear fit of the observations. Note that the fit provides a rough estimate of the tangent of the curve near the "center" of the x's, but cannot be considered very reliable unless the range of the x's is quite restricted.

Elasticities Vary When They Shouldn't



Linear fit to a log-linear model: The points in the figure represent elasticities implied by the fitted *linear* model at each of the observed x's. The horizontal line at $\beta = .5$ represents the true, constant elasticity for the model, and the two vertical lines indicate the mean (solid) and geometric mean (dotted) of the x's.

What Does Least Squares Do?

Several answers may be offered:

1. Least squares regression of y on X projects y onto the column space of X.

What Does Least Squares Do?

Several answers may be offered:

1. Least squares regression of y on X projects y onto the column space of X.
2. Least squares regression of y on X estimates the conditional expectation function $\mathcal{E}(Y|X)$.

What Does Least Squares Do?

Several answers may be offered:

1. Least squares regression of y on X projects y onto the column space of X.
2. Least squares regression of y on X estimates the conditional expectation function $\mathcal{E}(Y|X)$.
3. When the conditional expectation function is nonlinear in X, the LSE estimates the best approximation to $\mathcal{E}(Y|X)$ among the linear functions of X.

What Does Least Squares Do?

Several answers may be offered:

1. Least squares regression of y on X projects y onto the column space of X.
2. Least squares regression of y on X estimates the conditional expectation function $\mathcal{E}(Y|X)$.
3. When the conditional expectation function is nonlinear in X, the LSE estimates the best approximation to $\mathcal{E}(Y|X)$ among the linear functions of X.
4. Suppose the x_i's are generated randomly from some distribution, F, and that $E(y|x) = g(x)$, then $(\hat{a}, \hat{b})$ solves $\min \mathcal{E}_x(g(x) - a - bx)^2$, i.e., $\hat{a} + \hat{b}x$ is the best linear approximation to $g(x)$ in quadratic mean.

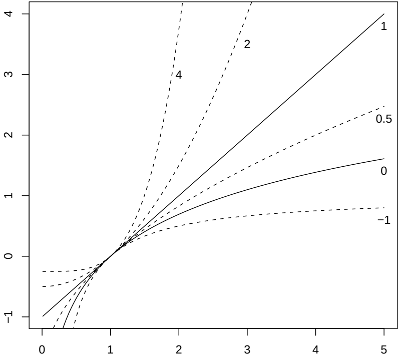
Box and Cox

Comic Opera by F.C. Burnand and Arthur Sullivan

The classical Box-Cox (1964) approach to dealing with choice of functional form in econometric models involves the family of power transformations

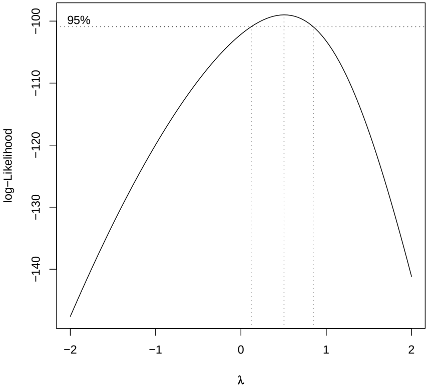
$$h(x,\lambda) = \begin{cases} \frac{x^\lambda - 1}{\lambda} & \lambda \neq 0 \\ \log x & \lambda = 0 \end{cases}$$

Some Box-Cox Transformations



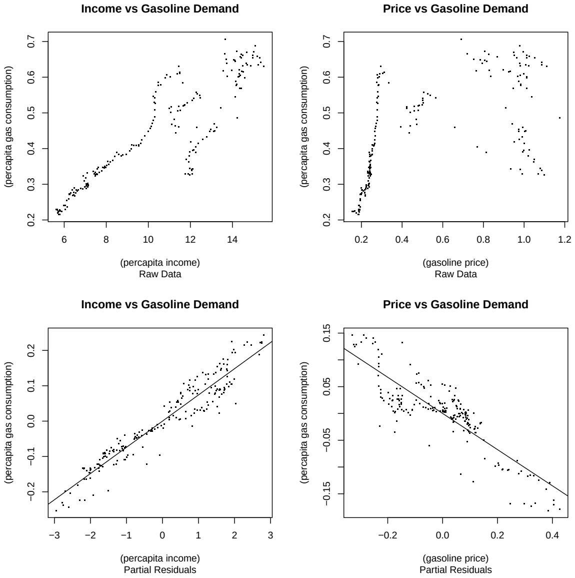
The Box-Cox Power Transformations: The Figure illustrates 6 versions of the Box-Cox Power family of transformations. Note that the log transformation fits nicely into the family with $\lambda = 0$.

Likelihood Ratio Inference About λ

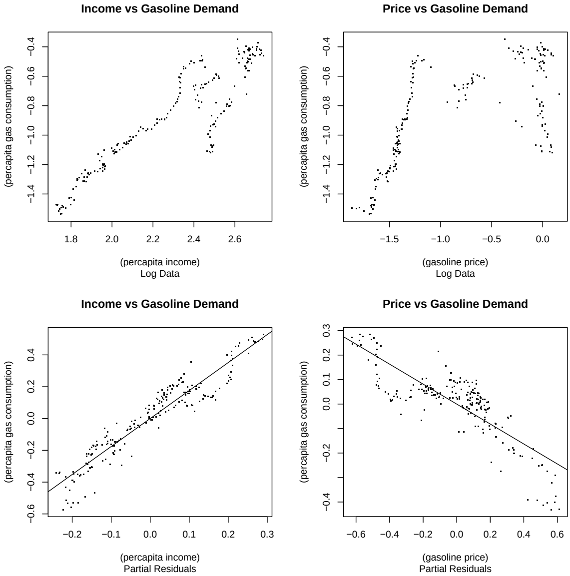


The Box-Cox Power Transformation: The Figure illustrates the profile log likelihood for a simple bivariate linear model, the confidence interval indicated for λ is based on the asymptotic theory of the likelihood ratio statistic.

Partial Residual Plots Unlogged



Partial Residual Plots Logged



Visualizing Multivariate Regression

When there are several explanatory variables, aka covariates, aka regressors, it is usual to have some way to visualize their separate effects. Consider the model

$$y = X\beta + z\gamma + u$$

and the least squares fitted values,

$$\hat{y} = X\hat{\beta} + z\hat{\gamma}.$$

Theorem (Gauss-Frisch-Waugh) $\hat{\gamma} = (z^\top M_X z)^{-1} z^\top M_X y$, where $M_X = I - P_X$, and $P_X = X(X^\top X)^{-1} X^\top$

Visualizing Multivariate Regression

Theorem (Gauss-Frisch-Waugh) $\hat{\gamma} = (z^\top M_X z)^{-1} z^\top M_X y$.

Proof Write

$$z^\top M_X \hat{y} = z^\top M_X X \hat{\beta} + z^\top M_X z \hat{\gamma}$$

but $M_X X = 0$, so it only remains to show that

$$z^\top M_X P_Z y = z^\top M_X y$$

where $Z = [X \vdots z]$.

Visualizing Multivariate Regression

Theorem (Gauss-Frisch-Waugh) $\hat{\gamma} = (z^\top M_X z)^{-1} z^\top M_X y$.

Proof Write

$$z^\top M_X \hat{y} = z^\top M_X X \hat{\beta} + z^\top M_X z \hat{\gamma}$$

but $M_X X = 0$, so it only remains to show that

$$z^\top M_X P_Z y = z^\top M_X y$$

where $Z = [X \vdots z]$. Note that

$$M_X P_Z = (I - P_X) P_Z = P_Z - P_X P_Z = P_Z - P_X.$$

Visualizing Multivariate Regression

Theorem (Gauss-Frisch-Waugh) $\hat{\gamma} = (z^\top M_X z)^{-1} z^\top M_X y$.

Proof Write

$$z^\top M_X \hat{y} = z^\top M_X X \hat{\beta} + z^\top M_X z \hat{\gamma}$$

but $M_X X = 0$, so it only remains to show that

$$z^\top M_X P_Z y = z^\top M_X y$$

where $Z = [X \vdots z]$. Note that

$$M_X P_Z = (I - P_X) P_Z = P_Z - P_X P_Z = P_Z - P_X.$$

In the last step note that $X^\top P_Z = (P_Z X)^\top = X^\top$, so $P_X P_Z = P_X$.

Visualizing Multivariate Regression

Theorem (Gauss-Frisch-Waugh) $\hat{\gamma} = (z^\top M_X z)^{-1} z^\top M_X y$.

Proof Write

$$z^\top M_X \hat{y} = z^\top M_X X \hat{\beta} + z^\top M_X z \hat{\gamma}$$

but $M_X X = 0$, so it only remains to show that

$$z^\top M_X P_Z y = z^\top M_X y$$

where $Z = [X \vdots z]$. Note that

$$M_X P_Z = (I - P_X) P_Z = P_Z - P_X P_Z = P_Z - P_X.$$

In the last step note that $X^\top P_Z = (P_Z X)^\top = X^\top$, so $P_X P_Z = P_X$. Finally, we can compute,

$$z^\top M_X P_Z y = z^\top (P_Z - P_X) y = z^\top y - z^\top P_X y = z^\top (I - P_X) y = z^\top M_X y.$$

The Barro Plot