effects when the CEF is linear. The second property tells us that we can expect regression estimates of a treatment effect to be close to those we'd get by matching on covariates and then averaging within-cell treatment-control differences, even if the CEF isn't linear.

Figure 2.1 documents the manner in which regression approximates the nonlinear CEF of log wages conditional on schooling. Although the CEF bounces around the regression line, this line captures the strong positive relationship between schooling and wages. Moreover, the regression slope is close to $E\{E[Y_i|X_i] - E[Y_i|X_i - 1]\}$; that is, the regression slope also comes close to the expected effect of a one-unit change in X_i on $E[Y_i|X_i]$.[16]

Bivariate Regression and Covariance

Regression is closely related to the statistical concept of *covariance*. The covariance between two variables, X_i and Y_i, is defined as

$$C(X_i, Y_i) = E\big[\big(X_i - E[X_i]\big)\big(Y_i - E[Y_i]\big)\big].$$

Covariance has three important properties:

(i) The covariance of a variable with itself is its variance; $C(X_i, X_i) = \sigma_X^2$.

(ii) If the expectation of either X_i or Y_i is 0, the covariance between them is the expectation of their product; $C(X_i, Y_i) = E[X_iY_i]$.

(iii) The covariance between linear functions of variables X_i and Y_i —written $W_i = a + bX_i$ and $Z_i = c + dY_i$ for constants a, b, c, d—is given by

$$C(W_i, Z_i) = bdC(X_i, Y_i).$$

The intimate connection between regression and covariance can be seen in a *bivariate regression model,* that is, a regression with one regressor, X_i, plus an intercept.[17] The bivariate regression slope and