



**FIGURE 3.18.** The same data set shown in Figure 3.17 is investigated further. Left: The blue dashed line is the least squares fit to the data. Since  $f(X)$  is in fact linear (displayed as the black line), the least squares regression line provides a very good estimate of  $f(X)$ . Right: The dashed horizontal line represents the least squares test set MSE, while the green solid line corresponds to the MSE for KNN as a function of  $1/K$  (on the log scale). Linear regression achieves a lower test MSE than does KNN regression, since  $f(X)$  is in fact linear. For KNN regression, the best results occur with a very large value of  $K$ , corresponding to a small value of  $1/K$ .