

Model Adequacy Checking – Part IV

Johns Hopkins Engineering

625.461 Statistical Models and Regression

Module 7 – Lecture 7E



Detection and Treatment of Outliers

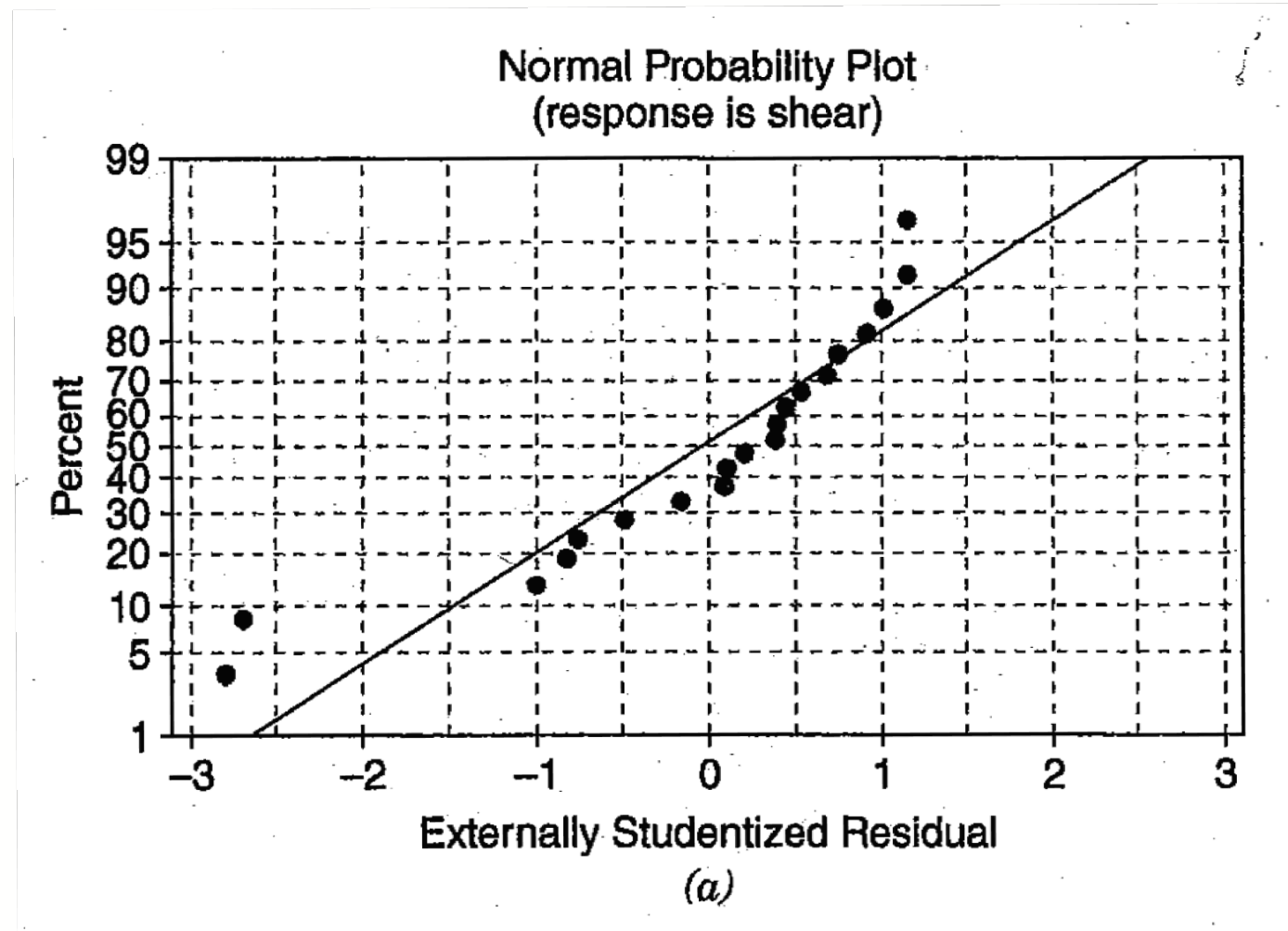
Residual plots and normal probability plots are helpful in identifying outliers.

The effect of outliers on the regression model may be easily checked by dropping these points and refitting the regression equation. If the statistics, t or F , R^2 , and the residual SS are not sensitive to the outliers, we would be happy.

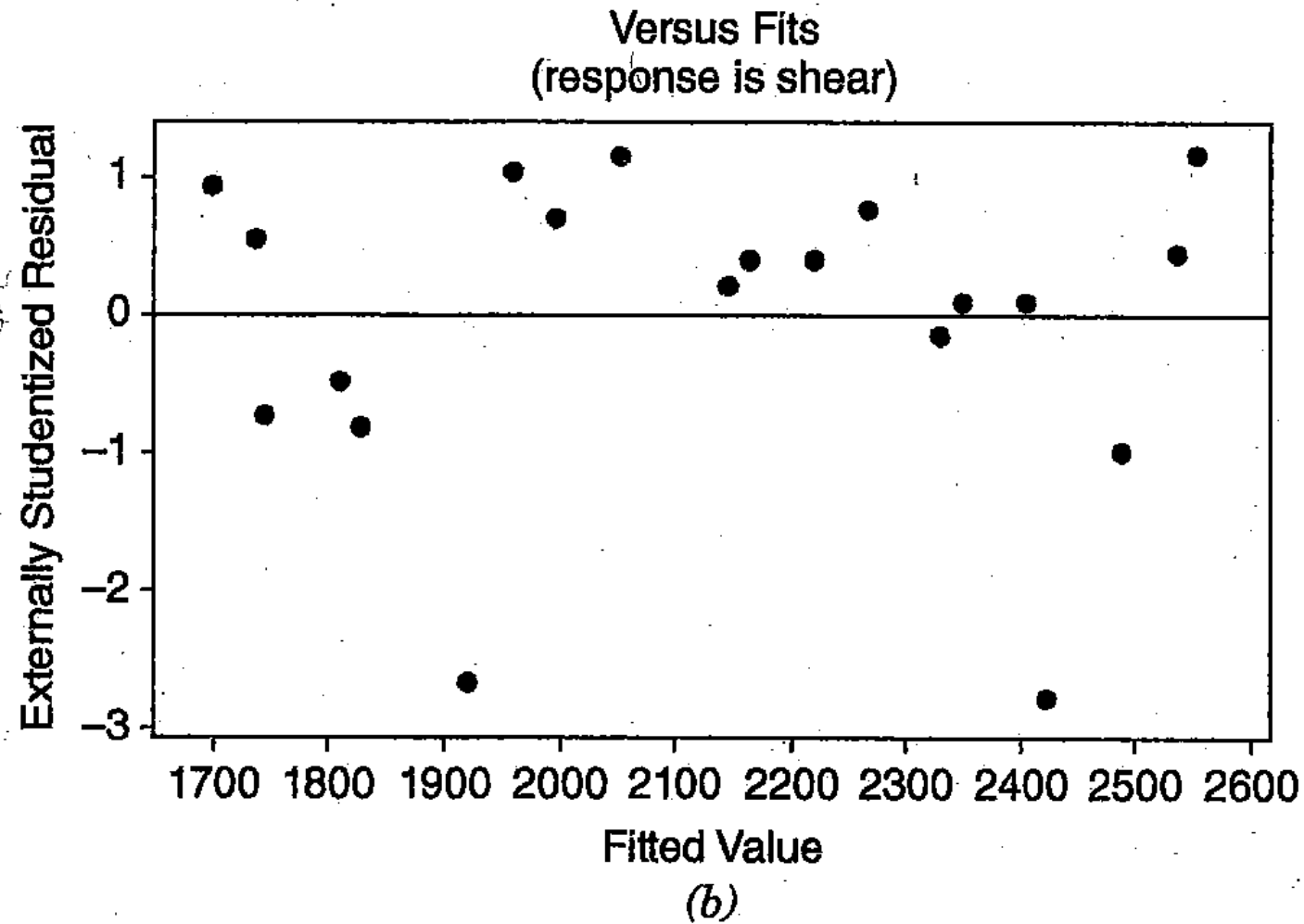
Detection and Treatment of Outliers

But, the regression relationship to be estimated should be based on all of the observations, unless **there is strong nonstatistical evidence that the outliers are “bad” values and should be discarded**

The Rocket Propellant Data (Ex. 4.7, page 153 of Textbook)



The Rocket Propellant Data (Ex. 4.7, page 153 of Textbook)



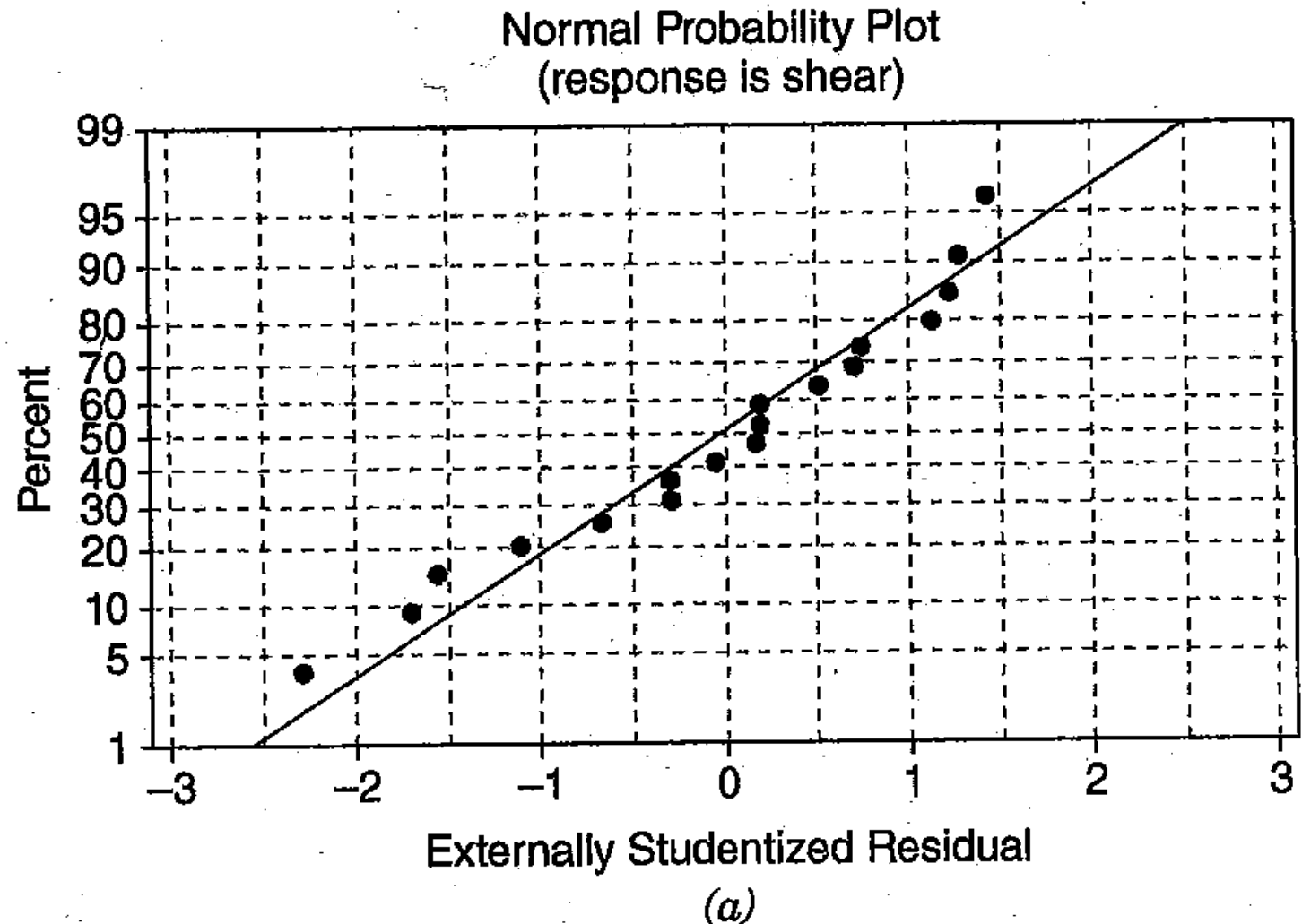
The Rocket Propellant Data (Ex. 4.7, page 153 of Textbook)

	Observations 5 and 6 IN	Observations 5 and 6 OUT
$\hat{\beta}_0$	2627.82	2658.97
$\hat{\beta}_1$	-37.15	-37.69
R^2	0.9018	0.9578
MS_{Res}	9244.59	3964.63
$se(\hat{\beta}_1)$	2.89	1.98

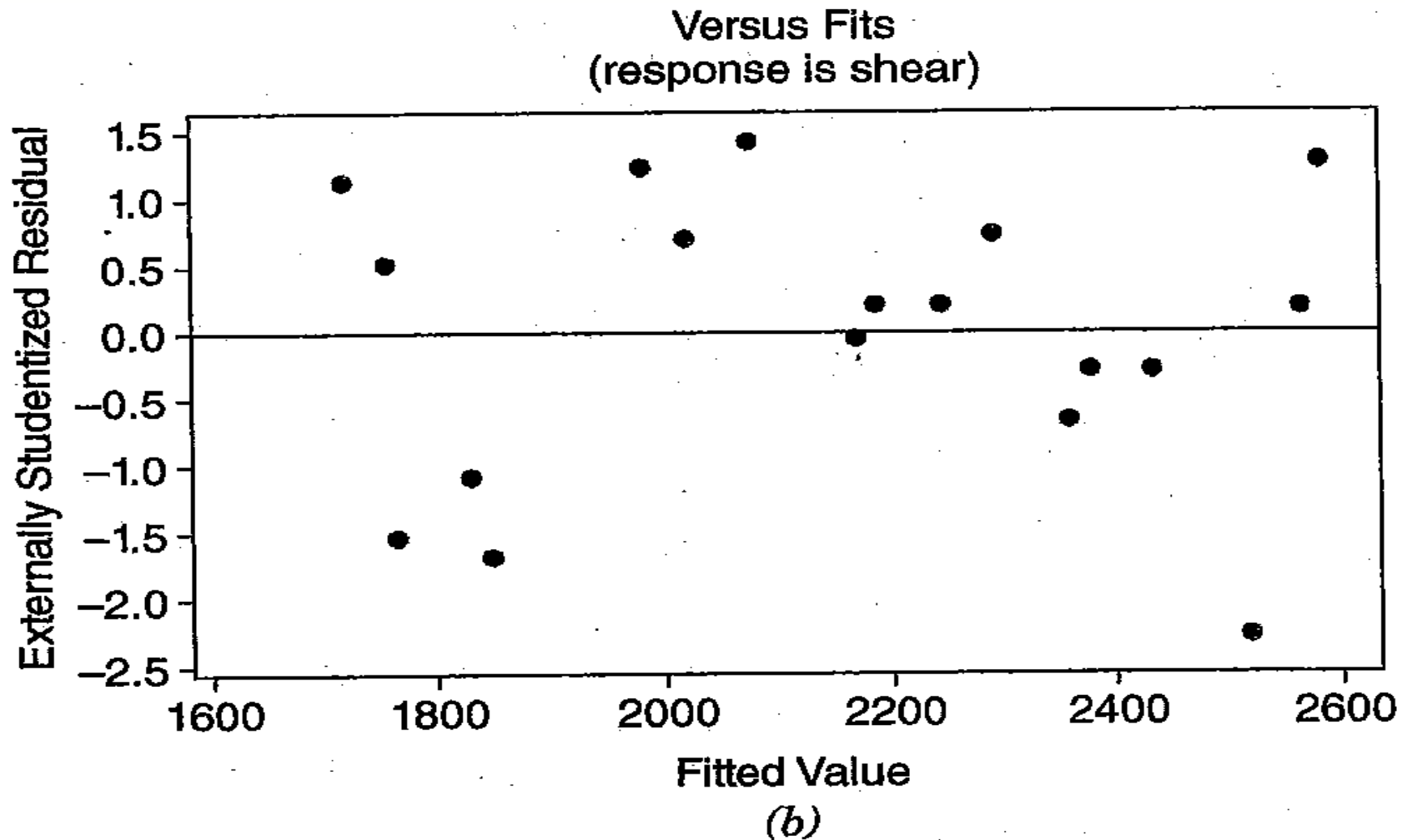
Deleting points 5 and 6 has almost no effect on the estimates of the regression coefficients. But, there has been a dramatic reduction in MS_{Res} , a moderate increase in R^2 , and approximately one-third reduction in the standard error of $\hat{\beta}_1$.

The Rocket Propellant Data (Ex. 4.7, page 153 of Textbook)

By deleting
points 5 and 6:



The Rocket Propellant Data (Ex. 4.7, page 153 of Textbook)



The Rocket Propellant Data (Ex. 4.7, page 153 of Textbook)

These two plots do not indicate any serious departures from assumptions.

Further examination of points 5 and 6 fails to reveal any reason for the unusually low propellant shear strengths obtained.

We should not discard these two points. Including them does not seriously limit the use of the model.



JOHNS HOPKINS
WHITING SCHOOL
of ENGINEERING