

Case selection

Lukas Lehotsky & Petr Ocelik

Should offer more
insight than cross-case

Case selection

Cross-case

- Sample selection
 - Random
 - Based on knowledge about population
- Sampling and analysis sequential tasks
- Large sample = easy elaboration of external validity
- Internal validity more concerning

Case study

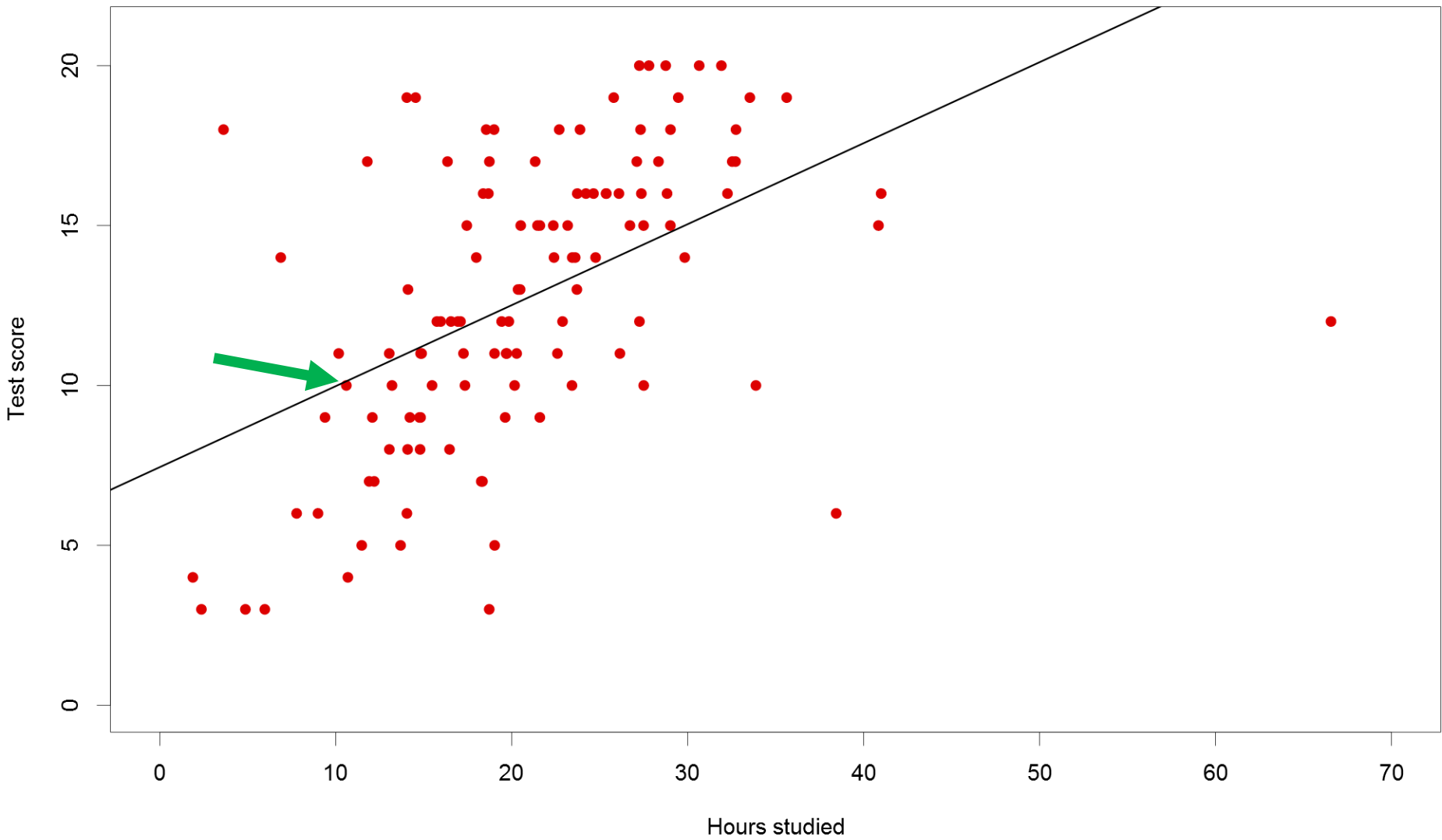
- Specific case selection prone to bias (cherry-picking)
- Sampling and analysis not separate from each other
- “Casing” – what is this case a case of?
- External validity more concerning

Case selection

- Does case study provide any advantage?
- What is the population? What is this case a case of?
- Am I interested in particular case? Why? What is it that I want to study?
- Do I want to test or build theory?

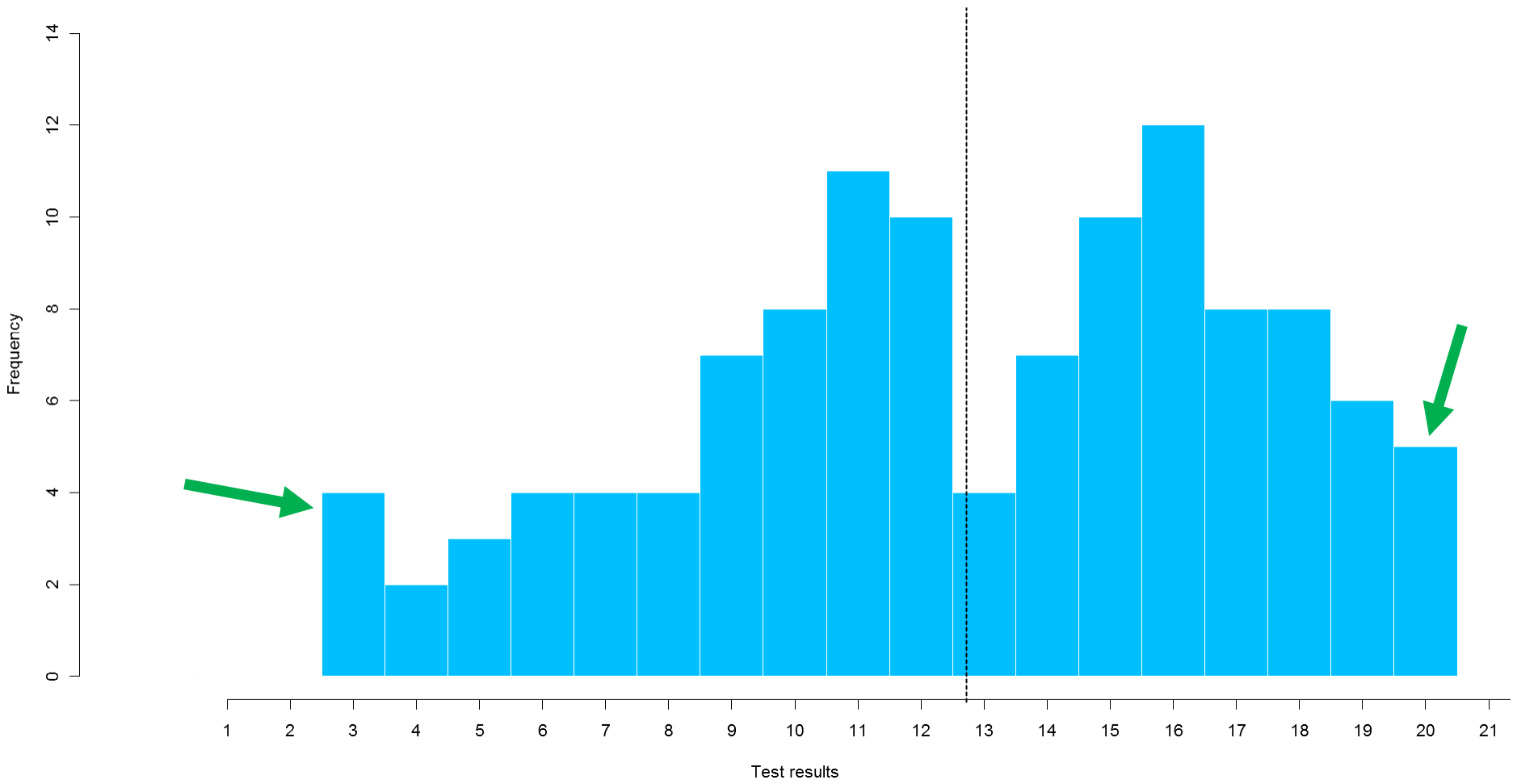
Normal (typical) case

- Represents typical relation between variables
- **Least residuals** – closest to the linear model prediction – outlier
- Representative by nature
- Theory testing



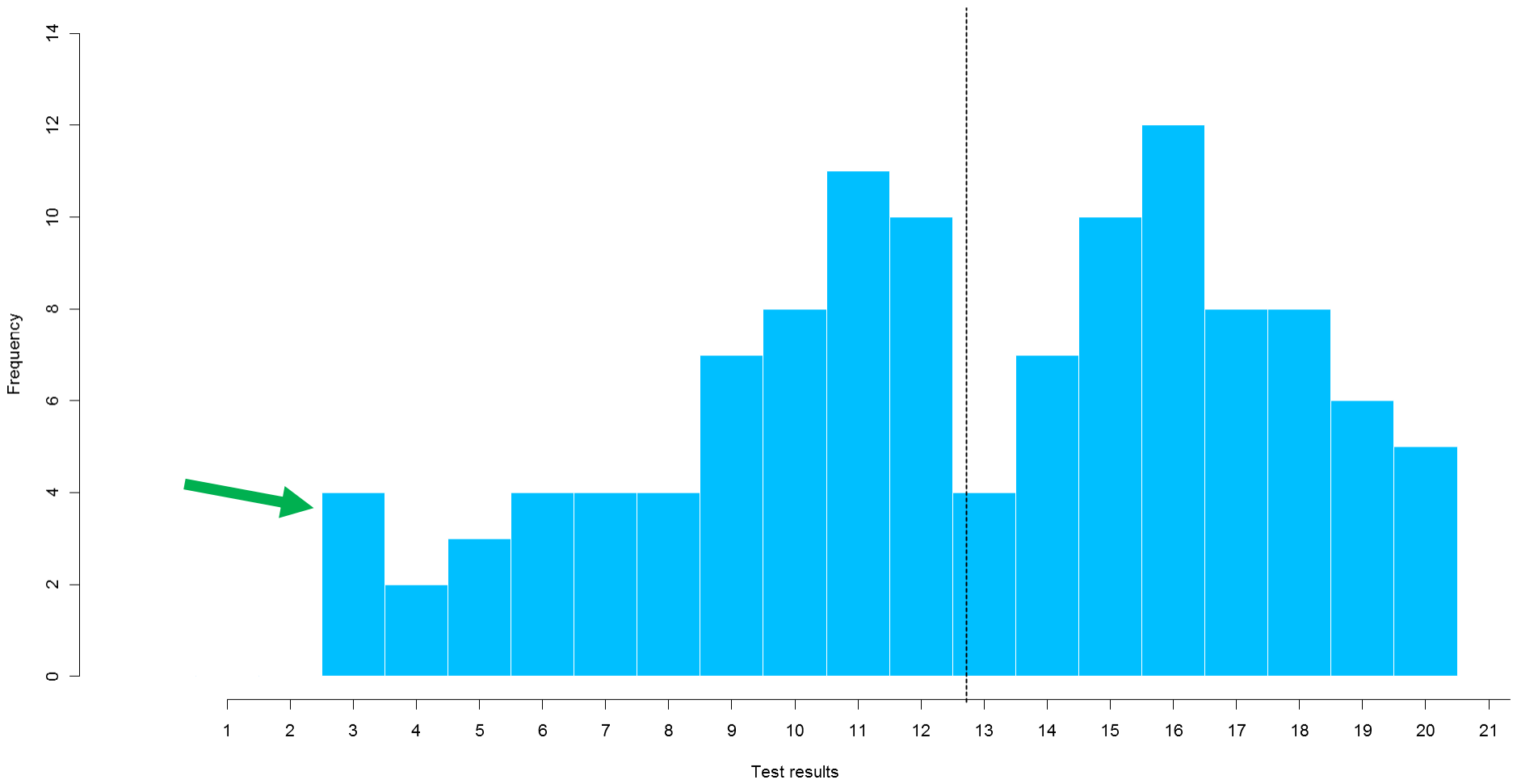
Diverse cases

- Two or more cases
- **Full scope of variation** of the variable/variables
- Standard-deviation
- Theory testing/building
- Represents full spectrum of variation (but not distribution)



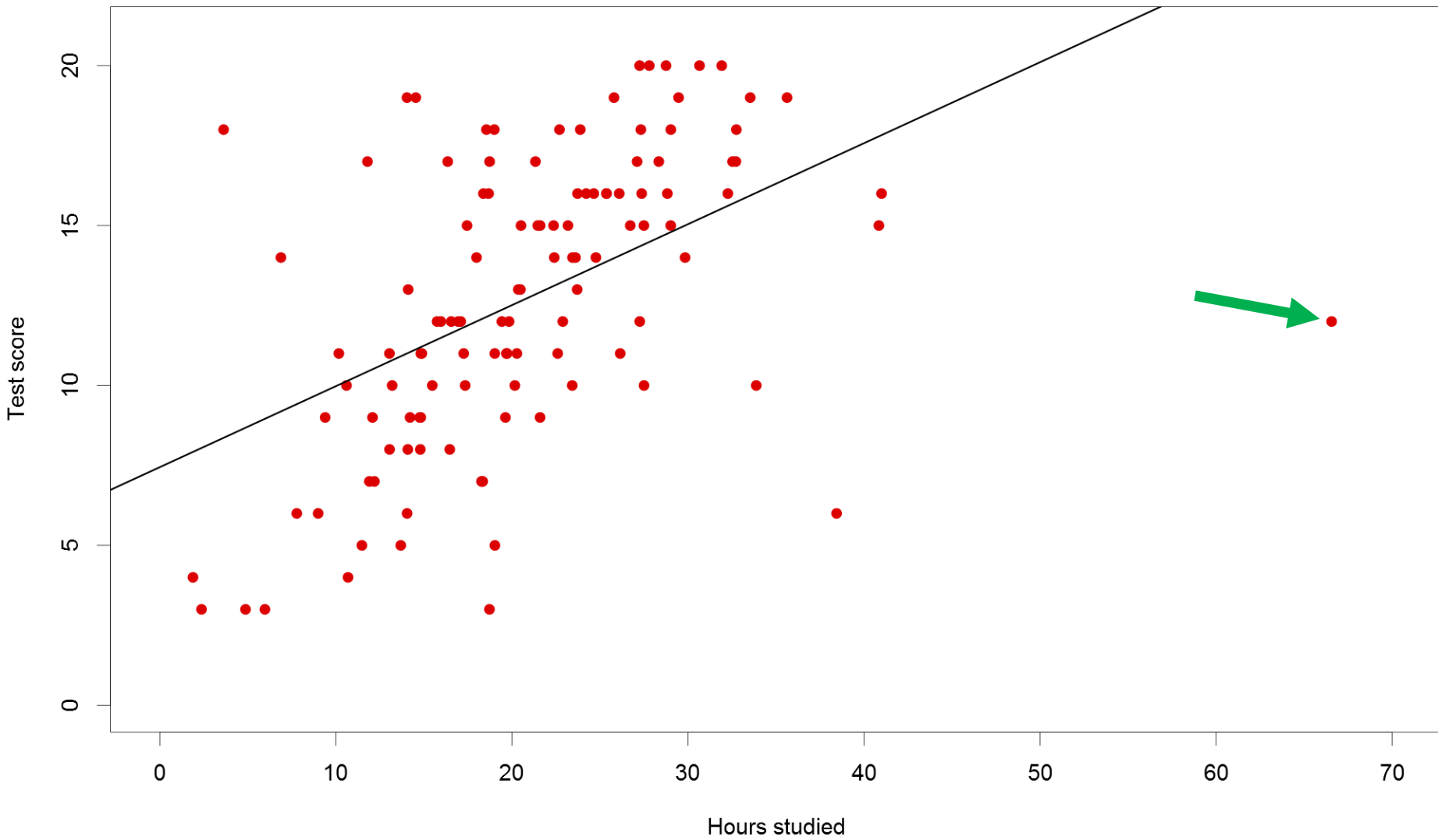
Extreme case

- One or more cases
- Extreme value of **variable/variables**
- Many standard deviations from the mean
- Theory building
- Representative in comparison with larger population



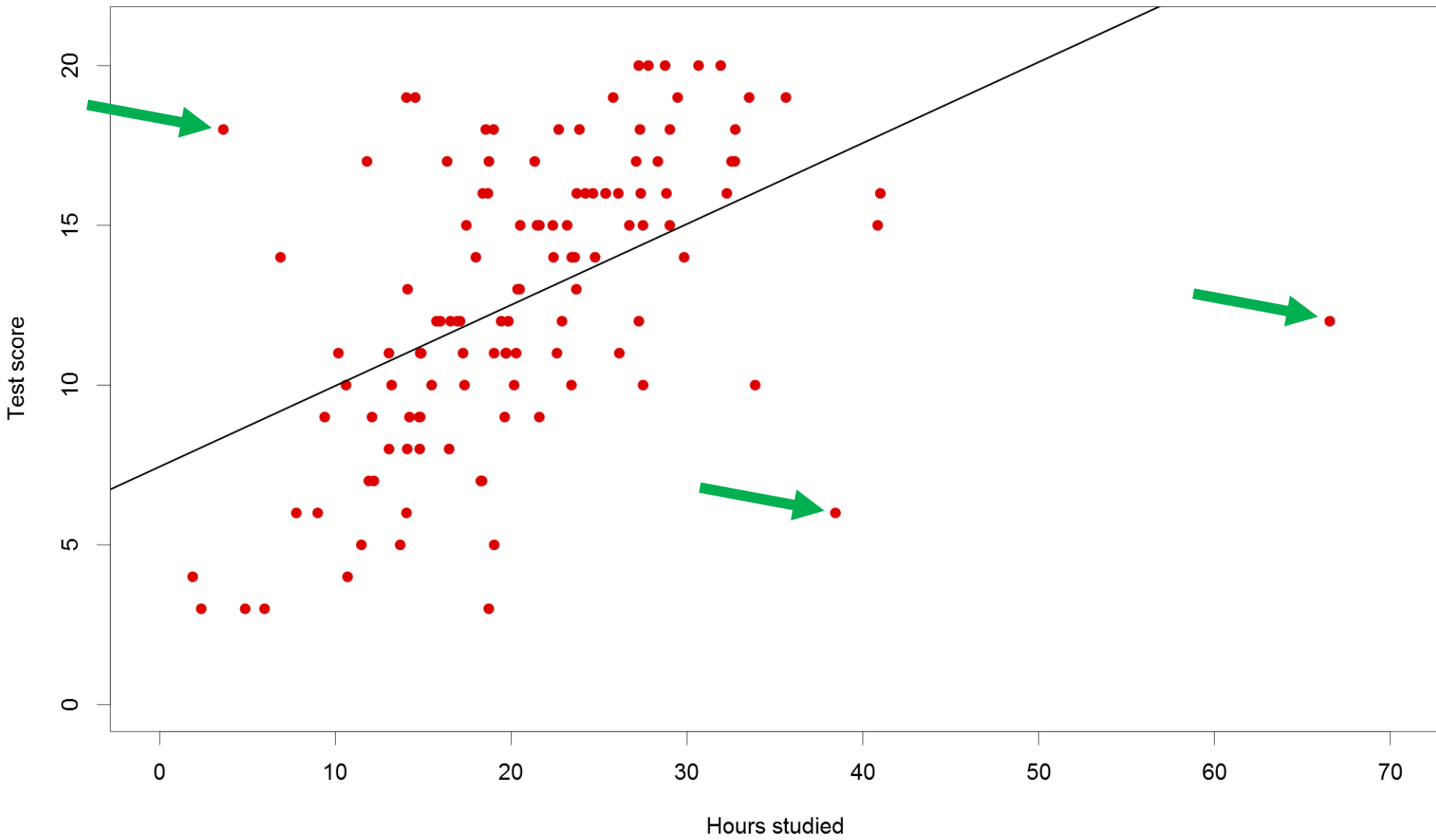
Deviant case

- One or more cases
- Outlier – the most distant case from the **linear model prediction** – highest residual
- Theory building, modification
- Unrepresentative by nature



Influential case

- One or more cases
- Similar to deviant case – seemingly disproves theory and has **high leverage – influences the model**
- Special configuration of variables – exception which (if explained) strengthens theory
- Theory testing, modification
- Representativeness irrelevant



Crucial case

- One or more cases
- Case which must fit the theory – proves validity of theory
 - Least-likely
 - Predicted to contradict the hypothesis, but does not
 - Confirms the theory
 - Most-likely
 - Predicted to support the hypothesis, but does not
 - Disconfirms the theory
 - Most-difficult test for an argument
- Theory testing
- Selecting on theoretical grounds

Pathway case

- One or more cases
- Case, where value of dependent variable is most likely caused by independent variable – apparent impact of indep. X on dependent Y
- Cross-tabulation or analysis of residuals (residuals of full model vs. model without causal variable)
- Theory testing – especially causal mechanisms

Most similar cases

- Two or more cases
- Comparative design
- All variables except those of interest are similar/same
- Selection from cross-case based on values of variables
- Theory testing and theory building

Most similar cases

Case	V1	V2	V3	V4	V5
Case 1	0	+	+	+	-
Case 2	+	+	+	-	-

Most different cases

- Two or more cases
- Comparative design
- All variables except those of interest are different
- Selection from cross-case based on values of variables
- Theory testing and theory building

Most diverse cases

Case	V1	V2	V3	V4	V5
Case 1	+	-	0	-	+
Case 2	+	+	+	-	-

Comparability

- Cases must be comparable for a comparative CS to be valid.
- Achieving comparability
 - Synchronical/spatial - more units in same time period
 - Diachronical - same unit in more time periods

Overview

Case type	Theory		#	Case selection
	Test	Build		
Normal	+		1	Min. ε of linear regression $Y=aX+bZ+\varepsilon$
Diverse	+	+	2	Max. range of variance of $X, Y, X/Y, \dots$
Extreme		+	1	Max. variance of Y
Deviant		+	1	Max. ε of linear regression $Y=aX+bZ+\varepsilon$
Crucial	+		1	ML predicts H1, proves to be H0 LL predicts H0, proves to be H1
Influential	+		1	Max. influence of a case on slope a of $Y=aX+bZ+\varepsilon$
Pathway	+		1	X_1 and not X_2 likely to have caused Y
Most similar	+	+	2	Cases similar on var. other than X_1/Y
Most different	+	+	2	Cases different on var. other than X_1/Y

