MacroEconometric Forecasting

Topic:

Types of experimental designs

Presented by Munisa Yashnarbekova

Basic principles

1. Formulate question/goal in advance
2. Comparison/control
3. Replication
4. Randomization
5. Stratification (aka blocking)
6. Factorial experiments

Example

Question: Does salted drinking water affect blood pressure (BP) in mice?

Experiment:

1. Provide a mouse with water containing 1% NaCl.
2. Wait 14 days.
3. Measure BP.

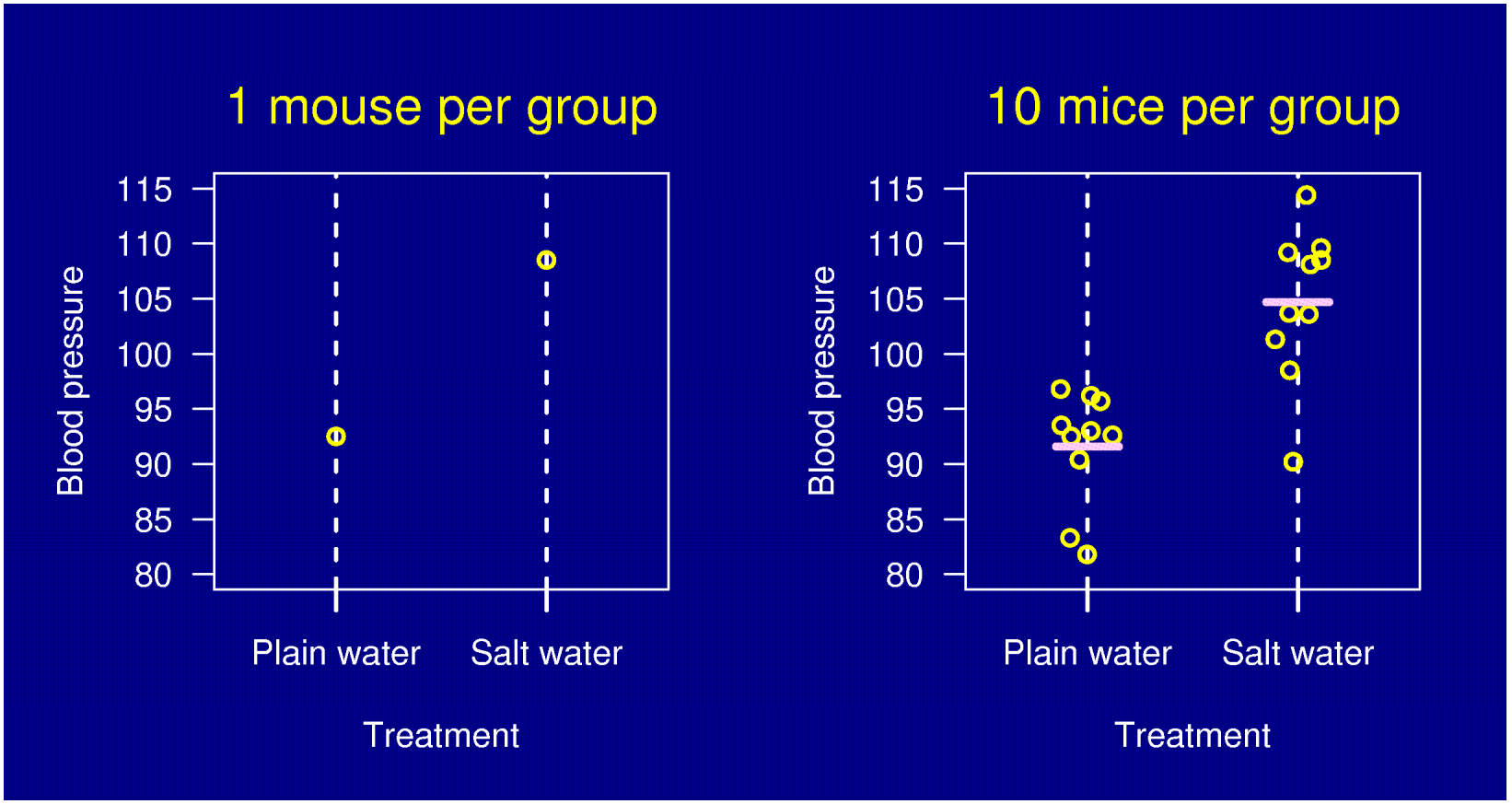
Comparison/control

Good experiments are comparative.

- Compare BP in mice fed salt water to BP in mice fed plain water.
- Compare BP in strain A mice fed salt water to BP in strain B mice fed salt water.

Ideally, the experimental group is compared to concurrent controls (rather than to historical controls).

Replication



Why replicate?

- Reduce the effect of uncontrolled variation (i.e., increase precision).
- Quantify uncertainty.

A related point:

An estimate is of no value without some statement of the uncertainty in the estimate.

Randomization

Experimental subjects (“units”) should be assigned to treatment groups at random.

At random does not mean haphazardly.

One needs to explicitly randomize using

- A computer, or
- Coins, dice or cards.

Why randomize?

- Avoid bias.
  - For example: the first six mice you grab may have intrinsicly higher BP.
- Control the role of chance.
  - Randomization allows the later use of probability theory, and so gives a solid foundation for statistical analysis.

Stratification

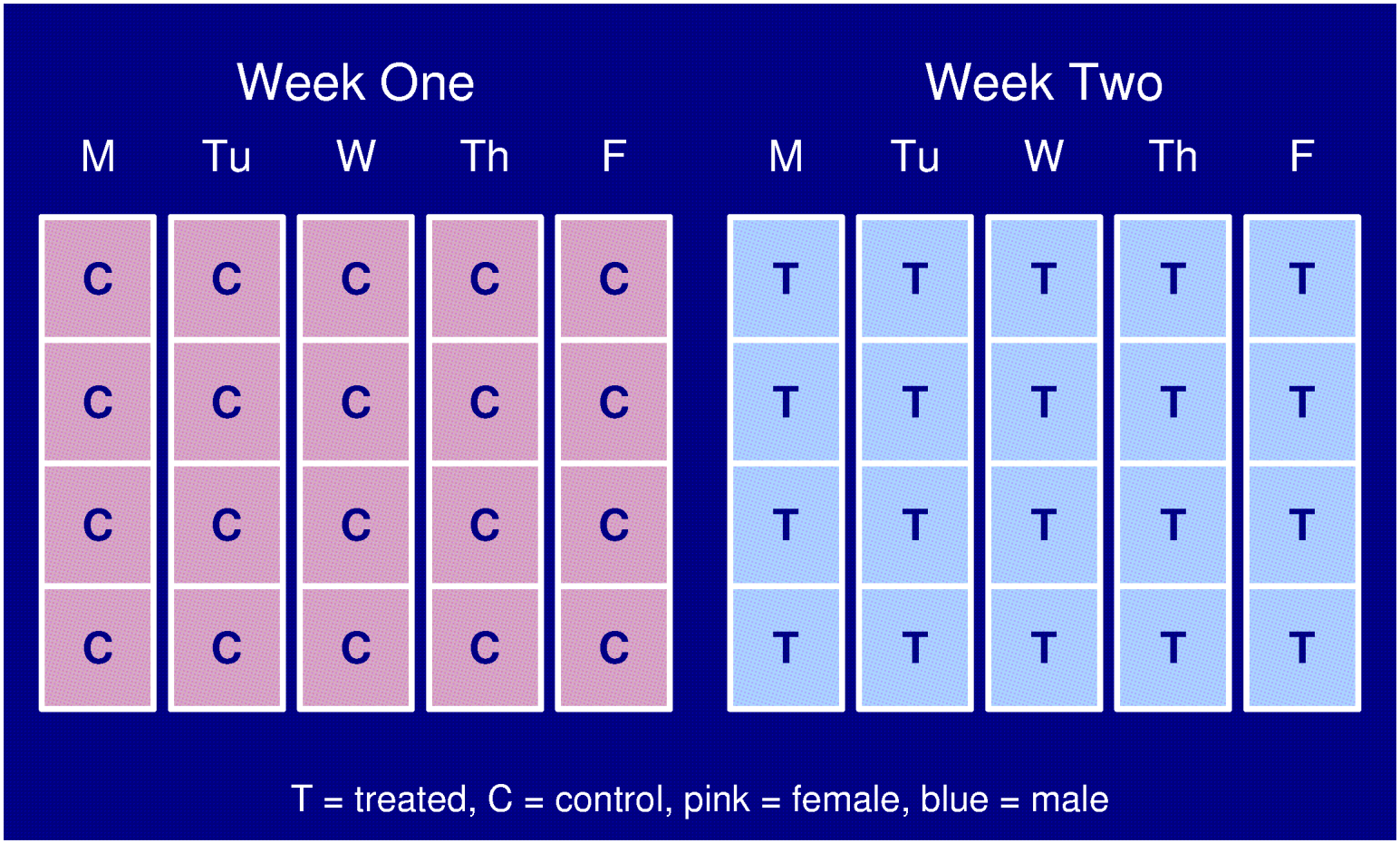
- Suppose that some BP measurements will be made in the morning and some in the afternoon.
- If you anticipate a difference between morning and afternoon measurements:
  - Ensure that within each period, there are equal numbers of subjects in each treatment group.
  - Take account of the difference between periods in your analysis.
- This is sometimes called “blocking”.

Example

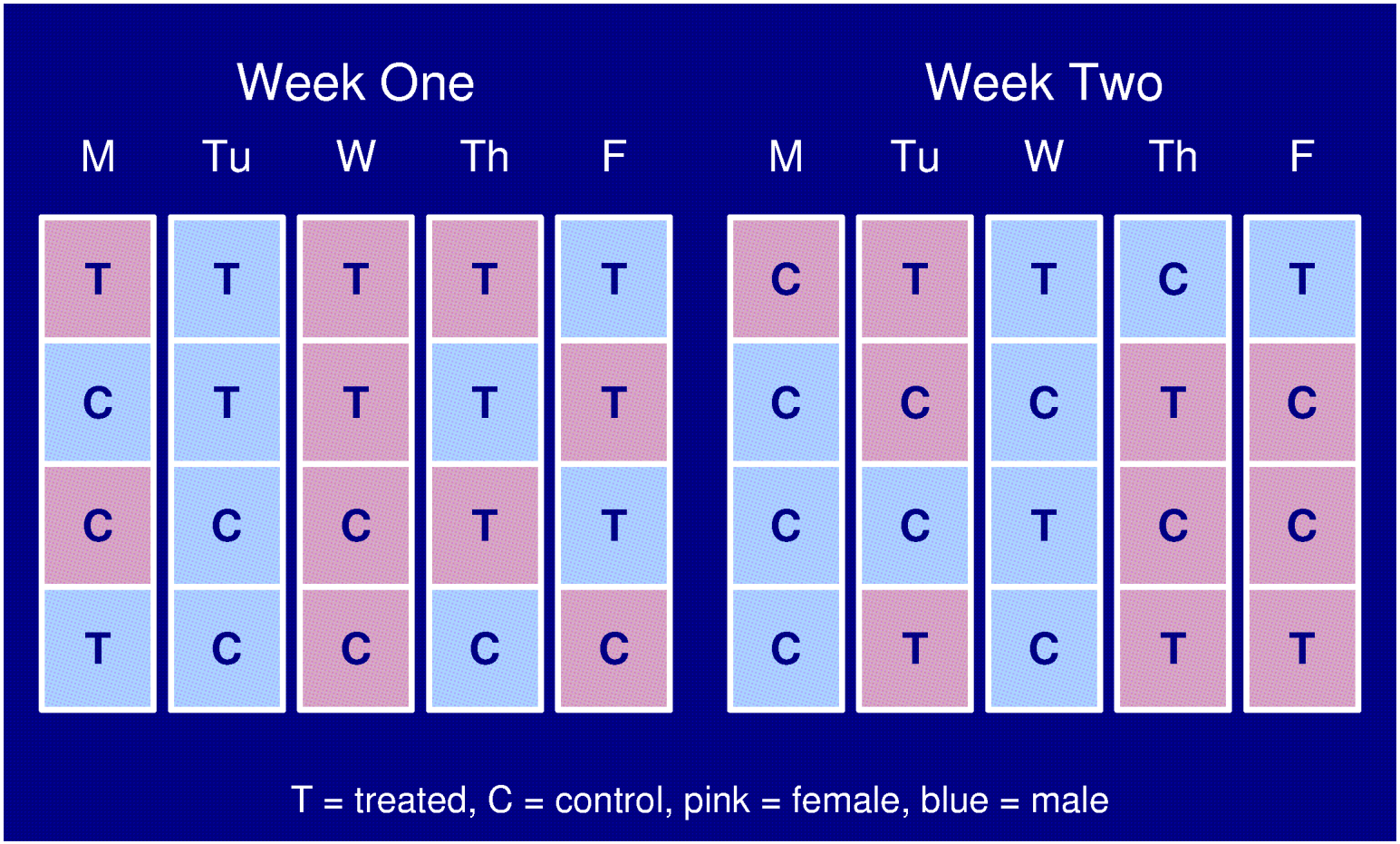
- 20 male mice and 20 female mice.
- Half to be treated; the other half left untreated.
- Can only work with 4 mice per day.

Question: How to assign individuals to treatment groups and to days?

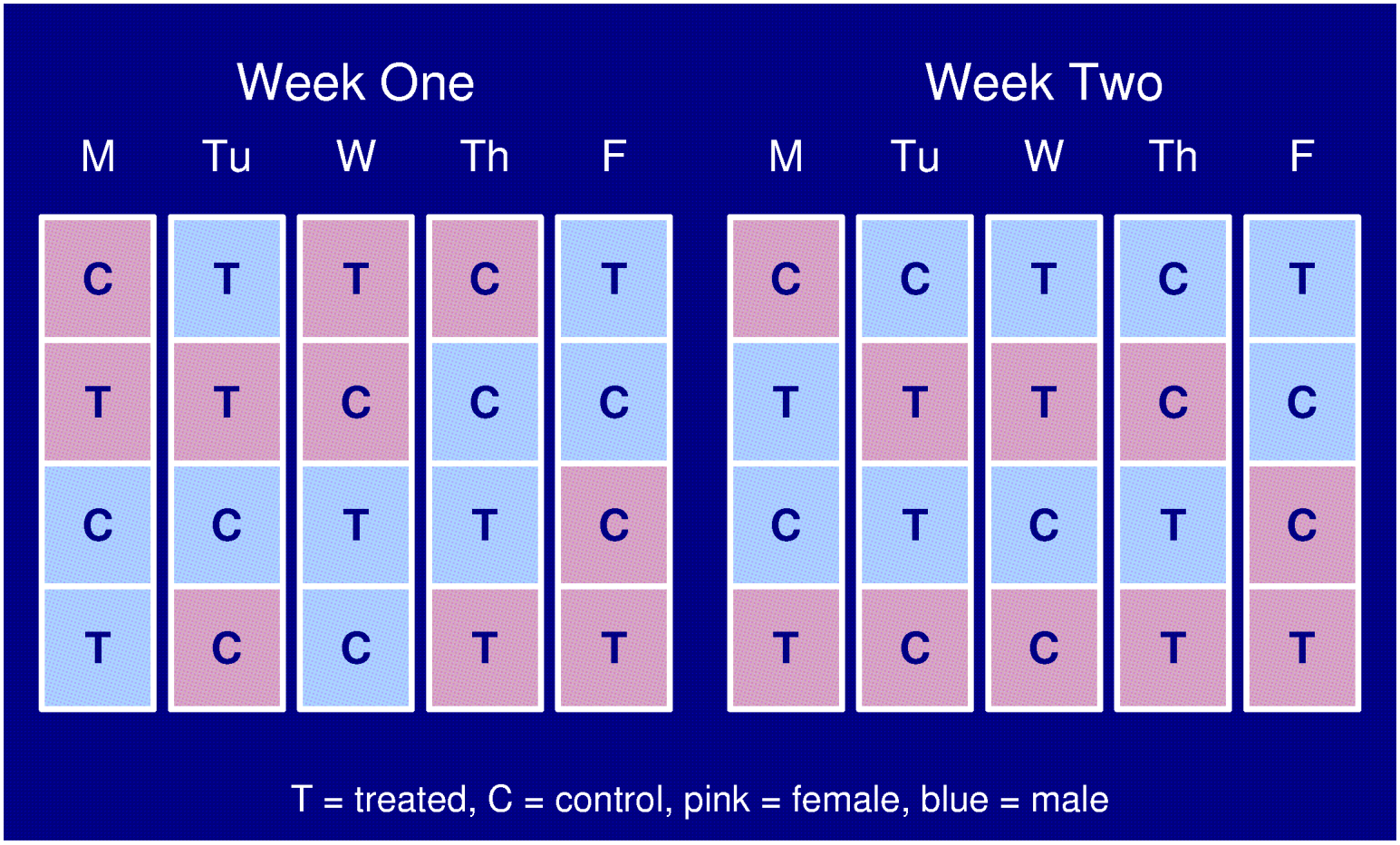
An extremely bad design



Randomized


Week One
M Tu W Th F
T T T T T
C T T T T
C C C T T
T C C C C
Week Two
M Tu W Th F
C T T C T
C C C T C
C C T C C
C T C T T
T = treated, C = control, pink = female, blue = male

A stratified design



Randomization and stratification

- If you can (and want to), fix a variable.
  - e.g., use only 8 week old male mice from a single strain.
- If you don’t fix a variable, stratify it.
  - e.g., use both 8 week and 12 week old male mice, and stratify with respect to age.
- If you can neither fix nor stratify a variable, randomize it.

Factorial experiments

Suppose we are interested in the effect of both salt water and a high-fat diet on blood pressure.

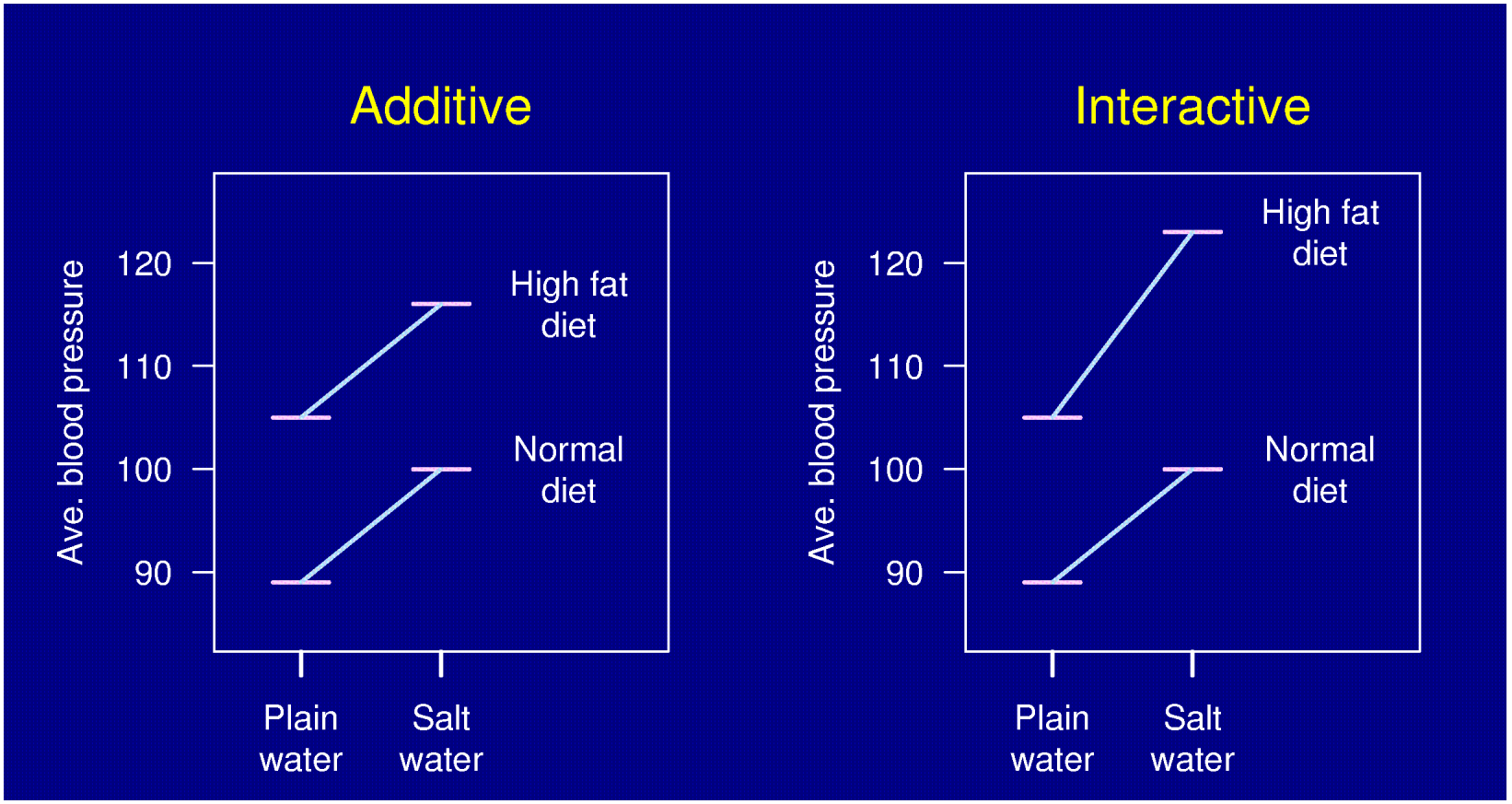
Ideally: look at all 4 treatments in one experiment.

Plain water	Normal diet
Salt water	High-fat diet

Why?

- We can learn more.
- More efficient than doing all single-factor experiments.

Interactions



Other points

- Blinding
  - Measurements made by people can be influenced by unconscious biases.
  - Ideally, dissections and measurements should be made without knowledge of the treatment applied.
- Internal controls
  - It can be useful to use the subjects themselves as their own controls (e.g., consider the response after vs. before treatment).
  - Why? Increased precision.

Other points

- Representativeness
  - Are the subjects/tissues you are studying really representative of the population you want to study?
  - Ideally, your study material is a random sample from the population of interest.

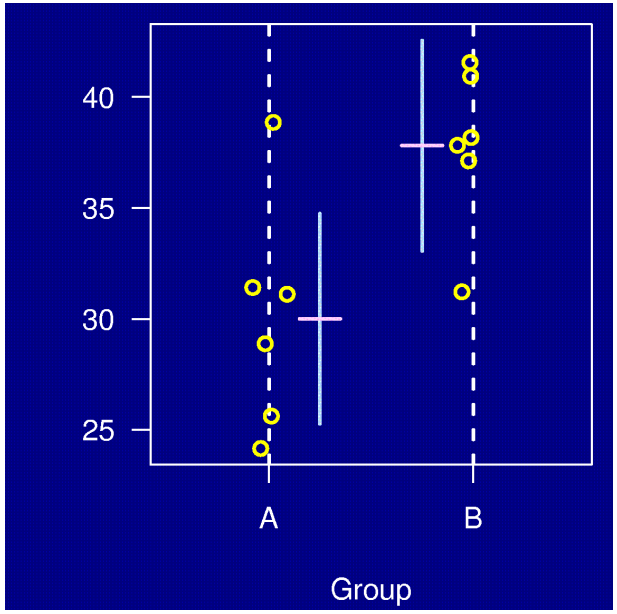
Summary

Characteristics of good experiments:

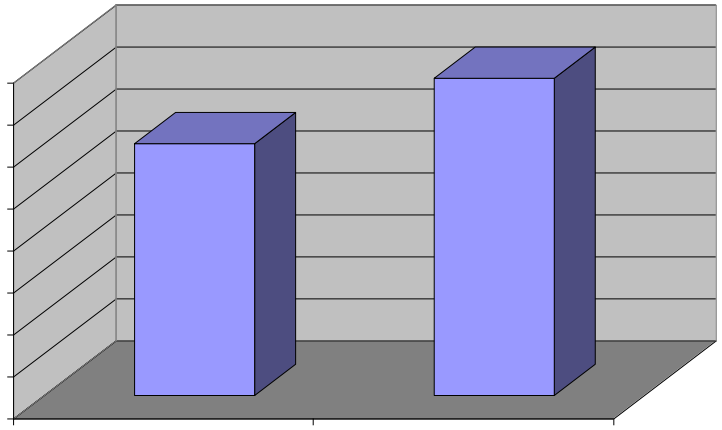
- Unbiased
  - Randomization
  - Blinding
- High precision
  - Uniform material
  - Replication
  - Blocking
- Simple
  - Protect against mistakes
- Wide range of applicability
  - Deliberate variation
  - Factorial designs
- Able to estimate uncertainty
  - Replication
  - Randomization

Data presentation

Good plot



Bad plot



Data presentation

Good table

Treatment	Mean	(SEM)

Bad table

Treatment	Mean	(SEM)

Sample size determination

Fundamental formula

Listen to the IACUC

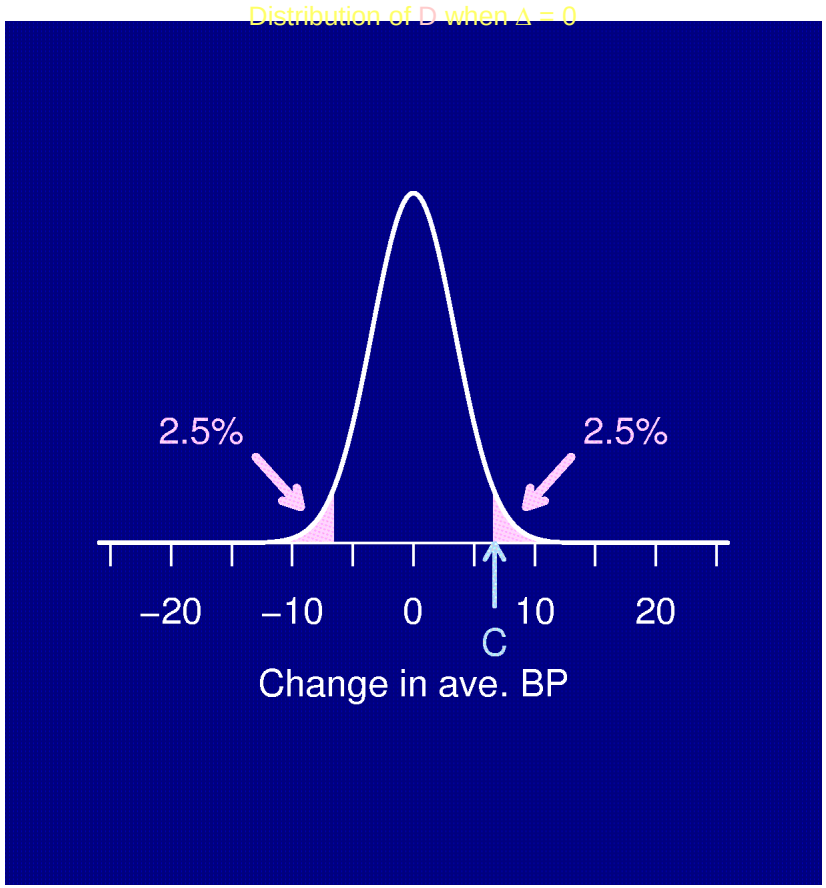
Too few animals → a total waste

Too many animals → a partial waste

Significance test

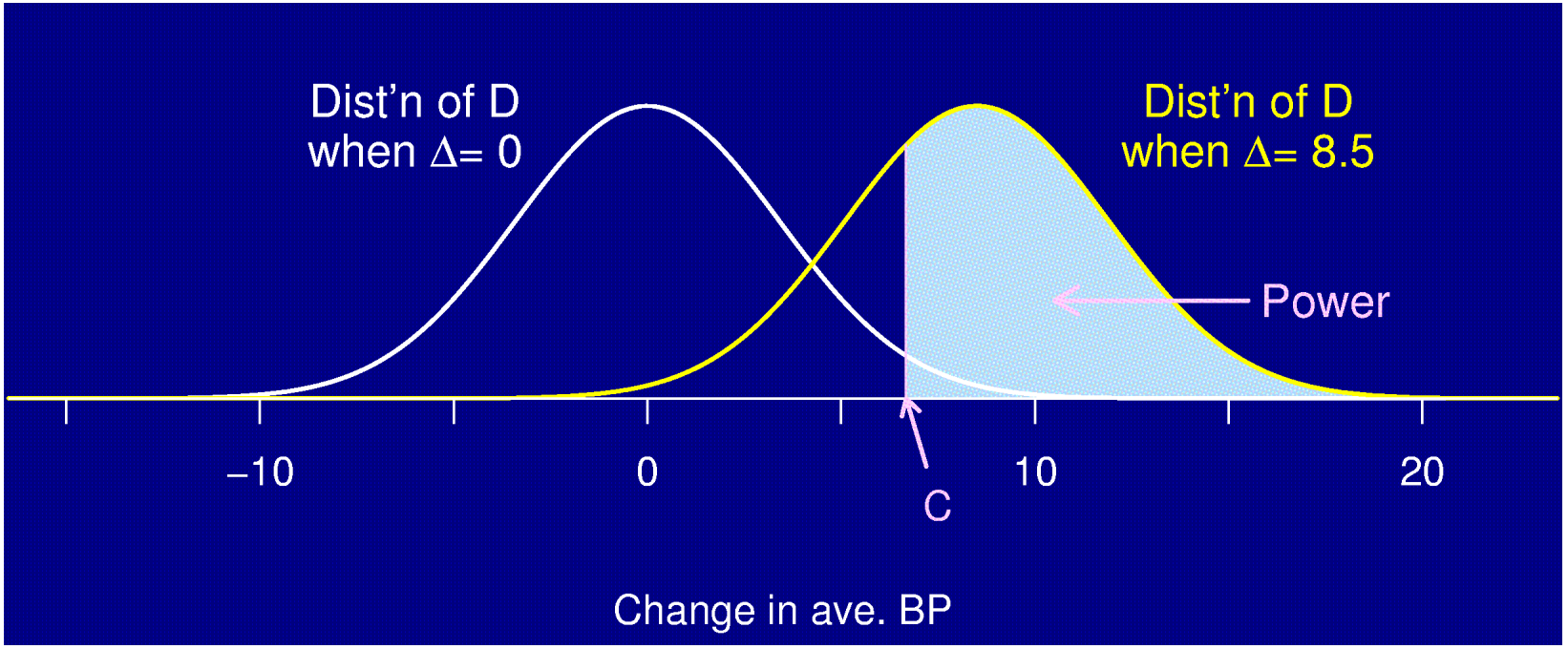
- Compare the BP of 6 mice fed salt water to 6 mice fed plain water.

- Δ = true difference in average BP (the treatment effect).
- H_0: $\Delta = 0$ (i.e., no effect)
- Test statistic, D.
- If |D| > C, reject H_0.
- C chosen so that the chance you reject H_0, if H_0 is true, is 5%



Statistical power

Power = The chance that you reject H_0 when H_0 is false (i.e., you [correctly] conclude that there is a treatment effect when there really is a treatment effect).



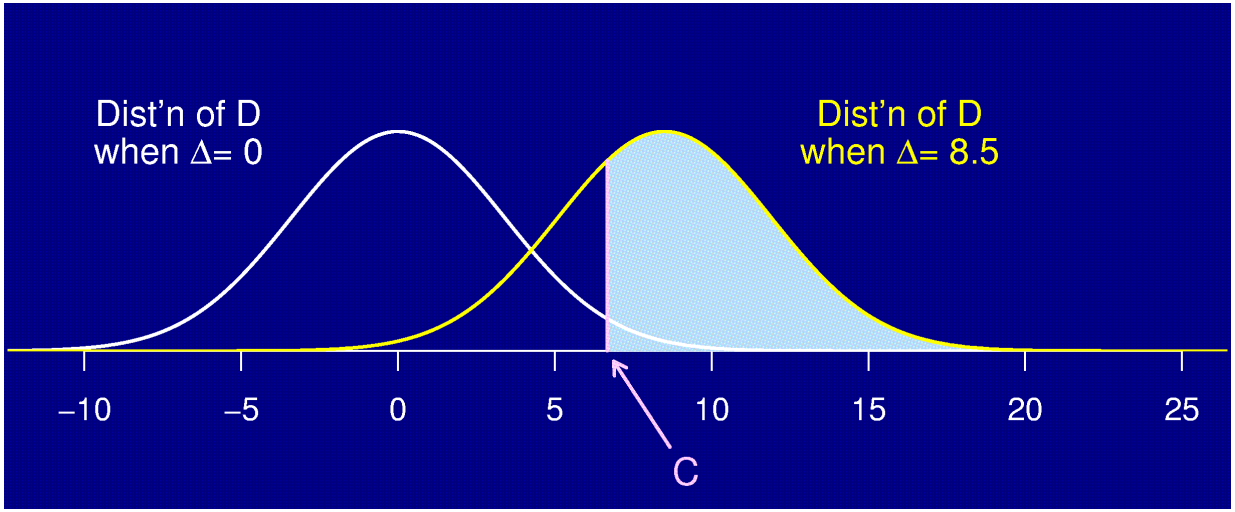
Power depends on…

- The structure of the experiment
- The method for analyzing the data
- The size of the true underlying effect
- The variability in the measurements
- The chosen significance level (α)
- The sample size

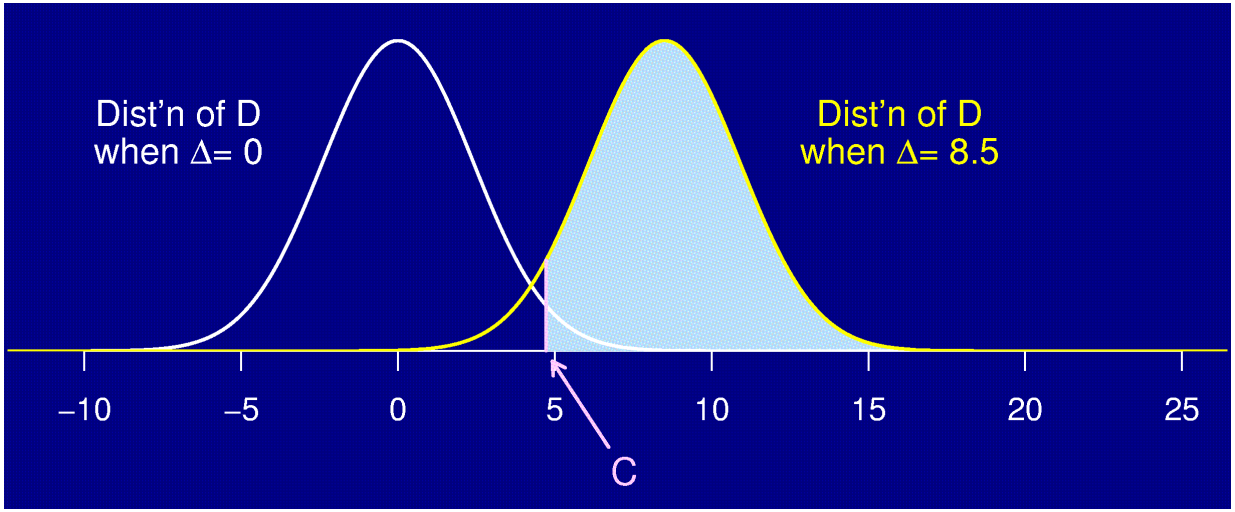
Note: We usually try to determine the sample size to give a particular power (often 80%).

Effect of sample size

6 per group:

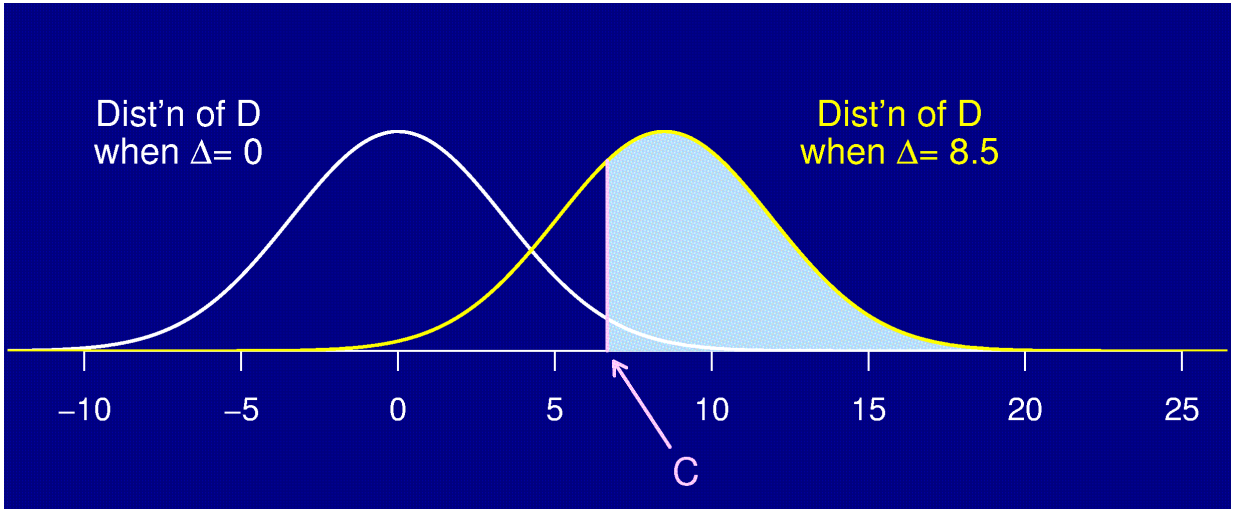


12 per group:

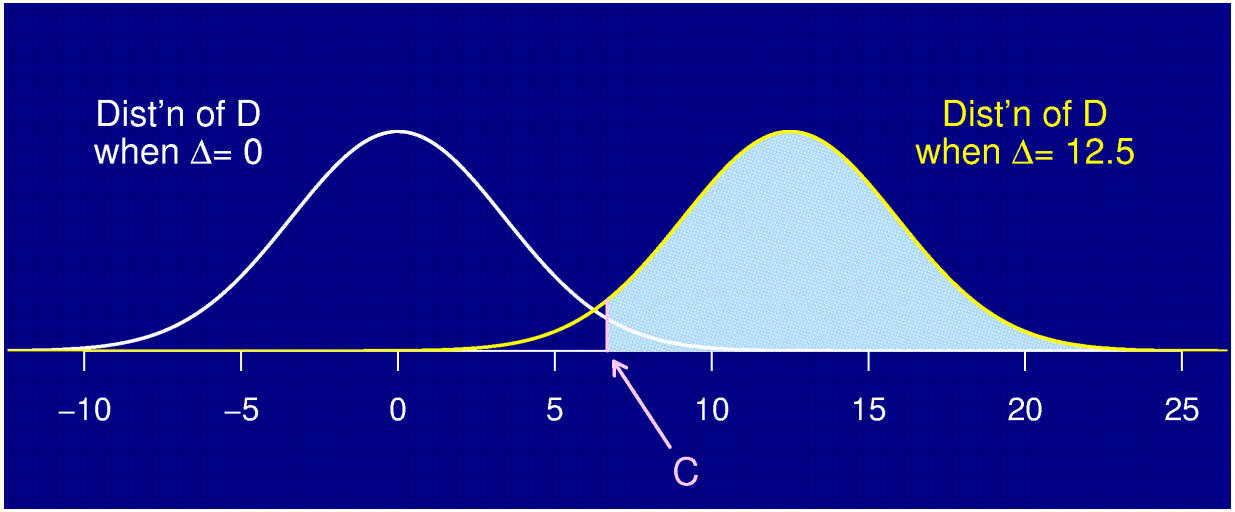


Effect of the effect

Δ = 8.5:



Δ = 12.5:



Various effects

- Desired power ↑ ⇒ sample size ↑
- Stringency of statistical test ↑ ⇒ sample size ↑
- Measurement variability ↑ ⇒ sample size ↑
- Treatment effect ↑ ⇒ sample size ↓

Determining sample size

The things you need to know:

- Structure of the experiment
- Method for analysis
- Chosen significance level, α (usually 5%)
- Desired power (usually 80%)

- Variability in the measurements
  - if necessary, perform a pilot study

- The smallest meaningful effect

A formula

$$n = \left(\sigma\right)^2 \left[\qquad 1-\beta \right]^2 \times 2$$

Censored

Reducing sample size

- Reduce the number of treatment groups being compared.
- Find a more precise measurement (e.g., average time to effect rather than proportion sick).
- Decrease the variability in the measurements.
  - Make subjects more homogeneous.
  - Use stratification.
  - Control for other variables (e.g., weight).
  - Average multiple measurements on each subject.

Final conclusions

- Experiments should be designed.
- Good design and good analysis can lead to reduced sample sizes.
- Consult an expert on both the analysis and the design of your experiment.

Reference and source

- Conceptual Econometrics Using R (ISSN Book 41) 1st Edition, by Hrishikesh D. Vinod (Editor)
- Principles of Macroeconometric Modeling (Volume 36) (Advanced Textbooks in Economics, Volume 36) by L.R. Klein, W. Welfe, et al. | Oct 5, 1999
- Macroeconomic Modeling and Macroeconometric Simulation: Illustrated with a developing economy Model (Macroeconometric model Book 1) Book 1 of 1: Macroeconometric model | by Kannapiran Arjunan | Jun 9, 2020
- Global and National Macroeconometric Modelling: A Long-Run Structural Approach by Anthony Garratt, Kevin Lee, et al. | May 4, 2012
- Simulation of a macroeconometric model with multiple time series considerations (Wayne economic papers) by Rosemary Rossiter | Jan 1, 1982