



Curtin University

# Null Hypothesis Significance Testing and the Problem of Underpowered Studies in Economics

Le (Lyla) Zhang, Curtin University  
(with Andreas Ortmann, UNSW)

2015 workshop in Experimental Methods:  
The replicability crisis in the social sciences and how to address it  
November, 2015

# Outline

## ❖ Null Hypothesis Significance Testing (NHST)

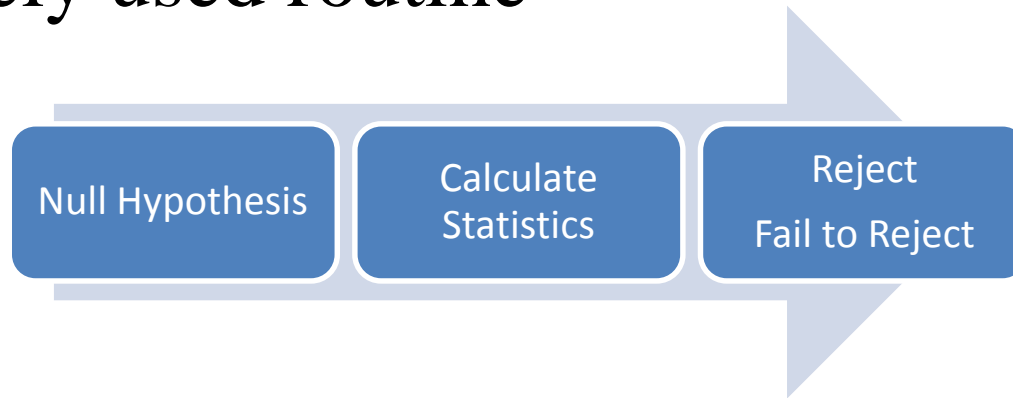
- Commonly Used Procedure
- Two Types of Errors

## ❖ The Statistical Power Analysis

- A Meta-analysis (to calculate effect size)
- Statistical power of dictator game experiments

# Null Hypothesis Significance Testing

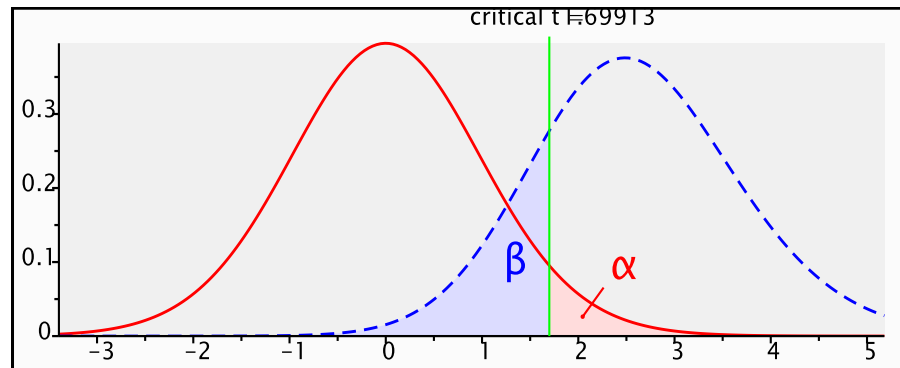
- Widely used routine



- Set “no treatment effect” as null hypothesis
- A common used (“conventional”) criterion:  
 $\alpha=5%$  (10%, 1%)

# Two Types of Errors

	Null is true ( $H_0$ )	Null is false ( $H_1$ )
Reject	$\alpha$ - Type I error false positive	$1-\beta$ (power)
Fail to reject	$1-\alpha$	$\beta$ - Type II error false negative



# Dictator Game Experiments



# Dictator Game Experiments



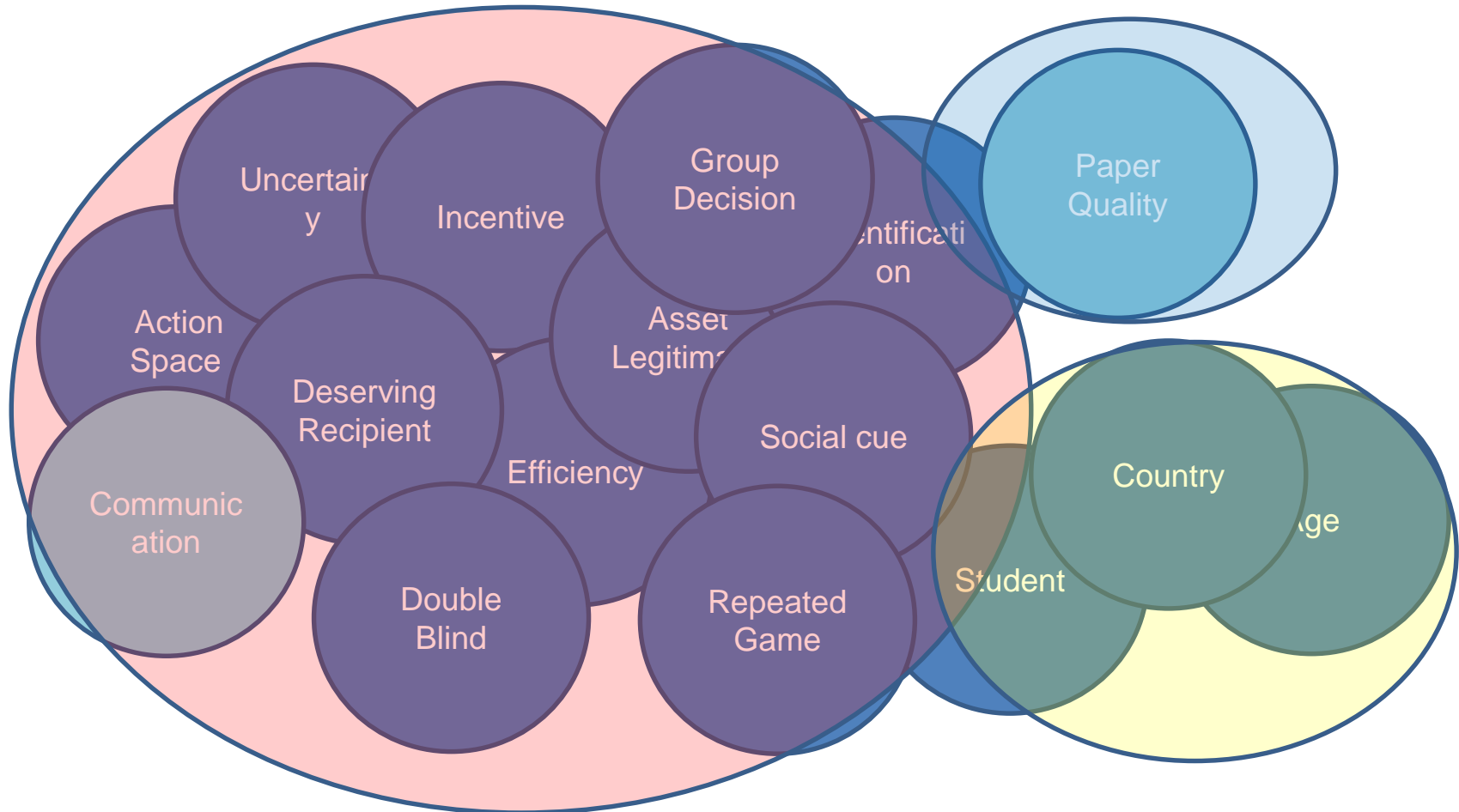
e.g., **\$10**



# Dictator Game Experiments

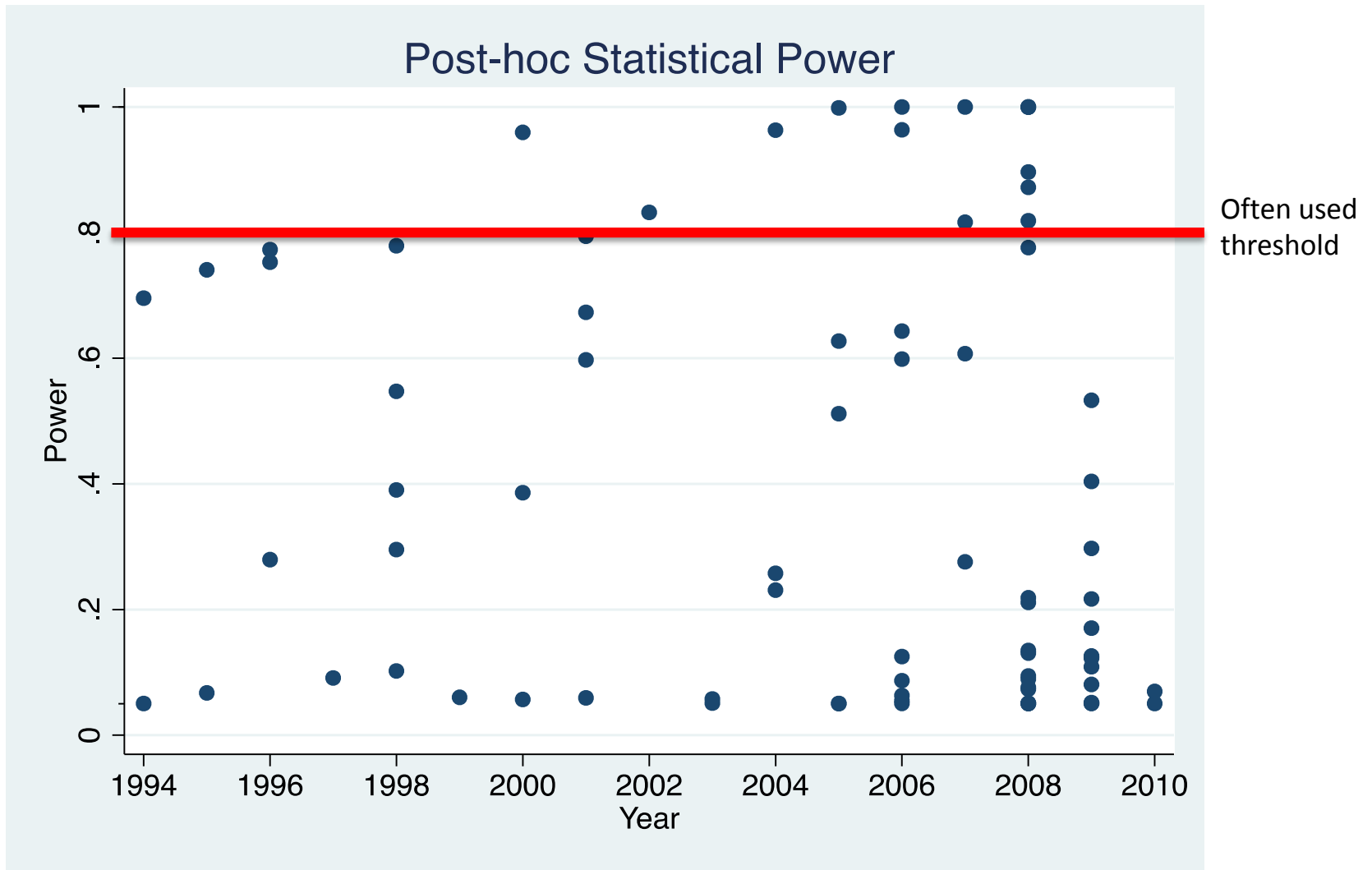
- Over the past 15 years, hundreds of dictator game experiments have been conducted (Engel, 2010; Zhang & Ortmann, 2014).
- These studies vary in experimental design variables (e.g., asset legitimacy, real money, etc) and substantial variables (e.g., country, student, age).
- Some of them are published, while others are not.

# A meta-analysis of dictator game experiments





# Dictator Game Experiments



# The severe situation of under-powered studies

- Large variations in statistical power of studies included in meta-analysis of DG game experiments (130 studies).  
(Min: 5%; Max: 100%; Median: 22.5%)
- The majority of them are under-powered (less likely to find an effect which exists).
- It depends on the sample size and the variables of interest (various design and implementation characteristics).

# Dictator Game Experiments

Large ES

- High statistical power

Medium  
ES

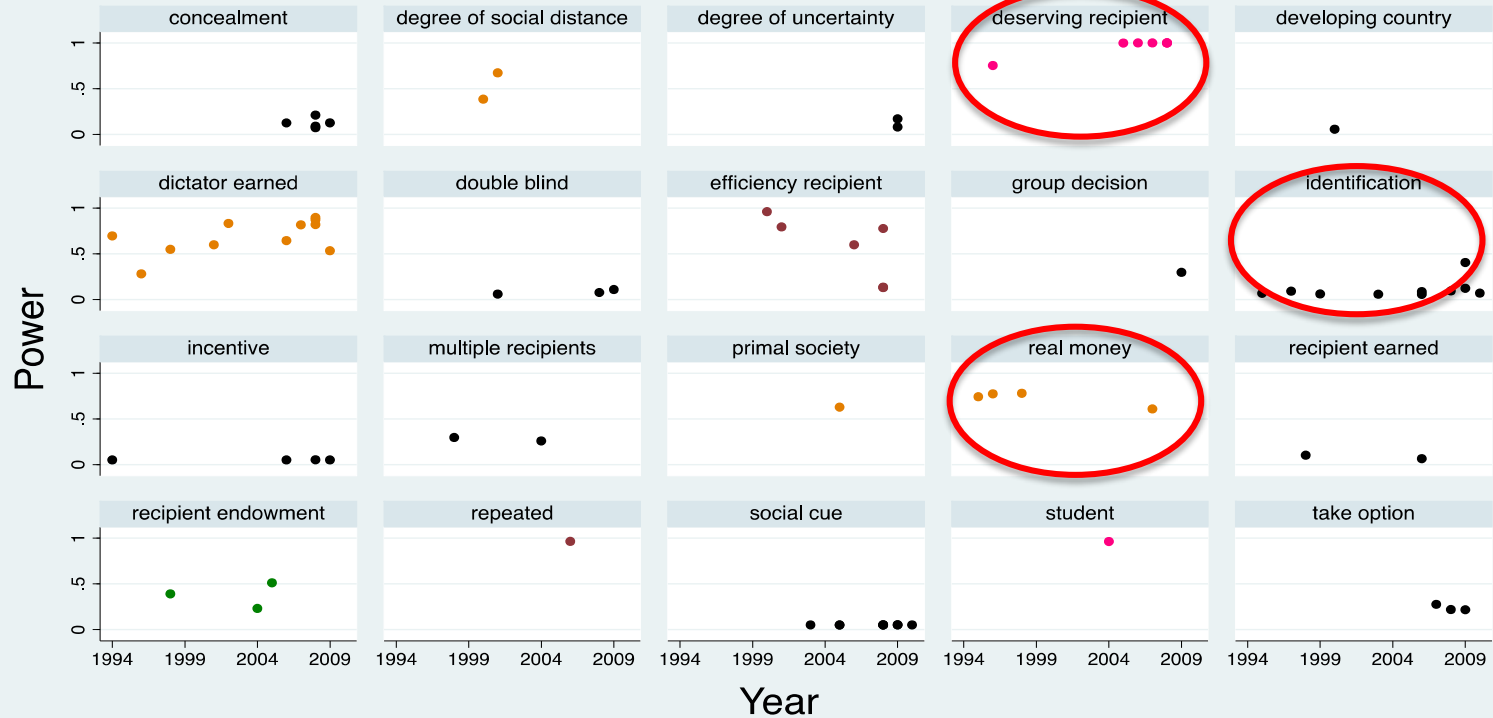
- Statistical power varies and it depends on sample size

Small ES

- Need a large sample to achieve the required statistical power

# Dictator Game Experiments

## Post-hoc Statistical Power



# What can we do?

- ❖ Rules of thumb: List et al (EE, 2010). However, it does not guarantee a high level of statistical power.
- ❖ Include a meta-analysis in the literature review, if possible.
- ❖ Use the average effect size in the meta-analysis for power analysis of future projects.
- ❖ It requires open data.
- ❖ If there is no extant study, pilot sessions would be helpful.

**Thank you!**