

Поведенческая и экспериментальная
экономика
Лекция 2. Основы экспериментальной
экономики

Елена Никишина
8 сентября 2017

Основные цели проведения экспериментов

(Roth, 1988)

- Тестирование моделей и теорий
 - Для оценки аккуратности и точности результатов модели
 - экспериментальная макроэкономика (денежная иллюзия [Fehr and Tyran (2001)], эффекты фискальной и монетарной политики [Riedl and Van Winden (2001), Bernasconi, Kirchkamp (2000)]);
 - экспериментальная экономика труда (выбор эффективной ставки заработной платы и уровня контроля [Falk et al., 2006]).
- Поиск фактов / генерирование новых данных
 - Фиксирование поведения, являющегося следствием определенной ситуации:
 - Замеры характеристик поведения в определенных ситуациях
(Пример: эксперименты на выявление рисковых и межвременных предпочтений)
 - Замеры характеристик поведения в зависимости от какого-либо события
(Пример: эксперименты на выявление эффекта владения новой информацией и т.д.)
- Рекомендации (policy implication)
 - Для выбора наиболее эффективного варианта политики:
 - Какой дизайн аукциона выбрать? [Kagel, Levin, 2001]
 - Как эффективнее потратить 1 млн \$ на борьбу с бедностью? Как повысить посещаемость школ в бедных странах?

Кроме того, эксперименты позволяют выявлять причинно-следственные связи.

Выявление причинно-следственных эффектов: рандомизация (1)

(Angrist, Pischke, 2008)

Вопрос: влияет ли на здоровье человека (Y_i) лечение в больнице?
Каким было бы здоровье человека, не лечившегося в больнице, если бы он всё же решил лечиться в больнице?

$D_i = \{0; 1\}$ - переменная, обозначающая то, лечился ли человек в больнице.

$D_i = 1$, если i -ый человек вошел в группу воздействия (treatment group), т.е. проходил лечение

$D_i = 0$, если i -ый человек вошел в контрольную группу (control group), т.е. не проходил лечение.

Для любого человека состояние здоровья в данном примере можно описать:

$$Y_i = \begin{cases} Y_{1i} & \text{если } D_i = 1 \\ Y_{0i} & \text{если } D_i = 0 \end{cases} = Y_{0i} + (Y_{1i} - Y_{0i})D_i$$

Выявление причинно-следственных эффектов: рандомизация (2)

$(Y_{1i} - Y_{0i})$ – эффект воздействия (treatment effect) для i -го человека.

НО:

1. Для каждого человека мы **не наблюдаем** и Y_{1i} , и Y_{0i} !
2. Интересен «средний» эффект воздействия.

Сравним средние значения показателя здоровья для тех, кто проходил лечение в больнице и нет:

$$E[Y_i | D_i = 1] - E[Y_i | D_i = 0] = E[Y_{1i} | D_i = 1] - E[Y_{0i} | D_i = 1]$$

Наблюдаемые различия в состоянии здоровья

Средний treatment effect для тех, кто подвергся лечению

Ненаблюдаемая
величина (здоровье того, кто
пошел в больницу в случае, если
бы он не пошел в больницу)

$$+ E[Y_{0i} | D_i = 1] - E[Y_{0i} | D_i = 0]$$

Selection bias

$$E[Y_{0i} | D_i = 1] - E[Y_{0i} | D_i = 0] - \textit{Selection bias (смещение отбора)}$$

Выявление причинно-следственных эффектов: рандомизация (3)

При случайном распределении D_i и независимости Y_{0i} и D_i можно заменить

$$E[Y_{0i}|D_i = 1] = E[Y_{0i}|D_i = 0] = E[Y_{0i}]$$

Таким образом,

$$E[Y_i|D_i = 1] - E[Y_i|D_i = 0] = E[Y_{1i}|D_i = 1] - E[Y_{0i}|D_i = 1] +$$

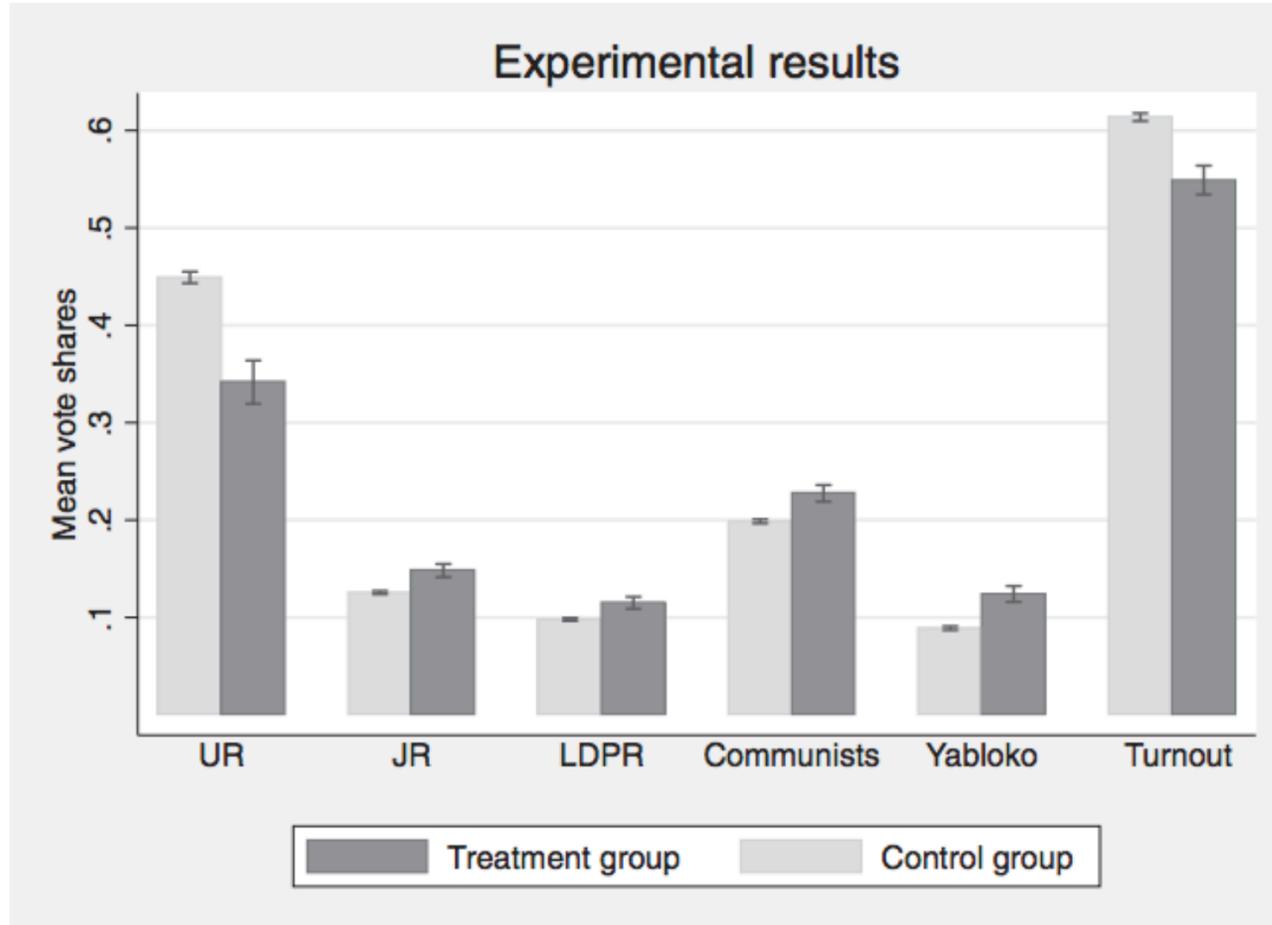
$$\underbrace{+ E[Y_{0i}|D_i = 1] - E[Y_{0i}|D_i = 0]}_{\text{Selection bias} = 0} = E[Y_{1i}|D_i = 1] - E[Y_{0i}|D_i = 1]$$

Selection bias = 0

Рандомизация играет ключевое значение для определения «чистого» эффекта.

Важность рандомизации: полевой эксперимент

Наблюдатели: выявление нарушений на выборах



Исследовательский вопрос:
как определить, были ли фальсификации
на выборах?

Полевой эксперимент:
Случайное распределение наблюдателей
по участкам.

Результат:
Наличие наблюдателей снижало долю
голосов за ЕР на 11 п.п.

Рандомизация позволяет говорить о
причинно-следственной связи.

Enikolopov et al. (2012) Field experiment estimate of electoral fraud in Russian parliamentary elections

Важность рандомизации: естественный эксперимент

Реформа образования в Китае

- Реформа образования в Китае с 2004 по 2010 гг.
 - Осуществлялась **не во всех** регионах одновременно.
 - Предполагала изменение текста школьных учебников и соответствующее изменение вступительного экзамена.
- **Результаты:**
 - По итогам реформы 20% школьников изменили свою позицию относительно демократии, роли государства в экономике, важности свободного рынка.

[Cantoni, 2017]



Требования к экспериментам

- Рандомизация
- Контроль прочих факторов
- Возможность повторить
- Внутренняя валидность
(особенно важна при тестировании моделей и теорий)
- Внешняя валидность
(особенно важна при поиске и генерировании новых данных)

Контроль в лабораторных экспериментах

на основе (Friedman and Sunder, 1994)

Непосредственный контроль (direct controls)

(измерение ключевых параметров эксперимента, выбранных в соответствии с дизайном эксперимента)

Два типа параметров:

- Контрольные переменные (controls)

остаются неизменными во время эксперимента (правила эксперимента (издержки, выгоды и т.д.); а также характеристики участника (в т.ч. пол, возраст, предпочтения и др.)

- Переменные воздействия (treatment variables)

изменяются во время эксперимента для определения эффекта воздействия (treatment effect)

Косвенный контроль (indirect controls):

- Блокировка (сохранение неизменными всех неконтролируемых параметров.
Пример: одна и та же лаборатория, один и тот же экспериментатор и т.д.)
- Рандомизация (разрушение максимального числа возможных связей между переменной воздействия (treatment variable) и ненаблюдаемыми характеристиками.
Пример: рандомизация распределения мест в лаборатории, ролей участников эксперимента, порядок тритментов внутри сессии и т.д.)

Валидность

- Оценка того, насколько используемые методы хорошо отражают действительность.
 - Внутренняя валидность (internal validity): возможность утверждать, что выявленная корреляция — это причинно-следственная связь.
 - Внешняя валидность (external validity): возможность обобщить выявленную связь на другие подобные ситуации / явления / объекты.

Сравнение типов экспериментов

	Лабораторные эксперименты	Полевые эксперименты	Естественные эксперименты
Внутренняя валидность	высокая	средняя	средняя
Внешняя валидность	низкая	средняя	высокая
Возможность повторить	высокая	средняя	низкая
Пример	«Ультиматум»	«Наблюдатели»	«Реформа образования»

Дизайн эксперимента (1)

- Between subjects design – разному типу воздействия подвергаются разные группы участников
(«контрольная группа и группа воздействия — отдельно)
- Within subjects design – каждый участник эксперимента подвергается всем типам воздействия.
(сам себе контрольная» группа).
- Часто в экспериментах комбинируются два типа дизайна.

Between subjects design

Предполагает:

- Рандомизацию участников между контрольной группой и группой воздействия (treatment group)
- Достаточно сильные выводы о различиях между группами.

НО:

- Статистическая мощность результата зависит от размера выборки (хотя наблюдения статистически одинаковы, данные достаточно шумные в силу гетерогенности индивидуальных характеристик)

Within subjects design

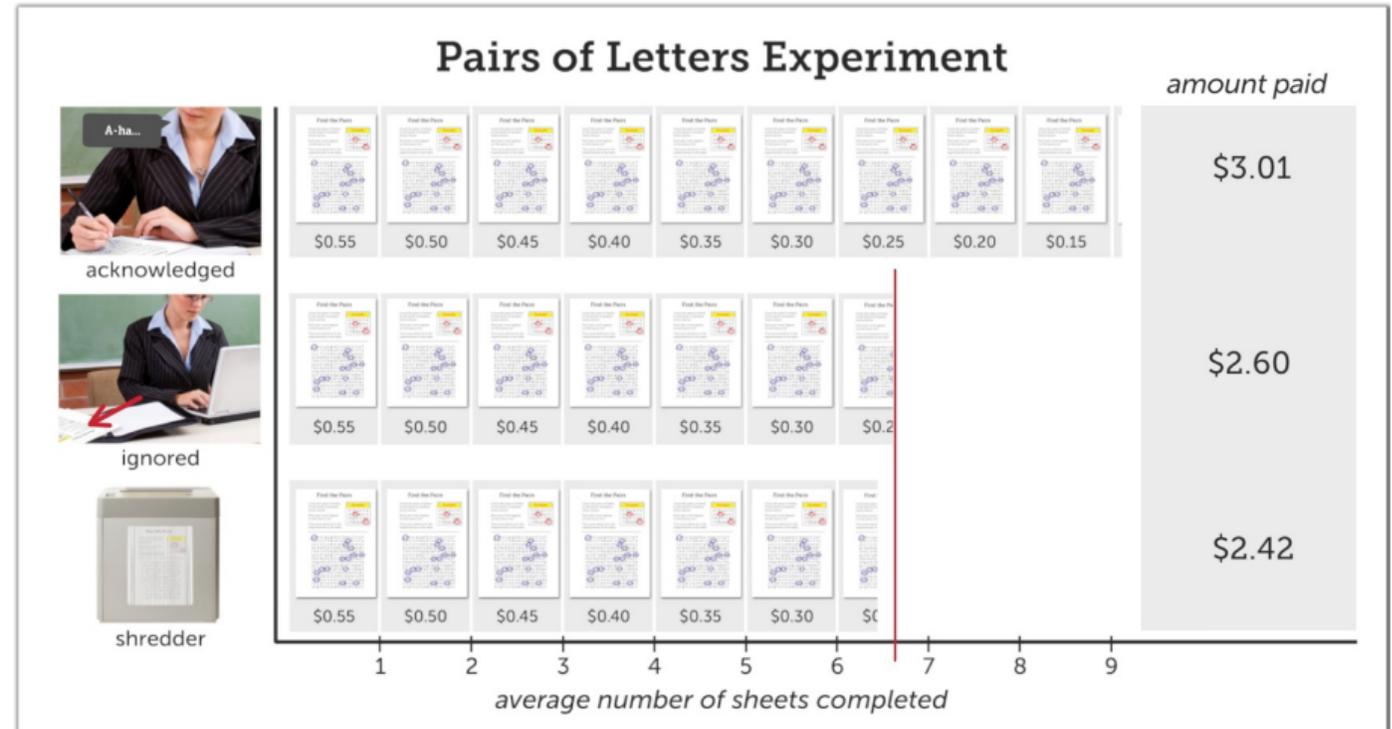
- Блокирует влияние индивидуальных характеристик
(ненаблюдаемая гетерогенность характеристик респондентов не зашумляет результат)
- Высокая статистическая мощность
(меньше шум; при равном количестве участников больше данных, по сравнению с дизайном between subjects.)
- Проблема:
 - «порядок» тритментов (А и В) может иметь значение
 - Решается за счет рандомизации порядка тритментов:
 - для части игроков - АВ;
 - для части – ВА.
 - сильнее выражен эффект экспериментатора

Within groups design.

Эксперимент о важности признания

Дизайн:

- Участники выполняют механическое задание, которое затем оплачивается.
- Участник сам решает, когда он останавливается в выполнении задачи.
- Три группы:
 - Экспериментатор игнорирует работу участников.
 - Экспериментатор пропускает работу участников через shredder.
 - Экспериментатор говорит участникам «ага» за каждое сделанное задание.
- Результаты:
 - Реакция экспериментатора увеличивает желание работать (и заработок)



Важно!

- При проведении экспериментов недостаточно сравнения средних.
- Необходимо проводить специальные статистические тесты (подробнее – см. курс Эконометрики)

Within / between subjects design:

Хоторнский эксперимент

Серия полевых экспериментов, проведенных на фабрике «Вестерн Электрикс» в США в 1924-1932 гг.

Исследовательский вопрос:

Увеличивает ли более яркое освещение производительность труда (производство реле)?

Дизайн:

- Контрольная группа (отсутствие изменений)
- Экспериментальная группа (изменение освещенности)

Результаты:

1. каждое изменение в уровне освещения приводило к росту производительности труда в экспериментальной группе, по сравнению с контрольной (between).
2. Со временем начался рост производительности в контрольной группе (within).
3. После отмены изменений производительность в экспериментальной группе вновь выросла (within).



Эффект экспериментатора (experimental demand effect)

[Zizzo, 2010]

— изменения в поведении участника эксперимента из-за его представления о надлежащем (ожидаемом, по его мнению) поведении в ходе эксперимента.

Примеры:

- Когнитивный эффект (влияние самого факта, что человек знает, что он участвует в эксперименте)
 - Хоторнский эффект (ведутся споры, есть ли он на самом деле [Levin, List, 2011]).
 - Эффект плацебо.
 - Эффект «глаз» в аудитории.
- Социальный эффект (влияние окружения, ожиданий относительно надлежащего поведения)
 - Контекст в инструкции («Диктатор» : «Помните, что он в Ваших руках»).
 - Догадки игрока, что именно проверяется в эксперименте.

Эффект экспериментатора: Пример: игра «Диктатор»

Дизайн эксперимента:

- Два игрока
 1. Активная роль (Диктатор)
 2. Пассивная роль (Жертва)
- Диктатор делит некоторую сумму M между собой и получателем.
- Жертва никак не может повлиять на Диктатора.

Результат эксперимента:

- средний размер дележа составляет 20% и может достигать 40-50% (противоречит «эгоистической модели».



Эффект экспериментатора:

- Эксперимент про дележ. Значит я должен делиться!

Как избавиться от эффекта экспериментатора?

- Скорректировать дизайн так, чтобы Диктатор мог не только делиться, но и отнимать.

Дизайн эксперимента (2)

- Объект наблюдения:
 - Отдельный человек
 - Группа из двух и более человек (например, для модификаций игры «Ультиматум»)
- Число повторений в экспериментальной сессии:
 - Один раунд (одноразовое взаимодействие) («Диктатор»)
 - Несколько раундов
 - Анонсированное число раундов.
 - «Случайное» число раундов.
- Выбор партнера:
 - Постоянный партнер (partner-design) – на протяжении всех раундов экспериментальной сессии партнеры не меняются.
 - Меняющийся партнер (strangers design) – каждый раунд игроки играют с новыми партнерами

Основные отличия экспериментов в экономике и психологии

1. Стимулы:

- Психология — «плоские» стимулы (не связаны с поведением).
- Экономика — дифференцирующие стимулы (связаны с поведением)

2. В экономических экспериментах никогда не обманывают (! no deception ! rule)

- Обман приводит к «потере контроля» над экспериментом: неясно, как участник воспринимает инструкцию.
- Обман снижает вероятность участия в последующих экспериментах.

St. Lawrence University Laboratory for Experimental Economics

Проведение эксперимента



Как подготовить эксперимент?

1. Сформулировать исследовательский вопрос
2. Решить, точно ли эксперимент – лучший способ ответа на этот вопрос?
3. Выбрать интересующие переменные.
4. Определиться с дизайном (within –between).
5. Определить число необходимых независимых наблюдений.
6. Определить число сессий.
7. Придумать дизайн эксперимента.
8. Написать инструкцию для каждого тритмента/контрольной группы.
9. Придумать вопросник.
10. Решить, будет ли эксперимент на компьютере, или нет.
11. Найти финансирование.

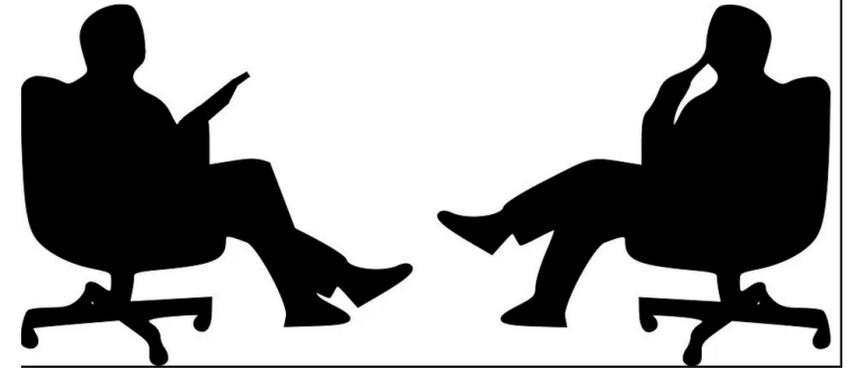
Как провести эксперимент

- Получение специального разрешения (IRB authorization в США, Франции)
- Проведение пилотного эксперимента
- Усовершенствование дизайна эксперимента/инструкции
- Проведение пилотного эксперимента
- Рекрутинг участников
- Проведение эксперимента
- Анализ данных

Инструкция: некоторые правила (1)

Инструкция:

- ...должна акцентировать внимание читателя только на необходимой информации.
 - ...должна использовать одни и те же слова при описании одних и тех же вещей в разных частях инструкции.
 - ...не должна оставлять недосказанностей.
 - ...должна создавать максимально «нейтральный контекст»:
 - без привязки к реальной экономической ситуации (например, эксперименты на коррупционное поведение).
 - без «ярлыков» при описании ролей участников.
 - ...должна по возможности нивелировать эффект экспериментатора.
-
- Дополнительно:
 - Включение иллюстративных примеров
 - ! Нужно избегать эффект якоря
 - Включение контрольных вопросов
 - Позволяет убедиться, что участники поняли правила;
 - Позволяет обобщить основное содержание инструкции.



Инструкция: некоторые правила (2)

- Во время эксперимента инструкции:
 - Раздаются всем участникам эксперимента;
 - Громко зачитываются «слово в слово».
- Никогда никакого обмана в экспериментах!
 - Этическая проблема.
 - Проблема потери контроля над экспериментом.
 - Проблема непубликуемости в экономических журналах.



Некоторые количественные параметры (правила «большого пальца»)

- Не менее **30** наблюдений в каждом экспериментальном условии, т.е.
 - 30 человек, если объект наблюдения – человек.
 - 30 групп, если объект наблюдения – группа (например, при сравнении разных дизайнов игры «Ультиматум»).
- Не менее **3** экспериментальных сессий.
- Оплата участников эксперимента – единого правила нет, но есть неписанные конвенции.

Выводы:

1. Проектирование и проведение эксперимента – искусство.
2. Качественный эксперимент – требует долгой и серьезной подготовки!

Литература

- Arifovic, J. and T.J. Sargent (2003), “Laboratory Experiments with an Expectational Phillips Curve” in D.E. Altig and B.D. Smith, eds., *Evolution and Procedures in Central Banking* (Cambridge: Cambridge University Press), 23-55.
- **Angrist, J. D., & Pischke, J. S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.**
- Ariely, D., Kamenica, E., & Prelec, D. (2008). Man's search for meaning: The case of Legos. *Journal of Economic Behavior & Organization*, 67(3), 671-677.
- Bernasconi, M. and O. Kirchkamp (2000), “Why do monetary policies matter? An experimental study of saving and inflation in an overlapping generations model”, *Journal of Monetary Economics* 46, 315-343.
- **Camerer, C. (2011). *The promise and success of lab-field generalizability in experimental economics: A critical reply to Levitt and List*.**
- Cantoni, D., Chen, Y., Yang, D. Y., Yuchtman, N., & Zhang, Y. J. (2017). Curriculum and Ideology. *Journal of Political Economy*, 125(2), 338-392.
- Enikolopov, R., Korovkin, V., Petrova, M., Sonin, K., & Zakharov, A. (2013). Field experiment estimate of electoral fraud in Russian parliamentary elections. *Proceedings of the National Academy of Sciences*, 110(2), 448-452.
- Falk, Armin and Michael Kosfeld (2006), “The Hidden Costs of Control”, *American Economic Review* 96(5): 1611-1630.
- Fehr, E. and J-F. Tyran (2001) “Does Money Illusion Matter?,” *American Economic Review* 91, 1239-62.
- Friedman, D., & Sunder, S. (1994). *Experimental methods: A primer for economists*. Cambridge University Press.
- **Levitt, S. D., & List, J. A. (2007). *On the generalizability of lab behaviour to the field*. *Canadian Journal of Economics/Revue canadienne d'économique*, 40(2), 347-370.**
- Levitt, S. D., AND J. A. List (2011) : “Was There Really a Hawthorne Effect at the Hawthorne Plant? An Analysis of the Original Illumination Experiments,” *American Economic Journal : Applied Economics*, 3(1), 224–38.
- Kagel, J. H., & Levin, D. (2001). Behavior in multi-unit demand auctions: experiments with uniform price and dynamic Vickrey auctions. *Econometrica*, 69(2), 413-454.
- Riedl, A., & Van Winden, F. (2001). Does the Wage Tax System Cause Budget Deficits A Macro-Economic Experiment. *Public Choice*, 109(3), 371-394.
- **Zizzo, D. (2010) : “Experimenter demand effects in economic experiments,” *Experimental Economics*, 13(1), 75–98.**