

Мониторинг и оценка результатов

Практика оценки влияния



Практика оценки влияний

Оценка социальных программ, проектов или стратегий:

Систематическая оценка для определения релевантности, эффективности, результативности, влияния и устойчивости запланированного, продолжающегося или завершенного вмешательства.

- Оценка потребностей
- Оценка оцениваемости и оценка плана оценки
- Оценка процесса, оценка внедрения
- Отслеживающее исследование, среднесрочные оценки, и т. д.
- Оценка влияния (например, “детальная”) и мета-оценка

Что такое оценка влияния?

Оценки влияния измеряют конкретные результаты, которые могут быть приписаны конкретному вмешательству или программе. Это достигается путем сравнения результатов, когда вмешательство применяется к результатам, при которых не было произведено само вмешательство.

Соответствующее сравнение показывает, что могло бы случиться при отсутствии вмешательства. Сравнивая результаты для этих двух групп, оценка влияния должна обеспечивать прямые доказательства того, насколько вмешательство изменяет результаты.

Инициатива разработки оценки влияний (DIME), Всемирный банк (2007)

Что такое мониторинг на основе результатов?

- Мониторинг результатов – это постоянный процесс сбора и анализа информации для сравнения того, как выполняется проект, программа или стратегия по отношению к ожидаемым результатам. (Иногда еще называется менеджментом на основе результатов, менеджментом для результатов и измерением результатов работы.)
- Типичные черты
 - Нет необходимости в группах сравнения.
 - Отсутствует внутренняя причинная связь (но основывается на системе результатов или логических моделях)
 - Цели результатов четкие и измеримые.
 - Отличается, но дополняет оценку влияния

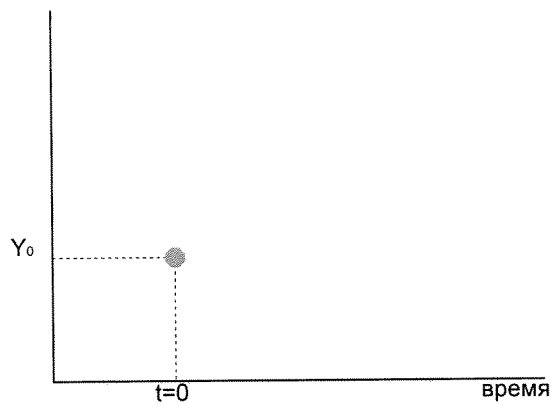
Виды оценки влияния

- Заметьте, что практически все виды оценки влияния имеют сравнительную группу и исходные данные.
- Наиболее распространены количественные оценки влияния
 - Смотрите распечатку
- Если не существуют исходных данных, смотрите работу Бамбергера (Bamberger) о *воссоздании исходных данных*

Когда целесообразно использовать оценку влияния?

- Подобные проекты, программы или стратегии, но с противоречивыми или разнящимися результатами
- Распределение ресурсов с альтернативными вариантами инвестирования
- Выявление проблем, связанных с вопросом (например, использование наркотиков, процент отсева учащихся из образовательного учреждения)
- Пилотные проекты (однако, если уже существует научная основа для подтверждения причинно-следственных связей, “детальная” оценка влияния может быть излишней)
 - Если подходят другие доказательства эффективности, а потенциальные выгоды велики, насколько больше необходимо доказательств?
 - Разница между мониторингом результатов и оценкой влияния
- Повышение прозрачности, надежности или подотчетности
- Обсуждение: существует много случаев, когда оценка влияния не подходит. Примеры?

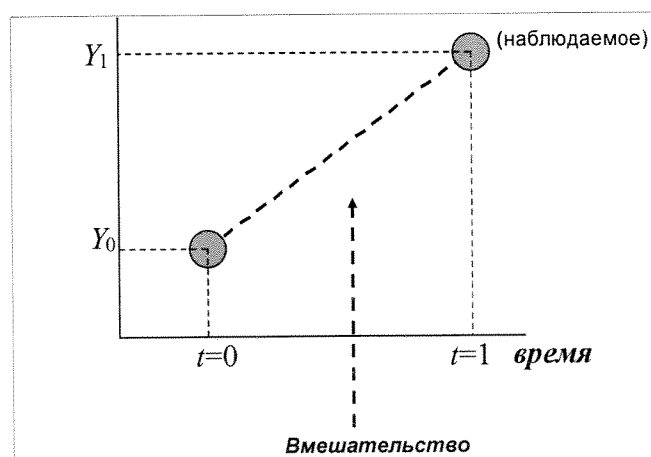
Мы наблюдаем показатель результатов



WORLD BANK INSTITUTE
Promoting knowledge and learning for a better world

7

И его значение увеличивается после вмешательства:



WORLD BA
Promoti

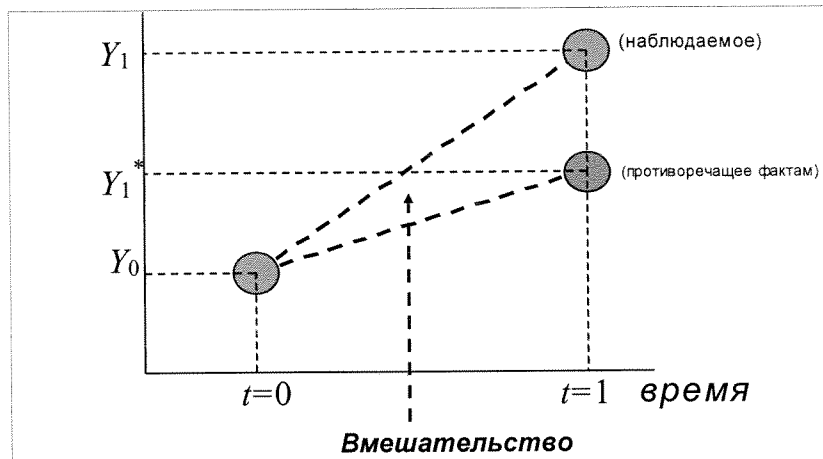
8

WORLD BANK INSTITUTE

Promoting knowledge and learning for a better world

© 2007. The World Bank Group. All rights reserved.

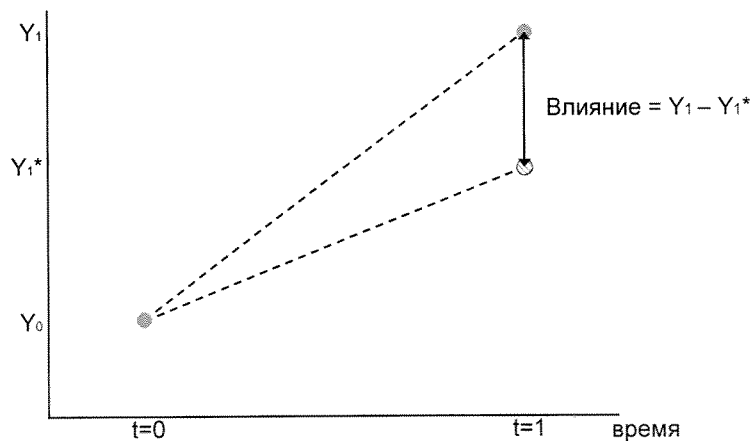
Наличие “идеального”, контрафакта...



WORLD BANK INSTITUTE
Promoting knowledge and learning for a better world

9

Позволяет нам оценить настоящее влияние



WORLD BANK INSTITUTE
Promoting knowledge and learning for a better world

10

WORLD BANK INSTITUTE

Promoting knowledge and learning for a better world

© 2007. The World Bank Group. All rights reserved.

Основы оценки программ

- Оценка программ – это комбинация техник, используемых для определения того, «срабатывает» ли вмешательство.
- Для определения причинного влияния принятых мер нам необходимы знания о контрафакте, то есть о том, что могло бы случиться в случае отсутствия вмешательства?
- Настоящий контрафакт не поддается наблюдению: Исследования по данным наблюдения или ассоциации не оценят причинную связь
- Основной целью всех программ / методов оценки влияния является построение или “имитация” контрафакта.

Продолжение

- Контроль направлен на то, чтобы оценить, что могло бы случиться с проектной совокупностью, если бы вмешательство не состоялось
- Разница между изменением t_1 для проекта и t_1 для контрольной группы является оценкой влияния
- Это называется баллом изменения

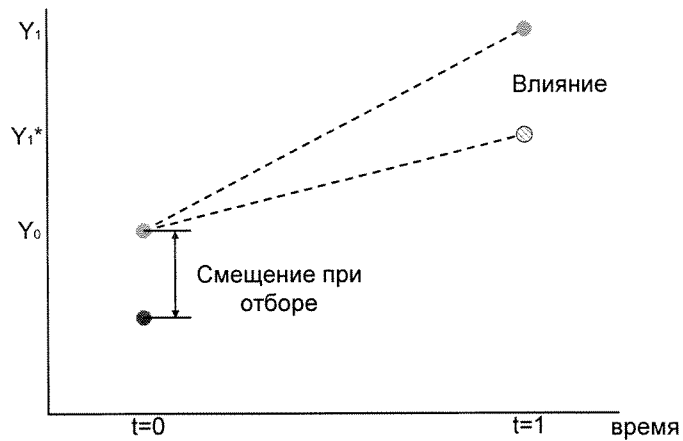
Продолжение

- Действительность балла изменения как оценки влияния зависит от того, насколько близко соответствует группа контроля группе исследования по основным характеристикам, относящимся к влиянию, которое изучается.

Особые проблемы при создании контрольных групп

- Смещение при отборе
 - Проектные сообщества или физические лица целенаправленно отбираются для определения
 - Целевых групп, которые вероятнее всего будут успешными
 - Наибеднейших и самых уязвимых групп населения (менее вероятно, что многие из них будут успешными)
 - Участники отбираются самостоятельно
 - Часто вероятнее всего, что люди / сообщества, которые сами изъявляют желание участвовать, будут успешными

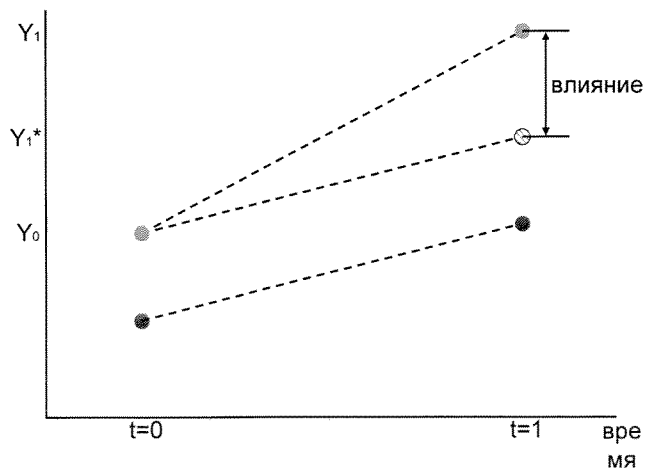
Смещение при отборе



WORLD BANK INSTITUTE
Promoting knowledge and learning for a better world

15

«Разница в разнице» требует, чтобы смещение было добавочным и неизменным во времени



WORLD BANK INSTITUTE
Promoting knowledge and learning for a better world

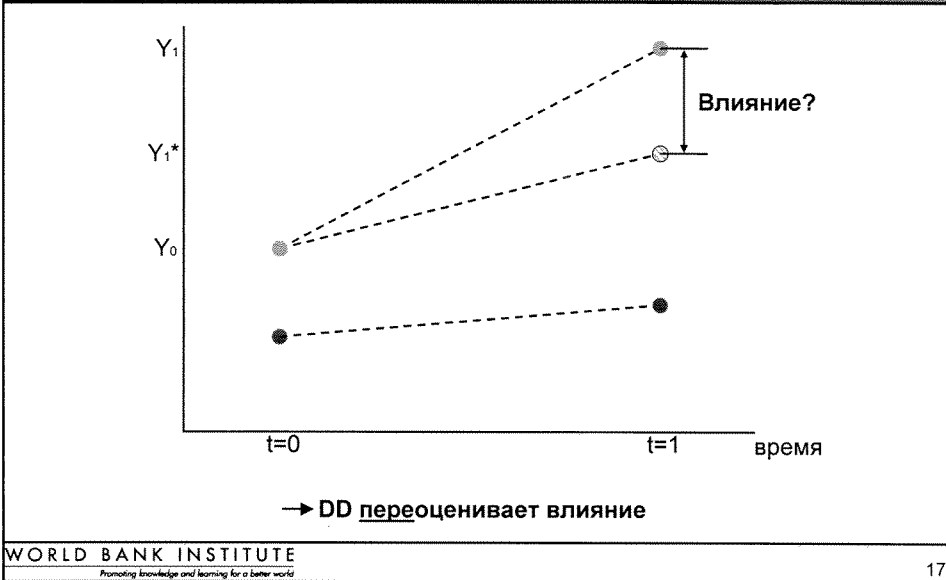
16

WORLD BANK INSTITUTE

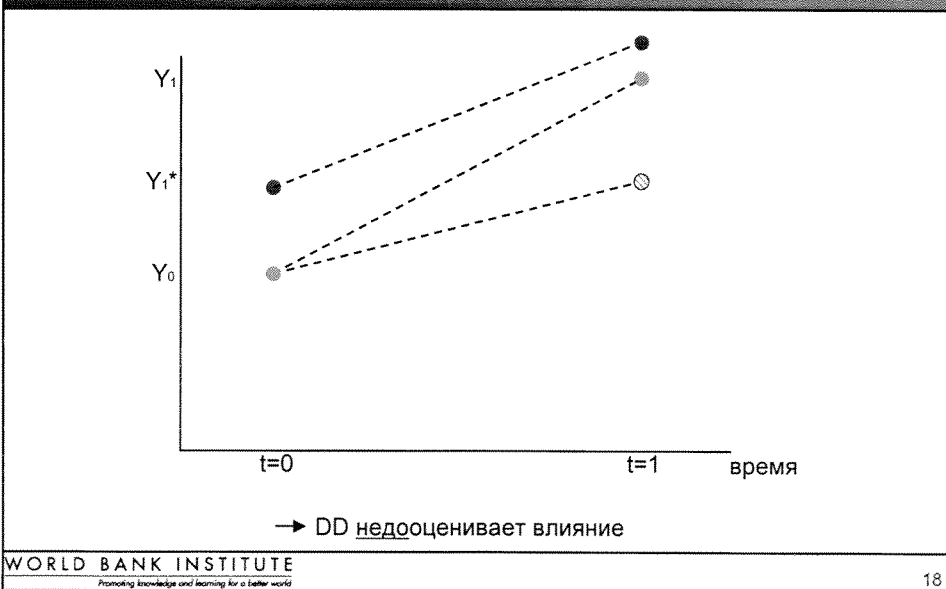
Promoting knowledge and learning for a better world

© 2007. The World Bank Group. All rights reserved.

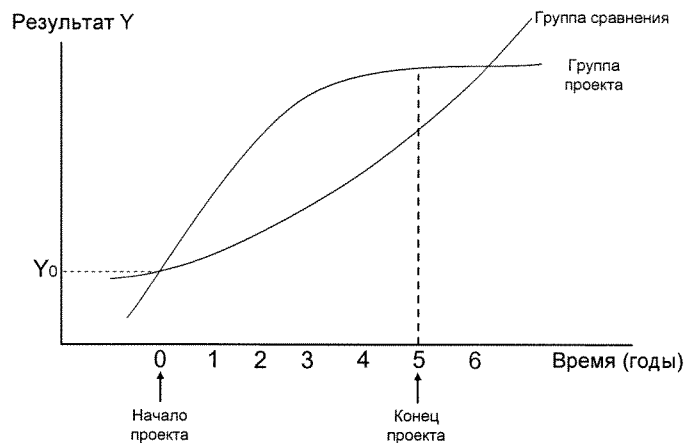
Метод оказывается неудачным, если сравнительная группа находится на различной траектории



Или...



Когда результаты не аддитивны и инвариантны по времени



WORLD BANK INSTITUTE
Promoting knowledge and learning for a better world

19

Некоторые наиболее часто используемые подходы к построению сравнительных групп

- Рандомизированное распределение мер
- Канал
- Согласование наблюдаемых переменных
 - На основе суждений (качественные методы согласования, но включая природные эксперименты, например, близнецы; географическая изоляция)
 - Согласование баллов склонности

WORLD BANK INSTITUTE
Promoting knowledge and learning for a better world

20

WORLD BANK INSTITUTE

Promoting knowledge and learning for a better world

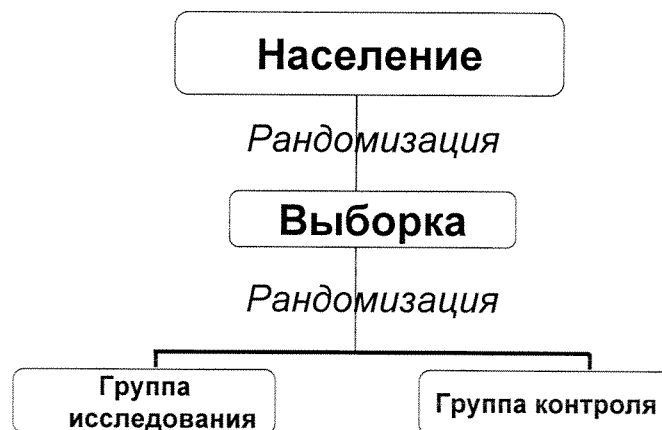
© 2007. The World Bank Group. All rights reserved.

Рандомизация

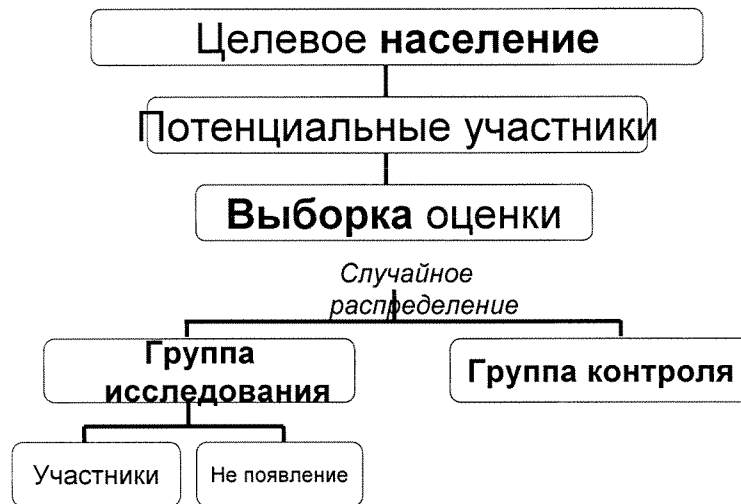
Рандомизация включает следующее в контексте социальных программ:

- Равную вероятность отбора единиц в программу
- Случайное распределение (при принятии решения куда поместить программу) и случайный отбор (при принятии решения, кто будет участвовать в программе) для преодоления смещения при отборе

Создание рандомизированной оценки



Создание рандомизированной оценки



На основе данных

Орра (Ort) (1999)

WORLD BANK INSTITUTE
Promoting knowledge and learning for a better world

23

Проблемы

- Сокращение (высокий процент отсева из групп контроля или групп исследования)
- Группа контроля может отказываться сотрудничать или отвечать
- Эффект Хоторна (угроза внутренней достоверности)
- Анализ частичного равновесия (угроза внешней достоверности). Может быть тяжело экстраполировать результаты пилотного эксперимента на национальную программу
- Этика и справедливость случайного распределения / отбора
- Политическая осуществимость случайного распределения и отбора для социальных программ

WORLD BANK INSTITUTE
Promoting knowledge and learning for a better world

24

WORLD BANK INSTITUTE

Promoting knowledge and learning for a better world

© 2007. The World Bank Group. All rights reserved.

Канальный подход

- Что такое «канальный» подход? (Использует дизайн проекта / программы)
 - Кандидаты на участие, которые еще не получили программу, формируют сравнительную группу
 - Предполагает внешнее распределение среди кандидатов
 - Отображает латентный отбор в программу
- Проблемы:
 - Аналогичны существующим для рандомизированных контрольных образцов, которые обсуждались выше в данной презентации

Согласование шкалы склонности

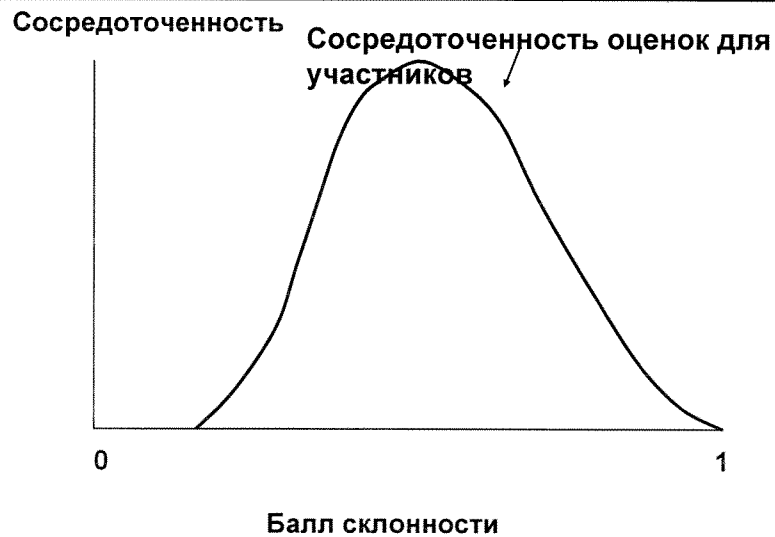
Что такое «согласование баллов склонности» или PSM?

- PSM использует статистические методы для присвоения баллов каждому участнику и неучастнику, которые основываются на наблюдаемых характеристиках. Эти баллы – называемые баллами склонности - впоследствии используются для определения группы контроля.
- Обычно используется, когда не существует группы сравнения, хотя может быть полезным для начального согласования

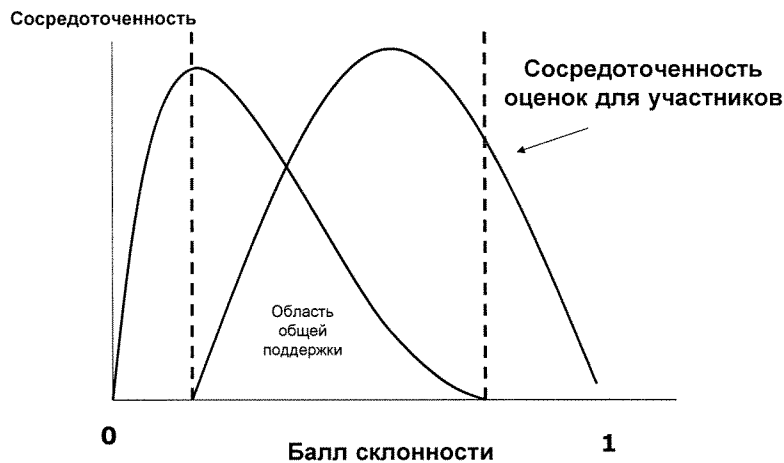
Шаги при согласовании баллов

1. Репрезентативные и сравнимые опросы неучастников и участников.
2. Составьте две выборки и оцените модель логит (или пробит) участия в программе.
3. Ограничьте выборки для обеспечения **общей поддержки** (важный источник смещения при наблюдательном изучении)
4. Найдите для каждого участника выборку неучастника, который имеет схожие баллы склонности.
5. Сравните индикаторы результата. Разница является **оценкой выгоды** для такого наблюдения вследствие программы.
6. Подсчитайте среднее значение этих отдельных выгод для получения средней общей выгоды.

Распределение баллов склонности



Распределение шкалы склонности: участники – неучастники



WORLD BANK INSTITUTE
Promoting knowledge and learning for a better world

29

Проблемы и ограничения PSM

- Требуется значительные выборки и хорошие данные
- Предполагает отсутствие смещения при отборе на основе ненаблюдаемых характеристик.
- Нахождение «области общей поддержки» может составить проблему, поскольку может быть незначительное количество участников из «исследуемой» и «неисследуемой» групп с одинаковыми баллами склонности

WORLD BANK INSTITUTE
Promoting knowledge and learning for a better world

30

WORLD BANK INSTITUTE

Promoting knowledge and learning for a better world

© 2007. The World Bank Group. All rights reserved.

Другие проблемы “детальной” оценки влияния

- **Заражение** (например, эффект влияния мест “исследования” на места “контроля”; введение других программ либо на месте “контроля”, либо на месте “исследования”, которые могут повлиять на соответствующие результаты)
- **Бюджет, время и политические ограничения**, включая административные вопросы (например, своевременность отчетных результатов)
- **Могут существовать разные результаты в зависимости от того, когда были собраны данные “до” или “после”**
 - Например, влияние вмешательств могли не иметь достаточного времени для самопроявления; или данные “до” были собраны слишком поздно.

Другие проблемы “детальной” оценки влияния

- **Техническая сложность и релевантность**
 - Вероятная неспособность покрытия важных вопросов (например, процессы изменения, устойчивость влияния, незапланированные последствия, и т. д.).
 - Вероятная сложность изменения инструментов исследования в начале изучения; определенная жесткость процедур
 - Дефицит (или отсутствие) информации о контекстуальных факторах (мотивация, самоуверенность, политические / социальные факторы, и т. д.), что влияет на результаты
 - Вероятная неприменимость для всех социальных вмешательств (например, отсутствие группы контроля; трудности при сборе данных от целевых бенефициантов на чувствительные темы (например, наркоманы) или от удаленных сельских жителей

ИСТОЧНИКИ

В этой презентации в основном используются работы, выполненные источниками Всемирного банка, включая:

- Мартина Равайона (Martin Ravallion), DEC
- Эммануеля Скуфиаса и Маркуса Голдстейна (Emmanuel Skoufias and Markus Goldstein), PREM
- Себастиана Мартинеза и Пола Дж. Гертлера (Sebastian Martinez and Paul J. Gertler), HDN
- Ховарда Уайта (Howard White), IEG
- Инициативы разработки оценки влияний (DIME)