Прикладная эконометрика

Задание 1. Часть 2.

Макеева Наталья э-401, Петраков Сергей э-401

**Рандомизация**

5. Автор делает рандомизацию по стратам, чтобы учесть их характерные (идиосинкразические) особенности и уменьшить величину необъясненной дисперсии эффекта. Интерес состоит в том, чтобы агитаторы посещали именно те участки, где находится максимально возможный потенциал для получения голосов.

Если просто назначить тритмент случайно, то это эквивалентно тому, что давать лекарство всем людям, тогда как действительно положительный эффект будет только для больных. Также и в данном исследовании, есть потребность в том, чтобы тритмент раздавался не случайно, а с поправкой на то, чтобы это было осмысленно с идейной точки зрения. В данном случае – агитаторы ресурс, и направить их надо в те места, где успех достигается с большей вероятностью.

Добиться цели автору помогают следующие шаги:

1. Стратификация – процесс вычисления потенциала для получения голосов и ранжирование участков в стратах по территориям по показателю потенциала
2. Рандомизация – случайное распределение на контроль и воздействие участков
3. Распределение агитаторов в те участки воздействия, где вероятность успеха больше

6. Автор отмечает 2 необычных аспекта рандомизации:

1. Не использовать все случайно тритменты, которые были изначально назначено до распределения агитаторов. Автор отмечает, что основа их подхода – это условный тритмент (conditional treatment), то есть необходимость использования тритмента по назначению, что добавляет некоторый trade off между рациональностью использования ресурсов и чистотой эксперимента.
2. Не было привычного разделения на этап определения выборки и этап разбиения её на группу контроля и воздействия. В рамках данного исследования нет чёткой границы между этими этапами. В данной работе вероятность включения в группу воздействия зависит от того, какие участки уже в неё включены, так как есть некоторое ограничение на численность жителей которое отводится на каждую отдельную страту. Автор отмечает, что финальные результаты робастны к этим ограничениям.

**Результаты**

3. Важность такой фильтрации в том, что эта бинарная переменная определяет территории в данных, в рамках которых был проведён тритмент, который определяется по результатам рандомизации. В разделе "Identification of Territories Which Followed the Randomization Plan" автор говорит о 2 причинах, по которым некоторые территории не были включены в list of allocated precincts:

1. Нет доступа к листу из-за того, что административные данные не зарегистрированы на сайте
2. Локальные власти автономно решали вопрос покрытия опросами территорий, таким образом, делая рандомизацию невозможной

6. На исследовательский вопрос отвечает коэффициент из таблицы 5 при переменной treatment, который показывает эффект от отнесения участка к исследуемой группе на результат на избирательном участке при прочих равных (с учётом фиксированных эффектов от страт) (1 модель). Во второй модели «при прочих равных» соответствует учёту фиксированных эффектов от страт, а также потенциальных результатов голосования на участке (явка), в 3 модели в «прочие равные» входит также контроль на список параметров, который включает 19 переменных, они всячески контролируют возрастные, демографические и экономические показатели.

Проинтерпретируем полученный результат для 3 модели (1 раунд): агитация в округе данного избирательного участка увеличила долю голосов за Франсуа Оланда на 0.44 п.п. единиц на избирательном участке, соответствующему этому региону.

7. Автор включает контрольные переменные для того, чтобы показать несмещённость полученной оценки. Как можно убедиться оценки устойчивы к модификациям модели, обращая внимание на стандартные ошибки можно заключить, что оценки изменяются статистически незначимо от между моделями. Для результатов таблицы 5 это справедливо на 1% уровне значимости, а для таблицы 10 на 10% уровне значимости. Стандартные ошибки ещё более устойчивы к разным спецификациям модели и практически неизменны во времени.

Оценка называется within estimator, так как оценивается модель с фиксированными эффектами. Именно такую модель стоит оценивать, поскольку разумной предпосылкой для данного исследования является наличие фиксированных, отличительных особенностей для разных территорий Франции, и, чтобы данный факт корректно учесть, стоит использовать именно эту модель.

Оценка не within estimator предполагает оценку в отсутствии предпосылки фиксированных эффектов, например, случайных эффектов, и тогда корректней использовать модель со случайными эффектами.

8. Результаты прошлых выборов не являются плохим контролем, так как плохой контроль – это ситуация, при которой тритмент влияет на контрольную переменную, в рамках исследования агитирование происходило в 2012 год, а предыдущие выборы в 2007. Будущее на прошлое повлиять не может.

9. У плацебо теста и balance of covariates разные предназначения:

В случае плацебо теста – это проверка эффективности тритмента путём рассмотрения прошлых данных, тогда как баланс ковариатов осуществляется для того, чтобы проверить, что в разделённой на контроль и воздействие выборке нет существенных отличий распределения ковариатов, то есть наблюдается их баланс.

10.

1. В рамках статьи была представлена статистика по агитаторам, довольно много персонифицированной информации, которую можно было бы использовать. В качестве контроля в регрессиях использовались демографические и экономические данные по населению территорий, однако возникает вопрос: «Как исследователи контролируют эффект индивидуальных особенностей агитаторов?». Иными словами, может быть при прочих равных не факт агитации производит впечатление, а то, кто это делает, условно проконтролировать на то, что такую агитацию проводит молодая девушка по сравнению с другими агитаторами. Возможно это может внести существенную долю в результаты.
2. Можно использовать данную базу данных для расчётов в исследованиях, связанных с изменениями в структуре избирательной массы и её характеристиках (если они есть), предпосылкой к такому исследованию служат работы, посвящённые постепенному снижению явки на выборах в разных странах за период 20 – начала 21 века. Вероятно, подробная статистика по избирателям позволит найти интересные закономерности в политических решениях общества (в том числе и неучастия в выборах).

11. Мы использовали робастные стандартные ошибки (НАС1), что позволило реплицировать такие же ошибки, которые получили авторы статьи.

**Репликации таблиц**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Table 5—Impact on Hollande's vote share** | | | | | | | | | |
|  | | | | | | | | | |
|  | First round | | | Second round | | | Average of first and second rounds | | |
|  | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
|  | | | | | | | | | |
| treatment | 0.0063 | 0.0050 | 0.0044 | 0.0048 | 0.0053 | 0.0046 | 0.0056 | 0.0049 | 0.0043 |
|  | (0.0023) | (0.0019) | (0.0019) | (0.0028) | (0.0019) | (0.0018) | (0.0024) | (0.0017) | (0.0016) |
|  |  |  |  |  |  |  |  |  |  |
| Strata fixed effects | X | X | X | X | X | X | X | X | X |
| Control for past outcome and PO |  | X | X |  | X | X |  | X | X |
| Additional controls |  |  | X |  |  | X |  |  | X |
| Mean in control group | 0.3157 | 0.2994 | 0.2994 | 0.5757 | 0.5597 | 0.5597 | 0.4457 | 0.4295 | 0.4295 |
| Observations | 3,397 | 2,665 | 2,665 | 3,397 | 2,665 | 2,665 | 3,397 | 2,665 | 2,665 |
| R2 | 0.0028 | 0.5161 | 0.5279 | 0.0011 | 0.6324 | 0.6448 | 0.0020 | 0.6450 | 0.6549 |
|  |  |  |  |  |  |  |  |  |  |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Table 10—Placebo-Impact on Royal’s vote share in 2007** | | | | | | | | | |
|  | | | | | | | | | |
|  | First round | | | Second round | | | Average of first and second rounds | | |
|  | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
|  | | | | | | | | | |
| treatment | 0.0050 | 0.0015 | 0.0011 | 0.0009 | -0.0013 | -0.0008 | 0.0029 | -0.0002 | -0.0001 |
|  | (0.0026) | (0.0023) | (0.0023) | (0.0029) | (0.0024) | (0.0023) | (0.0026) | (0.0020) | (0.0020) |
|  |  |  |  |  |  |  |  |  |  |
| Strata fixed effects | X | X | X | X | X | X | X | X | X |
| Control for past outcome |  | X | X |  | X | X |  | X | X |
| Additional controls |  |  | X |  |  | X |  |  | X |
| Mean in control group | 0.274 | 0.2619 | 0.2619 | 0.5146 | 0.5055 | 0.5055 | 0.3943 | 0.3837 | 0.3837 |
| Observations | 2,665 | 2,140 | 2,140 | 2,665 | 2,140 | 2,140 | 2,665 | 2,140 | 2,140 |
| R2 | 0.0018 | 0.3710 | 0.4068 | 0.00004 | 0.4954 | 0.5182 | 0.0006 | 0.5252 | 0.5500 |
|  |  |  |  |  |  |  |  |  |  |