

Лабораторная Работа №3 Фоминых Илья БПТ181

Задание 1.

1. Приведите пример модерации. Для этого сформулируйте подходящую исследовательскую гипотезу. Запишите, какая переменная выступает зависимой, какая – ключевым предиктором. Что является модератором, какую роль он выполняет во взаимосвязи ключевых переменных?

Допустим, мы выдвигаем исследовательскую гипотезу, что интенсивность политического насилия (**Violence**) способна повышать вероятность того, что в конфликт вмешается международное сообщество (бинарная переменная (**UN**)). Но связь политического насилия (**Violence**) и вероятности вмешательства международного сообщества (**UN**) зависит от стадии экономического цикла, в котором находится мир (**Kondratieff**). Мы предполагаем, что в случае роста мировой экономики страны будут внимательнее относиться к проблемам насилия в других странах и вмешаются с большей вероятностью, в противоположность сценарию, в котором они сами страдают от рецессии. Таким образом, (**Kondratieff**) может изменять характер взаимосвязи между (**Violence**) и (**UN**)), что делает её модератором (при этом Насилие (**Violence**) выступает предиктором, а вмешательство (**UN**) - зависимой переменной). Необходимо заметить, что мы не включали контрольные переменные.

2. Приведите пример медиации. Для этого сформулируйте подходящую исследовательскую гипотезу. Запишите, какая переменная выступает зависимой, какая – ключевым предиктором. Что является медиатором, какую роль он выполняет во взаимосвязи ключевых переменных?

Допустим, мы выдвигаем исследовательскую гипотезу, что экономический рост (**Inc**) в развивающихся странах может зависеть от политической стабильности. Но он (**Inc**) зависит от политической стабильности (**Stab**) не напрямую, а постольку, поскольку политическая стабильность способствует повышению уровня образования (**Edu**) и снижению транзакционных издержек для крупного бизнеса (**Cost**), что и делает его медиатором, определяя его роль во взаимосвязи ключевых переменных. В таком случае, переменной-медиатором будет выступать ключевой предиктор (**Stab**), зависимой переменной будет экономический рост (**Inc**), а остальные переменные (строго в этой спецификации) также будут предикторами. Необходимо заметить, что мы не включали контрольные переменные.

Задание 2. Напишите reflection paper по статье Mozaffar S., Scarritt J., Galaich G. (2003). «Electoral Institutions, Ethnopolitical Cleavages, and Party Systems in Africa's Emerging Democracies».

1. Сжато и максимально четко опишите основную идею статьи, исследовательский вопрос.

Статья посвящена изучению вопроса, как именно электоральные институты и этнополитические расколы влияют на партийную структуру развивающихся африканских стран.

Формализуя это подробнее, авторам статьи интересно, как этнополитическая ситуация в стране и электоральные институты (например, число мандатов) влияют на устройство партийных систем в африке.

2. Какие методы применяют авторы в статье? Проследите связь между теоретической и эмпирической частью. Насколько выбор методов соответствует поставленным задачам? Запишите спецификации моделей, тестируемые в исследовании.

Авторы строят OLS регрессию, включающую предикторы и переменные взаимодействия, чтобы протестировать 4 возможных спецификации модели при том условии, что сначала зависимой переменной выступает индекс эффективного числа партий, рассчитанный по всем партиям, участвовавшим в выборах, а затем - по победившим партиям (суммарно 8 моделей). Спецификации авторских моделей будут выглядеть так (симметричны во второй таблице при расчете с обращением ко второй зависимой переменной):

$$y_1 = \beta_0 + \beta_1 \cdot x_{i1} + \beta_2 \cdot x_{i2} + \beta_3 \cdot x_{i2} \cdot x_{i5}$$

$$y_2 = \beta_0 + \beta_1 \cdot x_{i3} + \beta_2 \cdot x_{i4} + \beta_3 \cdot x_{i4} \cdot x_{i3}$$

$$y_3 = \beta_0 + \beta_1 \cdot x_{i1} + \beta_2 \cdot x_{i2} + \beta_3 \cdot x_{i2} \cdot x_{i5} + \beta_4 \cdot x_{i3} + \beta_5 \cdot x_{i4} + \beta_6 \cdot x_{i4} \cdot x_{i3}$$

$$y_4 = \beta_0 + \beta_1 \cdot x_{i1} + \beta_2 \cdot x_{i2} + \beta_3 \cdot x_{i2} \cdot x_{i5} + \beta_4 \cdot x_{i3} + \beta_5 \cdot x_{i4} + \beta_6 \cdot x_{i4} \cdot x_{i3} \cdot x_{i1}$$

Где x_1 - District magnitude, x_2 - Proximity of Presidential and legislative elections, x_3 - Ethnopolitical Group fragmentation, x_4 - Ethnopolitical Group concentration x_5 - Effective number of Candidates.

Теоретическая часть указывает на конструктивистский подход авторов к пониманию этничности (и следовательно - этнополитических расколов), что предопределяет характер переменных, отвечающих за их отображение. Необходимо учесть тот факт, что специфицированная авторами модель подходит для анализа эффекта модерации, что в целом отвечает исследовательскому вопросу статьи (проверка совместного/раздельного влияния указанных предикторов на зависимую переменную).

3. Найдите ошибки, допущенные авторами в эмпирической части, и обоснуйте свой ответ. Если Вы нашли ошибки в спецификациях моделей, то исправьте их. Если Вы обнаружили ошибки в интерпретации, то также скорректируйте ее. Что еще не хватает в статье для полноценного анализа совместного эффекта переменных?

Так как авторы пытаются обратиться к вопросу, как именно электоральные институты и этнополитические расколы предопределяют структуру партийной системы (совместно или раздельно), перед ними встает проблема поиска подходящей зависимой переменной, отражающей партийную структуру той или иной страны. Для этого они решают применить индекс Ласло-Таагаперы, применяемый в электоральных исследованиях. Но этот индекс используется не для операционализации партийной системы: он чувствителен не к числу партий, участвующих в выборах, а к конкурентности принимающих участие партий (Туровский, 2012). Так, участие в выборах большого числа мелких партий практически не влияет на значение индекса: больше всего он изменяется от изменения пропорции голосования между участниками выборов. Это может быть критично для авторов статьи: если рассчитать ЭЧП на основе индекса Ласло-Таагаперы, а не Хуана-Молинара, берущего

поправку на наличие наибольшей партии, от исследователя ускользнут ситуации, в которых, например, две партии заняли в парламенте 90% мест и 10 % мест, и в которой 1 партия заняла 90% мест, а другие 10% заняли еще 10 партий.

В спецификации части моделей действительно допущены ошибки. (Спецификация 3й и 4й моделей является неверной, так как пропускает промежуточные переменные взаимодействия (например попарные произведения). Мы не станем полностью записывать правильную спецификацию модели для случая модерации, так как в таком контексте количество коэффициентов будет очень большим (из-за учета ВСЕХ переменных взаимодействия, которых выходит более 30, общий принцип будет включать в себя сначала поиск переменных попарного взаимодействия, потом - по 3, потом - по 4, и только потом переменную с 5), но запишем сокращенно:

$$y = \beta_0 + \beta_1 \cdot x_{i1} + \beta_2 \cdot x_{i2} + \beta_3 \cdot x_{i3} + \beta_4 \cdot x_{i4} + \beta_5 \cdot x_{i5} + \beta_6 \cdot x_{i1} \cdot x_{i2} + \dots + \beta_k \cdot x_{i1} \cdot x_{i2} \cdot x_{i3} \cdot x_{i4} \cdot x_{i5}$$

Полученные коэффициенты интерпретируются неправильно - автор интерпретирует их как marginal effect: например, он пишет "When group concentration is high, at 2.56, group fragmentation tends to increase the effective number of electoral parties by 0.04, and the effective number of legislative seats by 0.02 что неверно так как не учитывает изменения переменных взаимодействия - этот же принцип касается других интерпретаций. Более того, из указанной выше проблемы пропуска значительного числа предикторов невозможно правильно рассчитать предельные эффекты (на авторских моделях) и, соответственно, получить верные интерпретации тех коэффициентов, которые указаны авторами.

Для полноценного анализа авторам стоило рассчитать и проиллюстрировать доверительные интервалы для предельных эффектов (не говоря про расчет самих предельных эффектов). Помимо этого, в модели пропущены контрольные переменные, что не позволяет контролировать смещение эффектов и снижает долю необъясненной вариации зависимых переменных. Также было бы полезно построить корреляционную матрицу.

4. Если Вы обнаружили в статье какие-то еще ограничения (как в теоретической, так и эмпирической части), обозначьте их.

Хотелось бы указать на то, что в обеих моделях сильно меняется значимость коэффициентов регрессии (например, District magnitude сначала оказывается незначимым, а в 4й спецификации оказывается значимым при небольших значениях p) и в целом на невысокие уровни pvalue. Помимо этого - многие рассчитанные коэффициенты демонстрируют высокие значения стандартной ошибки, что также снижает их объяснительную силу.

Дополнительно стоит отметить в целом небольшой объем выборки - всего 62 кампании (и рассчитанных по им индексам эффективного числа партий). И хотя объем не критически мал, к результатам исследований на такой выборке все равно можно подойти скептически. Более того, учитывая тот факт, что страны в выборке - одни и те же, а число лет между выборами в этих странах - разное, реальностью становится возникновение скошенного распределения (ЭЧП считается от выборов) и вытекающей из него проблемы гетероскедастичности.

Другая проблема возникает, если мы говорим об операционализации переменной электоральных институтов: при исследовании сложных стран, существующих в том числе в Африке, переменная среднее число мандатов (District Magnitude) может начать сильно отклоняться в одну из сторон. Например, если есть столичный регион с 20 мандатами, и

5 мелких с 2, а в другой стране - по 15 мандатов на 2 региона, переменная "среднее число мандатов" вряд ли сможет адекватно отобразить их влияние на партийную структуру.