Федеральное государственное автономное образовательное учреждение высшего профессионального образования «Национальный исследовательский университет «Высшая школа экономики»

Факультет экономических наук

Образовательная программа «Экономика»

**Домашнее задание №1**

Часть 2

«Исследование эффекта пандемии COVID-19 на рост ВВП на душу населения»

Выполнили:

Пешков Максим БЭК181

Дусенбаев Батыржан БЭК1811

Преподаватель:

Ратникова Татьяна Анатольевна

Москва, 2021

Оглавление

[Введение и напоминание результатов первой части 4](#_Toc85653674)

[Введение 4](#_Toc85653675)

[Результаты первой части 4](#_Toc85653676)

[Переменные и их эндогенность 5](#_Toc85653677)

[Тестирование инструментов для эндогенных переменных 5](#_Toc85653678)

[Эндогенность уровня открытия бизнеса 5](#_Toc85653679)

[Тестирование релевантности 5](#_Toc85653680)

[Тестирование валидности 7](#_Toc85653681)

[Сравнительный анализ оценок 9](#_Toc85653682)

[Тесты Хаусмана для выбора моделей 9](#_Toc85653683)

[Эндогенность уровня безработицы 10](#_Toc85653684)

[Тестирование релевантости инструментов 10](#_Toc85653685)

[Тестирование валидности инструментов 11](#_Toc85653686)

[Сравнительный анализ оценок 12](#_Toc85653687)

[Тесты Хаусмана для выбора моделей 13](#_Toc85653688)

[Тестирование эндогенности уровня инфляции 14](#_Toc85653689)

[Тестирование релевантности инструментов 14](#_Toc85653690)

[Тестирование валидности инструментов 15](#_Toc85653691)

[Сравнительный анализ оценок 16](#_Toc85653692)

[Тесты Хаусмана для выбора модели 17](#_Toc85653693)

[Оценивание инвариантной переменной (количество лет среднего образования) 18](#_Toc85653694)

[Random effect модель 18](#_Toc85653695)

[FE+ BE модель (состоятельная, но не эффективная) 19](#_Toc85653696)

[Хаусман-Тейлор модель 19](#_Toc85653697)

[Тестирование на валидность инструментов 20](#_Toc85653698)

[Тестирование чувствительности модели Хаусмана-Тейлора в зависимости от различных сверхидентифицирующих условий 21](#_Toc85653699)

[Без одного слабого инструмента 21](#_Toc85653700)

[Без двух слабых инструментов 22](#_Toc85653701)

[Без одного значимого экзогенного фактора 23](#_Toc85653702)

[Без одного инвариантного по времени экзогенного регрессора 24](#_Toc85653703)

[Сравнительный анализ оценок моделей 25](#_Toc85653704)

[Динамические модели 26](#_Toc85653705)

[Ареллано-Бонд модель 26](#_Toc85653706)

[Системный GMM 27](#_Toc85653707)

[Сравнение оценок моделей при предположении экзогенности переменных 28](#_Toc85653708)

[Сравнение оценок с учетом эндогенности регрессоров 29](#_Toc85653709)

[Сравнение оценок с учетом влияния возможного годового тренда 30](#_Toc85653710)

[Тестирование валидности инструментов 31](#_Toc85653711)

[Ареллан-Бонд модели 31](#_Toc85653712)

[Системной GMM модели 31](#_Toc85653713)

[Тестирование на стационарность ошибок в лучшей модели 32](#_Toc85653714)

[Levin-Lin-Chu тесты 32](#_Toc85653715)

[Harris-Tzavalis тесты 33](#_Toc85653716)

[Im-Pesaran-Shin тесты 33](#_Toc85653717)

[Вывод по лучшей динамической модели 34](#_Toc85653718)

[Заключение 34](#_Toc85653719)

[Код 34](#_Toc85653720)

# Введение и напоминание результатов первой части

## Введение

## Результаты первой части

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|  | pool | fe | re | fe\_t | re\_t | re\_rob | pcpsar | scc\_re |
| VARIABLES | growth\_gdp | growth\_gdp | growth\_gdp | growth\_gdp | growth\_gdp | growth\_gdp | growth\_gdp | growth\_gdp |
|  |  |  |  |  |  |  |  |  |
| rule\_of\_law | 0.108 | 0.255 | 0.0544 | 0.790 | 0.0748 | 0.0748 | 0.0212 | 0.0748 |
|  | (0.213) | (1.147) | (0.306) | (1.167) | (0.314) | (0.308) | (0.238) | (0.183) |
| inflation | 0.126\*\*\* | -0.00646 | 0.0604 | -0.0487 | 0.0318 | 0.0318 | 0.120\*\* | 0.0318 |
|  | (0.0426) | (0.0581) | (0.0472) | (0.0587) | (0.0482) | (0.0708) | (0.0571) | (0.0276) |
| startbusin | 0.0295 | 0.0824 | 0.0411\* | 0.0713 | 0.0407 | 0.0407 | 0.0332 | 0.0407 |
|  | (0.0183) | (0.0540) | (0.0248) | (0.0538) | (0.0255) | (0.0301) | (0.0220) | (0.0261) |
| urban | -0.0567\*\*\* | -0.152 | -0.0587\*\*\* | -0.314 | -0.0601\*\*\* | -0.0601\*\*\* | -0.0603\*\*\* | -0.0601\*\*\* |
|  | (0.0121) | (0.219) | (0.0178) | (0.266) | (0.0181) | (0.0178) | (0.0139) | (0.0112) |
| log\_unemp | -0.654\*\*\* | -0.502 | -0.550\*\* | -0.216 | -0.482\* | -0.482 | -0.539\*\* | -0.482\*\*\* |
|  | (0.185) | (0.525) | (0.250) | (0.567) | (0.261) | (0.347) | (0.241) | (0.0552) |
| year2015 |  |  |  | -0.166 | -0.176 | -0.176 | -0.0927 | -0.176\*\*\* |
|  |  |  |  | (0.358) | (0.354) | (0.304) | (0.126) | (0.0159) |
| year2016 |  |  |  | -0.196 | -0.294 | -0.294 | -0.394\*\* | -0.294\*\*\* |
|  |  |  |  | (0.370) | (0.353) | (0.351) | (0.154) | (0.0225) |
| year2017 |  |  |  | 0.795\* | 0.546 | 0.546\* | 0.359\*\* | 0.546\*\*\* |
|  |  |  |  | (0.408) | (0.358) | (0.314) | (0.179) | (0.0560) |
| year2018 |  |  |  | 0.741 | 0.401 | 0.401 | 0.103 | 0.401\*\*\* |
|  |  |  |  | (0.453) | (0.364) | (0.340) | (0.196) | (0.0646) |
| year2019 |  |  |  | 0.0335 | -0.298 | -0.298 | -0.343\* | -0.298\*\*\* |
|  |  |  |  | (0.494) | (0.367) | (0.365) | (0.194) | (0.0753) |
| year2020 | -6.851\*\*\* | -6.955\*\*\* | -6.919\*\*\* | -6.617\*\*\* | -6.907\*\*\* | -6.907\*\*\* | -7.032\*\*\* | -6.907\*\*\* |
|  | (0.306) | (0.319) | (0.279) | (0.517) | (0.366) | (0.538) | (0.172) | (0.0696) |
| schooling | -0.262\*\* |  | -0.204 |  | -0.184 | -0.184 | -0.175\* | -0.184\*\* |
|  | (0.112) |  | (0.165) |  | (0.165) | (0.151) | (0.0939) | (0.0694) |
| dummy\_eu | 0.565\*\* |  | 0.463 |  | 0.391 | 0.391 | 0.341 | 0.391 |
|  | (0.271) |  | (0.400) |  | (0.406) | (0.443) | (0.247) | (0.390) |
| dummy\_oecd | 0.415 |  | 0.292 |  | 0.248 | 0.248 | 0.172 | 0.248 |
|  | (0.298) |  | (0.444) |  | (0.446) | (0.411) | (0.260) | (0.356) |
| Constant | 6.550\*\*\* | 6.762 | 5.233\* | 18.82 | 5.151\* | 5.151\* | 5.867\*\* | 5.151\* |
|  | (2.008) | (14.46) | (2.775) | (18.21) | (2.800) | (3.008) | (2.880) | (2.263) |
| Observations | 294 | 294 | 294 | 294 | 294 | 294 | 294 | 294 |
| R-squared | 0.693 | 0.731 |  | 0.745 |  |  | 0.759 |  |
| Number of country\_id |  | 42 | 42 | 42 | 42 | 42 | 42 |  |
| Number of groups |  |  |  |  |  |  |  | 42 |

## Переменные и их эндогенность

# Тестирование инструментов для эндогенных переменных

## Эндогенность уровня открытия бизнеса

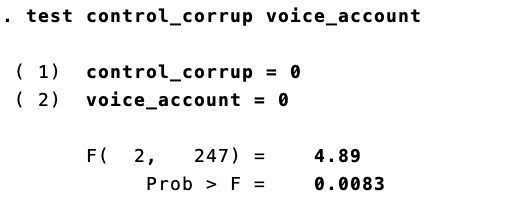
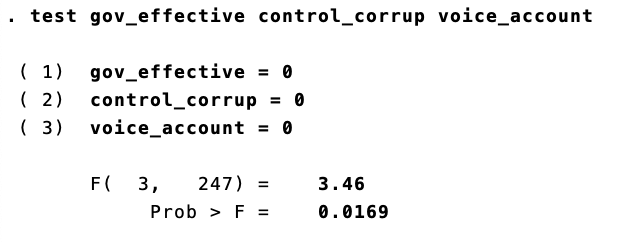
### Тестирование релевантности

Изображение выглядит как текст, квитанция, снимок экрана

Автоматически созданное описание

//F>30 - большой исходя из современных статистик

//control\_corrup voice\_account-сильные инструменты,gov\_effective-слабый

//gov\_effective control\_corrup voice\_account - инструменты

Изображение выглядит как стол

Автоматически созданное описание

### Тестирование валидности

Изображение выглядит как стол

Автоматически созданное описание

//такой набор невалидный

(дальше валидный)

Изображение выглядит как текст, снимок экрана, квитанция

Автоматически созданное описание

Изображение выглядит как текст, квитанция, снимок экрана

Автоматически созданное описание

(валидный и после re регрессии)

### Сравнительный анализ оценок

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) |
|  | fe | re | fe\_iv | re\_iv |
| VARIABLES | growth\_gdp | growth\_gdp | growth\_gdp | growth\_gdp |
|  |  |  |  |  |
| rule\_of\_law | 0.144 | 0.174 | -0.129 | -0.122 |
|  | (1.137) | (0.318) | (1.375) | (0.430) |
| urban | -0.104 | -0.0698\*\*\* | -0.309 | -0.0885\*\*\* |
|  | (0.211) | (0.0188) | (0.610) | (0.0261) |
| startbusin | 0.0755 | 0.0461\* | 0.172 | 0.149\* |
|  | (0.0532) | (0.0268) | (0.274) | (0.0861) |
| o.schooling | - |  | - |  |
|  |  |  |  |  |
| o.dummy\_eu | - |  | - |  |
|  |  |  |  |  |
| o.dummy\_oecd | - |  | - |  |
|  |  |  |  |  |
| covid\_year | -7.008\*\*\* | -6.979\*\*\* | -7.032\*\*\* | -7.170\*\*\* |
|  | (0.310) | (0.274) | (0.319) | (0.315) |
| schooling |  | -0.115 |  | 0.142 |
|  |  | (0.181) |  | (0.290) |
| dummy\_eu |  | 0.165 |  | 0.411 |
|  |  | (0.430) |  | (0.532) |
| dummy\_oecd |  | 0.137 |  | -0.280 |
|  |  | (0.496) |  | (0.658) |
| Constant | 2.967 | 4.025 | 9.898 | -5.892 |
|  | (13.77) | (3.002) | (23.81) | (8.493) |
|  |  |  |  |  |
| Observations | 294 | 294 | 294 | 294 |
| R-squared | 0.730 |  |  |  |
| Number of country | 42 | 42 | 42 | 42 |

### Тесты Хаусмана для выбора моделей

Изображение выглядит как стол

Автоматически созданное описание

Изображение выглядит как стол

Автоматически созданное описание

Везде выбираем без инструментов (модели эквивалентны)

## Эндогенность уровня безработицы

### Тестирование релевантости инструментов

Изображение выглядит как текст, квитанция, снимок экрана

Автоматически созданное описание

Изображение выглядит как текст

Автоматически созданное описание

### Тестирование валидности инструментов

Изображение выглядит как стол

Автоматически созданное описание

### Сравнительный анализ оценок

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) |
|  | fe | re | fe\_iv | re\_iv |
| VARIABLES | growth\_gdp | growth\_gdp | growth\_gdp | growth\_gdp |
|  |  |  |  |  |
| rule\_of\_law | 0.463 | 0.121 | 1.238 | -0.337 |
|  | (1.139) | (0.315) | (1.286) | (0.391) |
| urban | 0.0302 | -0.0494\*\*\* | -0.170 | -0.0194 |
|  | (0.182) | (0.0188) | (0.225) | (0.0240) |
| log\_unemp | -0.378 | -0.519\* | -3.193\* | -1.775\*\*\* |
|  | (0.511) | (0.266) | (1.705) | (0.632) |
| o.schooling | - |  | - |  |
|  |  |  |  |  |
| o.dummy\_eu | - |  | - |  |
|  |  |  |  |  |
| o.dummy\_oecd | - |  | - |  |
|  |  |  |  |  |
| covid\_year | -6.943\*\*\* | -6.883\*\*\* | -6.601\*\*\* | -6.861\*\*\* |
|  | (0.316) | (0.270) | (0.388) | (0.279) |
| schooling |  | -0.288\* |  | -0.432\*\* |
|  |  | (0.167) |  | (0.188) |
| dummy\_eu |  | 0.275 |  | 0.809 |
|  |  | (0.428) |  | (0.512) |
| dummy\_oecd |  | 0.408 |  | 0.610 |
|  |  | (0.472) |  | (0.506) |
| Constant | 0.142 | 9.108\*\*\* | 19.61 | 10.61\*\*\* |
|  | (13.73) | (1.511) | (18.35) | (1.735) |
|  |  |  |  |  |
| Observations | 294 | 294 | 294 | 294 |
| R-squared | 0.728 |  |  |  |
| Number of country | 42 | 42 | 42 | 42 |
|  |  |  |  |  |

### Тесты Хаусмана для выбора моделей

Изображение выглядит как стол

Автоматически созданное описаниеИзображение выглядит как стол

Автоматически созданное описание

## Тестирование эндогенности уровня инфляции

### Тестирование релевантности инструментов

Изображение выглядит как текст, снимок экрана, квитанция

Автоматически созданное описание

Изображение выглядит как стол

Автоматически созданное описание

### Тестирование валидности инструментов

Изображение выглядит как стол

Автоматически созданное описание

### Сравнительный анализ оценок

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) |
|  | fe | re | fe\_iv | re\_iv |
| VARIABLES | growth\_gdp | growth\_gdp | growth\_gdp | growth\_gdp |
|  |  |  |  |  |
| rule\_of\_law | 0.353 | 0.382 | 0.715 | 0.845 |
|  | (1.134) | (0.301) | (1.229) | (0.546) |
| urban | 0.0588 | -0.0649\*\*\* | -0.0564 | -0.0849\*\*\* |
|  | (0.180) | (0.0176) | (0.214) | (0.0290) |
| inflation | -0.00425 | 0.0566 | 0.277 | 0.431 |
|  | (0.0570) | (0.0480) | (0.252) | (0.295) |
| o.schooling | - |  | - |  |
|  |  |  |  |  |
| o.dummy\_eu | - |  | - |  |
|  |  |  |  |  |
| o.dummy\_oecd | - |  | - |  |
|  |  |  |  |  |
| covid\_year | -6.993\*\*\* | -6.853\*\*\* | -6.715\*\*\* | -6.594\*\*\* |
|  | (0.316) | (0.274) | (0.410) | (0.357) |
| schooling |  | -0.239 |  | -0.309 |
|  |  | (0.161) |  | (0.230) |
| dummy\_eu |  | 0.123 |  | 0.578 |
|  |  | (0.410) |  | (0.670) |
| dummy\_oecd |  | 0.363 |  | 0.620 |
|  |  | (0.464) |  | (0.674) |
| Constant | -2.587 | 8.550\*\*\* | 4.885 | 8.976\*\*\* |
|  | (13.37) | (1.455) | (15.45) | (2.055) |
|  |  |  |  |  |
| Observations | 294 | 294 | 294 | 294 |
| R-squared | 0.728 |  |  |  |
| Number of country | 42 | 42 | 42 | 42 |

### Тесты Хаусмана для выбора модели

Изображение выглядит как стол

Автоматически созданное описаниеИзображение выглядит как стол

Автоматически созданное описание

# Оценивание инвариантной переменной (количество лет среднего образования)

В нашей работе мы будем исследовать эффект переменной – количества лет среднего образования, являющейся инвариантной по времени для каждой страны. Мы считаем эту переменную эндогенной, то есть коррелированной с индивидуальными эффектами стран, так как на количество лет влияет внутреннее законодательство стран, которые определяют это количество самостоятельно. Рассмотрим несколько моделей для определения этого эффекта.

## **Random effect модель**

Оценим модель со случайным индивидуальным эффектом. В этой модели мы предполагаем некоррелированность индивидуальных эффектов с объясняющими переменными. Мы будем учитывать все доступные нам изменяющиеся и инвариантные по времени регрессоры.

Изображение выглядит как стол

Автоматически созданное описание

Рассмотрим полученной оценки регрессии. показывает, насколько модель хорошо объясняет изменения зависимой переменной в каждой стране на протяжении наблюдаемого периода. Для нашей модели значение = 0.7311 за 7 лет. – насколько модель хорошо объясняет изменения зависимой переменной между всеми странами за 7 лет и составляет 0.5703. является взвешенным средним между двумя предыдущими и равен примерно 69%.

Мы предполагаем, что корреляция ошибок и регрессоров равна нулю. Также можно заметить, что неоднородность временных тактов выше неоднородности стран в 2 раза (). 21% разброса приходится на индивидуальные эффекты.

Как видно из таблицы для модели Random effect коэффициент перед не значим и равен -0.09. Такая незначимость может быть обоснована тем, что в модели предполагается некоррелированность регрессоров и ошибок. Далее предлагается оценить влияние в 2 итерации, в которой нет необходимости в этой строгой предпосылке.

## **FE+ BE модель (состоятельная, но не эффективная оценка)**

В первой итерации мы построили FE регрессию изменяющихся переменных на экономический рост стран за 7 лет. После чего сохранили остатки и оценили between регрессию инвариантных регрессоров на полученные остатки. Данная модель дает состоятельные, но неэффективные оценки, а также является аналогом FE модели, позволяющая оценить влияние инвариантных переменных.

Изображение выглядит как текст, снимок экрана, квитанция

Автоматически созданное описание

Как видно на рисунке оцененная регрессия Between имеет не настолько сильную объясняющую силу, в виду того, что большая значимость регрессоров приходится на изменяющиеся переменные.

Тем не менее, здесь можно определить влияние , где коэффициент перед интересующим регрессором равен 0.93, причем он является значимым. Стандартная ошибка для коэффициента перед равна 0.36, что больше, чем было в RE модели, это же косвенно подтверждает и то, что оценки этой модели состоятельны, но не эффективны.

Таким образом, эта модель позволила оценить влияние инвариантной переменной, однако она не учитывает возможную корреляцию этой переменной с индивидуальным эффектом, из-за чего оценки являются неэффективными. Для получения более корректных оценок воспользуемся моделью Хаусмана-Тейлора.

## Хаусмана-Тейлора модель

Данная модель позволяет инструментировать как экзогенные, так и эндогенные переменные, а также получать оценку для инвариантных по времени регрессоров. В нашей модели мы использовали в качестве эндогенных меняющихся во времени регрессоров те, которые рассматривали ранее , а также в качестве экзогенных как изначально экзогенные переменные, так и инструменты к эндогенным, которые мы также тестировали на релевантность и валидность ранее. Также в нашей модели в качестве эндогенной инвариантной по времени переменной использовали , а в качестве экзогенных использовали дамми-переменные на принадлежность к Евросоюзу, а также к ОЭСР, так как при внешние шоки слабо влияют на вход или выход и принадлежность к этим организациям.

Изображение выглядит как стол

Автоматически созданное описание

Данная модель предполагает независимость и одинаковое распределение ошибок, а также аналогично RE позволяет оценивать разброс по индивидуальным наблюдениям и временным тактам, которые равны 3.4 и 1.5, показывающее больший разброс по странам, чем по времени (что отличается от выводов по RE модели). При этом 82% приходится на индивидуальные эффекты.

Оцененная регрессия позволила посчитать эффект инвариантной переменной на экономический рост, который равен 0.93 и не является значимым. При этом стандартная ошибка равна 1.09, что является достаточно большим показателем и может быть вызвано неправильно подобранными инструментами, поэтому необходимо провести тестирование на их валидность.

### Тестирование на валидность инструментов

Для проверки валидности инструментов в модели Хаусмана-Тейлора будем использовать тест Саргана-Хансена, где нулевая гипотеза заключается в валидности инструментов. Тестовую статистику можно наблюдать ниже.

Изображение выглядит как текст

Автоматически созданное описание

P-value статистики Саргана-Хансена является высоким показателем и для любого адекватного уровня значимости нулевая гипотеза не отвергается, значит инструменты валидны.

Тем не менее, оценка является незначимой и менее эффективной, чем модель FE+BE, что может быть вызвано слабой релевантностью инструментов, поэтому далее будет проведен анализ чувствительности модели Хаусмана-Тейлора в зависимости от различных сверхиндентифирующих условий, то есть в зависимости от различного набора инструментов.

## Исследование чувствительности модели Хаусмана-Тейлора в зависимости от различных сверхидентифицирующих условий

Для исследования чувствительности оценки Хаусмана-Тейлора от набора инструментов предлагается рассмотреть следующие варианты: убрать один и два слабых изменяющихся инструмента, убрать один значимых экзогенный инструмент, убрать один инвариантный по времени инструмент. После каждого варианта будут рассмотрены оценки перед и проведены тестирования на валидности инструментов.

### Без одного слабого инструмента

В этом случае мы убрали , так как мы выявили, что это был слабый и нерелевантный инструмент для изменяющихся по времени эндогенных переменных.

Изображение выглядит как стол

Автоматически созданное описание

При оценивании с таким набором инструментов видно, что стандартная ошибка перед выросла до 1.42, а также вырос и сам коэффициент до 1.57, что повлияло на значимость и p-value уменьшилось до 0.268, но этот коэффициент также является незначимым. Проведем тестирование на валидность такого набора инструмента.

Изображение выглядит как текст

Автоматически созданное описание

Как можно видеть набор инструментов так же валиден, и нулевая гипотеза в Тесте Саргана не отвергается. Таким образом, уменьшение набора экзогенных переменных на один слабый инструмент не сильно меняет оценки коэффициентов и их значимость.

### Без двух слабых инструментов

Изображение выглядит как стол

Автоматически созданное описание

В данном случае мы убрали , так как они менее всего были значимыми и сильными инструментами для меняющихся по времени эндогенных переменных. Результаты, представленные выше, не сильно отличаются от предыдущих, а именно коэффициент перед равен 1.55, а стандартная ошибка равна 1.43, такой результат показывает, что и этот инструмент был слабым и позволял улучшить оценку. Однако проверим такой набор инструментов на валидность благодаря тесту Саргана.

Изображение выглядит как текст

Автоматически созданное описание

Как можно видеть из приведенной статистики следует, что нулевая гипотеза не отвергается и инструменты валидны, что было ожидаемо, ведь удаление слабых инструментов привносит небольшой вклад в получение оценок, поэтому предлагается изучить эффект от удаления сильного инструмента.

### Без одного значимого экзогенного фактора

В данной модели была убрана дамми-переменная, отвечающая за 2020 год, во время которого из-за пандемии COVID-19 экономический рост в различных странах снизился на несколько процентов. Во всех моделях эта переменная была значимой и помогала объяснить сильное падение в 2020 году. Фактически эта переменная отображает экзогенный шок и поэтому является априори экзогенной и коррелирует с различными показателями, так как в 2020 году не только экономический рост снизился, но и выросла инфляция и безработица, из-за чего использование этой переменной как инструмент для эндогенных позволит более верно оценить коеффициенты и эта дамми будет являться сильным инструментом, ведь сможет объяснить большую часть разброса, вызванного пандемией. Результаты модели Хаусмана-Тейлора, если исключить этот инструмент, представлены ниже.

Изображение выглядит как стол

Автоматически созданное описание

Исходя из результатов таблицы видно, что отсутствие сильного инструмента сильно влияет на оценку коэффициента перед , а именно увеличивает его в 17 раз, а стандартную ошибку увеличивает до 6.7. При это коэффициент является значимым на уровне значимости 1% и показывает, что увеличение количества лет в средней школе на 1.7 лет увеличивает рост на 0.1%, что кажется не таким сильным эффектом, но вполне ожидаемым. Тем не менее, из-за того, что оценки сильно изменились, то предлагается теперь проверить валидность инструментов, которые используются в модели.

Изображение выглядит как текст

Автоматически созданное описание

Тест Саргана-Хансена на валидность инструментов показывает, что нулевая гипотеза отвергается на уровне значимости в 10%, хотя и не отвергается на уровне значимости в 1% и 5%. Таким образом, такой набор слабо валидный, что и ведет к искажению оценок.

Тем не менее, использование такой модели показало, что при удалении сильного инструмента оценки качественно меняются и могут привести к потери валидности вовсе, из-за чего будет неверно использовать такую модель. Проверим, будут ли такие же выводы, если удалить инвариантный экзогенный регрессор (инструмент для эндогенной инвариантной переменной).

### Без одного инвариантного по времени экзогенного регрессора

В этом случае мы предлагаем исследовать как изменятся коэффициенты, если убрать одну из предполагаемых экзогенных инвариантных переменных – дамми на принадлежность к ОЭСР. И хотя мы предполагаем, что эта переменная экзогенна, но есть различные причины предполагать, что она эндогенна, так как в ОЭСР входят более развитые страны, поэтому этот показатель может коррелировать с индивидуальным эффектом и не являться хорошим инструментом. Результаты полученной модели представлены ниже.

Изображение выглядит как стол

Автоматически созданное описание

Как можно видеть, оценки коэффициента перед не сильно изменилась, до того, как мы не убирали дамми-переменную на вхождение в ОЭСР, а именно коэффициент равен 1.3, а стандартная ошибка 1.33. Такой результат может быть вызван тем, что удаленная переменная была слабым инструментом. Тем не менее, нужно также проверить валидность инструментов благодаря тесту Саргана.

Изображение выглядит как текст

Автоматически созданное описание

Статистика теста Саргана-Хансена показывает, что нулевая гипотез не отвергается, значит, инструменты валидны.

Таким образом, удаление слабого инвариантного экзогенного фактора слабо влияет на эффективность оценки в модели Хаусмана-Тейлора, а также не влияет на валидность инструментов.

## Сравнительный анализ оценок моделей

Ниже представлена общая таблица с оценками коэффициентом при использовании различных методов подсчета влияния инвариантного регрессора в панельных данных с учетом возможной его корреляции с индивидуальными эффектами.

Из таблицы видно, что оценки для в модели FE+BE и модели Хаусмана-Тейлора со всеми инструментами похожи, хотя она не является самой эффективной. При этом модель RE является самой эффективной, при которой стандартная ошибка наименьшая, что подтверждает наш выбор этой модели в первой части домашнего задания. При тестировании чувствительности модели Хаусмана-Тейлора к выбору инструментов обнаружилось, что удаление слабых инструментов слабо влияет как на оценку, так и на эффективность модели, при этом набор инструментов был всегда валиден. Однако при удалении сильно инструмента набор становился слабо валидным, а оценка и стандартная ошибка выросли в 17 и 6 раз, что показывает важность выбора инструментов для модели Хаусмана-Тейлора.

Таким образом, можно заключить, что наиболее подходящая модель для оценивания влияния количества лет в средней школе для наших данных является RE модель, являющаяся наиболее эффективной, хотя коэффициент и является менее значимым. Модель Хаусмана-Тейлора является хорошим способом для оценки эффекта инвариантной переменной, но требует тщательного подбора инструментов, которые в нашей работе являются несильно релевантными и поэтому она показала менее эффективные оценки.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|  | re1 | be | ht4 | ht5 | ht1 | ht2 | ht3 |
| VARIABLES | growth\_gdp | residfe | growth\_gdp | growth\_gdp | growth\_gdp | growth\_gdp | growth\_gdp |
|  |  |  |  |  |  |  |  |
| rule\_of\_law | -0.262 |  | 3.132 | -0.347 | -0.318 | 0.322 | 0.357 |
|  | (0.926) |  | (2.201) | (1.153) | (1.157) | (1.164) | (1.146) |
| urban | -0.0629\*\*\* |  | -1.641\*\*\* | -0.136 | -0.109 | -0.131 | -0.131 |
|  | (0.0188) |  | (0.275) | (0.0869) | (0.0702) | (0.0819) | (0.0822) |
| startbusin | 0.0173 |  | -0.0141 | 0.0631 | 0.0572 | 0.0605 | 0.0616 |
|  | (0.0293) |  | (0.0864) | (0.0472) | (0.0458) | (0.0463) | (0.0463) |
| inflation | 0.0485 |  | 0.167\* | -0.0103 | -0.00893 | -0.0100 | -0.00845 |
|  | (0.0499) |  | (0.0963) | (0.0555) | (0.0554) | (0.0559) | (0.0552) |
| log\_unemp | -0.312 |  | -2.141\*\* | -0.304 | -0.279 | -0.536 | -0.547 |
|  | (0.294) |  | (0.905) | (0.510) | (0.507) | (0.506) | (0.502) |
| regulatory\_qual | 1.305\* |  | 2.242 | 1.237 | 1.237 |  |  |
|  | (0.670) |  | (1.964) | (1.077) | (1.067) |  |  |
| voice\_account | -0.570 |  | -0.867 | -1.350 | -1.117 | -1.100 | -1.018 |
|  | (0.577) |  | (2.584) | (1.210) | (1.155) | (1.195) | (1.121) |
| pol\_stability | 0.195 |  | -0.253 | 0.296 | 0.204 | 0.111 |  |
|  | (0.391) |  | (1.193) | (0.628) | (0.616) | (0.625) |  |
| gov\_effective | 0.172 |  | 5.963\*\*\* | 1.976\* | 1.897\* | 2.502\*\* | 2.504\*\* |
|  | (0.902) |  | (1.976) | (1.093) | (1.084) | (1.087) | (1.087) |
| control\_corrup | -0.394 |  | -2.806 | -0.922 | -0.960 | -0.942 | -0.962 |
|  | (0.668) |  | (1.984) | (0.983) | (0.963) | (0.992) | (0.983) |
| schooling | -0.0910 | 0.931\*\* | 17.68\*\*\* | 1.309 | 0.907 | 1.574 | 1.550 |
|  | (0.173) | (0.360) | (6.765) | (1.338) | (1.093) | (1.421) | (1.438) |
| dummy\_eu | 0.553 | 1.617 | 13.77 | 1.510 | 1.406 | 2.090 | 2.092 |
|  | (0.429) | (1.048) | (12.67) | (1.547) | (1.499) | (1.801) | (1.802) |
| dummy\_oecd | 0.165 | 1.687 | -1.750 |  | -0.797 | -1.186 | -1.230 |
|  | (0.512) | (1.059) | (12.71) |  | (1.606) | (1.845) | (1.822) |
| year2020 | -6.866\*\*\* |  |  | -6.897\*\*\* | -6.904\*\*\* | -6.876\*\*\* | -6.875\*\*\* |
|  | (0.281) |  |  | (0.280) | (0.279) | (0.285) | (0.285) |
| Constant | 5.668\*\* | -11.25\*\*\* | -64.67 | -7.972 | -4.904 | -9.863 | -9.728 |
|  | (2.887) | (3.702) | (63.94) | (11.14) | (9.400) | (11.57) | (11.60) |
|  |  |  |  |  |  |  |  |
| Observations | 294 | 294 | 294 | 294 | 294 | 294 | 294 |
| R-squared |  | 0.308 |  |  |  |  |  |
| Number of country | 42 | 42 | 42 | 42 | 42 | 42 | 42 |

# Динамические модели

Роль экономического роста в предыдущем году является значимой, ведь он отражает в себе уровни экономического развития в прошлых периодах, которые влияют на сегодняшний рост. Именно поэтому для оценивания факторов, влияющих на экономический рост необходимо использовать динамические модели. В нашей работе такими будут модели Ареллано-Бонда и системного GMM, которые также могут учитывать эндогенность регрессоров, и в конце будет проведен сравнительный анализ оценок различных моделей.

## Ареллано-Бонд модель

Модель Ареллано-Бонда использует двух ступенчатую процедуру GMM, благодаря которому можно также учитывать эндогенность коэффициентов. Для начала предлагается оценить модель без учета эндогенности, предполагая экзогенность регрессоров, полученные результаты представлены ниже.

Как можно заметить в таблице обозначено количество инструментов равное 26, отображающий тот факт, что при подсчете ковариационной матрицы используются как сами переменные, так и их лаги первого порядка, а также лаги первого и второго порядка зависимой переменной. Также из этой таблицы видно, что уменьшилось количество лет наблюдений – стало 5 лет, что является последствием применения лагов второго порядка и уменьшения количества лет на 2 года.

Если смотреть на коэффициент перед лагом экономического роста, то он равен 0.1, хотя и незначим. Это является хорошим результатом, так как проблема единичного корня становится менее опасной. Добавление лага в целом слабо изменило значимость и оценку коэффициентов регрессоров, что показывает наличие зависимости экономического роста от предыдущего показателя и подтверждает необходимость использования динамических моделей.

**Изображение выглядит как стол

Автоматически созданное описание**

## Системный GMM (Метод Бланделла-Бонда)

Этот метод является некоторым усовершенствованием метода Ареллано-Бонда, учитывающий возможную автокорреляцию ошибок. Учитывая результаты первой части домашнего задания, мы обнаружили в наших данных автокорреляцию, поэтому предполагаем, что этот метод будет более подходящим. Его называют системным, так как он учитывает возможное отличие в уровнях и помогает лучше учитывать изменения в индивидуальных эффектах. Результаты модели при использовании стандартных инструментов – лагов переменных, представлена ниже.

Из этой таблицы можно видеть, что увеличилось количество инструментов до 31, что связано с оценивание двух уравнений в системе и необходимости использования больших инструментов в зависимости от уровня. А также можно заметить, что для этой процедуры достаточно использовать только лаг первого порядка и поэтому количество лет равно 6. Причем инструменты для уравнения в уровнях – лаг разности экономического роста, а для уравнения в разностях – первые разности всех регрессоров.

Оценка коэффициента перед лагом экономического роста выросла до 0.21 и стала значима, так как стандартные ошибки уменьшились относительно модели Ареллано-Бонда. Оценка меньше 1 является хорошим признаком отсутствия наличия единичного корня, что позволяет более правильно оценивать краткосрочные и долгосрочные эффекты. В целом многие коэффициенты в этой модели стали значимыми, а также уменьшились стандартные ошибки, что говорит о том, что эта модель более эффективна.

**Изображение выглядит как стол

Автоматически созданное описание**

## Сравнение оценок моделей при предположении экзогенности переменных

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) | (5) |
|  | pool | fe | re | ab | bb |
| VARIABLES | growth\_gdp | growth\_gdp | growth\_gdp | growth\_gdp | growth\_gdp |
|  |  |  |  |  |  |
| rule\_of\_law | 0.450 | -0.245 | -0.262 | -3.927\*\*\* | -3.465\*\*\* |
|  | (0.799) | (1.371) | (0.926) | (0.888) | (0.930) |
| urban | -0.0613\*\*\* | -0.249 | -0.0629\*\*\* | -0.367\* | -0.0465 |
|  | (0.0126) | (0.222) | (0.0188) | (0.223) | (0.0439) |
| startbusin | 0.0101 | 0.0728 | 0.0173 | -0.0430 | 0.00991 |
|  | (0.0225) | (0.0554) | (0.0293) | (0.0343) | (0.0418) |
| inflation | 0.101\*\* | -0.0136 | 0.0485 | 0.0100 | -0.00605 |
|  | (0.0464) | (0.0594) | (0.0499) | (0.0343) | (0.0296) |
| log\_unemp | -0.468\*\* | -0.445 | -0.312 | 0.555 | 1.561\*\*\* |
|  | (0.225) | (0.562) | (0.294) | (0.700) | (0.537) |
| regulatory\_qual | 1.422\*\*\* | 0.927 | 1.305\* | 3.384\*\*\* | 3.560\*\*\* |
|  | (0.500) | (1.204) | (0.670) | (1.084) | (0.956) |
| voice\_account | -0.468 | -1.947 | -0.570 | 3.975\*\* | 3.577\*\*\* |
|  | (0.418) | (1.601) | (0.577) | (1.770) | (1.291) |
| pol\_stability | 0.0646 | 0.194 | 0.195 | 1.022 | 1.455\*\* |
|  | (0.298) | (0.732) | (0.391) | (0.806) | (0.665) |
| gov\_effective | -0.889 | 2.479\*\* | 0.172 | 2.074\*\* | 1.317 |
|  | (0.795) | (1.223) | (0.902) | (1.001) | (0.993) |
| control\_corrup | -0.320 | -1.211 | -0.394 | -1.445 | -2.638\*\*\* |
|  | (0.527) | (1.239) | (0.668) | (0.883) | (0.930) |
| schooling | -0.166 |  | -0.0910 |  |  |
|  | (0.118) |  | (0.173) |  |  |
| dummy\_eu | 0.619\*\* |  | 0.553 |  |  |
|  | (0.288) |  | (0.429) |  |  |
| dummy\_oecd | 0.231 |  | 0.165 |  |  |
|  | (0.354) |  | (0.512) |  |  |
| year2020 | -6.812\*\*\* | -6.822\*\*\* | -6.866\*\*\* | -6.376\*\*\* | -6.729\*\*\* |
|  | (0.305) | (0.323) | (0.281) | (0.334) | (0.286) |
| L.growth\_gdp |  |  |  | 0.101 | 0.216\*\* |
|  |  |  |  | (0.118) | (0.0935) |
| Constant | 7.189\*\*\* | 14.20 | 5.668\*\* |  |  |
|  | (2.099) | (14.71) | (2.887) |  |  |
|  |  |  |  |  |  |
| Observations | 294 | 294 | 294 | 210 | 252 |
| R-squared | 0.703 | 0.739 |  |  |  |
| Number of country |  | 42 | 42 | 42 | 42 |

Выше представленная таблица показывает оценки Pooled, FE, RE, Arellano-Bond, Bladella-Bond моделей. Сравнивая эти модели между собой можно заметить, что использование динамических моделей позволяет обнаружить больше значимых коэффициентов, а также учитывать влияние лага зависимой переменной. Причем более эффективной моделью является метод системного GMM, так как в нем у большинства переменных стандартные ошибки стали меньше, в том числе и у лага первого порядка экономического роста, сделав его значимым. При этом значения коэффициентов у динамических моделей в 3 раза выше, чем у панельных моделей. Такой эффект может быть связан с тем, что учитывает автокорреляция и лаг экономического роста. Тем не менее, для года пандемии оценки остаются практически одинаковыми и значимыми, что подтверждает правильность оценивания в динамических моделях. Однако если сравнивать RE и метод системного GMM, то первый является более эффективным, что возможно связано с учетом индивидуальных эффектов и эффектов инвариантных переменных, что уменьшает проблему эндогенности. Действительно, в предыдущих разделах мы показали, что проблема эндогенности свойственна нашим данным, а в динамических моделях можно учитывать это, что позволит получить лучше оценки.

## Сравнение оценок с учетом эндогенности регрессоров

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|  | fe | ab | ivab | bb | ivbb | ivabr | ivbbr |
| VARIABLES | growth\_gdp | growth\_gdp | growth\_gdp | growth\_gdp | growth\_gdp | growth\_gdp | growth\_gdp |
|  |  |  |  |  |  |  |  |
| rule\_of\_law | -0.245 | -3.927\*\*\* | -8.330\*\*\* | -3.465\*\*\* | -1.248 | -8.330 | -1.248 |
|  | (1.371) | (0.888) | (2.498) | (0.930) | (0.950) | (5.192) | (5.950) |
| urban | -0.249 | -0.367\* | -0.364 | -0.0465 | -0.00690 | -0.364 | -0.00690 |
|  | (0.222) | (0.223) | (0.245) | (0.0439) | (0.0201) | (0.458) | (0.363) |
| startbusin | 0.0728 | -0.0430 | 0.0182 | 0.00991 | -0.0157 | 0.0182 | -0.0157 |
|  | (0.0554) | (0.0343) | (0.0239) | (0.0418) | (0.0192) | (0.0723) | (0.435) |
| inflation | -0.0136 | 0.0100 | -0.0117 | -0.00605 | 0.00631 | -0.0117 | 0.00631 |
|  | (0.0594) | (0.0343) | (0.0212) | (0.0296) | (0.0145) | (0.0811) | (0.256) |
| log\_unemp | -0.445 | 0.555 | -0.0816 | 1.561\*\*\* | 1.110\*\*\* | -0.0816 | 1.110 |
|  | (0.562) | (0.700) | (0.491) | (0.537) | (0.282) | (1.136) | (6.127) |
| regulatory\_qual | 0.927 | 3.384\*\*\* | 2.828\*\*\* | 3.560\*\*\* | 2.854\*\*\* | 2.828 | 2.854 |
|  | (1.204) | (1.084) | (0.985) | (0.956) | (0.548) | (2.055) | (6.986) |
| voice\_account | -1.947 | 3.975\*\* | 1.875 | 3.577\*\*\* | -1.921\*\* | 1.875 | -1.921 |
|  | (1.601) | (1.770) | (1.441) | (1.291) | (0.785) | (2.828) | (17.69) |
| pol\_stability | 0.194 | 1.022 | 0.337 | 1.455\*\* | 0.293 | 0.337 | 0.293 |
|  | (0.732) | (0.806) | (0.841) | (0.665) | (0.357) | (1.762) | (6.301) |
| gov\_effective | 2.479\*\* | 2.074\*\* | 3.376\*\*\* | 1.317 | 3.953\*\*\* | 3.376\* | 3.953 |
|  | (1.223) | (1.001) | (0.877) | (0.993) | (1.019) | (1.973) | (6.036) |
| control\_corrup | -1.211 | -1.445 | -1.631 | -2.638\*\*\* | -2.784\*\*\* | -1.631 | -2.784 |
|  | (1.239) | (0.883) | (1.493) | (0.930) | (0.743) | (2.977) | (13.69) |
| year2020 | -6.822\*\*\* | -6.376\*\*\* | -6.715\*\*\* | -6.729\*\*\* | -6.847\*\*\* | -6.715\*\*\* | -6.847\*\*\* |
|  | (0.323) | (0.334) | (0.139) | (0.286) | (0.113) | (0.397) | (0.777) |
| L.growth\_gdp |  | 0.101 | -0.0304 | 0.216\*\* | 0.218\*\*\* | -0.0304 | 0.218 |
|  |  | (0.118) | (0.0478) | (0.0935) | (0.0355) | (0.171) | (0.571) |
| Constant | 14.20 |  |  |  |  |  |  |
|  | (14.71) |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
| Observations | 294 | 210 | 210 | 252 | 252 | 210 | 252 |
| R-squared | 0.739 |  |  |  |  |  |  |
| Number of country | 42 | 42 | 42 | 42 | 42 | 42 | 42 |

В приведенной таблице сравниваются оценки моделей FE, Arellano-Bond, Системного GMM, а также Arellano-Bond и Системного GMM с учетом эндогенности переменных, которые мы исследовали в прошлых разделах нашей работы, и с использованием робастных стандартных ошибок.

Из таблицы видно, что динамические модели лучше, чем FE, показывая более эффективные и значимые оценки. Также для каждого из метода видно, что использование инструментов ведет к более эффективным оценкам, а также увеличению значимых коэффициентов. При этом в каждом из методов коэффициент перед лагом первого порядка экономического роста меньше одного, что также является хорошим признаком. Тем не менее, значимые оценки для лага наблюдаются только при использовании системного GMM как с учетом, так и без учета эндогенности переменных (более верное использование инструментов). При этом все методы показывают почти одинаковые оценки и высокую значимость коэффициента перед годом пандемии. Тем не менее, для других коэффициентов наибольшая значимость наблюдается при использовании инструментированного системного GMM. Однако получившиеся коэффициенты являются противоречивыми, например, уровень безработицы положительно значим для уровня экономического роста. Такие противоречивые коэффициенты могут быть вызваны слабо релевантным набором инструментов, которые мы используем в нашей работе. Однако при использовании робастных ошибок наблюдаются логичные коэффициенты. Тем не менее, использование робастных ошибок привело к тому, что пропала значимость практически у всех коэффициентов, кроме года пандемии. При этом использование робастных ошибок ведет к уменьшению эффективности и стандартные ошибки растут в несколько раз. Такое поведение может быть вызвано тем, что робастные ошибки могут учитывать возможные «аутлаеры», которые могут быть вызваны в том числе кризисом во время пандемии, из-за чего и увеличились стандартные ошибки.

Таким образом, использование инструментов в динамических моделях для учета эндогенности позволяет получить более эффективные оценки, а использование робастных ошибок наоборот ведет к уменьшению эффективности. При этом лучшей моделью является метод системного GMM с использованием инструментов. Однако эти модели не учитывают линейный тренд экономического роста в странах, который мы выявили в первой части домашнего задания и использовали годовые дамми-переменные, являющиеся значимыми в регрессиях, поэтому для учета этого эффекта включим константу года в наши модели и проанализируем изменения.

## 

## Сравнение оценок с учетом влияния возможного годового тренда

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|  | fe | ab | ivab | bb | ivbb | ivabr | ivbbr |
| VARIABLES | growth\_gdp | growth\_gdp | growth\_gdp | growth\_gdp | growth\_gdp | growth\_gdp | growth\_gdp |
|  |  |  |  |  |  |  |  |
| rule\_of\_law | 0.0207 | -4.172\*\*\* | -6.503\*\* | -2.189\*\*\* | 1.003 | -6.503 | 1.003 |
|  | (1.395) | (0.874) | (2.777) | (0.795) | (1.846) | (20.19) | (5.106) |
| urban | -0.416 | -0.965\*\*\* | -0.294 | -0.261\*\*\* | -0.0538\* | -0.294 | -0.0538 |
|  | (0.276) | (0.364) | (0.485) | (0.0501) | (0.0317) | (5.751) | (0.356) |
| startbusin | 0.0649 | -0.0311 | -0.0217 | -0.0633\* | -0.0996\*\*\* | -0.0217 | -0.0996 |
|  | (0.0559) | (0.0382) | (0.0276) | (0.0338) | (0.0196) | (0.817) | (0.895) |
| inflation | -0.0190 | -0.00292 | -0.0486\*\*\* | -0.0173 | -0.0329\* | -0.0486 | -0.0329 |
|  | (0.0597) | (0.0350) | (0.0182) | (0.0279) | (0.0189) | (0.358) | (0.264) |
| log\_unemp | -0.207 | 0.691 | 1.258\*\*\* | -0.291 | 1.938\*\*\* | 1.258 | 1.938 |
|  | (0.609) | (0.727) | (0.428) | (0.508) | (0.453) | (4.904) | (6.867) |
| regulatory\_qual | 0.867 | 3.107\*\*\* | 1.993\* | 2.678\*\*\* | 2.375\*\*\* | 1.993 | 2.375 |
|  | (1.205) | (1.107) | (1.073) | (0.662) | (0.723) | (9.557) | (4.225) |
| voice\_account | -1.986 | 5.555\*\*\* | 2.719 | 2.561\* | -1.821\*\* | 2.719 | -1.821 |
|  | (1.601) | (1.539) | (2.108) | (1.328) | (0.799) | (22.54) | (13.74) |
| pol\_stability | 0.354 | 1.085 | -0.0797 | 0.0659 | 0.409 | -0.0797 | 0.409 |
|  | (0.749) | (0.801) | (1.329) | (0.713) | (0.690) | (14.00) | (8.597) |
| gov\_effective | 2.453\*\* | 1.855\* | 3.024\*\* | 2.259\*\* | 3.401\*\*\* | 3.024 | 3.401 |
|  | (1.223) | (1.014) | (1.316) | (1.018) | (1.109) | (12.65) | (6.577) |
| control\_corrup | -1.264 | -1.888\* | -2.789\* | -2.397\*\*\* | -3.713\*\*\* | -2.789 | -3.713 |
|  | (1.240) | (0.984) | (1.557) | (0.864) | (1.188) | (19.90) | (18.08) |
| year | 0.0986 | 0.296\*\* | 0.257\*\* | 0.0983\* | 0.241\*\*\* | 0.257 | 0.241 |
|  | (0.0965) | (0.115) | (0.104) | (0.0538) | (0.0632) | (1.313) | (0.791) |
| covid\_year | -7.019\*\*\* | -6.744\*\*\* | -7.419\*\*\* | -6.287\*\*\* | -7.303\*\*\* | -7.419\*\*\* | -7.303\*\*\* |
|  | (0.376) | (0.358) | (0.155) | (0.306) | (0.195) | (2.096) | (1.793) |
| L.growth\_gdp |  | 0.122 | -0.103\*\* | 0.0992 | 0.158\*\* | -0.103 | 0.158 |
|  |  | (0.118) | (0.0526) | (0.0875) | (0.0706) | (0.646) | (0.446) |
| Constant | -172.0 | -524.1\*\* | -493.8\*\*\* | -173.3 | -477.0\*\*\* | -493.8 | -477.0 |
|  | (182.8) | (211.9) | (182.8) | (108.0) | (126.6) | (2,272) | (1,516) |
|  |  |  |  |  |  |  |  |
| Observations | 294 | 210 | 210 | 252 | 252 | 210 | 252 |
| R-squared | 0.740 |  |  |  |  |  |  |
| Number of country | 42 | 42 | 42 | 42 | 42 | 42 | 42 |

В приведенной таблице сравниваются оценки моделей FE, Arellano-Bond, Системного GMM, а также Arellano-Bond и Системного GMM с учетом эндогенности переменных, которые мы исследовали в прошлых разделах нашей работы, и с использованием робастных стандартных ошибок, а также с учетом годового линейного тренда экономического роста в рассматриваемых странах.

Сравнивая с результатами без учета возможного годового линейного тренда, можно заметить, что результаты несильно изменились. Во-первых, использование динамических моделей лучше, чем FE. Во-вторых, инструментированные методы показывают более эффективные и значимые оценки. В-третьих, метод системного GMM является более эффективным, чем Ареллано-Бонда, что может быть связано с тем, что он учитывает возможную автокорреляцию ошибок. В-четвертых, использование робастных ошибок приводит к значимому уменьшению эффективности и падению значимости почти у всех коэффициентов, кроме года пандемии, значимый во всех моделях. В-пятых, можно увидеть, что предположение о наличии линейного тренда оправдалась, так как в большинстве моделей он значим. В-шестых, инструментированный метод системного GMM позволяет получать идентичные в абсолютном значении, но более эффективные оценки, чем у Ареллано-Бонда, поэтому является наилучшей моделью.

Однако для каждой из динамической модели, в которой мы инструментировали, необходимо проверить релевантность и валидность инструментов, что убедиться в верности полученных выводов.

## Тестирование адекватности инструментов

### Ареллан-Бонд модели

Для проверки на валидность инструментов необходимо также провести тест Саргана и проверить необходимые предпосылки модели о том, что существует автокорреляция первого порядка, но не существует автокорреляция второго порядка.

Как видно из картинки ниже нулевая гипотеза в тесте Саргана не отвергается, а значит, инструменты являются валидными. Тем не менее, тестирование на автокорреляцию показало, что нулевая гипотеза об отсутствии автокорреляции первого и второго порядка не отвергаются, то есть инструменты являются слабо релевантными для модели Ареллано-Бонда и могу вести к неправильным и неэффективным оценкам, из-за чего видимо мы и наблюдали отрицательный коэффициент перед лагом первого порядка экономического роста.

Исходя из этого тестирование, можем заключить, что эта модель не является адекватной при существующем наборе инструментов.

**Изображение выглядит как текст

Автоматически созданное описание**

### Системной GMM модели

Сравнительный анализ оценок моделей показал, что этот метод является наилучшим по эффективности и значимости оценок. Однако необходимо проверить адекватность инструментов для него. Тест Саргана позволяет проверить валидность инструментов, а тестирование на наличие автокорреляции первого порядка и отсутствия второго позволяет проверить релевантность инструментов.

Изображение выглядит как текст, стол

Автоматически созданное описание

Из приведенных статистик следует, что этот набор инструментов валиден, так как не отвергается нулевая гипотеза в тесте Саргана, а также необходимые предпосылки про автокорреляции полностью подтвердились, а именно гипотеза об отсутствии автокорреляции первого порядка отвергается на уровне значимости 5%, а гипотеза об отсутствии автокорреляции второго порядка не отвергается на уровне значимости больше 20%.

Таким образом, инструменты для метода системного GMM являются релевантными и валидными, что позволяет быть уверенными в эффективных оценках. Тем не менее, необходимо проверить еще одну предпосылку для утверждения полной адекватности модели, а именно об отсутствии единичного корня и проверки стационарности.

## Тестирование на стационарность ошибок в лучшей модели

Для адекватной динамической модели необходимо, чтобы остатки и первые разности были стационарными, так как иначе при наличии единичных корней вычисленные оценки методами Ареллано-Бонда или системного GMM будут не интерпретируемы и невозможно будет оценить краткосрочный и долгосрочный эффекты.

Для тестирования стационарности существуют различные тесты, каждый из который имеет свои асимптотические статистики, но у всех у них нулевая гипотеза заключается в наличии единичных корней. Проведем 3 теста, которые лучше всего по асимптотическим предположениям о данных подходят под нашу модель.

### Levin-Lin-Chu тесты

**Изображение выглядит как стол

Автоматически созданное описание**

**Изображение выглядит как стол

Автоматически созданное описание**

Как можно видеть из рисунков выше как для первых разностей, так и для остатков нулевая гипотеза отвергается, то есть они стационарны. Тем не менее, этот тест используется для данных, при которых количество временных тактов стремится к бесконечности, а у нас их всего лишь 5 используется в модели. Поэтому используем более подходящие тесты.

### Im-Pesaran-Shin тесты

**Изображение выглядит как стол

Автоматически созданное описание**

**Изображение выглядит как стол

Автоматически созданное описание**

Проведенные тесты показывают, что нулевые гипотезы об наличии единичных корней отвергаются в обоих случаях. Однако этот тест предполагает, что количество наблюдений и временных тактов стремится к бесконечности, то есть он лучше применим при больших макропанелях. Для нашего случая с короткими панельными данными этот тест не самый лучший, поэтому посмотрим на следующий тест.

### Harris-Tzavalis тесты

**Изображение выглядит как стол

Автоматически созданное описание**

**Изображение выглядит как стол

Автоматически созданное описание**

Проведенные тесты показывают, что нулевые гипотезы в обоих случаях отвергаются, значит остатки и первые разности стационарны. Причем этот тест идеально подходит под наши условия, так как он предполагает фиксированное количество временных тактов и в разы превосходящее количество индивидуальных наблюдений.

Таким образом, для модели системного GMM выполнено условие об отсутствии единичного корня, а значит ее и правда можно считать адекватной, ведь инструменты валидны и основные предположения выполнены.

## Вывод по лучшей динамической модели

Исходя из нашего анализа оценок и тестирования на адекватность инструментов и стационарности остатков и первых разностей, мы пришли к выводу, что наилучшей динамической моделью является метод Бланделла-Бонда (системный GMM) с учетом эндогенности различных переменных и линейного годового тренда. Результаты этой модели можно увидеть ниже.

Большинство полученных оценок значимы, что ожидаемо, ведь эта модель эффективна и ее стандартные ошибки достаточно малы. При этом коэффициент перед лагом экономического роста меньше единицы и равен 0.15, показывающий, что рост на 1% в прошлом году приводит к росту на 0.15% в этом году, такое смещение может быть вызвано годом пандемии, для которой коэффициент также значим и равен -7.3, отображающий краткосрочный эффект на экономический рост (уменьшение в 7.3%), а долгосрочный эффект равен уменьшение на 8.23%, что с современными условиями вполне вероятно. Тем не менее, в этой модели появились и противоречивые коэффициенты, такие как безработица, уровень урбанизации, уровня контроля коррупции и свободы слова, что может быть вызвано слабо релевантным набором инструментов и не учетом других факторов. И даже несмотря на это, можно увидеть, что выбранная модель показывает и верные оценки по знаку для инфляции, качества государственного управления и года пандемии.

Таким образом, хотя эта модель показала для некоторых переменных противоречивые коэффициенты, но они являются эффективными благодаря возможности учета эндогенности и автокорреляции ошибки, которую мы определили в первой части.

Изображение выглядит как стол

Автоматически созданное описание

# Заключение

В ходе данной работы мы научились:

* тестировать инструменты для эндогенных переменных на релевантность и валидность
* выбирать между моделями с инструментами и без
* оценивать эффект инвариантной переменной в панельных данных
* учитывать возможную эндогенность инвариантного регрессора благодаря методу Хаусмана-Тейлора
* оценивать динамические модели методами Ареллано-Бонда и Бланделла-Бонда
* учитывать эндогенность в динамических моделях
* тестировать в динамических моделях адекватность инструментов
* выбирать лучшую динамическую модель
* тестировать на стационарность и адекватность динамическую модель

# Код