

ЛЕКЦИОННЫЕ И МЕТОДИЧЕСКИЕ МАТЕРИАЛЫ

Введение в эконометрический анализ панельных данных¹⁾

Ратникова Т.А.

Курс лекций посвящен одному из наиболее востребованных современных инструментов количественного анализа статистической информации в экономике – анализу панельных данных, которые представляют собой прослеженные во времени пространственные²⁾ выборки индивидуумов, домохозяйств, предприятий, регионов, стран и т.п.

Панельные данные применяются в мире в эмпирических исследованиях экономических явлений с 60-х гг. XX в. Их использование дает ряд существенных преимуществ в оценивании параметров регрессионных зависимостей, поскольку они сочетают в себе возможности как анализа временных рядов, так и анализа пространственных наблюдений. С помощью панельных данных становится возможным изучение таких хронических проблем общества, как бедность, безработица, преступность, а также последствий проведения различных государственных социально-экономических и политических программ.

В курсе излагаются базовые концепции анализа панельных данных и принципы построения наиболее востребованных моделей. Одна из лекций посвящена анализу и интерпретации моделей, построенных по реальным российским панельным данным – данным РМЭЗ (Российского мониторинга экономического состояния и здоровья населения).

В этом номере журнала публикуются четыре первые лекции курса. В первой из них излагаются сведения общего порядка, обсуждаются источники панельных данных и некоторые особенности их использования. Во второй лекции дается представление о простейших

Ратникова Т.А. – к.ф.-м.н., доцент кафедры математической экономики и эконометрики ГУ ВШЭ.

¹⁾ Автор выражает глубокую признательность всем, кто оказывал содействие в постановке этого курса и совершенствовании текста: Г.Г. Канторовичу, С.А. Айвазяну, В.А. Бессонову, Ф. Гарду, В.Е. Гимпельсону, Б. Дормонт, Э.Б. Ершову, П.К. Катяшеву, Е.В. Коссовой, М. Морель, А.А. Пересецкому, И.Г. Поспелову.

Работа над материалами курса была поддержана грантом НФПК – Национального фонда подготовки кадров в рамках программы «Совершенствование преподавания социально-экономических дисциплин в вузах» инновационного проекта развития образования. Материалы курса сертифицированы НКСУМ – Независимым комитетом по сертификации учебных материалов.

²⁾ В англоязычной литературе для этого понятия используется термин «cross-section data». Словосочетание «пространственная выборка», может быть, не самый адекватный, но довольно часто употребляемый перевод этого понятия.

моделях анализа панельных данных и некоторых наиболее употребительных методах их оценивания. Третья лекция посвящена обсуждению свойств оценок и методов тестирования спецификации. В четвертой лекции продолжается обсуждение тестов на спецификацию и рассматривается пример практического приложения изложенных методов к анализу данных РМЭЗ.

Лекция 1.

1. Введение

1.1. История создания микроэконометрики

Сравнительно недавно эмпирические исследования в эконометрике были обогащены возможностью анализа новых источников данных: пространственных выборок объектов (индивидуумов, домохозяйств, предприятий и т.п.), наблюдаемых в течение некоторого периода времени. Такие пролонгированные пространственные выборки, где каждый объект наблюдается многократно (например, ежегодно) на протяжении отрезка времени, получили название *панельных данных*.

По словам нобелевского лауреата 2000 г. Джеймса Хекмана [20], создание подобных баз данных – это главное достижение XX в. Использование этих источников открыло новые перспективы в развитии экономической науки и математических методов, обслуживающих ее.

Смысл высказывания Хекмана состоит в следующем.

Ранние эконометрические модели, опиравшиеся на данные пространственных выборок или временных рядов, носили агрегированный характер и описывали поведение усредненных объектов, для которых Альфред Маршалл ввел специальные термины: «репрезентативный потребитель» или «репрезентативная фирма». Со временем выяснилось, что эти модели часто оказывались не слишком эффективными инструментами для анализа экономических явлений и выработки рекомендаций по социально-экономической политике. Очень часто ни значения, ни знаки коэффициентов, посчитанных по регрессиям для агрегированных временных рядов, не соответствовали предположениям экономической теории, так как возникало серьезное смещение агрегирования. Об этом писали и Тейл в 1954 г., и Грин в 1964, и Фишер в 1969 г.

Одним из решений проблемы виделась разработка программы сбора комбинированных макро- и микроданных, сторонником которой выступил Оркутт в 1964 г. Усилия Оркутта послужили толчком, который привел в движение силы, создавшие современную *микроэконометрику*, одним из разделов которой служит *анализ панельных данных*.

Основной источник микроэкономических данных – национальные репрезентативные опросы, проведение которых является весьма дорогостоящей акцией. Чтобы инициировать такого рода деятельность, необходимо наличие серьезных мотивов. Этими мотивами явились, во-первых, потребность в исследованиях, выявляющих причины социально-экономических проблем, способных нарушить стабильность уклада общественной жизни, а во-вторых, спрос на социальные программы, адресованные непосредственно тем или иным специфическим проблемным группам.

Основной целью моделей, создаваемых на базе микроданных в 60–70-х гг. прошлого века, было изучение старой политики в новых условиях или предсказание возможных эффектов новой, никогда ранее не проводимой политики. Особенно пристальное внимание экономистов в эти годы занимал рынок труда. Попытки использования неоклассической теории для его описания вызвали потребность в данных индивидуального уровня и методах анализа и интерпретации зависимостей, получаемых на основании этих данных.

Когда быстро растущий уровень развития вычислительной техники позволил оперативно оценивать сотни разнообразных регрессионных моделей, появился спрос на методы выявления среди множества этих часто взаимно противоречивых результатов таких, которые поддавались бы прозрачной экономической интерпретации. Помимо этого отобранные модели должны были при минимальной размерности вмещать в себя все богатство и разнообразие информации, поставляемой новым типом данных.

Теперь, в начале XXI в., можно констатировать (опять же по словам Хекмана), что развитие микроэконометрики привело к ряду важных эмпирических открытий.

Наиболее важное открытие – это очевидность того, что неоднородность и многообразие (экономических агентов и явлений) пронизывают экономическую жизнь, а следовательно, они должны непременно учитываться в эконометрических моделях.

Второй важный результат – появление новых моделей экономических явлений – моделей анализа панельных данных, которые предоставляют разнообразные возможности учета неоднородности.

1.2. Описание наиболее употребимых источников панельных данных

Панельные обследования в той или иной форме проводятся практически во всех экономически развитых странах, однако впервые сбор панельных данных начался в США.

Сегодня наиболее востребованными можно назвать базы NLS – National Longitudinal Surveys of Labor Market Experience и PSID – University of Michigan's Panel Study of Income Dynamics. О них следует сказать несколько слов, поскольку примеры анализа этих данных часто используются в различных учебниках и научных публикациях [13].

База NLS содержит данные по различным сегментам рабочей силы: мужчины, которым в 1966 г. исполнилось от 45 до 59 лет, юноши, которым в 1966 г. исполнилось от 14 до 24 лет, женщины, которым в 1967 г. исполнилось от 30 до 44 лет, девушки, которым в 1968 г. исполнилось от 14 до 24 лет, и молодежь обоих полов, которым в 1979 г. исполнилось от 14 до 21 года. Первые 4 сегмента периодически опрашивались в течение 15 лет. Последний сегмент продолжает наблюдаться. Перечень наблюдаемых характеристик насчитывает 1000 наименований с точки зрения рыночного предложения рабочей силы.

База PSID возникла в 1968 г. с ежегодного сбора данных репрезентативной национальной выборки, охватывающей около 6000 семей и 15000 индивидуумов, и пополняется до сих пор. Данные содержат около 5000 характеристик, включая занятость, доход, переменные человеческого капитала, жилищные условия, мобильность и т.п.

В России сбор панельных данных начался в 90-е гг. XX в.

Примерами панельных данных о российской экономике являются RLMS (Russia Longitudinal Monitoring Survey) или в русской аббревиатуре РМЭЗ – Российский мониторинг экономического положения и здоровья населения, Российский экономический тренд (доступные бесплатно по Интернету) и Российский экономический барометр – платная база данных. На РМЭЗ имеет смысл остановиться особо, поскольку эти данные очень широко используются исследователями и в России, и за рубежом.

РМЭЗ³⁾ представляет собой серии общенациональных репрезентативных опросов, регулярно проводимых с 1992 г. с целью систематического наблюдения воздействия российских реформ на динамику экономического благосостояния домохозяйств и отдельных индивидов. Опросы проводятся международным консорциумом организаций при участии Института социологии РАН. База данных РМЭЗ представляет результаты опросов свыше 10 тыс. человек и содержит информацию о размерах, источниках и структуре доходов и расходов домохозяйств, занятости, распределении времени, уровне образования, состоянии здоровья и других характеристик (всего свыше 500 показателей).

Собираемая информация имеет двухуровневую структуру.

1. Информация индивидуального уровня – индивидуальные файлы содержат:

- данные из всех взрослых и детских анкет;
- общую статистическую информацию (регион, тип населенного пункта и т.п.) для каждого человека, участвовавшего в исследовании;
- некоторые сводные индивидуальные индексы (образование, профессиональная группа и т.п.);
- показатели участия данного человека в предыдущих и последующих волнах исследования.

2. Информация уровня домохозяйства (семьи) – семейные файлы включают в себя:

- данные семейных анкет;
- общую статистическую информацию (регион, тип населенного пункта и т.п.) для каждой семьи;
- показатели участия данной семьи в предыдущих волнах исследования.

А вот как выглядит неполный перечень исследований, в которых были использованы данные РМЭЗ [9].

1. Анализ сберегательного поведения российских домохозяйств.
2. Незанятость в России: вынужденная или добровольная.
3. Субъективные и объективные оценки здоровья населения.
4. Бедность в России: масштабы и структурные особенности.
5. Измерение продолжительности бедности в России.
6. Экономический анализ причин вторичной занятости.
7. Микроэкономический анализ динамических изменений на российском рынке труда.
8. Распространенность курения в России.

³⁾ Подробная информация о РМЭЗ и первичные данные представлены на сайте: <http://www.epc.unc.edu/projects/rlms/home.html>.

9. Проблема алкоголизма в России.
10. Рабочее время как ресурс благосостояния.
11. Динамика среднего класса в России 1990-х гг.
12. Экономическая эффективность высшего образования.
13. Финансовое поведение домохозяйств: сбережение, инвестирование, кредитование, страхование.
14. Толерантность и динамика социального самочувствия в современном российском обществе.
15. Гендерные аспекты инвестиций в человеческий капитал в современной России.
16. Мобильность населения по доходам как механизм изменения неравенства.
17. Роль государства и семьи в экономической поддержке пожилых людей в Российской Федерации.
18. Человеческий капитал в России: модели текущих и пожизненных расходов.
19. Сравнительная ценность различных форм человеческого капитала в России.
20. Эволюция социального самочувствия россиян и особенности социально-экономической адаптации.
21. Трудовая незащищенность и задолженность по заработной плате в Российской Федерации.
22. Социально-экономические факторы феминизации бедности в России.
23. Женщины в сфере занятости и на рынке труда в российской экономике.
24. Анализ затрат домохозяйств на здравоохранение.
25. Экономический статус и здоровье человека.
26. Интерпретация скачка смертности в России.
27. Доходы и занятость.

1.3. Преимущества использования панельных данных

Пролонгированная, или панельная, совокупность данных представляет собой пространственную выборку объектов, прослеживаемую во времени, и, таким образом, предоставляет множество наблюдений над каждым отдельным объектом. Панели можно создавать, объединяя вместе готовые временные ряды (как правило, так строятся панели стран и регионов).

Основные преимущества данных этого типа в следующем:

- 1) они предоставляют исследователю большое количество наблюдений, увеличивая число степеней свободы и снижая зависимость между объясняющими переменными, а следовательно, стандартные ошибки оценок;
- 2) они позволяют анализировать множество экономических вопросов, которые не могут быть адресованы к временным рядам и пространственным данным в отдельности;
- 3) они позволяют предотвратить смещение агрегированности, неизбежно возникающее как при анализе временных рядов (где рассматривается временная эволюция усредненного «репрезентативного» объекта), так и при анализе перекрестных данных (где не учитываются ненаблюдаемые индивидуальные характеристики объектов и предполагается *однородность*, всех коэффициентов регрессии);

4) они дают возможность проследить индивидуальную эволюцию характеристик всех объектов выборки во времени;

5) они решают проблему поиска «хороших» инструментов при оценивании моделей с эндогенными (т.е. коррелированными со случайными ошибками) регрессорами;

6) они дают возможность избежать ошибок спецификации, возникающих от невключения в модель существенных переменных.

Поясним все вышесказанное следующими примерами.

Трудности с выводами о динамике изменения каких-либо объектов из пространственных наблюдений хорошо иллюстрируются следующей ситуацией на рынке труда.

Рассмотрим влияние профсоюзных объединений на экономическое поведение рынка.

Одна группа экономистов, которая намеревается интерпретировать наблюдаемые различия между фирмами, где есть профсоюз и где его нет, полагает, что союзы и коллективно осуществляемые процессы фундаментально меняют ключевые аспекты соотношений занятости: компенсацию, внутреннюю и внешнюю мобильность труда, порядок работы и окружение. Другая группа экономистов рассматривает эффекты от объединения как иллюзорные попытки противостояния совершенной конкуренции, условиям которой достаточно близко удовлетворяет реальный мир. Эти экономисты полагают, что наблюдаемые различия существуют главным образом благодаря различиям, предшествующим объединению или возникшим после. Профсоюзы не способствуют повышению заработной платы в долгосрочном периоде потому, что фирмы реагируют на это повышение повышением требований к качеству работников. Если одни полагают, что коэффициент при фиктивной переменной, отражающей статус участия в профсоюзе в уравнении заработной платы, есть мера эффекта от объединения, то другие считают, что этот коэффициент просто отражает уровень квалификации работника.

Модели, основанные только на пространственных данных, обычно не могут позволить выбрать верную гипотезу из этих двух, так как оценки отражают меж-индивидуальные различия только в данный момент. При использовании панельных данных можно различить эти две ситуации, изучая разницу в заработной плате работника, движущегося от фирмы без профсоюза к фирме с профсоюзом. Если эффекта от участия в профсоюзе нет, то не будет меняться и заработная плата, и наоборот. Прослеживая данные фирмы до создания в них профсоюза и после, можно сконструировать модель, измеряющую эффект от деятельности профсоюза.

Рассмотрим пример абстрактной модели с распределенными лагами:

$$(1.1) \quad y_t = \sum_{\tau=0}^n \beta_{\tau} x_{t-\tau} + u_t, \quad t = 1, \dots, T.$$

Как правило, в таких моделях возникает проблема квазимультиколлинеарности между $(n + 1)$ объясняющими переменными $x_t, x_{t-1}, \dots, x_{t-n}$. Таким образом, нет достаточной информации, чтобы получить точные оценки некоторых коэффициентов при лаговых переменных без априорного предположения о том, что они являются функциями небольшого числа параметров.

Когда есть панельные данные, мы можем использовать индивидуальные различия в величинах x , чтобы снизить проблему мультиколлинеарности. Более того, доступность пространственных массивов данных позволяет использовать различные предварительные ограничения на коэффициенты при лаговых регрессорах $\{\beta_\tau\}$.

Помимо того, что панельные данные дают возможность конструировать и тестировать более сложные поведенческие модели, чем чистые пространственные данные или временные ряды, использование панельных данных позволяет снижать размерность моделей и дает средство разрешения некоторых ключевых эконометрических проблем. Например, такой проблемой является понимание, заключается ли причина наблюдаемого эффекта в пропущенных (неверно измеренных, ненаблюдаемых) переменных, которые коррелированы с объясняющими переменными.

Рассмотрим в качестве примера простую модель:

$$(1.2) \quad y_{it} = \alpha + X'_{it}\beta + Z'_{it}\gamma + u_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, T,$$

где X'_{it} и Z'_{it} – векторы-строки объясняющих переменных, β, γ – векторы коэффициентов, случайная ошибка u_{it} подчиняется обычным предположениям теоремы Гаусса – Маркова.

Если модель (1.2) верно специфицирована, то МНК дает несмещенную и состоятельную оценку α, β и γ .

Предположим, что переменные X'_{it} наблюдаемы, а Z'_{it} ненаблюдаемы, и $\text{cov}(X_{it}, Z_{it}) \neq 0$. Тогда оценки коэффициентов регрессии y на X будут смещены. Однако, если доступны повторяющиеся наблюдения для групп индивидуумов, они могут позволить нам выявить нежелательный (смещающий оценки при переменных X_{it}) эффект от невключения Z и устранить его. Пусть $Z_{it} = Z_i$ для $\forall t$. Мы можем перейти к первым разностям по времени:

$$y_{it} - y_{i,t-1} = (X'_{it} - X'_{i,t-1})\beta + (u_{it} - u_{i,t-1}), \quad i = 1, \dots, N, \quad t = 2, \dots, T.$$

Мы также можем взять отклонение от среднего:

$$y_{it} - \bar{y}_t = (X'_{it} - \bar{X}'_t)\beta + (u_{it} - \bar{u}_t), \quad i = 1, \dots, N, \quad t = 1, \dots, T,$$

где $\bar{y}_t = \frac{1}{N} \sum_{i=1}^N y_{it}$ и т.д.

Теперь оценка МНК $\hat{\beta}$ будет несмещенной (и этому не препятствует автокоррелированность случайных ошибок в преобразованных моделях).

Если бы мы имели только пространственные данные ($T = 1$) для ($Z_{it} = Z_i$) или только временной ряд ($N = 1$) для ($Z_{it} = Z_t$), такое бы было невозможным. Часто в таких случаях приходится использовать метод инструментальных пере-

менных с инструментом, коррелирующим с X , но некоррелированным с Z и u . Найти такой инструмент, как правило, довольно сложно.

В работе Макарди [21] по жизненным циклам в предложении труда мужчин есть хорошая иллюстрация вышеизложенного. При определенных упрощающих предположениях Макарди показал, что функция предложения труда может быть записана в виде (1.2), где y – логарифм рабочих часов; X – логарифм реальной ставки заработной платы; Z – логарифм предельной полезности начального благосостояния работника. Z является ненаблюдаемой переменной и обуславливается суммарной величиной заработной платы работника и дохода от собственности за всю его жизнь к моменту начала наблюдения. Поэтому $Z_{it} = Z_i$. В этой задаче не только X коррелирует с Z , но и любая другая экономическая переменная (образование и т.п.). Следовательно, нельзя оценить β состоятельно из пространственных данных, но переходом к первым разностям по времени в панельных данных получают состоятельные оценки.

1.4. Проблемы использования панельных данных

1.4.1. Гетерогенное смещение

Привлекательность панельных данных проистекает из теоретической возможности элиминировать в регрессионной модели влияние некоторых специфических трудноизмеряемых факторов, например политики.

Если данные генерируются простым контролируемым экспериментом, то могут быть применены стандартные статистические методы. К несчастью, большая часть панельных данных поступает из очень сложных процессов повседневной экономической жизни. Типичное предположение, что y генерируется параметрической функцией распределения вероятностей $P(y|\theta)$, где θ – m -мерный действительный вектор, один и тот же для всех индивидуумов и во все времена, может быть нереальным. Игнорирование таких гетерогенных параметров может привести к несостоятельности оценок.

Рассмотрим следующую модель:

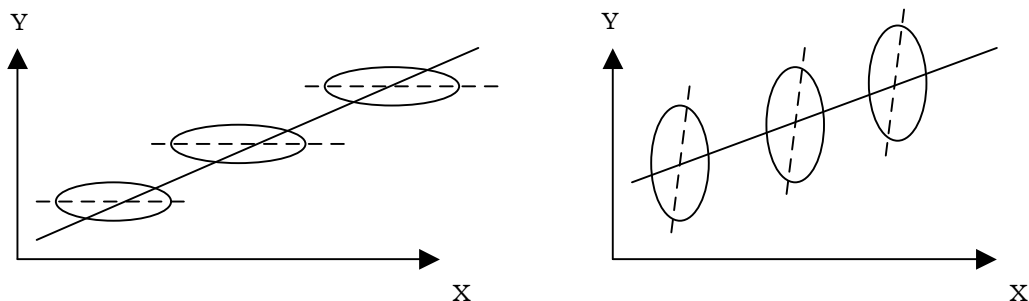
$$(1.3) \quad y_{it} = \alpha_i + \beta_i X_{it} + u_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, T,$$


где X – единственная экзогенная переменная; случайная ошибка u_{it} подчиняется обычным предположениям теоремы Гаусса – Маркова.


Параметры α_i и β_i могут быть различны для различных индивидуумов, хотя и оставаться постоянными во времени. Следовательно, будут встречаться различные выборочные распределения, которые могут серьезно смещать регрессию y_{it} на X_{it} , оцененную по всем NT-наблюдениям и игнорирующую индивидуальную неоднородность коэффициентов модели (1.3).

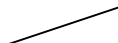
Вышесказанное можно проиллюстрировать следующими примерами:

1) гетерогенный (неодинаковый) для различных индивидуумов свободный член и гомогенный (одинаковый) наклон: $\alpha_i \neq \alpha_j, \beta_i = \beta_j$ для $\forall i, j$



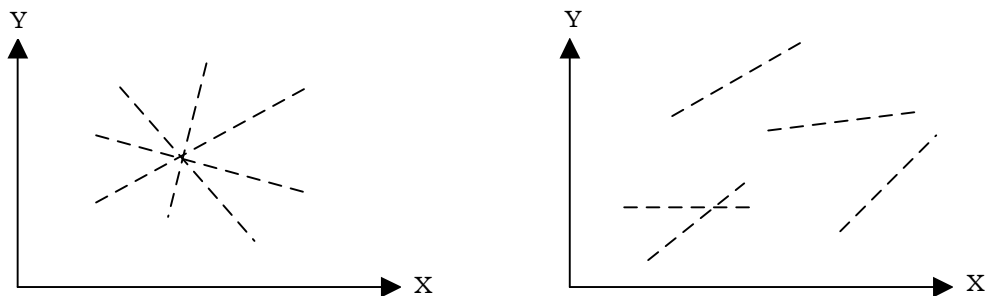
Здесь:  - диаграммы рассеяния для отдельных индивидуумов во времени;

 - индивидуальные регрессии;

 - регрессия по всем NT наблюдениям.

Во всех этих ситуациях сквозная регрессия⁴⁾, игнорирующая гетерогенность константы, является смещенной, причем направление смещения не может быть диагностировано априорно.

2) И свободный член, и наклон гетерогенны: существуют такие i, j , для которых $\alpha_i \neq \alpha_j$, $\beta_i \neq \beta_j$.



На первом рисунке изображена ситуация, когда сквозная регрессия приводит к бессмысленному результату, так как индивидуальные направления (коэффициенты наклона) существенно различаются. Во втором случае некий смысл сквозной регрессии имеется, но приводит к ложным результатам о криволинейности сквозного соотношения.

Аналогичные примеры можно привести в случае, когда свободный член и наклон изменяются со временем и одинаковы для индивидуумов.

⁴⁾ Так, начиная с этого момента, мы будем переводить англоязычный термин «pooled», под которым подразумевается регрессия, оцененная без учета особой (панельной) структуры данных.

1.4.2. Смещение самоотбора

Другой распространенный источник смещения – неслучайная выборка. Например, известный факт, что в данных РМЭЗ практически нет наблюдений, относящихся к индивидуумам из высокодоходных групп населения. Когда такие неполные данные используются в качестве зависимой (объясняемой) переменной, это может повлечь за собой смещение самоотбора. Чтобы это продемонстрировать, рассмотрим пример с пространственными данными. Зависимость ищется в виде:

$$y_i = X_i' \beta + u_i, \quad i = 1, \dots, N, \quad E(u_i) = 0, \quad D(u_i) = \sigma_u^2 I,$$

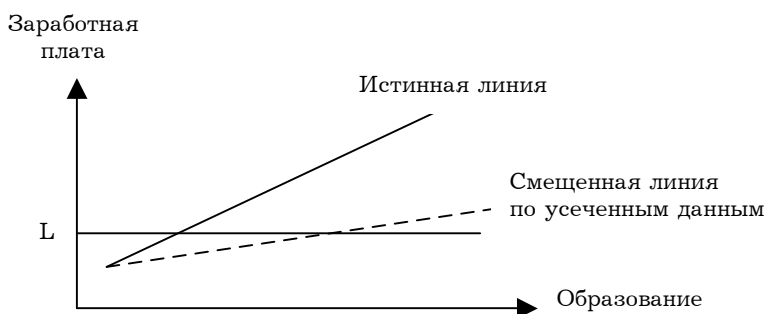
где y – заработная плата, X – набор экзогенных переменных, включая образование, интеллект и т.д.; I – единичная диагональная матрица.

Причем

при $y_i = \beta' X_i + u_i \leq L$ – индивидуумы включаются в выборку;

при $y_i > L$ – исключаются.

Для простоты теперь предположим, что все экзогенные переменные принимают одни и те же значения для всех наблюдений, кроме образования (которое измеряется как продолжительность обучения).



Из приведенного схематического рисунка видно, что линия регрессии, построенная по усеченным данным, будет иметь меньший угол наклона, чем ее аналог, который мог бы быть получен по полной выборке. Таким образом, влияние образования оказывается недооцененным. Это происходит оттого, что в данных выборок такого типа появляется корреляция между объясняемой переменной y_i и случайной ошибкой u_i , что ведет к недооценке или переоценке влияния экзогенных переменных.

Смещение самоотбора при анализе панельных данных часто является следствием истощения выборки, т.е. постепенного убывания числа объектов наблюдения. Истощение панели – это типичное явление. Панели домохозяйств могут истощаться из-за перемещений, распадов семей, а также из-за отказов участвовать в опросах в дальнейшем. Если выбытие происходит по случайным причинам,

смещения самоотбора может и не быть, но если существуют некие скрытые закономерности, то смещение неизбежно. Например, при повышении уровня доходов у домохозяйства могут пропасть стимулы участвовать в опросе, и тогда в выборке будут оставаться низкодоходные слои населения, что сделает выборку нерепрезентативной.

Перечисленные проблемы могут быть разрешены с помощью некоторых специальных приемов, которые подробно будут изложены в лекции 10. Это может быть переход или к несбалансированным панелям, где разные индивидуумы наблюдаются в течение различного числа тактов времени, или к панелям с замещением, где выбывшие объекты заменяются новыми, или использованием псевдопанелей, где в качестве объектов наблюдения выступают не отдельные индивидуумы, а группы индивидуумов со схожими (в некотором смысле) характеристиками. Хотя, конечно, это усложняет процесс оценивания.

Для решения проблемы самоотбора при исследовании пространственных выборок используют модель Хекмана. В настоящее время появились разработки, обобщающие эту модель для анализа панельных данных.

К часто встречающимся недостаткам панелей можно отнести также немногочисленность наблюдений, составляющих временные ряды для отдельных индивидуумов.

Лекция 2.

2. Простейшие модели анализа панельных данных

Изучение моделей анализа панельных данных мы начнем, введя для простоты следующие предположения:

- будем рассматривать только статические модели, в которых матрица регрессоров не содержит столбцов лаговых значений зависимой переменной;
- будем рассматривать только сбалансированные панели, т.е. те, в которых все индивидуумы наблюдаются одинаковое число временных тактов;
- будем рассматривать панели с короткими временными рядами, что очень часто встречается на практике;
- для того чтобы отразить временной эффект, будем использовать аддитивные фиктивные переменные, которые будут включены в число столбцов матрицы регрессоров;
- сосредоточим свои усилия на изучении возможностей учета специфического индивидуального эффекта, под которым будем подразумевать ненаблюдаемые и неизменяемые со временем характеристики объектов выборки.

2.1. Спецификация моделей

Модель сквозной регрессии.

Уравнение модели в покомпонентной записи:

$$y_{it} = X_{it}'b + a + \varepsilon_{it}.$$

Потребуем, чтобы наша модель удовлетворяла следующим основным предположениям:

- X'_i – вектор-строка значений детерминированных (пока) регрессоров;
- a и вектор-столбец b – коэффициенты регрессии, одинаковые для всех наблюдений;
- ε_i нормальны и удовлетворяют условиям классической линейной регрессионной модели, в том числе условию некоррелированности с X'_i .

Эта модель является самой ограничительной из возможных, так как предписывает одинаковое поведение всем объектам выборки во все моменты времени. Если эти предположения выполняются, то параметры модели могут быть состоятельно оценены с помощью метода наименьших квадратов (МНК). Соответствующая оценка в матричной форме записи будет иметь вид: $\hat{\beta}_{\text{МНК}} = (X'X)^{-1}X'Y$, где

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_i \\ \vdots \\ y_N \end{bmatrix}, \quad y_i = \begin{bmatrix} y_{i1} \\ \vdots \\ y_{iT} \end{bmatrix}, \quad X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_i \\ \vdots \\ X_N \end{bmatrix}, \quad \hat{\beta} = \begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix}.$$

Модель регрессии с детерминированным индивидуальным эффектом (fixed effect model – FE-модель).

Уравнение модели в покомпонентной записи:

$$y_{it} = X'_{it}b + a_i + \varepsilon_{it}.$$

Модельные предположения соответствуют предыдущему случаю во всем, кроме того, что касается свободного члена a_i , который теперь принимает различные значения для каждого объекта выборки. Смысл a_i в том, чтобы отразить влияние пропущенных или ненаблюдаемых переменных, характеризующих индивидуальные особенности исследуемых объектов, не меняющиеся со временем. Например, при изучении панели предприятий под a_i можно подразумевать влияние качества менеджмента.

В матричном виде эта модель записывается так:

$$\begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix}_{(T,1)} = \begin{bmatrix} X_1 \\ \vdots \\ X_N \end{bmatrix}_{(T,K)} \cdot \underset{(K,1)}{b} + \overbrace{\begin{bmatrix} \bar{i}_T & 0 & \dots & 0 \\ 0 & \bar{i}_T & & \vdots \\ \vdots & & \ddots & \\ 0 & \dots & & \bar{i}_T \end{bmatrix}}^{N \text{ векторов}} \cdot \begin{bmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_N \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_N \end{bmatrix}_{(T,1)}, \text{ где } \bar{i}_T = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}_{(T,1)}.$$

Обозначим через $A = [a_1 \dots a_N]'$ вектор констант, соответствующих детерминированным индивидуальным эффектам, а через Z – матрицу фиктивных переменных, стоящую перед вектором A , тогда:

$$(2.1) \quad y = X \cdot b + Z \cdot A + \varepsilon.$$

$\begin{matrix} (NT,1) & & (NT,K) & (K,1) & (NT,N) & (N,1) & (NT,1) \end{matrix}$

Поскольку модель не содержит общей (одинаковой для всех наблюдений) константы, матрица $(X \ Z)$ будет полного ранга, и эту модель тоже можно оценивать МНК. Соответствующая оценка

$$\hat{\beta}_{LSDV} = \begin{pmatrix} XX' & XZ' \\ Z'X & ZZ' \end{pmatrix}^{-1} \begin{pmatrix} XY' \\ Z'Y' \end{pmatrix}$$

получила в литературе название «оценки МНК в регрессии с фиктивными переменными» (LSDV).

Эта модель является довольно гибкой, так как, в отличие от предыдущей модели, она позволяет учитывать индивидуальную гетерогенность объектов выборки. Однако за эту гибкость часто приходится расплачиваться потерей значимости оценок (из-за увеличения их стандартных ошибок), так как приходится оценивать N лишних параметров. Кроме того, необходимость обращать матрицу высокой размерности $(N + K)$ вызывает вычислительные трудности. Но это последнее затруднение легко обходится, как будет показано ниже.

Модель регрессии со случайным индивидуальным эффектом (random effect model – RE-модель).

В матричной записи уравнение модели имеет вид

$$y = X \cdot b + u,$$

$\begin{matrix} (N,1) & & (NT,K) & (K,1) & (NT,1) \end{matrix}$

где $u_{it} = \alpha_i + \varepsilon_{it}$,

$$(2.2) \text{ или } y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_i \\ \vdots \\ y_N \end{bmatrix}_{(T,1)} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_i \\ \vdots \\ X_N \end{bmatrix}_{(T,K)} \cdot \underset{(K,1)}{b} + \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_i \\ \vdots \\ u_N \end{bmatrix}_{(T,1)}, \text{ где } \begin{cases} u - \text{нормально распределен} \\ X - \text{детерминированная матрица} \\ E(u) = 0, \text{ поскольку } E(\alpha) = 0, E(\varepsilon) = 0, \\ E(uu') = \Omega \neq \sigma_u^2 \cdot I_{NT} \\ E(u_i u_{i'}) = \delta_{ii'} \sigma_\alpha^2 + \delta_{ii'} \delta_{ii'} \sigma_\varepsilon^2 \end{cases}$$

$\sigma_\alpha^2, \sigma_\varepsilon^2$ – дисперсии случайных компонент α_i и ε_{it} ; $\delta_{ii'} = \begin{cases} 1, & i = i' \\ 0, & i \neq i' \end{cases}$ – символ Кронекера;

I_{NT} – единичная диагональная матрица.

Смысл α_i , также как и в предыдущем случае, состоит в том, чтобы отразить влияние пропущенных или ненаблюдаемых переменных, характеризующих индивидуальные особенности исследуемых объектов. Но теперь эти индивидуальные различия носят случайный характер, в среднем нивелируются, и их теоретические дисперсии предполагаются одинаковым для всех объектов выборки и равными σ_α^2 .

Эта модель является компромиссом между двумя предыдущими, поскольку она менее ограничительная, чем первая модель, и позволяет получать более статистически значимые оценки, чем вторая.

Если сформулированные предположения выполняются, оценки обобщенного метода наименьших квадратов (GLS или в русской аббревиатуре – ОМНК) этой модели $\hat{b}_{GLS} = (X' \Omega^{-1} X)^{-1} X' \Omega^{-1} y$ будут несмещенными.

Именно эта модель и будет служить в дальнейшем предметом нашего изучения.

2.2. Оценивание модели со случайным индивидуальным эффектом

2.2.1. Операторы «between» (B) и «within» (W)

Этими операторами удобно пользоваться при обращении с данными, имеющими двойные индексы. Они позволяют удобным образом разлагать векторы наблюдений на две взаимно ортогональные компоненты, что значительно упрощает процесс получения аналитических выражений для оценок моделей, сформулированных в предыдущем параграфе.

Но прежде, чем будут даны определения этих операторов, нам понадобится ввести понятие *кронекерова произведения матриц*.

а) Кронекерова произведение матриц.

Определение. Пусть A и B матрицы порядков (m, n) и (p, q) . Их кронекеровым произведением называется матрица порядка (mp, nq) вида: $A \otimes B = (a_{ij}, B)$.

$$\text{Пример: } I_N \otimes I_T = \begin{bmatrix} 1 \cdot I_T & 0 \cdot I_T & \dots & 0 \cdot I_T \\ 0 \cdot I_T & 1 \cdot I_T & \dots & \\ \vdots & & \ddots & \\ 0 \cdot I_T & \dots & & 1 \cdot I_T \end{bmatrix} = I_{(NT, NT)}.$$

Свойства кронекерова произведения:

$$(A + B) \otimes C = (A \otimes C) + (B \otimes C);$$

$$AB \otimes CD = (A \otimes C)(B \otimes D);$$

$$A \otimes (B + C) = (A \otimes B) + (A \otimes C);$$

$$(A \otimes B)' = A' \otimes B';$$

$$\alpha A \otimes \beta B = \alpha \beta (A \otimes B), \quad \alpha, \beta \in R;$$

$$(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}.$$

б) Операторы в пространстве R^T .

Это пространство образовано векторами наблюдений над i -ым объектом выборки.

Рассмотрим следующие вектора и матрицы:

$$y_i = \begin{bmatrix} y_{i1} \\ \vdots \\ y_{it} \\ \vdots \\ y_{iT} \end{bmatrix}, \quad \bar{i}_T = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}, \quad J_T = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & \ddots & & \\ \vdots & & & \\ 1 & & & 1 \end{bmatrix},$$

где y_i – вектор наблюдений объясняемой переменной для i -го индивидуума в течении T периодов времени.

Можно убедиться, что $\frac{J_T}{T} = \bar{i}_T (\bar{i}_T' \bar{i}_T)^{-1} \bar{i}_T'$. Это оператор проецирования на единичный вектор \bar{i}_T . Подействуем этим оператором на вектор y_i :

$$\frac{J_T}{T} \cdot y_i = \begin{bmatrix} \frac{1}{T} & \dots & \frac{1}{T} \\ \vdots & \ddots & \vdots \\ \frac{1}{T} & \dots & \frac{1}{T} \end{bmatrix} \cdot \begin{bmatrix} y_{i1} \\ \vdots \\ y_{it} \\ \vdots \\ y_{iT} \end{bmatrix} = \begin{bmatrix} \frac{1}{T} \sum_{t=1}^T y_{it} \\ \vdots \\ \frac{1}{T} \sum_{t=1}^T y_{it} \end{bmatrix} = \begin{bmatrix} y_{i\bullet} \\ \vdots \\ y_{i\bullet} \end{bmatrix} = (y_{i\bullet})_{(T,1)},$$

где $y_{i\bullet} = \frac{1}{T} \sum_{t=1}^T y_{it}$ – среднее индивидуальное по времени для i -го индивидуума.

Можно ввести оператор вычисления отклонений от этого среднего:

$$\left(I_T - \frac{J_T}{T} \right)_{(T,T)} \cdot y_{(T,1)} = \left(I_T - \frac{J_T}{T} \right) \cdot \begin{bmatrix} y_{i1} \\ \vdots \\ y_{it} \\ \vdots \\ y_{iT} \end{bmatrix} = \begin{bmatrix} y_{i1} - y_{i\bullet} \\ \vdots \\ y_{it} - y_{i\bullet} \\ \vdots \\ y_{iT} - y_{i\bullet} \end{bmatrix} = \begin{bmatrix} y_{it} - y_{i\bullet} \\ \vdots \\ y_{it} - y_{i\bullet} \end{bmatrix}.$$

Свойства операторов $\frac{J_T}{T}$ и $\left(I_T - \frac{J_T}{T} \right)$:

$$\frac{J_T}{T} = \left(\frac{J_T}{T} \right)' = \frac{J_T}{T} \cdot \frac{J_T}{T}, \quad \left(I_T - \frac{J_T}{T} \right) = \left(I_T - \frac{J_T}{T} \right)' = \left(I_T - \frac{J_T}{T} \right) \cdot \left(I_T - \frac{J_T}{T} \right).$$

Матрицы операторов, обладающие такими свойствами, называются симметричными и идемпотентными, а сами операторы – проекторами.

в) Операторы «between» (B) и «within» (W).

Эти операторы вычисляют векторы средних значений y и векторы отклонений от их средних в пространстве всех наблюдений R^{NT} .

Оператор «between» имеет вид: $B_{(NT,NT)} = \left(I_N \otimes \frac{J_T}{T} \right)$.

Рассмотрим вектор $y_{(NT,1)}$ и действие на него оператора B :

$$B_{(NT,NT)} \cdot y_{(NT,1)} = \left(I_N \otimes \frac{J_T}{T} \right) y_{(NT,1)} = \begin{bmatrix} \frac{J_T}{T} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \frac{J_T}{T} & \dots & \vdots \\ \vdots & & \ddots & \\ \mathbf{0} & \dots & & \frac{J_T}{T} \end{bmatrix}_{(T,T)} \cdot \begin{bmatrix} y_{(T,1)} \\ y_{(T,1)} \\ \vdots \\ y_{(T,1)} \end{bmatrix} = \begin{bmatrix} \frac{J_T}{T} y_1 \\ \vdots \\ \frac{J_T}{T} y_i \\ \vdots \\ \frac{J_T}{T} y_N \end{bmatrix}_{(T,1)} = \begin{bmatrix} y_{i\bullet} \\ \vdots \\ y_{i\bullet} \\ \vdots \\ y_{N\bullet} \\ \vdots \\ y_{N\bullet} \end{bmatrix} \left. \begin{array}{l} T \text{ раз} \\ \vdots \\ T \text{ раз} \\ \vdots \\ T \text{ раз} \end{array} \right\}$$

Таким образом, $B \cdot y$ есть вектор средних индивидуальных значений y , повторенных T раз для каждого индивидуума.

Оператор «within» имеет вид: $W_{(NT,NT)} = I_N \otimes \left(I_T - \frac{J_T}{T} \right)$.

$$W_{(NT,NT)} \cdot y_{(NT,1)} = \left(I_N \otimes \left(I_T - \frac{J_T}{T} \right) \right) y = \begin{bmatrix} \left(I_T - \frac{J_T}{T} \right)_{(T,1)} y_1 \\ \vdots \\ \left(I_T - \frac{J_T}{T} \right)_{(T,1)} y_i \\ \vdots \\ \left(I_T - \frac{J_T}{T} \right)_{(T,1)} y_N \end{bmatrix} = \begin{bmatrix} \left. \begin{matrix} y_{11} - y_{1\bullet} \\ \vdots \\ y_{1T} - y_{1\bullet} \end{matrix} \right\} \text{для 1-го инд.} \\ \vdots \\ \left. \begin{matrix} y_{i1} - y_{i\bullet} \\ \vdots \\ y_{iT} - y_{i\bullet} \end{matrix} \right\} \text{для 2-го инд.} \\ \vdots \\ \left. \begin{matrix} y_{N1} - y_{N\bullet} \\ \vdots \\ y_{NT} - y_{N\bullet} \end{matrix} \right\} \text{для } N\text{-го инд.} \end{bmatrix}.$$

Таким образом, $W \cdot y$ есть вектор отклонений индивидуальных наблюдений от своих средних значений по времени.

Можно самостоятельно убедиться, что операторы обладают следующими свойствами:

$$B = B' = B^2, \quad W = W' = W^2, \quad W + B = I_{NT}, \quad WB = BW = 0.$$

Оба эти оператора являются ортогональными проекторами, более того, ортогональными дополнениями, а их матрицы неотрицательно определены.

Теперь мы можем разложить вектор y на две ортогональные компоненты:

$$y = By + Wy, \quad \text{поскольку } \text{cov}(By, Wy) = BV(y)W' = 0.$$

Если вектор y центрирован относительно выборочного среднего $y_{..} = \frac{1}{NT} \sum_i \sum_t Y_{it}$, то $\frac{1}{NT} \sum_i \sum_t y_{it}^2$ есть выборочная дисперсия, а $\sum_i \sum_t y_{it}^2 = y'y$ – сумма квадратов отклонений y (TSS). Можно разложить TSS y на две компоненты:

$$y'y = y'By + y'Wy,$$

т.е. известное соотношение дисперсионного анализа (после деления на NT):

$$\text{общая (выборочная) дисперсия} = \text{межгрупповая дисперсия («between»)} + \\ + \text{внутригрупповая дисперсия («within»),}$$

где $y'By$ – отражает не зависящие от времени различия между объектами;

$y'Wy$ – отражает временные флуктуации индивидуальных наблюдений вокруг среднего по времени значения.

Аналогичное разложение может быть применено к матрице $\begin{matrix} X \\ (NT, K) \end{matrix}$

$$X'X = X'BX + X'WX.$$

2.2.2. Оценки «between» и «within»

В анализе панельных данных принято вычислять оценки коэффициентов несколькими способами и путем сравнения полученных результатов выбирать спецификацию, наиболее адекватную данным.

Оценка «between»:

$$\hat{b}_b = (X'BX)^{-1} X'By$$

получается, если применить МНК к преобразованному под действием оператора «between» уравнению регрессии $By = BXb + Bu$.

Оценка «within»:

$$\hat{b}_w = (X'WX)^{-1} X'Wy,$$

получается, если применить МНК к преобразованному под действием оператора «within» уравнению регрессии $Wy = WXb + Wu$.

Пользуясь предположениями модели со случайным индивидуальным эффектом мы можем найти аналитические выражения для математических ожиданий и ковариационных матриц полученных оценок:

$$E(\hat{b}_b) = E\left((X'BX)^{-1} X'BY\right) = (X'BX)^{-1} X'B \cdot E(Y) = (X'BX)^{-1} X'BXb = b,$$

$$V(\hat{b}_b) = (X'BX)^{-1} X' B \Omega BX (X'BX)^{-1},$$

где Ω – ковариационная матрица случайного возмущения. (При вычислении ковариационных матриц оценок здесь и в дальнейшем используется следующий простой результат: если u – случайный вектор, A – постоянная матрица, то $V(Au) = AV(u)A'$.)

Совершенно аналогично можно получить

$$E(\hat{b}_w) = b,$$

$$V(\hat{b}_w) = (X'WX)^{-1} X'W \Omega WX (X'WX)^{-1}.$$

В сущности обе эти оценки являются результатом использования МНК, но только применительно не к исходным, а к преобразованным данным.

Оценку «within» \hat{b}_w часто называют оценкой \hat{b}_{FE} , т.е. оценкой модели с детерминированным индивидуальным эффектом. Почему это возможно?

Обратимся к уравнению модели с детерминированным индивидуальным эффектом (2.1), переписав его с помощью кронекерова произведения матриц

$$\begin{aligned}
 y &= \underset{(NT,1)}{X} \cdot \underset{(NT,K)}{b} + \underset{(K,1)}{Z} \cdot \underset{(NT,N)}{A} + \underset{(N,1)}{\varepsilon} = \\
 &= \underset{(NT,K)}{X} \cdot \underset{(K,1)}{b} + \underset{(NT,N)}{(I_N \otimes \bar{I}_T)} \underset{(N,1)}{A} + \underset{(NT,1)}{\varepsilon}.
 \end{aligned}$$

Доказать, что оценка вектора коэффициентов \widehat{b}_{FE} в модели с детерминированным эффектом совпадает с оценкой \widehat{b}_w в модели со случайным эффектом, можно с помощью теоремы Фриша–Вау–Ловелла.

Теорема Фриша–Вау–Ловелла [18, 27].

Оценка МНК $\widehat{b}_{(*)}$ в модели (*) $y = Xb + Zc + u$ совпадает с оценкой МНК $\widehat{b}_{(**)}$ в модели (**) $M_Z y = M_Z Xb + M_Z u$, где $M_Z = I - Z(Z'Z)^{-1}Z'$ есть проектор на подпространство, ортогональное подпространству, натянутому на столбцы матрицы Z .

Доказательство.

Подействуем оператором M_Z на (*):

$$M_Z y = M_Z Xb + M_Z Zc + M_Z u = M_Z Xb + M_Z u,$$

так как $M_Z Zc = (I - Z(Z'Z)^{-1}Z')Zc = (Z - Z(Z'Z)^{-1}Z'Z)c = 0$.

Следовательно, модели совпали после преобразования, а значит, совпадают и оценки $\widehat{b}_{(*)} = \widehat{b}_{(**)} = (X'M_Z X)^{-1} X'M_Z Y$, ч.т.д.

В нашем случае роль Z играет $(I_N \otimes i_T)$, а роль M_Z играет $W = I_N \otimes \left(I_T - \frac{J_T}{T} \right)$.

Если теперь покомпонентно записать результаты воздействия оператора W на уравнения обеих моделей (как с детерминированным, так и со случайным индивидуальным эффектами), то эти результаты оказываются идентичными:

$$y_{it} - y_{i\cdot} = (X'_{it} - X'_{i\cdot})\beta + (\varepsilon_{it} - \varepsilon_{i\cdot}), \quad i = 1, \dots, N, \quad t = 1, \dots, T,$$

а следовательно, идентичными будут и оценки.

В дальнейшем изложении будут использоваться, помимо введенных оценок \widehat{b}_w и \widehat{b}_B , еще две традиционные оценки.

Оценка МНК: $\widehat{b}_{MНК} = (X'X)^{-1} X'Y$.

Она является несмещенной в рамках предположений модели со случайным индивидуальным эффектом, так как $E(\widehat{b}_{MНК}) = b$, и обладает ковариационной матрицей:

$$V(\widehat{b}_{MНК}) = (X'X)^{-1} X'\Omega X (X'X)^{-1}.$$

Оценка обобщенного МНК: $\hat{b}_{\text{ОМНК}} = (X' \Omega^{-1} X)^{-1} X' \Omega^{-1} y$ опять несмещена, $E(\hat{b}_{\text{ОМНК}}) = b$, $V(\hat{b}_{\text{ОМНК}}) = (X' \Omega^{-1} X)^{-1}$.

Если модель со случайным индивидуальным эффектом верно специфицирована, то последняя оценка является наилучшей в классе линейных несмещенных оценок. Это утверждение следует понимать в том смысле, что матрица $V(\tilde{\beta}) - V(\hat{\beta}_{\text{ОМНК}})$, где $V(\tilde{\beta})$ – ковариационная матрица произвольной линейной несмещенной оценки, будет являться неотрицательно определенной.

2.3. Ковариационная матрица случайного возмущения в модели со случайным индивидуальным эффектом

Чтобы исследовать эффективность всех оценок и построить оценку обобщенного МНК, нам необходимо знать ковариационную матрицу случайного возмущения Ω . Изучим поподробнее ее структуру.

Согласно (2.2)

$$\Omega = E(uu') = E \begin{bmatrix} u_1 \\ \vdots \\ u_N \end{bmatrix} \begin{bmatrix} u_1 & \dots & u_N \\ (\sigma, T) & & (\sigma, T) \end{bmatrix} = E \begin{bmatrix} u_1 u_1' & u_1 u_2' & \dots & u_1 u_N' \\ (\sigma, T) & (\sigma, T) & & (\sigma, T) \\ \vdots & \vdots & \ddots & \vdots \\ u_N u_1' & \dots & & u_N u_N' \\ (\sigma, T) & & & (\sigma, T) \end{bmatrix}.$$

В рассматриваемой модели $E(u_i u_{i'}) = \delta_{ii'} \sigma_a^2 + \delta_{ii'} \delta_{ii'} \sigma_\varepsilon^2$, следовательно, для двух разных индивидуумов $i \neq j$ и $E(u_i u_j') = 0$, а для одного индивидуума

$$i=j \text{ и } E(u_i u_i') = \begin{bmatrix} \sigma_a^2 + \sigma_\varepsilon^2 & \sigma_a^2 & \dots & \sigma_a^2 \\ \sigma_a^2 & \ddots & & \vdots \\ \vdots & & \ddots & \\ \sigma_a^2 & \dots & & \sigma_a^2 + \sigma_\varepsilon^2 \end{bmatrix} = \sigma_a^2 J_T + \sigma_\varepsilon^2 I_T = \Sigma_{(T,T)}.$$

$$\text{Следовательно, } \Omega = \begin{bmatrix} \Sigma_{(T,T)} & 0 & \dots & 0 \\ 0 & \Sigma_{(T,T)} & & \vdots \\ \vdots & & \ddots & \\ 0 & \dots & & \Sigma_{(T,T)} \end{bmatrix} = I_N \otimes \Sigma = I_N \otimes (\sigma_a^2 J_T + \sigma_\varepsilon^2 I_T).$$

Можно выразить Ω через матрицы операторов B и W :

$$\begin{aligned}\Sigma &= \sigma_\alpha^2 J_T + \sigma_\varepsilon^2 I_T = \sigma_\varepsilon^2 \left(I_T + \frac{T\sigma_\alpha^2 J_T}{\sigma_\varepsilon^2 T} + \frac{J_T}{T} - \frac{J_T}{T} \right) = \\ &= \sigma_\varepsilon^2 \left(I_T - \frac{J_T}{T} + \left(\frac{T\sigma_\alpha^2}{\sigma_\varepsilon^2} + 1 \right) \frac{J_T}{T} \right) = \sigma_\varepsilon^2 \left(\left(I_T - \frac{J_T}{T} \right) + \frac{1}{\theta^2} \frac{J_T}{T} \right),\end{aligned}$$

где $\theta^2 = \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + T\sigma_\alpha^2}$.

Тогда: $\Omega = I_N \otimes \sigma_\varepsilon^2 \left(\left(I_T - \frac{J_T}{T} \right) + \frac{1}{\theta^2} \frac{J_T}{T} \right) = \sigma_\varepsilon^2 \left(W + \frac{1}{\theta^2} B \right)$

и $\Omega^{-1} = \frac{1}{\sigma_\varepsilon^2} (W + \theta^2 B)$.

Теперь можно преобразовать некоторые выражения, полученные ранее.

Подставив выражение для Ω и воспользовавшись свойствами операторов B и W , можно значительно упростить вид ковариационных матриц оценок:

$$V(\hat{b}_B) = (X'BX)^{-1} X' B \Omega B X (X'BX)^{-1} = (\sigma_\varepsilon^2 + T\sigma_\alpha^2) (X'BX)^{-1},$$

$$V(\hat{b}_W) = (X'WX)^{-1} X' W \Omega W X (X'WX)^{-1} = \sigma_\varepsilon^2 (X'WX)^{-1}.$$

Можно показать, что оценка $\hat{b}_{ОМНК} = (X' \Omega^{-1} X)^{-1} X' \Omega^{-1} y$ эквивалентна оценке обычного метода наименьших квадратов, если последний применить к преобразованным данным $y_{it} - ay_{i\cdot}$, с $a = 1 - \sqrt{\theta^2}$.

$$\begin{aligned}\text{Действительно, } \sigma_\varepsilon^2 \Omega^{-1} &= W + \theta^2 B = (I_{NT} - B + \theta^2 B) = I_{NT} - (1 - \theta^2) B = \\ &= (I_{NT} - (1 - \theta) B) (I_{NT} - (1 - \theta) B) \Rightarrow \\ \hat{b}_{ОМНК} &= (X' \Omega^{-1} X)^{-1} X' \Omega^{-1} y = \\ &= (X' (I_{NT} - (1 - \theta) B) (I_{NT} - (1 - \theta) B) X)^{-1} X' (I_{NT} - (1 - \theta) B) (I_{NT} - (1 - \theta) B) y = \\ &= (\tilde{X} \tilde{X}')^{-1} \tilde{X}' \tilde{y},\end{aligned}$$

где $\tilde{X} = (I_{NT} - (1 - \theta) B) X = X_{it} - aX_{i\cdot}$, $\tilde{y} = y_{it} - ay_{i\cdot}$.

2.4. Интерпретация параметра θ^2

Параметр $\theta^2 = \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + T\sigma_\alpha^2}$ можно интерпретировать как комбинацию «*between*» и «*within*» дисперсий компоненты u_{it} случайного возмущения u . Поскольку $u_{it} = \alpha_i + \varepsilon_{it}$, дисперсия «*within*» u_{it} вычисляется следующим образом:

$$D_w(u_{it}) = D(u_{it} - u_{i\cdot}) = D(\varepsilon_{it} - \varepsilon_{i\cdot}) = \sigma_\varepsilon^2 \left(1 - \frac{1}{T}\right),$$

так как

$$\begin{aligned} D(\varepsilon_{it} - \varepsilon_{i\cdot}) &= E(\varepsilon_{it} - \varepsilon_{i\cdot})^2 = \sigma_\varepsilon^2 + E\frac{1}{T^2} \left(\sum_t \varepsilon_{it}\right)^2 - \frac{2}{T} E\left(\varepsilon_{it} \sum_t \varepsilon_{it}\right) = \sigma_\varepsilon^2 \left(1 + \frac{1}{T} - \frac{2}{T}\right) = \\ &= \sigma_\varepsilon^2 \left(1 - \frac{1}{T}\right), \text{ а дисперсия «between» } D_b(u_{it}) = D(u_{i\cdot}) = D(\alpha_i) + D(\varepsilon_{i\cdot}) = \sigma_\alpha^2 + \frac{\sigma_\varepsilon^2}{T}, \end{aligned}$$

так как α и ε независимы.

$$\text{Следовательно, } \theta^2 = \frac{1}{T-1} \frac{D(u_{it} - u_{i\cdot})}{D(u_{i\cdot})}.$$

Таким образом, параметр θ^2 отражает отношение внутригрупповой дисперсии к межгрупповой, нормированное на $T - 1$.

2.5. Оценивание параметра θ^2

Как обычно, в эконометрике оценки дисперсий случайных возмущений σ_α^2 и σ_ε^2 выводятся из анализа остатков $\hat{u} = y - X\hat{b}$. Существуют методы получения оценок σ_α^2 и σ_ε^2 из остатков сквозной модели. Эти оценки будут состоятельными, но иногда встречаются трудности такого рода, как отрицательные значения $\hat{\sigma}_\alpha^2$ или $\hat{\theta}^2 > 1$ при конечных T . Поэтому на практике существует более простое решение, которое состоит в использовании остатков, полученных при «*between*» и «*within*»-оценивании, что позволяет получать состоятельные оценки.

Рассмотрим остатки «*between*»-регрессии

$$\hat{u}_B = By - BX\hat{b}_B$$

и покажем, что дисперсия компоненты $u_{i\cdot}$ случайного возмущения Bi может быть состоятельно оценена с помощью суммы квадратов остатков регрессии «*between*» следующим образом:

$$(2.3) \quad \hat{\sigma}_{u_B}^2 = \frac{\hat{u}_B' \hat{u}_B}{N-K}.$$

$$\begin{aligned} \hat{u}_B &= By - BX\hat{b}_B = By - BX(X'BX)^{-1}X'By = (I - BX(X'BX)^{-1}X')By = \\ &= (I - BX(X'BX)^{-1}X')B(Xb + u) = (I - BX(X'BX)^{-1}X'B)Bu, \end{aligned}$$

так как $B = B^2$.

Тогда

$$\begin{aligned} E(\hat{u}_B' \hat{u}_B) &= Etr(u'B(I - P_{BX})Bu) = tr(B(I - P_{BX})BE(uu')) = \\ &= tr(B(I - P_{BX})B\Omega) = tr\left(B(I - P_{BX})B\sigma_\varepsilon^2\left(W + \frac{1}{\theta^2}B\right)\right) = \\ &= tr\left(\frac{\sigma_\varepsilon^2}{\theta^2}B(I - P_{BX})B\right) = \frac{\sigma_\varepsilon^2}{\theta^2}tr((I - P_{BX})B) = \frac{\sigma_\varepsilon^2}{\theta^2}rank(B - P_{BX}) = \\ &= \frac{\sigma_\varepsilon^2}{\theta^2}(N - K) = (\sigma_\varepsilon^2 + T\sigma_\alpha^2)(N - K), \end{aligned}$$

где $P_{BX} = BX(X'BX)^{-1}X'B$.

Следовательно, $E(\hat{\sigma}_{u_B}^2) = \sigma_\varepsilon^2 + T\sigma_\alpha^2$.

Аналогичным образом можно вычислить оценку дисперсии компоненты $u_{it} - u_i$ случайного возмущения Wu

$$(2.4) \quad \hat{\sigma}_{u_W}^2 = \frac{\hat{u}_W' \hat{u}_W}{NT - N - K},$$

анализируя остатки «within»-регрессии, и показать, что $E(\hat{\sigma}_{u_W}^2) = \sigma_\varepsilon^2$.

Тогда в качестве оценки параметра $\hat{\theta}^2$ может быть использовано отношение $\hat{\sigma}_{u_W}^2$ к $\hat{\sigma}_{u_B}^2$, и можно показать, что при больших значениях N эта оценка будет состоятельна:

$$p \lim_{N \rightarrow \infty} \hat{\theta}^2 = \frac{p \lim_{N \rightarrow \infty} \hat{\sigma}_{u_W}^2}{p \lim_{N \rightarrow \infty} \hat{\sigma}_{u_B}^2} = \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + T\sigma_\alpha^2} = \theta^2.$$

2.6. Реализуемый обобщенный МНК

В заглавие этого параграфа вынесено название метода оценивания параметров регрессии в случае, когда ковариационная матрица случайной ошибки неизвестна. Однако часто пользуются предположением, что известна структура этой матрицы, т.е. форма ее зависимости от одного или нескольких (обычно немногих) параметров, которые полагаются неизвестными и подлежащими оцениванию.

Такая ситуация имеет место в модели со случайным индивидуальным эффектом. Ковариационная матрица $\Omega = \sigma_\varepsilon^2 \left(W + \frac{1}{\theta^2} B \right)$ известна с точностью до двух параметров: σ_ε^2 и θ^2 .

Оценка реализуемого ОМНК (feasible GLS) коэффициентов регрессионной модели со случайным индивидуальным эффектом имеет вид:

$$\hat{b}_{\text{РОМНК}} = (X' \hat{\Omega}^{-1} X)^{-1} X' \hat{\Omega}^{-1} y,$$

где $\hat{\Omega} = \hat{\sigma}_\varepsilon^2 \left(W + \frac{1}{\hat{\theta}^2} B \right)$, $\hat{\sigma}_\varepsilon^2 = \hat{\sigma}_{u_w}^2 = \frac{\hat{u}_w' \hat{u}_w}{NT - N - K}$, $\hat{\theta}^2 = \frac{\hat{\sigma}_{u_w}^2}{\hat{\sigma}_{u_B}^2}$.

Все эти оценки будут состоятельными.

Лекция 3.

3. Сравнение оценок

3.1. Декомпозиция оценок

Удобно объединить полученные результаты единым параметрическим представлением и построить класс оценок, который впоследствии будет удобно анализировать.

Оценки класса μ : $\hat{b}(\mu) = \mu \underset{(K,1)}{\hat{b}_w} + (I - \mu) \underset{(K,K)}{\hat{b}_B}$.

Очевидно, что оценки этого класса представляют собой взвешенную сумму оценок \hat{b}_w и \hat{b}_B , где μ – квадратная матрица. Поскольку $\text{cov}(\hat{b}_w, \hat{b}_B) = 0$ (в этом нетрудно убедиться самостоятельно), мы имеем дело с ортогональным разложением. Последнее обстоятельство служит причиной широкого использования этой формы параметризации оценок.

При

$$\begin{aligned}\mu &= I_K, \quad \widehat{b}(\mu) = \widehat{b}_w, \\ \mu &= 0, \quad \widehat{b}(\mu) = \widehat{b}_b, \\ \mu &= (X'X)^{-1} X'WX, \quad \widehat{b}(\mu) = \widehat{b}_{\text{МНК}}, \\ \mu &= (X'WX + \theta^2 X'BX)^{-1} X'WX, \quad \widehat{b}(\mu) = \widehat{b}_{\text{ОМНК}}.\end{aligned}$$

Легко показать, что в рамках предположения модели со случайным индивидуальным эффектом все оценки являются несмещенными:

$$E(\widehat{b}(\mu)) = \mu E(\widehat{b}_w) + (I - \mu) E(\widehat{b}_b) = \mu b + (I - \mu)b = b.$$

Такая параметризация позволяет легко анализировать, насколько разные виды оценок хорошо отражают структуру имеющихся данных.

3.2. Асимптотические свойства оценок при $N \rightarrow \infty, T \rightarrow \infty$

Если помимо предположений модели со случайным индивидуальным эффектом сделать дополнительные предположения при $N \rightarrow \infty, T \rightarrow \infty$,

$$\begin{aligned}\frac{X'BX}{NT} &\rightarrow B_{xx} \text{ — положительно определенная матрица,} \\ \frac{X'WX}{NT} &\rightarrow W_{xx} \text{ — положительно определенная матрица,}\end{aligned}$$

то при этих условиях оценки $\widehat{b}_{\text{МНК}}, \widehat{b}_b, \widehat{b}_w, \widehat{b}_{\text{ОМНК}}$ сходятся по вероятности к неизвестному истинному значению b , и, таким образом, все они являются состоятельными.

Это обстоятельство легко продемонстрировать. Поскольку уже показана несмещенность оценок, для доказательства состоятельности достаточно убедиться в том, что дисперсии оценок стремятся к нулю при $N \rightarrow \infty, T \rightarrow \infty$.

В самом деле:

$$\begin{aligned}V(\widehat{b}_{\text{МНК}}) &= (X'X)^{-1} X'\Omega X (X'X)^{-1} = \frac{\sigma_\varepsilon^2}{NT} \left(\frac{X'X}{NT} \right)^{-1} \left(\frac{X'WX}{NT} + \frac{\sigma_\varepsilon^2 + T\sigma_\alpha^2}{\sigma_\varepsilon^2} \frac{X'BX}{NT} \right) \left(\frac{X'X}{NT} \right)^{-1} = \\ &= \frac{\sigma_\varepsilon^2}{NT} \left(\frac{X'X}{NT} \right)^{-1} + \frac{\sigma_\alpha^2}{N} \left(\frac{X'X}{NT} \right)^{-1} \frac{X'BX}{NT} \left(\frac{X'X}{NT} \right)^{-1}\end{aligned}$$

и при $N \rightarrow \infty, T \rightarrow \infty$ ковариационная матрица этой оценки сходится к нулевой матрице. Однако наличие второго слагаемого делает скорость сходимости пропорциональной N .

Совершенно аналогичные выводы можно сделать относительно ковариационной матрицы оценки \hat{b}_b :

$$V(\hat{b}_b) = (\sigma_\varepsilon^2 + T\sigma_\alpha^2)(X'BX)^{-1} = \frac{\sigma_\varepsilon^2 + T\sigma_\alpha^2}{NT} \left(\frac{X'BX}{NT} \right)^{-1} = \frac{\sigma_\varepsilon^2}{NT} \left(\frac{X'BX}{NT} \right)^{-1} + \frac{\sigma_\alpha^2}{N} \left(\frac{X'BX}{NT} \right)^{-1}.$$

Ковариационная матрица оценки \hat{b}_w

$$V(\hat{b}_w) = \sigma_\varepsilon^2 (X'WX)^{-1} = \frac{\sigma_\varepsilon^2}{NT} \left(\frac{X'WX}{NT} \right)^{-1}$$

тоже сходится к нулевой матрице, но скорость сходимости теперь уже будет пропорциональна NT .

То же самое можно отнести к $V(\hat{b}_{\text{ОМНК}})$:

$$V(\hat{b}_{\text{ОМНК}}) = (X'\Omega^{-1}X)^{-1} = \frac{\sigma_\varepsilon^2}{NT} \left(\frac{X'WX}{NT} + \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + T\sigma_\alpha^2} \frac{X'BX}{NT} \right)^{-1}.$$

При этом следует заметить, что второе слагаемое в скобке при $T \rightarrow \infty$ обращается в ноль (так как $\theta^2 = \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + T\sigma_\alpha^2} \rightarrow 0$), и $V(\hat{b}_{\text{ОМНК}})$ оказывается асимптотически эквивалентной $V(\hat{b}_w)$. Асимптотически эквивалентными (и это легко показать самостоятельно) при $T \rightarrow \infty$ оказываются и сами оценки $\hat{b}_{\text{ОМНК}}$ и \hat{b}_w .

Если учесть еще, что $\varepsilon \sim N(0, \sigma_\varepsilon^2 I)$, то

$$\begin{aligned} \sqrt{NT}(\hat{b}_{\text{ОМНК}} - b) &\sim N(0, \sigma_\varepsilon^2 W_{XX}^{-1}) \\ \sqrt{NT}(\hat{b}_w - b) &\sim N(0, \sigma_\varepsilon^2 W_{XX}^{-1}). \end{aligned}$$

Итак, все оценки оказались состоятельными, $\hat{b}_{\text{ОМНК}}$ и \hat{b}_w – асимптотически эквивалентны, но $\hat{b}_{\text{МНК}}$ и \hat{b}_b являются асимптотически менее эффективными, чем $\hat{b}_{\text{ОМНК}}$ и \hat{b}_w .

Что касается оценки реализуемого обобщенного МНК $\hat{b}_{\text{РОМНК}}$, то при обсуждаемых условиях $\hat{b}_{\text{РОМНК}} \rightarrow \hat{b}_{\text{ОМНК}}$, так как $\hat{\theta}^2$ сходится по вероятности к θ^2 . В свою очередь, $\hat{b}_{\text{ОМНК}}$ и \hat{b}_w – асимптотически эквивалентны.

Таким образом, оказывается, что удобнее всего использовать оценку \hat{b}_w , поскольку ее получение требует менее трудоемких вычислений.

3.3. Асимптотические свойства оценок при $N \rightarrow \infty$ и конечных T

Это типичная практическая ситуация, когда число индивидуумов в выборке значительно превосходит количество временных периодов, в течении которых велись наблюдения.

Пусть выполнены все требования модели со случайным индивидуальным эффектом и следующие предположения:

$$\frac{X'BX}{N} \rightarrow B_{XX}^T \text{ — положительно определенная матрица,}$$

$$\frac{X'WX}{N} \rightarrow W_{XX}^T \text{ — положительно определенная матрица}$$

при $N \rightarrow \infty$ и конечных значениях T .

При этих предположениях все оценки сходятся по вероятности к теоретическому значению b .

При конечных значениях T параметр $\hat{\theta}^2$ больше не стремится к нулю, поэтому оценка «within» теряет свои хорошие свойства и больше не эквивалентна оценке ОМНК асимптотически. Оценка ОМНК $\hat{b}_{\text{ОМНК}}$ и сходящаяся к ней оценка реализуемого ОМНК $\hat{b}_{\text{РОМНК}}$, напротив, становятся эффективными по сравнению с \hat{b}_w , $\hat{b}_{\text{МНК}}$ и \hat{b}_B .

3.4. Свойства оценок при конечных значениях N и T

3.4.1. Сравнительная эффективность оценок

Когда N и T конечны, при условии выполнения предположений модели со случайным индивидуальным эффектом можно установить, что

$$V(\hat{b}_{\text{ОМНК}}) \leq \begin{cases} V(\hat{b}_{\text{МНК}}) \\ V(\hat{b}_B) \\ V(\hat{b}_w) \end{cases}.$$

Как уже пояснялось, эта запись означает, что, например, разность $V(\hat{b}_{\text{ОМНК}}) - V(\hat{b}_B)$ является отрицательно полуопределенной матрицей. Можно также показать, что $V(\hat{b}_{\text{МНК}}) < V(\hat{b}_B)$:

$$V(\hat{b}_B) - V(\hat{b}_{\text{МНК}}) = \frac{\sigma_\varepsilon^2}{\theta^2} \left\{ (X'BX)^{-1} - (X'X)^{-1} (X'BX + \theta^2 X'WX) (X'X)^{-1} \right\} >$$

$$> \frac{\sigma_\varepsilon^2}{\theta^2} \left\{ (X'BX)^{-1} - (X'X)^{-1} (X'BX + X'WX) (X'X)^{-1} \right\} = \frac{\sigma_\varepsilon^2}{\theta^2} \left\{ (X'BX)^{-1} - (X'X)^{-1} \right\} > 0.$$

Последняя строка выкладки есть следствие того, что $X'X \geq X'BX$ (т.е. разность $X'X - X'BX = X'(I - B)X = X'WX$ в силу идемпотентности W – неотрицательно определенной матрица).

Свойства $\hat{b}_{\text{РОМНК}}$ наиболее сложно установить в случае конечных N и T , так как эта оценка получается в несколько этапов. Но, если случайные возмущения нормально распределены, в работе [32] было показано, что $\hat{b}_{\text{РОМНК}}$ несмещена, если $N \geq K+5$ (K – число регрессоров) и $T \geq 2$. Было также показано, что при $N \geq K + 10$ и $T \geq 2$

$$V(\hat{b}_{\text{РОМНК}}) \leq \begin{cases} V(\hat{b}_B) \\ V(\hat{b}_W) \end{cases}.$$

В случае маленьких выборок предпочтительнее пользоваться $\hat{b}_{\text{РОМНК}}$.

3.4.2. Сравнение оценок при конечных значениях N и T в зависимости от структуры дисперсий наблюдений

Выше были сформулированы результаты для случая детерминированных регрессоров X (2.2). Однако все выводы остаются справедливыми и при переходе к случайным регрессорам, если дополнительно потребовать независимость объясняющих переменных X и случайного возмущения u и требования к распределениям переформулировать в терминах условных распределений (при фиксированном X).

При конечных значениях N и T в зависимости от структуры ковариационной матрицы объясняющих переменных различные оценки могут быть очень близки между собой. Чтобы это продемонстрировать, рассмотрим модель со случайным эффектом и единственной объясняющей переменной:

$$\underset{(NT,1)}{y} = \underset{(NT,1)}{x} \underset{(1,1)}{b} + u.$$

Воспользуемся μ -параметризацией оценок: $\hat{b}(\mu) = \mu \hat{b}_W + (1 - \mu) \hat{b}_B$.

Здесь μ будет просто скаляром, и тогда:

$$\hat{b}(1) = \hat{b}_W, \quad \hat{b}(0) = \hat{b}_B, \quad \hat{b}\left(\frac{x'Wx}{x'x}\right) = \hat{b}_{\text{МНК}}, \quad \hat{b}\left(\frac{x'Wx}{x'Wx + \theta^2 x'Bx}\right) = \hat{b}_{\text{ОМНК}}.$$

Можно рассмотреть два предельных случая:

- если $\frac{x'Bx}{x'x} \rightarrow 1$ (по вероятности) (т.е. $\frac{x'Wx}{x'x} \rightarrow 0$, так как $x'x = x'Bx + x'Wx$),

тогда $\mu = 0$ и $\hat{b}_{\text{МНК}} \cong \hat{b}_B \cong \hat{b}_{\text{ОМНК}}$. Эта ситуация соответствует исчезающе малым различиям между наблюдениями, относящимися к одному и тому же индивидууму;

• если $\frac{x'Wx}{x'x} \rightarrow 1$ (по вероятности) (т.е. $\frac{x'Bx}{x'x} \rightarrow 0$), тогда $\mu = 1$ и $\hat{b}_{\text{МНК}} \cong \hat{b}_w \cong \hat{b}_{\text{ОМНК}}$, что соответствует ситуации, когда межиндивидуальные различия малы.

Итак, подведем итог.

Если гипотеза о независимости объясняющих переменных X и случайного возмущения u выполняется, все полученные оценки состоятельны, и исследователю остается изучать эффективность оценок. В таком случае, при $N \rightarrow \infty$, $T \rightarrow \infty$ \hat{b}_w , $\hat{b}_{\text{РОМНК}}$, $\hat{b}_{\text{ОМНК}}$ асимптотически эквивалентны и обладают большей эффективностью, чем другие оценки. Когда же только $N \rightarrow \infty$, \hat{b}_w теряет свои хорошие свойства, и единственной доступной хорошей оценкой остается $\hat{b}_{\text{РОМНК}}$, которая асимптотически эквивалентна $\hat{b}_{\text{ОМНК}}$.

4. Тестирование спецификации

Многообразие методов оценивания моделей панельных данных дает возможность выявить ошибки спецификации путем сравнительного анализа полученных оценок.

Возникновение методики тестирования спецификации связано с именем Мундлака, в 1978 г. подвергнувшего критике формулировку модели со случайным индивидуальным эффектом. Идеи Мундлака были развиты Хаусманом и впоследствии вылились в создание системы тестов.

4.1. Критика Мундлаком спецификации модели со случайным индивидуальным эффектом

Эдвард Мундлак [28], занимаясь сравнительным анализом моделей со случайным и детерминированным индивидуальным эффектом, пришел к выводу о некорректности формулировки модели со случайным эффектом. Основанием к этому послужило то обстоятельство, что эта модель не учитывает возможную корреляцию между индивидуальным эффектом и объясняющими переменными. Есть основания полагать, что обычно такая корреляция существует.

Рассмотрим, например, оценивание производственной функции по данным о фирмах. Выпуск каждой фирмы y_{it} может определяться ненаблюдаемым качеством менеджмента α_i . Фирма с более эффективным руководством более эффективно использует ресурсы. В такой ситуации α_i и X_i не могут быть независимыми. Игнорирование этого обстоятельства приведет к смещению оценок.

С точки зрения Мундлака, различия в оценках моделей со случайным и детерминированным эффектом часто связаны просто с некорректной формулировкой модели со случайным эффектом.

Сопоставим эти две модели.

Модель со случайным эффектом:

$$\begin{aligned} y_{it} &= X'_{it}b + u_{it} \\ u_{it} &= \alpha_i + \varepsilon_{it} \end{aligned}$$

где α_i и ε_{it} – случайные возмущения, независимые между собой и не зависящие от X_{it} . В этом случае $\hat{b}_{\text{ОМНК}}$ является наилучшей оценкой в классе линейных и несмещенных оценок.

Модель с детерминированным эффектом:

$$y_{it} = X'_{it}b + \alpha_i + \varepsilon_{it},$$

где ε_{it} случайны, независимы от X_{it} , α_i детерминированы. В этом случае \hat{b}_w является наилучшей оценкой в классе линейных и несмещенных оценок.

При $N \rightarrow \infty$, $T \rightarrow \infty$ эти модели асимптотически эквивалентны, и также эквивалентны оценки $\hat{b}_{\text{ОМНК}}$ и \hat{b}_w .

Мундлак показал, что если учесть корреляцию α_i и X_i в модели со случайным эффектом, эта модель становится эквивалентной модели с детерминированным эффектом, а оценка $\hat{b}_{\text{ОМНК}}$ – эквивалентной оценке \hat{b}_w , и не только асимптотически. В противном же случае, т.е. при игнорировании имеющейся корреляции α_i и X_i , оценка $\hat{b}_{\text{ОМНК}}$ является смещенной. Перефразируя высказывание Мундлака, можно сказать, что различие между оценками $\hat{b}_{\text{ОМНК}}$ и \hat{b}_w является вымышленным и вызвано некорректной спецификацией, игнорирующей корреляцию между индивидуальным эффектом и объясняющими переменными.

Рассмотрим теперь формулировку модели со случайным индивидуальным эффектом, предложенную Мундлаком:

$$\begin{cases} y_{it} = X'_{it}b + \alpha_i + \varepsilon_{it} \\ E(\alpha_i | X_{it}, t = 1, \dots, T) \neq 0 \end{cases}$$

где α_i и ε_{it} случайны, и ε_{it} не зависит от X_{it} .

Налагаются еще некоторые требования:

- X_{it} – нормальны, независимы, одинаково распределены и $E(X_{it}) = 0$;
- α_i и ε_{it} тоже нормальны, независимы, одинаково распределены и $E(\alpha_i) = 0$, $E(\varepsilon_{it}) = 0$ для всех t и i ;
- $E(\alpha_i | X_i)$ есть вспомогательная линейная регрессионная функция:

$$\alpha_i = X'_{i*} \cdot \gamma + w_i, \text{ где } w_i \sim N(0, \sigma_w^2 I) \text{ и не зависит от } X_{it} \text{ и } \varepsilon_{it}, X_{i*} = \frac{\bar{i}'_t X_i}{T}.$$

Недостатком этой формулировки является жесткое требование одинаковой распределенности X_{it} при любом t . При нарушении этого требования нельзя гарантировать независимость между w_i и X_{it} , и ставится под сомнение правильность формулировки Мундлака. Можно устранить эту проблему, как показал Чемберлен [14], рассмотрев более общую регрессионную зависимость для условного математического ожидания:

$$E(\alpha_i | X_{it}, t = 1, \dots, T) = \sum_{t=1}^T X'_{it} \gamma_{(1,K)(K,1)}.$$

Теперь модель сложной ошибки будет содержать $2K$ объясняющих переменных и случайные компоненты w_i и ε_{it} , независимые от X_{it} и X_i :

$$\begin{cases} y_{it} = X'_{it} b_{(1,K)(K,1)} + X_{it} \gamma_{(1,K)(K,1)} + v_{it}, \\ v_{it} = w_i + \varepsilon_{it} \end{cases},$$

т.е. для любого индивидуума будет иметь место регрессия:

$$y_{it} = X'_{it} b_{(NT,K)(K,1)} + \bar{i}'_T X'_{it} \gamma_{(T,1)(1,K)(K,1)} + v_{it} = X'_i b_{(NT,K)(K,1)} + \frac{\bar{i}'_T \bar{i}'_T}{T} X_i \gamma_{(T,1)} + v_i,$$

а общий вид регрессии для всей панели будет:

$$y_{(NT,1)} = X_{(NT,K)(K,1)} b_{(NT,K)(K,1)} + \left(I_N \otimes \frac{J_T}{T} \right)_{(NT,K)(K,1)} X_{(NT,K)(K,1)} \gamma_{(NT,1)} + v_{(NT,1)}$$

или

$$y = Xb + BX\gamma + v,$$

где $E(v) = 0$, $E(vv') = V = \sigma_\varepsilon^2 \left(W + \frac{1}{\tau^2} B \right)$, $\tau^2 = \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + T\sigma_w^2}$.

Таким образом, теперь имеются две конкурирующие гипотезы для модели со случайным индивидуальным эффектом:

$$H_0 : \begin{cases} y = Xb + u \\ E(uu') = \Omega \end{cases},$$

где $u_{it} = \alpha_i + \varepsilon_{it}$ и α_i и X_{it} — независимы, $\text{cov}(\alpha_i, X_{it}) = 0$.

$$H_M : \begin{cases} y = Xb + BX\gamma + v \\ E(vv') = V \end{cases},$$

где $v_{it} = w_i + \varepsilon_{it}$ и α_i и X_{it} — коррелируют, $\text{cov}(\alpha_i, X_{it}) \neq 0$ — это гипотеза Мундлака.

Оценки коэффициентов при гипотезе H_M обобщенным МНК:

$$\begin{bmatrix} \hat{b}_M \\ \hat{\gamma}_M \end{bmatrix} = \left[\begin{pmatrix} X' \\ X'B \end{pmatrix} V^{-1} (X \ BX) \right]^{-1} \begin{bmatrix} X' \\ X'B \end{bmatrix} V^{-1} y.$$

После подстановки выражения для $V^{-1} = \frac{1}{\sigma_e^2} (W + \tau^2 B)$ получаем:

$$\begin{aligned} \hat{b}_M &= (X'WX)^{-1} X'Wy = \hat{b}_W \\ \hat{\gamma}_M &= (X'BX)^{-1} X'By - (X'WX)^{-1} X'Wy = \hat{b}_B - \hat{b}_W. \end{aligned}$$

Если верна гипотеза Мундлака, то \hat{b}_W является несмещенной и эффективной, в то время как $\hat{b}_{МНК}$, \hat{b}_B и $\hat{b}_{ОМНК}$ будут смещены. Например, $E(\hat{b}_B) = b + \gamma$. Но асимптотически при $N \rightarrow \infty$, $T \rightarrow \infty$ $\hat{b}_{ОМНК}$ эквивалентна \hat{b}_W и является состоятельной.

Теперь очевидно, что модель сложной ошибки, где случайный эффект коррелирован с объясняющими переменными, эквивалентна модели с детерминированным эффектом, для которой \hat{b}_W является наилучшей в классе линейных несмещенных оценок по теореме Гаусса–Маркова. Отсюда следует, что *при отсутствии оснований считать α_i и X_{it} независимыми, лучшая оценка для b есть \hat{b}_W .*

4.2. Тесты Хаусмана на ошибки спецификации

Тестируется независимость α_i и X_{it} , т.е. H_0 против H_M или, иначе говоря, гипотеза о том, что $\gamma = 0$ в модели Мундлака.

4.2.1. Принцип тестов Хаусмана

Рассмотрим общий принцип тестов Хаусмана, поскольку они применяются не только в анализе панельных данных. Эти тесты дают возможность сделать выбор между оценкой состоятельной и эффективной при гипотезе H_0 (модель правильно специфицирована), но несостоятельной при гипотезе H_A (модель неправильно специфицирована) и оценкой, состоятельной при обеих гипотезах.

Итак,

- \hat{b}_0 – оценка состоятельная и асимптотически эффективная при H_0 ,
- $\hat{b}_{(K,1)}$ – оценка асимптотически смещенная при H_A ,
- \hat{b}_1 – оценка состоятельная и при H_0 , и при H_A .

Если справедлива H_0 , то при $N \rightarrow \infty$

$$\begin{aligned}\sqrt{N}(\hat{b}_0 - b) &\sim N(0, V(\hat{b}_0)) \\ \sqrt{N}(\hat{b}_1 - b) &\sim N(0, V(\hat{b}_1))\end{aligned}$$

т.е. отклонения оценок \hat{b}_0 и \hat{b}_1 от теоретического значения вектора параметров b подчиняются центрированному нормальному закону.

Рассмотрим разность оценок: $\hat{q} = \hat{b}_1 - \hat{b}_0$.

При H_0 случайная величина $m = N\hat{q}(V(\hat{q}))^{-1}\hat{q} \sim \chi_k^2$ (асимптотически), где $V(\hat{q}) = V(\hat{b}_1) - V(\hat{b}_0)$. (Последнее равенство действительно справедливо при H_0 , это строго доказывается, так как $\sqrt{N}(\hat{b}_0 - b)$ и $\sqrt{N}\hat{q}$ асимптотически некоррелированы, а \hat{b}_0 – асимптотически эффективна, но здесь это доказательство приводиться не будет.)

Если m велика, тогда H_0 отвергается, и в этом случае следует использовать оценку \hat{b}_1 . На практике при вычислении m используется оценка $\hat{V}(\hat{q})$, которая асимптотически стремится к $\hat{V}(\hat{q})$.

4.2.2. Приложение теста Хаусмана к модели сложной ошибки

H_0 : модель сложной ошибки с $\text{cor}(\alpha_i, X_{it}) = 0$ верно специфицирована.

H_A : модель сложной ошибки неверно специфицирована, так как $\text{cor}(\alpha_i, X_{it}) \neq 0$.

Тогда

$$\begin{aligned}\hat{b}_{\text{ОМНК}} &\text{ – оценка состоятельная и асимптотически эффективная при } H_0, \\ \hat{b}_{(K,1)} &\text{ – оценка асимптотически смещенная при } H_A, \\ \hat{b}_W &\text{ – оценка состоятельная и при } H_0, \text{ и при } H_A. \\ \hat{b}_{(K,1)}\end{aligned}$$

Рассмотрим сразу несколько версий теста Хаусмана, введя

$$\begin{aligned}\hat{q}_1 &= \hat{b}_W - \hat{b}_{\text{ОМНК}}, \\ \hat{q}_2 &= \hat{b}_W - \hat{b}_B, \\ \hat{q}_3 &= \hat{b}_{\text{ОМНК}} - \hat{b}_B.\end{aligned}$$

Воспользовавшись μ -параметризацией, можно разложить $\hat{b}_{\text{ОМНК}}$ на взвешенную сумму \hat{b}_W и \hat{b}_B :

$$\hat{b}_{\text{ОМНК}} = \mu\hat{b}_W + (I - \mu)\hat{b}_B \quad \text{с} \quad \mu = (X'WX + \theta^2 X'BX)^{-1} X'WX.$$

Следовательно,

$$\begin{aligned}\hat{q}_1 &= \hat{b}_W - \hat{b}_{ОМНК} = (I - \mu)(\hat{b}_W - \hat{b}_B) = (I - \mu)\hat{q}_2 \\ \hat{q}_3 &= \hat{b}_{ОМНК} - \hat{b}_B = \mu(\hat{b}_W - \hat{b}_B) = \mu\hat{q}_2.\end{aligned}$$

В силу «пропорциональности» \hat{q}_1 , \hat{q}_2 и \hat{q}_3 все они дают одну и ту же статистику: если соответствующие матрицы не вырождены

$$m = \hat{q}_1(V(\hat{q}_1))^{-1}\hat{q}_1 = \hat{q}_2(V(\hat{q}_2))^{-1}\hat{q}_2 = \hat{q}_3(V(\hat{q}_3))^{-1}\hat{q}_3,$$

и если верна H_0 , то $m \sim \chi_K^2$.

Чтобы провести этот тест, достаточно знать \hat{b}_W и \hat{b}_B , так как $\hat{q}_2 = \hat{b}_W - \hat{b}_B$, и $V(\hat{q}_2) = V(\hat{b}_W) + V(\hat{b}_B)$ (в последнем легко убедиться, проведя самостоятельно несложные выкладки).

Таким образом, этот тест выявляет значимость различий \hat{b}_W и \hat{b}_B . Чем значимее различия, тем больше оснований принять модель Мундлака:

$$Y = Xb + BX\gamma + v, \quad \text{где} \quad \hat{\gamma} = \hat{b}_B - \hat{b}_W = -\hat{q}_2,$$

следовательно, это тест на значимость оценки $\hat{\gamma}$.

4.3. Тесты на существование и независимость индивидуального эффекта

Во всех эмпирических приложениях вопрос существования индивидуально-го эффекта предшествует вопросу о его независимости от регрессоров. Рассмотрим простейшую процедуру тестирования наличия индивидуального эффекта.

В модели $Y = Xb + u$, где $u \sim N(0, \Omega)$, $\Omega = \sigma_\varepsilon^2 \left(W + \frac{1}{\theta^2} B \right)$, $\theta^2 = \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + T\sigma_\alpha^2}$ тестирование наличия индивидуального эффекта равносильно тестированию гипотезы о том, что $\sigma_\alpha^2 = 0$ (так как $E(\alpha_i) = 0$ уже заложено в модели).

Тест построен на сравнении дисперсий остатков регрессий «*within*» и «*between*». Если верна гипотеза нормальности случайного возмущения u , то:

$$\hat{\sigma}_{uw}^2 \sim \frac{\sigma_\varepsilon^2}{NT - N - K} \chi_{NT - N - K}^2, \quad \hat{\sigma}_{ub}^2 \sim \frac{\sigma_\varepsilon^2 + T\sigma_\alpha^2}{N - K} \chi_{N - K}^2.$$

Остатки являются независимыми, что легко показывается, так как

$$\begin{aligned}\hat{u}'_B \hat{u}_B &= u'(I - P_{BX})u, \quad P_{BX} = BX(X'BX)^{-1}X'B, \\ \hat{u}'_W \hat{u}_W &= u'(I - P_{WX})u, \quad P_{WX} = WX(X'WX)^{-1}X'W,\end{aligned}$$

$(I - P_{BX})(I - P_{WX}) = 0$, а квадратичные формы, матрицы которых удовлетворяют этому соотношению ортогональности, являются статистически независимыми.

$$\text{Следовательно, } \frac{\hat{\sigma}_{u_B}^2 / (\sigma_\varepsilon^2 + T\sigma_\alpha^2)}{\hat{\sigma}_{u_W}^2 / \sigma_\varepsilon^2} \sim \frac{\chi_{N-K}^2 / (N-K)}{\chi_{NT-N-K}^2 / (NT-N-K)} = F(N-K, NT-N-K).$$

Значит, если верна гипотеза $H_0: \sigma_\alpha^2 = 0$, то

$$\frac{\hat{\sigma}_{u_B}^2}{\hat{\sigma}_{u_W}^2} \sim F(N-K, NT-N-K).$$

Этот тест является разновидностью известного теста Бройша–Пагана.

На практике для больших панелей ($N > 150$) обычно берут $F(\infty, \infty) = 1$, а уровень значимости 5% или 1%. Если $\hat{\sigma}_{u_B}^2 > \hat{\sigma}_{u_W}^2$ значимо, то гипотеза об отсутствии индивидуального эффекта отвергается.

Для больших выборок этот тест осуществляется также с помощью метода Лагранжа, и тогда в качестве тестовой статистики используется множитель Лагранжа:

$$LM = \frac{NT}{2(T-1)} \left[\frac{T^2 \sum_{i=1}^N (\hat{u}_{i\cdot})^2}{\sum_{i=1}^N \sum_{t=1}^T (\hat{u}_{it})^2} - 1 \right]^2,$$

подчиняющийся асимптотически при H_0 χ^2 – распределению с одной степенью свободы.

Здесь под \hat{u} понимаются остатки сквозной регрессии.

Замечание. Результаты теста существенно зависят от спецификации регрессионного уравнения.

Гипотеза об отсутствии случайного индивидуального эффекта редко отвергается, если модель сформулирована в темпах роста. Если даже этот эффект наблюдался в уровнях, он исчезает в темпах роста. Например, пусть Y и X измеряются в логарифмах: $y_{it} = X_{it}b + \alpha_i + \varepsilon_{it}$. Эта же модель в темпах роста: $\dot{y}_{it} = y_{it} - y_{it-1} = \dot{X}_{it}b + \varepsilon_{it} - \varepsilon_{it-1}$ уже не содержит α_i . Оценка «within» в темпах роста трудна в интерпретации, поэтому не представляет большого интереса.

4.4. О правомочности теста Хаусмана

Проиллюстрируем применимость теста Хаусмана к модели сложной ошибки следующей обобщающей таблицей:

$N \rightarrow \infty, T$ -фиксировано	I	II	III
Гипотеза	$E(u_{it} X_{it}) = 0$, т.е.	$E(\alpha_i X_{it}) \neq 0$,	$E(\alpha_i X_{it}) \neq 0$,
Оценка	$\text{cov}(X_{it}, \alpha_i) = 0$	$E(\varepsilon_{it} X_{it}) = 0$	$E(\varepsilon_{it} X_{it}) \neq 0$
	$\text{cov}(X_{it}, \varepsilon_{it}) = 0$		
$\hat{b}_{МНК}$	состоятельна, неэффективна	несостоятельна	несостоятельна
\hat{b}_B	состоятельна, неэффективна	несостоятельна	несостоятельна
\hat{b}_W	состоятельна, неэффективна	состоятельна, эффективна	несостоятельна
$\hat{b}_{ОМНК}$	состоятельна, эффективна	несостоятельна	несостоятельна

Ситуация III имеет место, когда модель неверно специфицирована, например, пропущена важная объясняющая переменная. В этом случае тест Хаусмана неприменим. Чтобы уменьшить последствия ошибок спецификации, необходим сравнительный анализ оценок, полученных различными методами.

Лекция 4.

5. Классификация моделей анализа панельных данных

5.1. Схема используемых моделей

Пусть у нас есть выборка наблюдений характеристик N индивидуумов в течении T периодов времени, обозначенная (y_i, x_{kit}) , где $i = \overline{1, N}$; $t = \overline{1, T}$; $k = \overline{1, K}$. Y предполагается случайным исходом некоторого эксперимента с распределением вероятности, условным по векторам характеристик x и параметров θ $f(y|x, \theta)$.

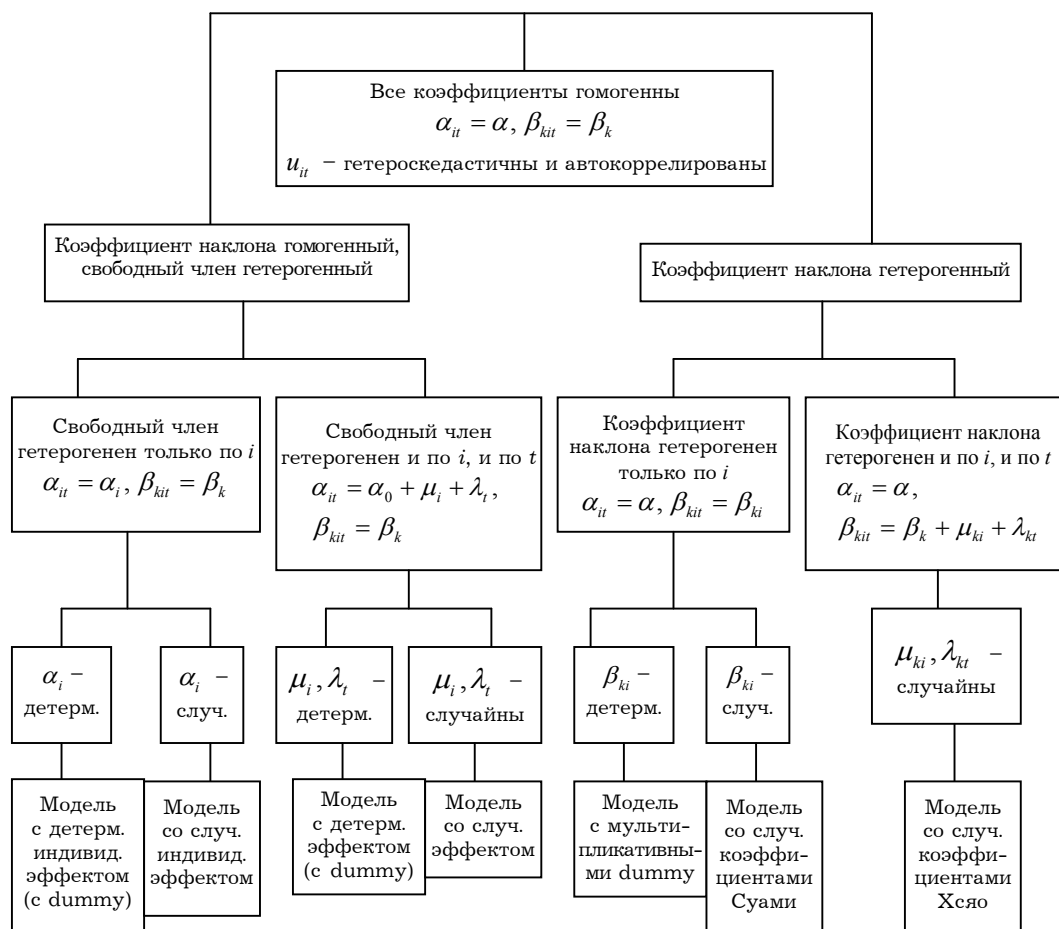
Когда применяются панельные данные, одна из поставленных целей – использовать всю имеющуюся информацию, чтобы сделать выводы о θ . Если, например, выбрана простая линейная зависимость y от x , чтобы применить сквозное оценивание МНК по всем NT-наблюдениям, необходимо допущение об одинаковости регрессионных параметров θ для всех объектов выборки во все периоды времени. Если это допущение не верно, сквозная регрессия приведет к ложным заключениям. Следовательно, первый шаг на пути полной эксплуатации имеющихся данных – **тестирование постоянства параметров по всем i и t** .

Самая общая спецификация уравнения имеет вид:

$$y_{it} = \alpha_{it} + \sum_{k=1}^K x_{kit} \beta_{kit} + u_{it}, \quad i = \overline{1, N}, \quad t = \overline{1, T}.$$

Очевидно, что такая модель не поддается оцениванию. Необходимо наложить ограничения на коэффициенты уравнения.

Чтобы уяснить полную картину вариантов моделей анализа панельных данных в зависимости от характера коэффициентов, рассмотрим следующую схему.



5.2. Модель анализа ковариаций

В предыдущих параграфах мы рассматривали модель с индивидуальным эффектом, причем этот эффект отражался только свободным членом регрессии, коэффициент наклона мы полагали постоянным.

В этом параграфе будут изложены общие принципы тестирования постоянства коэффициентов регрессии.

Обычно используемая для того чтобы различить влияние как качественных, так и количественных переменных, линейная модель постулируется так:

$$y_{it} = \alpha_{it} + x'_{it}\beta_{it} + u_{it}, \quad i = \overline{1, N}, \quad t = \overline{1, T},$$

где α_{it} – скаляр,

$$\beta_{it} = (\beta_{1it}, \beta_{2it}, \dots, \beta_{kit})' - 1 \times K, \quad x'_{it} = (x_{1it}, \dots, x_{kit}) - 1 \times K, \quad E(u_{it}) = 0, \quad D(u_{it}) = \sigma_u^2,$$

где $1 \times K$ – k -мерный вектор.

Тестируются два аспекта:

- гомогенность (однородность) наклона;
- гомогенность (однородность) свободного члена.

Процедура состоит из трех основных шагов.

1. Проверка, являются или нет наклон и свободный член одновременно гомогенными для различных индивидуумов в разные моменты времени.

2. Проверка, является или нет наклон регрессии одним и тем же для всех наблюдений.

3. Проверка, является или нет свободный член регрессии одним и тем же для всех наблюдений.

Очевидно, если принята гипотеза об общей гомогенности (1), то дальнейшее тестирование излишне. Если (1) отвергается, то тестируется (2). Если (2) не отвергается, переходят к (3).

Для начала предположим, что параметры не зависят от времени, но различаются между индивидуумами. Следовательно, для i -го индивидуума имеет место регрессия:

$$(4.1) \quad y_{it} = \alpha_i + x'_{it}\beta_i + u_{it}, \quad i = \overline{1, N}, \quad t = \overline{1, T}.$$

Здесь могут быть наложены три типа ограничений:

H1: $y_{it} = \alpha_i + x'_{it}\beta + u_{it}$, т.е. наклоны одинаковы;

H2: $y_{it} = \alpha + x'_{it}\beta_i + u_{it}$, т.е. свободные члены одинаковы;

H3: $y_{it} = \alpha + x'_{it}\beta + u_{it}$, т.е. и наклон, и свободный член одни и те же.

Тип ограничений H2 используется очень редко, поэтому в дальнейшем будет игнорироваться. Гипотеза H3 соответствует сквозной регрессии. Гипотеза H1 соответствует модели с детерминированным индивидуальным эффектом.

$$\text{Пусть } y_{i\bullet} = \frac{1}{T} \sum_{t=1}^T y_{it}, \quad x_{i\bullet} = \frac{1}{T} \sum_{t=1}^T x_{it}.$$

Оценки МНК β_i и α_i в модели без ограничений:

$$\left\{ \begin{array}{l} \hat{\beta}_i = W_{xx,i}^{-1} W_{xy,i} \\ \hat{\alpha}_i = y_{i\bullet} - \hat{\beta}_i' x_{i\bullet} \\ i = \overline{1, N} \end{array} \right. \quad \text{где} \quad \begin{array}{l} W_{xx,i} = \sum_{t=1}^T (x_{it} - x_{i\bullet})(x_{it} - x_{i\bullet})' = x'_{it} \left(I_T - \frac{J_T}{T} \right) x_{it} \\ W_{xy,i} = \sum_{t=1}^T (x_{it} - x_{i\bullet})(y_{it} - y_{i\bullet})' = x'_{it} \left(I_T - \frac{J_T}{T} \right) y_{it} \\ W_{yy,i} = \sum_{t=1}^T (y_{it} - y_{i\bullet})^2 = y'_{it} \left(I_T - \frac{J_T}{T} \right) y_{it}. \end{array}$$

Это оценки группы «*within*». Сумма квадратов остатков модели без ограничений: $S_0 = \sum_{i=1}^N RSS_i$, где $RSS_i = W_{yy,i} - W'_{xy,i} W_{xx,i}^{-1} W_{xy,i}$.

Регрессия МНК в модели с ограничениями Н1 порождает следующие оценки параметров:

$$\left\{ \begin{array}{l} \hat{\beta}_W = W_{xx}^{-1} W_{xy} \\ \hat{\alpha}_i = y_{i\cdot} - \hat{\beta}'_W x_{i\cdot}, \\ \quad i = \overline{1, N} \end{array} \right. \quad \text{где} \quad \begin{array}{l} W_{xx} = \sum_{i=1}^N W_{xx,i} = x' W x \\ W_{xy,i} = \sum_{i=1}^N W_{xy,i} = x' W y \\ W_{yy,i} = \sum_{i=1}^N W_{yy,i} = y' W y. \end{array}$$

Оценка $\hat{\beta}_W$ – это уже известная оценка «*within*», записанная в несколько иных, чем в п. 2.2.2 обозначениях.

Сумма квадратов остатков модели с ограничениями: $S_1 = W_{yy} - W'_{xy} W_{xx}^{-1} W_{xy}$.

МНК – оценивание сквозной регрессии или модели с ограничением Н3 порождает оценки параметров:

$$\left\{ \begin{array}{l} \hat{\beta} = T_{xx}^{-1} T_{xy} \\ \hat{\alpha} = y_{\cdot\cdot} - \beta' x_{\cdot\cdot}, \end{array} \right. \quad \text{где} \quad \begin{array}{l} T_{xx} = \sum_{i=1}^N \sum_{t=1}^T (x_{it} - x_{\cdot\cdot})(x_{it} - x_{\cdot\cdot})' = x' T^* x \\ T_{xy} = \sum_{i=1}^N \sum_{t=1}^T (x_{it} - x_{\cdot\cdot})(y_{it} - y_{\cdot\cdot})' = x' T^* y, \\ T_{yy} = \sum_{i=1}^N \sum_{t=1}^T (y_{it} - y_{\cdot\cdot})^2 = y' T^* y \\ y_{\cdot\cdot} = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T y_{it} \quad x_{\cdot\cdot} = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T x_{it}. \end{array}$$

Сумма квадратов остатков модели с ограничением Н3: $S_3 = T_{yy} - T'_{xy} T_{xx}^{-1} T_{xy}$.

Для проверки ограничения Н1 используется F-тест:

$$F_1 = \frac{(S_1 - S_0) / [(N-1)K]}{S_0 / [NT - N(K+1)]} \stackrel{H_1}{\sim} F((N-1)K, NT - N(K+1)).$$

Для проверки ограничения Н3 используется F-тест:

$$F_3 = \frac{(S_3 - S_0) / [(N-1)(K+1)]}{S_0 / [NT - N(K+1)]} \stackrel{H_3}{\sim} F((N-1)(K+1), NT - N(K+1)).$$

Логика исследования такова: если гипотеза Н3 отвергается, то проверяется гипотеза Н1, если Н1 отвергается, можно проверить гипотезу Н2 или оценивать регрессию без ограничений. Если же гипотезу Н1 нет оснований отвергнуть, про-

веряется гипотеза о гомогенности свободного члена при условии, что гипотеза о гомогенности наклона выполнена, т.е.

$$H4: \alpha_1 = \alpha_2 = \dots = \alpha_N \mid \beta_1 = \beta_2 = \dots = \beta_N.$$

Для проверки ограничения H4 используется F-тест:

$$F_4 = \frac{(S_3 - S_1)/(N-1)}{S_1/[N(T-1)-K]} \stackrel{H4}{\sim} F(N-1, N(T-1)-K).$$

Аналогично можно исследовать модель, где предполагается, что коэффициенты ведут себя одинаково для всех индивидуумов, но изменяются со временем.

6. Приложение:

Оценивание уравнения заработной платы по данным РМЭЗ

6.1. Постановка задачи

В качестве примера практического приложения рассмотренных моделей оценим уравнение заработной платы на основании данных РМЭЗ – Российского мониторинга экономического положения и здоровья населения. РМЭЗ представляет собой единственное в России представительное панельное обследование семей. Особенности данных РМЭЗ для изучения рынка труда см. в [3, 4, 6–9].

Индивидуальная анкета РМЭЗ содержит вопросы, ответы на которые предоставляют широкий спектр информации о заработной плате, в частности о сумме зарплаты, полученный в прошлом месяце, о наличии и величине задолженности и т.д., а также данные о социально-демографических характеристиках респондентов. Сбор данных проводится в последнем квартале каждого года.

Целью нашего исследования будет изучение эффекта инвестиций в человеческий капитал (в данном случае отраженных переменной $educ_{it}$) на заработную плату в экономике переходного периода. Включение остальных переменных призвано предотвратить смещение ошибок спецификации.

В панель, которую мы будем исследовать, вошли данные 1994, 1996, 1998, 2000 гг. Эти годы относятся к периоду высокой инфляции в российской экономике. В отдельные месяцы этого периода месячные темпы роста потребительских цен достигали 10–15% (например, конец 1994 или 1998 гг.). При этом в разных регионах темпы роста номинальной заработной платы могли заметно различаться. Поэтому номинальная заработная плата в нашем исследовании дефлирована с помощью официальных месячных индексов потребительских цен для соответствующих регионов. Деноминация 1996 г. тоже учтена.

Оцениваемое уравнение имеет вид:

$$l\text{wage}_{it} = b_0 + b_1 \text{educ}_{it} + b_2 \text{age}_{it} + b_3 \text{age2}_{it} + b_4 \text{stagna}_{it} + b_5 \text{gen}_i + b_6 \text{marst}_{it} + b_7 \text{city}_{it} + b_8 \text{isco}_1_{it} + b_9 \text{isco}_2_{it} + \dots + b_{14} \text{isco}_7_{it} + b_{15} \text{isco}_8_{it} + \varepsilon_{it},$$

где $l\text{wage}_{it}$ – логарифм месячной заработной платы;

$educ_{it}$ – продолжительность образования (в годах);

age_{it} – возраст;
 $age2_{it}$ – квадрат возраста;
 $stagna_{it}$ – стаж на данном месте работы;
 gen_i – пол;
 $marst_{it}$ – семейный статус;
 $city_{it}$ – тип места проживания (город=1 или село=0);
 $isco_1_{it}$ – $isco_8_{it}$ – дамми-переменные для профессиональных групп по классификации ISCO-88, $isco_9$ (неквалифицированные рабочие) – референтная группа для сравнений.

Оценивание произведено в пакете «СТАТА».

5.2. Модель с индивидуальными эффектами

Сквозное оценивание уравнения нашей модели, игнорирующее панельную природу данных, приводит к следующим результатам:

```

Number of obs =      8005
F(15, 7989) =      18.35
Prob > F      =      0.0000
R-squared     =      0.0333
Adj R-squared =      0.0315
Root MSE     =      2.9936
  
```

lwage	Coef.	Std. Err.	t	P>t
educ	-.0070401	.0143194	-0.49	0.623
age	.0839767	.0189821	4.42	0.000
age2	-.0010822	.0002239	-4.83	0.000
stagna	.043348	.0226577	1.91	0.056
gen	-.3003566	.0811726	-3.70	0.000
marst	-.2823525	.0393991	-7.17	0.000
city	.516275	.0777583	6.64	0.000
isco_1	-.9586997	.215312	-4.45	0.000
isco_2	.6404035	.139976	4.58	0.000
isco_3	.4757471	.1342797	3.54	0.000
isco_4	.5639411	.1601308	3.52	0.000
isco_5	.073882	.1522852	0.49	0.628
isco_6	-.1159434	.4916549	-0.24	0.814
isco_7	.5972895	.1347492	4.43	0.000
isco_8	.4241841	.1303195	3.25	0.001
_cons	8.071219	.4005636	20.15	0.000

Значения коэффициентов детерминации (R-squared и Adj R-squared) невелики, что, во-первых, типично для такого рода данных, а во-вторых, связано с отсутствием данных об отраслях, где заняты респонденты. Однако F-тест показывает значимость зависимости в целом. Интересующий нас коэффициент оказывается незначим, но это может быть вызвано смещением оценки, связанным с пропуском существенных переменных и учетом гетерогенности выборки.

Попробуем исправить ситуацию, оценив регрессию со случайным индивидуальным эффектом, что позволит нам учесть гетерогенность выборки в ковариационной матрице случайных ошибок:

```

Random-effects GLS regression           Number of obs       8005
Group variable (i) : aid_i             Number of groups =  3538
R-sq:  within = 0.1212                 Obs per group: min =   1
      between = 0.0023                   avg =   2.3
      overall = 0.0287                   max =   4
Random effects u_i ~ Gaussian          Wald chi2(15) =    354.02
corr(u_i, X)      (assumed)           Prob > chi2      =    0.0000

-----+-----
lwage      Coef.   Std. Err.      z    P>z
-----+-----
educ     -.0234971  .0177263    -1.33  0.185
age        .061222    .0237307     2.58   0.010
age2       -.0010815   .0002816    -3.84   0.000
stagna     .0463718   .0259667     1.79   0.074
gen        -.2052513   .1066411    -1.92   0.054
marst      -.3730998   .045139     -8.27   0.000
city       .5546138   .1034034     5.36   0.000
isco_1     -1.408905  .2413125    -5.84   0.000
isco_2     .7235324   .1675966     4.32   0.000
isco_3     .5017355   .1588268     3.16   0.002
isco_4     .6695941   .1912666     3.50   0.000
isco_5     -.0584244   .1806597    -0.32   0.746
isco_6     .1134344   .548757      0.21   0.836
isco_7     .7832161   .1590831     4.92   0.000
isco_8     .4717575   .1554644     3.03   0.002
_cons     9.226788   .500212     18.45   0.000
-----+-----
sigma_u    1.343514
sigma_e    1.9862585
rho        .31390445   (fraction of variance due to u_i)

```

Прокомментируем особенности, связанные с оцениванием этой модели. В исследуемой нами панели участвует 3538 индивидуумов (Number of groups), но не для всех из них оказывается доступной вся запрашиваемая информация. Если в каком-то году респондентом пропущен ответ хотя бы на один вопрос, программа игнорирует все наблюдения, относящиеся к этому году, поэтому в среднем индивидуумы наблюдаются не 4 года, а около двух с половиной лет. Еще одна особенность – появление трех разных коэффициентов детерминации. В данном случае их нет смысла интерпретировать, поскольку регрессия оценивается с помощью обобщенного МНК, а, значит, R^2 не может служить адекватной мерой качества регрессии, но в регрессии с детерминированным эффектом R^2 опять приобретает смысл. О том, что регрессия в целом значима, свидетельствует высокое значение статистики Вальда (Wald chi2=354). В регрессии с индивидуальным эффектом F-тест на значимость регрессии в целом заменяется при работе в пакете «STATA» тестом Вальда. Результаты оценивания коэффициентов регрессии несколько отличаются от предыдущего случая, а именно, значимость переменной *gen* упала. Интересующий нас эффект по-прежнему незначим.

Напомним, что, оценивая последнюю регрессию, мы исходили из предположения о некоррелированности индивидуального эффекта и независимых переменных, но это предположение не очень обосновано. В ненаблюдаемый индивидуальный эффект входят различные компоненты, например способности респондента, которые, как правило, коррелируют с образованием и профессиональной группой. Другая ненаблюдаемая компонента индивидуального эффекта – это отрасль, где занят респондент, может коррелировать и с возрастом, и с образованием, и со стажем, поскольку, например, в прибыльных топливно-энергетических отраслях аккумулируются наиболее молодые и энергичные индивидуумы. Из всего вышесказанного следует, что модель с детерминированным индивидуальным эффектом может более адекватна данным, чем две предыдущие. Оценим теперь ее:

```
Fixed-effects (within) regression      Number of obs   =   8005
Group variable (i) : aid_i             Number of groups =   3538
R-sq:  within = 0.6454                 Obs per group:  min =    1
      between = 0.0471                   avg   =   2.3
      overall  = 0.0001                  max   =    4
                                         F(14,4453)     =   578.88
corr(u_i, Xb) = -0.9771                Prob > F       =    0.0000
```

lwage	Coef.	Std. Err.	t	P>t
educ	.0012534	.0304621	0.04	0.967
age	-.945309	.0464557	-20.35	0.000
age2	-.0004523	.0005528	-0.82	0.413
stagna	-.0160337	.0271686	-0.59	0.555
gen	(dropped)			
marst	-.2265007	.0486234	-4.66	0.000
city	(dropped)			
isco_1	-.6047924	.255228	-2.37	0.018
isco_2	-.247772	.2148656	-1.15	0.249
isco_3	-.0841697	.1913615	-0.44	0.660
isco_4	.1709404	.2302459	0.74	0.458
isco_5	-.231582	.2210071	-1.05	0.295
isco_6	.5360691	.6245203	0.86	0.391
isco_7	.3115733	.1920045	1.62	0.105
isco_8	.2398692	.1916611	1.25	0.211
_cons	43.50516	1.715893	25.35	0.000

```
sigma_u  13.433868
sigma_e  1.9862585
rho      .97860671  (fraction of variance due to u_i)
```

```
F test that all u_i=0:  F(3537, 4453) = 3.87  Prob > F = 0.0000
```

Большая часть вариации данных (98%) приходится на индивидуальные эффекты: $\rho = 0,9786$. Велика корреляция индивидуальных эффектов с регрессорами: $\text{corr}(u_i, Xb) = -0,9771$, что говорит в пользу этого метода оценивания. Коэффициенты при двух регрессорах gen_i и $city_{it}$ не оцениваются. Это происходит оттого, что эти переменные не меняются со временем. Слабо меняется со

временем и большая часть остальных переменных, в том числе и продолжительность образования, поскольку образование люди получают в основном в молодом возрасте. Хотя оценки коэффициентов в этой модели теперь освобождены от смещения гетерогенности, но слабая изменчивость данных по времени не позволяет получить значимые результаты.

6.3. Качество подгонки и выбор наиболее адекватной модели

Теперь уместно прокомментировать смысл трех коэффициентов детерминации. Здесь под R-sq понимаются квадраты выборочных коэффициентов корреляции между наблюдаемыми и оцененными значениями объясняемой переменной, заданными в соответствующей форме, а именно:

$$R_{within}^2(\hat{\beta}_W) = \text{corr}^2\{\hat{y}_{it}^W - \hat{y}_{i\cdot}^W, y_{it} - y_{i\cdot}\},$$

где $y_{i\cdot} = \frac{1}{T} \sum_{t=1}^T y_{it}$ – усредненные по времени для каждого i -го объекта значения зависимой переменной, $\hat{y}_{it}^W = X_{it}' \hat{\beta}_W$, $\hat{y}_{i\cdot}^W = X_{i\cdot}' \hat{\beta}_W$.

$$R_{between}^2(\hat{\beta}_W) = \text{corr}^2\{\hat{y}_{i\cdot}^W, y_{i\cdot}\},$$

$$R_{overall}^2(\hat{\beta}_W) = \text{corr}^2\{\hat{y}_{it}^W, y_{it}\}.$$

О качестве подгонки в этой модели следует судить по коэффициенту детерминации $R_{within}^2(\hat{\beta}_W) = \text{corr}^2\{\hat{y}_{it}^W - \hat{y}_{i\cdot}^W, y_{it} - y_{i\cdot}\} = 0,6454$. Это достаточно высокий показатель. Значимость регрессии в целом тоже велика: $F(14,4453) = 578,88$ и $\text{Prob} > F = 0,0000$, но все это достигается учетом индивидуальных эффектов и только. И здесь нам не удалось выявить влияние образования на заработную плату.

В самом конце таблицы результатов оценивания модели с детерминированными индивидуальными эффектами приводится тест на значимость детерминированных индивидуальных эффектов:

$$F \text{ test that all } u_{i=0}: F(3537, 4453) = 3.87 \quad \text{Prob} > F = 0.0000.$$

Это F-тест для проверки гипотезы H_4 из п. 5.2 (гипотезы о гомогенности свободного члена при условии, что гипотеза о гомогенности наклона выполнена).

Результаты свидетельствуют в пользу модели с детерминированными индивидуальными эффектами и против модели сквозной регрессии.

Теперь сделаем тест Бройша – Пагана, суть которого была изложена в п. 4.3, для осуществления выбора между сквозной регрессией и регрессией со случайным индивидуальным эффектом.

Breusch and Pagan Lagrangian multiplier test for random effects:

$$lwage[aid_i,t] = Xb + u[aid_i] + e[aid_i,t]$$

Estimated results:

	Var	sd = sqrt(Var)
lwage	9.252759	3.041835
e	3.945223	1.986258
u	1.80503	1.343514

Test: Var(u) = 0

$$\begin{aligned} \text{chi2}(1) &= 287.60 \\ \text{Prob} > \text{chi2} &= 0.0000 \end{aligned}$$

Значение статистики $\text{chi2}(1) = 287,60$ говорит в пользу регрессии со случайным индивидуальным эффектом.

И, наконец, тест Хаусмана убедительно демонстрирует, что обнаруженный случайный эффект сильно коррелирован с регрессорами, т.е. наиболее адекватна данным модель с детерминированными эффектами, что мы и предполагали.

Hausman specification test

	---- Coefficients ----		
	Fixed	Random	
lwage	Effects	Effects	Difference
educ	.0012534	-.0234971	.0247505
age	-.945309	.061222	-1.006531
age2	-.0004523	-.0010815	.0006292
stagna	-.0160337	.0463718	-.0624055
gen	9.909991	-.2052513	10.11524
marst	-.2265007	-.3730998	.1465992
isco_1	-.6047924	-1.408905	.804113
isco_2	-.247772	.7235324	-.9713044
isco_3	-.0841697	.5017355	-.5859052
isco_4	.1709404	.6695941	-.4986538
isco_5	-.231582	-.0584244	-.1731576
isco_6	.5360691	.1134344	.4226347
isco_7	.3115733	.7832161	-.4716428
isco_8	.2398692	.4717575	-.2318883

Test: Ho: difference in coefficients not systematic

$$\begin{aligned} \text{chi2}(14) &= (b-B)'[S^{(-1)}](b-B), S = (S_{fe} - S_{re}) = 8513.46 \\ \text{Prob} > \text{chi2} &= 0.0000 \end{aligned}$$

6.4. Модель с индивидуальными и временными эффектами

В нашем исследовании мы не учли, что 1990-е годы были очень динамичным периодом в российской экономике, а значит, следует ожидать существенных отличий в регрессионных параметрах, относящихся к разным годам, или, иначе

говоря, существенных временных эффектов. Будем трактовать эти эффекты как детерминированные поправки к свободному члену и оценим нашу модель с их учетом, введя соответствующие дамми-переменные $d96$, $d98$, $d00$:

```

Random-effects GLS regression           Number of obs   =    8005
Group variable (i) : aid_i             Number of groups =   3538
R-sq:  within = 0.9622                 Obs per group:  min =    1
        between = 0.8591                                     avg  =    2.3
        overall = 0.9151                                     max  =    4
Random effects u_i ~ Gaussian          Wald chi2(18)   = 133668.46
corr(u_i, X) = 0 (assumed)            Prob > chi2     =    0.0000

```

	lwage	Coef.	Std. Err.	z	P> z
educ		.037834	.0048219	7.85	0.000
age		.0597386	.0065354	9.14	0.000
age2		-.0007217	.0000777	-9.28	0.000
stagna		-.0051972	.0066988	-0.78	0.438
gen		-.433246	.0302027	-14.34	0.000
marst		.010381	.0117268	0.89	0.376
city		.5286354	.0296726	17.82	0.000
isco_1		.5958171	.0621485	9.59	0.000
isco_2		.3778088	.0444242	8.50	0.000
isco_3		.3796478	.0417913	9.08	0.000
isco_4		.3155396	.0505183	6.25	0.000
isco_5		.3258878	.047718	6.83	0.000
isco_6		.3552049	.141719	2.51	0.012
isco_7		.3109485	.0419713	7.41	0.000
isco_8		.3803242	.0411601	9.24	0.000
d96		1.055731	.0220666	47.84	0.000
d98		-5.648684	.0222827	-253.50	0.000
d00		-4.884359	.0217856	-224.20	0.000
_cons		9.852308	.1380617	71.36	0.000
sigma_u		.62073471			
sigma_e		.64533893			
rho		.48057385	(fraction of variance due to u_i)		

Очевидно, что значимость регрессии существенно возросла, о чем свидетельствует значение статистики Вальда: $Wald\ chi2(18) = 133668,46$. Временной эффект оказался очень существенным, причем если в 1996 г. заработная плата была значимо выше, чем в 1994, то в 1998 и 2000 гг. она значительно ниже. Это объясняется последствиями дефолта. И наконец, нам удалось получить значимую и положительную оценку эффекта образования. Но вот результаты теста Хаусмана:

Hausman specification test

$$chi2(17) = (b-B)'[S^{-1}](b-B), S = (S_{fe} - S_{re}) = 90.15$$

$$Prob>chi2 = 0.0000.$$

Они свидетельствуют о том, что эта модель все же не адекватна данным, и нужно опять использовать модель с детерминированными эффектами:

```
Fixed-effects (within) regression      Number of obs      = 8005
Group variable (i) : aid_i            Number of groups   = 3538
R-sq:  within = 0.9626                Obs per group: min = 1
      between = 0.7498                  avg               = 2.3
      overall = 0.8700                  max               = 4
                                         F(17,4450)        =6735.74
corr(u_i, Xb) = -0.0195                Prob > F           = 0.0000
```

-----	-----	-----	-----	-----
lwage	Coef.	Std. Err.	t	P> t
-----	-----	-----	-----	-----
educ 	.000039	.0098972	0.00	0.997
age	.038562	.0395535	0.97	0.330
age2	-.0010249	.0001796	-5.71	0.000
stagna	-.0017276	.0088323	-0.20	0.845
gen	(dropped)			
marst	.0230691	.0158567	1.45	0.146
city	(dropped)			
isco_1	.3033066	.083247	3.64	0.000
isco_2	.1202153	.0698838	1.72	0.085
isco_3	.1766294	.0622356	2.84	0.005
isco_4	.1893502	.0748159	2.53	0.011
isco_5	.1514229	.0718546	2.11	0.035
isco_6	.4854903	.2030174	2.39	0.017
isco_7	.219465	.0624326	3.52	0.000
isco_8	.2518914	.0622966	4.04	0.000
d96	1.102401	.0733646	15.03	0.000
d98	-5.484452	.1512189	-36.27	0.000
d00	-4.600193	.2241671	-20.52	0.000
_cons	11.9032	1.379644	8.63	0.000
-----	-----	-----	-----	-----
sigma_u	1.0577206			
sigma_e	.64533893			
rho	.72873059	(fraction of variance due to u_i)		
-----	-----	-----	-----	-----

F test that all u_i=0: F(3537, 4450) = 3.00 Prob > F = 0.0000

Мы получили, судя по $R\text{-sq}(\text{within}) = 0,9626$ и статистике $F(17,4450)=6735,74$, достаточно качественную модель, оценки которой свободны от гетерогенного смещения и смещения ошибки спецификации, имевшей место пока мы не учли временной эффект. Последнее обстоятельство хорошо отражается коэффициентом корреляции между индивидуальными эффектами и регрессорами: $\text{corr}(u_i, Xb) = -0,0195$. Это значение в 50 раз меньше, чем в модели, учитывающей только детерминированные индивидуальные эффекты. Теперь 70% разброса наблюдений приходится на индивидуальные эффекты (вместо бывших 98%) и оставшиеся 30% в основном объясняются временными эффектами. Значимого влияния половины регрессоров по-прежнему не удастся выявить из-за их слабой временной динамики. Исключения составляют только переменные, отве-

чающие за принадлежность к профессиональным группам. Интересующий нас эффект образования опять оказывается незначимым.

Еще раз подтверждаются данные о значительном падении уровня заработной платы в 1998 и 2000 гг. по сравнению с 1994 г., но результаты теста Вальда, сопоставляющего коэффициенты при дамми-переменных, соответствующих временным эффектам 1998 и 2000 гг., показывают, что уже в 2000 г. заработная плата стала значимо выше, чем в 1998 г.:

$$F(1, 4450) = 130,70 \quad \text{Prob} > F = 0,0000.$$

Мы учли временной эффект с помощью аддитивных дамми-переменных, но, может быть, в разные годы коэффициент наклона при образовании тоже был разным. Введение в регрессию дополнительных переменных, элиминирующих временной эффект в коэффициент наклона при переменной *educ*, приводит нас к такому результату:

```
Fixed-effects (within) regression      Number of obs   =      8005
Group variable (i) : aid_i            Number of groups =      3538
R-sq:  within = 0.9627                Obs per group:  min =         1
      between = 0.7413                  avg =         2.3
      overall  = 0.8664                 max =         4
                                         F(20,4447)     =    5735.33
corr(u_i, Xb) = -0.0282                Prob > F        =     0.0000
```

	Coef.	Std. Err.	t	P> t
lwage				
educ	-.0095355	.0110487	-0.86	0.388
deduc96	-.002887	.0078136	-0.37	0.712
deduc98	.0125607	.0084006	1.50	0.135
deduc00	.0225405	.0084271	2.67	0.008
d96	1.14552	.1242481	9.22	0.000
d98	-5.6317	.1836601	-30.66	0.000
d00	-4.867094	.2475536	-19.66	0.000
_cons	12.11421	1.380885	8.77	0.000
sigma_u	1.0754491			
sigma_e	.64480761			
rho	.73557272	(fraction of variance due to u_i)		

F test that all u_i=0: F(3537, 4447) = 3.01 Prob > F = 0.0000

Отсюда можно сделать окончательные выводы:

- инвестиции в человеческий капитал, в частности в образование, были обесценены в 1990-е гг., об этом говорят незначимые коэффициенты при переменных **educ**, **deduc96**, и **deduc98**;

- в 2000 г. тенденция начала меняться. Наличие образования вызывает ускоренный рост заработной платы, о чем свидетельствует значимый коэффициент при переменной **deduc00**. Этот коэффициент показывает, что в 2000 г. у индивидуумов с более высоким уровнем образования уровень заработной платы был значимо выше, чем у остальных;

• значимые коэффициенты при временных дамми-переменных показывают рост уровня заработной платы в 1996 г., существенный его спад в 1998 г. и небольшой подъем в 2000 г.

* *

*

СПИСОК ЛИТЕРАТУРЫ

1. *Анатольев С.* Курс лекций по эконометрике для продолжающих. Российская экономическая школа. 2002. (<http://www.nes.ru/Acad-year-2003/5th-module/econometrics-3-rus.htm>).
2. *Васькович Н., Гурова Е., Поляков К.* Регрессионная модель панельных данных с однофакторной случайной составляющей // Математические модели экономики: Сборник научных трудов. М.: МИЭМ, 2002.
3. *Гимпельсон В., Капелюшников Р., Ратникова Т.* Страх безработицы и гибкость заработной платы в России // Экономический журнал ВШЭ. Т. 7. № 3. 2003.
4. *Колеников С.* Прикладной эконометрический анализ в статистическом пакете STATA. М.: Российская экономическая школа, 2001.
5. *Магнус Я.Р., Катышев П.К., Пересецкий А.А.* Эконометрика. Начальный курс: Учебник. 5-е изд., испр. М.: Дело, 2004.
6. *Нестерова Д., Сабирьянова К.* Инвестиции в человеческий капитал в переходный период в России. Доклад на конференции EERC. 1999.
7. *Ратникова Т.А.* Анализ панельных данных в пакете STATA. Методические указания к компьютерному практикуму по курсу «Эконометрический анализ панельных данных». М.: ГУ ВШЭ, 2005.
8. *Роцин С.Ю.* Предложение труда в России: микроэкономический анализ экономической активности населения: Препринт WP3/2003/02. Серия «Проблемы рынка труда». М.: ГУ ВШЭ, 2003.
9. Список публикаций на основе данных Российского мониторинга экономического положения и здоровья населения (РМЭЗ). Материалы конференции «Российский мониторинг экономического положения и здоровья населения», 17 июня 2003 г.
10. *Arellano M., Bond S.R.* Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations // Review of Economic Studies. 1991. Vol. 58.
11. *Baltagi B.* Econometric Analysis of Panel Data. John Wiley & Sons, 1995.
12. *Baltagi B.H., Raj B.* A Survey of Recent Theoretical Developments in the Econometrics of Panel Data // Empirical Economics. 1992. Vol. 17.
13. *Chamberlain G.* Omitted Variable Bias in Panel Data. Estimating the Return to Schooling // Annales de l'INSEE. 1978. № 30/31.
14. *Chamberlain G.* Panel Data. Handbook of Econometrics / Ed. by Z. Griliches, M.D. Intriligator. 1984. Vol. II.
15. *Cheng H.* Analysis of Panel Data: 1st ed. Cambridge University Press, 1986.
16. *Cornwell C., Trumbull W.N.* Estimating the Economic Model of Crime with Panel Data // The Review of Economics and Statistics. 1994. Vol. 76. № 2.
17. *Dormont B.* Introduction à l'Économétrie des données de panel. Paris, 1989.

18. *Frisch R., Waugh F.V.* Partial Time Regressions as Compared with Individual Trends // *Econometrica*. 1933. Vol. 1.
19. *Hausman J.A., Taylor W.E.* Panel Data and Unobservable Individual Effects // *Econometrica*. Vol. 49.
20. *Heckman J.J.* Micro Data, Heterogeneity and Evaluation of Public Policy. Nobel Lecture // *Journal of Political Economy*. 2001. Vol. 109. № 4.
21. *Heckman J.J., Macurdy T.E.* The Review of Economic Studies // *Econometrics Issue*. 1980. Vol. 47. № 1.
22. *Greene W.H.* *Econometric Analysis*. 3rd ed. Prentice Hall, 1997. (Chapter 14.)
23. *Griliches Z.* Estimating the Return to Schooling: Some Econometric Problems // *Econometrica*. 1977. Vol. 45.
24. *Griliches Z., Hausman J.A.* Errors in Variables in Panel Data // *Econometrica*. 1986. Vol. 54.
25. *Kiefer N.M.* Population Heterogeneity and Inference from Panel Data on the Effects of Vocational Education // *Journal of Political Economy*. 1979. Vol. 87. № 5.
26. *Kim B.S., Maddala G.S.* Estimation and Specification Analysis of Models of Dependent Behavior Based on Censored Panel Data // *Empirical Economics*. 1992. Vol. 17.
27. *Lovell M.C.* Seasonal Adjustment of Economic Time Series // *Journal of the American Statistical Association*. 1963. № 58.
28. *Mundlak Y.* On the Pooling of Time Series and Cross-Section Data // *Econometrica*. 1978. Vol. 46.
29. *Sabirianova K.Z.* The Great Human Capital Reallocation: A Study of Occupational Mobility in Transitional Russia // *Journal of Comparative Economics*. 2002. № 30.
30. *Sevestre P., Trognon A.* A Note on Autoregressive Error Component Models // *Journal of Econometrics*. 1985. Vol. 28.
31. *Tekin E.* Employment, Wages and Alcohol Consumption in Russia: Evidence from Panel Data // *IZA Discussion Paper*. 2002. № 432.
32. *Trognon A.* Donnees individuelles temporelles. Polycope de l'ENSAE. Cours d'Econometrie II. 1987. Tomes 2 et 3.
33. *Verbeek M.* *A Guide to Modern Econometrics*. John Wiley & Sons, 2003.
34. *Verbeek M., Nijman Th.* Can Cohort Data Be Treated as Genuine Panel Data? // *Empirical Economics*. 1992. Vol. 17.
35. *Wansbeek T.J., Koning R.H.* Measurement Error and Panel Data // *Statistica Neerlandica*. 1989. Vol. 45.