Министерство образования и науки Российской Федерации

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ «САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н. Г. ЧЕРНЫШЕВСКОГО»

Кафедра теории функций и
стохастического анализа

# ОЦЕНИВАНИЕ РЕГРЕССИОННЫХ МОДЕЛЕЙ ДЛЯ ПАНЕЛЬНЫХ ДАННЫХ ОБОБЩЕННЫМ МЕТОДОМ МОМЕНТОВ

АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ


студента 4 курса 412 группы
направления 01.03.02 — Прикладная математика и информатика
механико-математического факультета
Донскова Игоря Вячеславовича


Научный руководитель
старший преподаватель ______________ А. Д. Луньков

Заведующий кафедрой
д. ф.-м. н. ______________ С. П. Сидоров

Саратов 2018

# ВВЕДЕНИЕ

**Актуальность темы.** В эконометрике до определенного момента активно развивались две методики, основанные на моделях перекрестных данных и моделях временных рядов. Каждый метод относился к определенному типу статистических данных. Классический регрессионный анализ применяется к перекрестным данным, для которых единицей наблюдения является некоторая пространственная, социальная или физическая единица, временные же ряды в качестве единицы выборки берут временной момент. Методология панельных данных синтезировала два вышеупомянутых метода. Она очень популярна, так как позволяет исследовать с помощью регрессионных моделей эффекты, присущие отдельным наблюдениям и строить более гибкие модели. Одним из современных методов оценивания моделей для панельных данных является обобщенный метод моментов. Он позволяет решать более масштабные задачи по сравнению с классическим методом моментов, давая однако некоторые погрешности в точности выполнения накладываемых условий.

Особенностью современных регрессионных моделей является наличие следующих составляющих: учет взаимосвязи между единицами наблюдения с помощью весовых матриц, учет как пространственных, так и временных эффектов, наличие обратных связей, учитываемых с помощью системы одновременных уравнений, наличие режимов переключения между видами моделей. Первой из упомянутых составляющих уделено внимание в этой работе.

**Целью бакалаврской работы** является оценивание параметров пространственных панельных регрессионных моделей, применительно к российским региональным данным, с помощью обобщенного метода моментов.

**Объект исследования** – панельные данные.

**Предмет исследования** – пространственные эконометрические модели для индекса цен на жилье, среднего дохода, прироста населения, плотности населения.

Для достижения поставленных целей в работе необходимо решить следующие **задачи**:

– рассмотреть модель парной регрессии и основные гипотезы связанные с этой моделью;

– определить основные гипотезы, лежащие в основе модели множественной регрессии, описать методику построения оценок её параметров;

— рассмотреть основные регрессионные модели для панельных данных;

— описать структуру моделей с фиксированным и случайным эффектом;

— изучить методы оценивания - в частности, обобщенный метод моментов;

— описать построение регрессионной модели для панельных данных, и методику ее оценивания;

— создать код, позволяющий оценить параметры пространственной регрессионной модели для панельных данных с помощью обобщенного метода моментов;

— провести анализ полученных результатов. Помимо прочего, будет создана программа, позволяющая оценивать параметры регрессионной модели фиксированных эффектов для искусственно сгенерированных данных.

**Практическая значимость.** Исследована зависимость индекса цен на жилье в регионах от индекса цен в соседствующих регионах, дохода, плотности населения и миграционного прироста внутри региона. Модель построена на основе данных, полученных с портала http://www.gks.ru/ и может быть полезна для прогнозирования цен на жилье, выявления доходов, наиболее существенно влияющих на эту цену. Эта же методика может быть применена при переходе на более низкий уровень, от регионов к районам, при наличии информации, и может быть полезна деятельности муниципалитета. Создан программный продукт и проанализированы по реальным современным социально-экономическим данным зависимости между вышеперечисленными показателями. Результатам дана содержательная интерпретация.

**Структура и содержание бакалаврской работы.** Работа состоит из введения, пяти разделов, заключения, списка использованных источников, содержащего 20 наименований, и двух приложений. Общий объем работы составляет 40 страниц.

## ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во **введении** обосновывается актуальность темы работы, формулируются цель работы и решаемые задачи.

В **первом** разделе рассматривается модель парной регрессии.

**Подгонка кривой.**

Ставиться задача подобрать («подогнать») функцию $Y = f(X)$ из параметрического семейства функций $f(X,\beta)$, «наилучшим» способом описывающую зависимость Y от X.

В качестве меры отклонения функции $f(X,\beta)$ от набора наблюдений можно взять:

1. сумму квадратов отклонений $F = \sum\limits_{t=1}^{n} (Y_t - f(Xt,\beta))^2$,
2. сумму модулей отклонений $F = \sum\limits^{n} t = 1|Y_t - f(Xt,\beta)|$,
3. $F = \sum\limits^{n} t = 1g(Y_t - f(X_t,\beta))$, где $g$ – любое преобразование отклонения $Y_t - f(X_t,\beta)$, входящего в функционал $F$.

**Линейная регрессионная модель с двумя переменными.**

Модель зависимости $Y_t$ от $X_t$ построим в виде

$Y_t = a + bX_t + \varepsilon_t$, *t = 1,. . . , n,*

где $X_t$ — неслучайная величина, а $Y_t$, $\varepsilon_t$ — случайные величины. $Y_t$ называется объясняемой (зависимой) переменной, а $X_t$ – объясняющей (независимой) переменной или *регрессором*. Уравнение, приведенное выше, также называется *регрессионным уравнением*.

**Основные гипотезы**

1. $Y_t$ = a + b$X_t$ + $\varepsilon_t$, $t = 1, ..., n$, — спецификация модели.
2. $X_t$ — детерминированная величина; вектор ($X_1$,..., $X_t$)$'$ неколлинеарен вектору $r = (1, ..., 1)'$.
3. E$\varepsilon_t$ = 0, E($\varepsilon_t^2$) = V($\varepsilon_t$) = $\sigma^2$ — не зависит от t.
4. E($\varepsilon_t\varepsilon_s$) = 0 при t $\neq$ s, некоррелированность ошибок для разных наблюдений.

Часто добавляется условие:

5. Ошибки $\varepsilon_t$, t = 1,..., n, имеют совместное нормальное распределение: $\varepsilon_t$ $\sim$ N(0, $\sigma^2$).

В этом случае модель называется нормальной линейной регрессионной.

Проинтерпретируем гипотезы, лежащие в основе линейной регрессионной модели.

1. Спецификация модели отражает представление о механизме зависимости $Y_t$ от $X_t$ и сам выбор объясняющей переменной $X_t$.

3-4. Эти условия в векторной форме могут быть записаны так:

$$E\varepsilon = 0, V(\varepsilon) = \sigma^2 I_n,$$

где $\varepsilon$ = ($\varepsilon_1$,..., $\varepsilon_n$)' $I_n$ — n×n единичная матрица, V($\varepsilon$) — n×n матрица ковариаций.

Условие E$\varepsilon$ = 0 означает, что $EY_t = a + bX_t$ , т. е. при фиксированном $X_t$ среднее ожидаемое значение $Y_t$ равно $a + bX_t$ .

Условие независимости дисперсии ошибки от номера наблюдения (от регрессора $X_t$): E($\varepsilon_t^2$) = V($\varepsilon_t$) = $\sigma^2$, t = 1,. . . , n, называется *гомоскедастичностью*.

Условие E($\varepsilon_t\varepsilon_s$) = 0, t $\neq$ s указывает на некоррелированность ошибок для разных наблюдений. В случае, когда это условие не выполняется, говорят об *автокорреляции ошибок*.

Далее в разделе рассматриваются методы оценивания параметров $a, b, \sigma^2$:

1. с помощью метода наименьших квадратов в предположении теоремы Гаусса-Маркова.

$$\hat{b} = \frac{n \sum X_t Y_t - ( \sum X_t)( \sum Y_t)}{n \sum X_t^2 - (\sum X_t)^2},$$

$$\hat{a} = \frac{1}{n} \sum Y_t - \frac{1}{n} \sum X_t \hat{b} = \overline{Y} - \overline{X}\hat{b}.$$

А также оценка $\sigma^2$:

$$\widehat{\sigma}^2 = \frac{1}{n-2} \sum e_t^2$$

2. с помощью метода максимального правдоподобия:

$$\hat{b}_{ML} = \frac{\sum x_t y_t}{\sum x_t^2}; \quad \hat{a}_{ML} = \overline{Y} - \hat{b}_{ML}\overline{X}; \quad \widehat{\sigma}^2_{ML} = \frac{1}{n} \sum e_t^2.$$

Во **втором** разделе рассмотрена модель множественной регрессии.

Она является обобщением модели с двумя переменными:

$$y_t = \beta_1 x_{t1} + \beta_2 x_{t2} + \cdots + \beta_k x_{tk} + \varepsilon_t, \quad t = 1, \ldots, \ n,$$

где $x_{tp}$ — значения регрессора $x_p$ в наблюдении $t$, а $x_{t1} = 1$, $t = 1, \ldots, n$.

**Основные гипотезы**

Гипотезы, лежащие в основе модели множественной регрессии, являются естественным обобщением модели парной регрессии:

1. $y_t = \beta_1 + \beta_2 x_{t2} + \cdots + \beta_k x_{tk} + \varepsilon_t, \quad t = 1, \ldots, \ n,$— спецификация модели.
2. $x_{t1}, \ldots, \ x_{tk}$— детерминированные величины. Векторы $x_s = (x_{1s}, \ldots, x_{ns})'$,

$s = 1, ..., k$ линейно независимы в $R^n$.

3-5 совпадают с гипотезами парной регрессионной модели.

В этом случае модель называется *нормальной линейной регрессионной*.

Далее в разделе оценивались следующие параметры:

$$\widehat{\beta}_{OLS} = (X'X)^{-1}X'y \text{ -- методом наименьших квадратов,}$$

$$\widehat{\sigma}^2 = \frac{e'e}{n-k} = \frac{\sum e_t^2}{n-k}$$

А также проверялась гипотеза *линейного ограничения общего вида* $H_0 : H\beta = r$, с помощью следующей статистики:

$$F = \frac{(H\widehat{\beta} - r)'(H(X'X)^{-1}H')^{-1}(H\widehat{\beta} - r)/q}{e'e/(n-k)} \sim F(q, n-k).$$

**Третий** раздел посвящен панельным данным и построению для них основных регрессионных моделей. Панельные данные состоят из наблюдений одних и тех же экономических единиц или объектов. Сбор данных проводится в последовательные моменты времени.

**Обозначение и основные модели**

Вводятся обозначения. Пусть $y_{it}$ – зависимая переменная для экономической единицы $i$ в момент времени $t$, $x_{it}$ – набор объясняющих (независимых) переменных (вектор размерности $k$) для той же единице и того же времени, и $\varepsilon_{it}$ – соответствующая ошибка, $i = 1, ..., n,\ t = 1, ..., T$. При переходе к векторам используются обозначения

$$y_i = \begin{bmatrix} y_{i1} \\ \vdots \\ y_{iT} \end{bmatrix}, \quad X_i = \begin{bmatrix} x'_{i1} \\ \vdots \\ x'_{iT} \end{bmatrix}, \quad \varepsilon_i = \begin{bmatrix} \varepsilon_{i1} \\ \vdots \\ \varepsilon_{iT} \end{bmatrix}.$$

Вводятся также "объединенные" наблюдения и ошибки:

$$y = \begin{bmatrix} y_1 \\ \vdots \\ y_T \end{bmatrix}, \quad X = \begin{bmatrix} x'_1 \\ \vdots \\ x'_T \end{bmatrix}, \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_T \end{bmatrix}.$$

(Здесь $y, \varepsilon - nT \times 1$ векторы, $X - nT \times k$ матрица.)

Простейшая модель – это обычная линейная модель регрессии в матричной форме

$$y = X\beta + \varepsilon, \tag{1}$$

которая, по существу, не учитывает панельную структуру данных. При этом предполагается, что все ошибки $\varepsilon_{it}$ некоррелированы между собой как по $i$, так и по $t$, и некоррелированы со всеми объясняющими переменными $x_{it}$. При выполнении этих предположений обычные МНК-оценки $\hat{\beta}_{OLS}$ являются состоятельными и эффективными.

Панельные данные позволяют учитывать индивидуальные различия между экономическими единицами. Одна из возможных реализаций этой идеи выглядит следующим образом:

$$y_{it} = \alpha_i + x'_{it}\beta + \varepsilon_{it}, \tag{2}$$

где величина $\alpha_i$ выражает индивидуальный эффект объекта $i$, не зависящий от времени $t$, при этом регрессоры $x_{it}$ не содержат константу.

В зависимости от предположений относительно характера величины $\alpha_i$ рассматриваются две модели.

Модель с фиксированным эффектом (fixed effect model): предполагается, что в соотношении (2) величины $\alpha_i$ являются неизвестными параметрами.

Модель со случайным эффектом (random effect model): предполагается, что в соотношении (2) $\alpha_i = \mu + u_i$, где $\mu$ — параметр, общий для всех единиц во все моменты времени, а $u_i$ — ошибки, некоррелированные с $\varepsilon_{it}$ и некоррелированные при разных $i$.

Методы оценивания опираются на понижение размерности вектора неизвестных переменных (удаление среднего), на классический и обобщенный МНК. В общем случае оцениваются вектор коэффициентов $\beta, \alpha_i$, дисперсия ошибки, дисперсия эффектов (в предположении случайного эффекта).

**Модель с фиксированным эффектом**

Модель с фиксированным эффектом (fixed effect model) описывается уравнением (2). Предполагается, что выполнены следующие условия:

1. ошибки $\varepsilon_{it}$ некоррелированы между собой по $i$ и $t$, $\mathrm{E}(\varepsilon_{it}) = 0$, $\mathrm{V}(\varepsilon_{it}) = \sigma^2_\varepsilon$;
2. ошибки $\varepsilon_{it}$ некоррелированы с регрессорами $x_{js}$ при всех $i, j, t, s$.

Если ввести фиктивные переменные для каждой экономической единицы: $d_{ij} = 1$, если $i = j$, и $d_{ij} = 0$, если $i \neq j$, то модель (2) может быть сведена к более привычному виду линейной регрессии

$$y_{it} = \sum_{j=1}^{n} \alpha_j d_{ij} + x'_{it}\beta + \varepsilon_{it} \tag{3}$$

Это спецификация модели с фиксированным эффектом.

Если объединить все фиктивные переменные в одну матрицу

$$D = \begin{bmatrix} \imath_T & 0 & \dots & 0 \\ 0 & \imath_T & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \imath_T \end{bmatrix} = I_n \otimes \imath_T,$$

где вектор $\imath_T = [1, \dots, 1]'$ имеет размерность $T$, а $I_n$ — единичная матрица размера $n$, и обозначить $\alpha = [\alpha_1, \dots, \alpha_n]'$, модель (3) можно по аналогии с соотношением (1) переписать в следующей матричной форме:

$$y = D\alpha + X\beta + \varepsilon.$$

Далее в разделе получены оценки параметров:

$$\hat{\beta} = (X'M_D X)^{-1} X' M_D y, \quad \hat{\alpha}_i = \overline{y}_i - \overline{x}'_i \hat{\beta}_{FE}, \ i = 1, ..., n$$

$$\hat{\sigma}^2_{\varepsilon} = \frac{1}{nT - n - k} \sum_{i=1}^{n} \sum_{t=1}^{T} (y_{it} - \overline{y}_i - (x_{it} - \overline{x}_i)' \hat{\beta}_{FE})^2.$$

**Модель со случайным эффектом**

Модель со случайным эффектом (random effect model) описывается уравнением

$$y_{it} = \mu + x'_{it}\beta + u_i + \varepsilon_{it}, \tag{4}$$

где $\mu$ - константа, а $u_i$ - случайная ошибка, инвариантная по времени для каждой экономической единицы. Предполагается, что выполнены следующие условия:

1-2. пересекаются с условиями модели с фиксированным эффектом;

3. ошибки $u_i$ некоррелированы, $E(u_i) = 0$, $V(u_i) = \sigma^2_u$;

4. ошибки $u_i$ некоррелированы с регрессорами $x_{ij}$ при всех $i, j, t$;

5. ошибки $u_i$ и $\varepsilon_{jt}$ некоррелированы при всех $i, j, t$.

Модель со случайным эффектом (4) можно рассматривать как линейную модель, в которой ошибка $\omega_{it} = u_i + \varepsilon_{it}$ имеет некоторую специальную структуру. Можно переписать соотношение (4) в виде

$$y_i = \mu\imath_T + X_i\beta + \omega_i$$

или, используя объединенные наблюдения, в матричной форме

$$y = \mu\imath_{nT} + X\beta + \omega.$$

Далее в разделе получены оценки параметров:

$$\begin{bmatrix} \widehat{\mu}_{GLS} \\ \widehat{\beta}_{GLS} \end{bmatrix} = \left( \begin{bmatrix} \imath'_{nT} \\ X' \end{bmatrix} (I_n \otimes \Sigma^{-1})[\imath_{nT} \quad X] \right)^{-1} \begin{bmatrix} \imath'_{nT} \\ X' \end{bmatrix} (I_n \otimes \Sigma^{-1})y.$$

$$\hat{\sigma}_u^2 = \hat{\sigma}_B^2 \frac{1}{T} \hat{\sigma}_\varepsilon^2.$$

В **четвертом** разделе описывается обобщенный метод моментов, который в настоящее время является одним из наиболее распространенных методов оценивания.

Предполагается, что модель включает переменные $y_i$, $x_i$, $z_i$, $\; i = 1, \ldots, n$, и пусть выполнены следующие равенства:

$$E\left(m_j\left(y_i,\ x_i,\ z_i,\ \theta\right)\right) = 0, \quad j = 1, \ldots l, \tag{5}$$

где $m_j\left(y_i,\ x_i,\ z_i,\ \theta\right)$ — некоторые известные скалярные функции, а $\theta$ — k-мерный вектор параметров. (В применении к моделям регрессии можно считать $y_i$ зависимой переменной, $x_i$ — набором регрессоров, $z_i$ — инструментальными переменными, т.е. дополнительные, не участвующие в модели переменные, некоррелированные со случайными ошибками.)

Равенства (5) называются моментными тождествами или условиями ортогональности. Определим вектор функцию:

$$g\left(y,\ X,\ Z,\ \theta\right) = \frac{1}{n}\sum_{i=1}^{n} m\left(y_i,\ x_i,\ z_i,\ \theta\right)$$

Необходимо построить оценку параметров $\theta$ таким образом, чтобы, говоря нестрого, вектор $g(\theta)$ был как можно ближе к нулю. Например, найти оценку $\widehat{\theta}$ путем решения задачи

$$g'(\theta)\,g(\theta)=\sum_{j=1}^{l} g_j^2(\theta)\to min. \tag{6}$$

Вместо минимизации суммы квадратов компонент вектора $g(\theta)$ можно было бы рассматривать более общую задачу, а именно,

$$g'(\theta)\,Sg(\theta)\to min, \tag{7}$$

где S — некоторая симметричная положительно определенная матрица (размера $l\times l$). Оценка, полученная решением задачи (7), называется оценкой обобщенного метода моментов или GMM-оценкой (Generalized Method of Moments, GMM): $\widehat{\theta}=\widehat{\theta}_{GMM}$.

Ясно, что разным весовым матрицам S соответствуют разные (состоятельные) оценки $\widehat{\theta}_{GMM}$. Можно показать, что для получения асимптотически оптимальной оценки (т. е. имеющей минимальную асимптотическую матрицу ковариаций, предел которой известен, а сама матрица нет) в качестве S надо взять матрицу, обратную матрице ковариаций вектора моментов, которая (при отсутствии корреляции между наблюдениями) выглядит следующим образом:

$$S^{opt}=\left(E(m\,(y_i,\ x_i,\ z_i,\ \theta)\,m'\,(y_i,\ x_i,\ z_i,\ \theta))\right)^{-1}. \tag{8}$$

На первом этапе находится оценка $\widehat{\theta}_{(0)}$ путем решения задачи (6) (т. е. с единичной весовой матрицей). Затем строится состоятельная оценка матрицы $S^{opt}$:

$$S_n^{opt}=\left(\frac{1}{n}\sum_{i=1}^{n} m\left(y_i,\ x_i,\ z_i,\ \widehat{\theta}_{(0)}\right)m'\left(y_i,\ x_i,\ z_i,\ \widehat{\theta}_{(0)}\right)\right)^{-1}$$

Наконец, решается задача (7) с S = $S_n^{opt}$ и в результате получается оценка $\widehat{\theta}_{GMM}$. Два последних шага можно повторить.

Можно показать, что построенная таким образом оценка $\widehat{\theta}_{GMM}$ является асимптотически нормальной:

$$\sqrt{n}\left(\widehat{\theta}_{GMM}-\widehat{\theta}\right)\xrightarrow{d}N\left(0,V\right).$$

$$\text{где } V = \left(DS^{opt}D'\right)^{-1}, \text{ и } D = E\left(\frac{\partial m\left(y_i,\ x_i,\ z_i,\ \theta\right)}{\partial \theta'}\right).$$

Этот метод применим к панельным данным.

**Пятый** раздел посвящен описанию эмпирической части.

В ходе работы были собраны данные по расстояниям между административными центрами субъектов Российской Федерации. Было взято восемьдесят регионов РФ (краев, областей и республик). Были исключены Чукотский автономный округ и Камчатский край т.к. для них невозможно рассчитать расстояние по дорогам. Посчитаны расстояния между ними и занесены в таблицу Excel. Расстояния высчитывались несколькими способами:

1. с помощью сайта https://www.avtodispetcher.ru/, на котором можно рассчитать расстояние не только по дорогам, но и по прямой;
2. посредством формулы для расстояния по большой дуге (Great-circle distance).

$$d = r\Delta\sigma, \tag{9}$$

где $r \approx 6371$ – средний радиус Земли (в км), а $\Delta\sigma$ – угловая разница, которая высчитывается по модифицированной формуле гаверсинусов:

$$\Delta\sigma = \arctan\frac{\sqrt{(\cos\phi_2\sin(\Delta\lambda))^2 + (\cos\phi_1\sin\phi_2 - \sin\phi_1\cos\phi_2\cos(\Delta\lambda))^2}}{\sin\phi_1\sin\phi_2 + \cos\phi_1\cos\phi_2\cos(\Delta\lambda)} \tag{10}$$

где $\lambda_1, \lambda_2$ – долгота и $\phi_1, \phi_2$ – широта двух точек в радианах, $\Delta\lambda$ – разность координат по долготе.

Формула (9) определяет расстояние между двумя точками на Земле используя их географические координаты, а именно долготы и широты. Кратчайшим расстоянием между ними является длина дуги круга, проведенного на сфере по этим двум точкам. Поскольку в расчете участвует радиус, а у Земли, как у не совсем правильной сферы, он разный, на северном полюсе (6356.752 км), а на экваторе (6378.137 км), то в расчетах берется среднее значение (6371.008 км), что дает погрешность около 0.5%.

Для полученных расстояний была построена обратная матрица, а также вычислены коэффициенты корреляции для среднего размера выборки (30 городов), что бы узнать зависимость изменения расстояний:

1. по формуле – прямое = 0,988426115
2. прямое – по дорогам = 0,979964874

3. по формуле – по дорогам = 0,951021915.

В результате видно, что коэффициент корреляции близок к единице, что свидетельствует о том, что расстояния коррелированы.

Эти данные будут использоваться для построения весовых матриц в задачах пространственной эконометрики. Такие матрицы являются необходимой составляющей в пространственных регрессионных моделях. Они позволяют описывать взаимосвязь между единицами наблюдения, в нашем случае между российскими регионами.

Создан код, позволяющий оценить параметры пространственной регрессионной модели для панельных данных с помощью обобщенного метода моментов, в бакалаврской работе разработана программа, позволяющая оценивать параметры модели фиксированных эффектов для искусственно сгенерированных данных.

В **заключении** приведены результаты бакалаврской работы.

## Основные результаты

1. Рассмотрена модель парной регрессии.

2. Определены основные гипотезы, лежащие в основе модели множественной регрессии, описана методика построения оценок ее параметров.

3. Рассмотрены основные регрессионные модели для панельных данных, описана методика оценивания.

4. Изучен один из наиболее распространенных методов оценивания - обобщенный метод моментов, заключающийся в таком выборе параметров, при котором некоторые соотношения для заданных наблюдений выполняются с минимальной возможной погрешностью.

5. По российским регионам был проведен предварительный анализ потенциальных весовых матриц. Было рассмотрено три вида весовых матриц: расстояние напрямую, по большой дуге и по дорогам. Корреляционный анализ установил значительную линейную зависимость между этими расстояниями. Это позволяет говорить о том, что можно пользоваться любой из этих матриц хотя бы в случае недостатка информации.

Создана программа, позволяющая оценивать параметры модели фиксированных эффектов для искусственно сгенерированных данных.