

ПРАВИТЕЛЬСТВО РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ
«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ
«ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»

Факультет «Санкт-Петербургская школа экономики и менеджмента»

Департамент экономики

Жданова Мария Олеговна

ГЕДОНИСТИЧЕСКОЕ ЦЕНООБРАЗОВАНИЕ НА РЫНКЕ ЖИЛЬЯ

Выпускная квалификационная работа – БАКАЛАВРСКАЯ РАБОТА

по направлению подготовки 38.03.01 «Экономика»

студентки группы № 146

(образовательная программа «Экономика»)

Рецензент
К.э.н., доцент
Боченина М.В.

Научный руководитель
К.э.н., доцент
Иванов М.А.

Санкт-Петербург

2015

Оглавление

ВВЕДЕНИЕ.....	4
ГЛАВА I. ТЕОРЕТИЧЕСКИЕ АСПЕКТЫ ГЕДОНИСТИЧЕСКОГО ЦЕНООБРАЗОВАНИЯ.....	6
1. 1. Особенности рынка жилья и актуальность проведения качественного анализа динамики цен на жилье.	6
1. 2. Гедонистические функции цен	8
1. 3. Оценка спроса.....	10
1. 4. Анализ благосостояния.....	13
ГЛАВА II. МЕТОДОЛОГИЯ ПОСТРОЕНИЯ ГЕДОНИСТИЧЕСКИХ УРАВНЕНИЙ И ИНДЕКСОВ	16
2. 1. Метод фиктивных переменных (time-dummy methods).....	16
2. 2. Метод подстановок (Imputation method)	18
2. 3. Построение индекса в пространстве характеристик (characteristics methods).....	20
2. 4. Использование непараметрических методов (nonparametric methods)	21
2. 5. Повторные продажи и гибридная гедонистическая модель (Repeat sales and hedonic hybrid methods).....	21
2. 6. Существующие исследования в России и выявленные проблемы при использовании гедонистического подхода	25
ГЛАВА III. ГЕДОНИСТИЧЕСКИЙ АНАЛИЗ ЦЕН НА РЫНКЕ ПЕРВИЧНОГО ЖИЛЬЯ В САНКТ-ПЕТЕРБУРГЕ	28
3. 1. Подготовка данных	28
3. 2. Построение простейших гедонистических моделей и тестирование базовых гипотез о влиянии отдельных характеристик квартиры на ее цену	29
3. 2. 1. Влияние классов квартиры на цену квадратного метра	29
3. 2. 2. Влияние типов и классов квартиры на цену квадратного метра	33
3. 3. 3. Влияние района на цену квадратного метра.....	41
3. 3. Построение полной гедонистической модели цены на рынке первичного жилья в Санкт-Петербурге	49
ЗАКЛЮЧЕНИЕ	58
СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ	59
ПРИЛОЖЕНИЕ	62
Приложение 1. Проверка нормальности распределения цены квадратного метра.....	62
Приложение 2. Проверка нормальности распределения логарифма цены квадратного метра	64
Приложение 3. Проверка гипотез по регрессионной модели №1	66
Приложение 4. Проверка гипотез по регрессионной модели №2	68
Приложение 5. Проверка гипотез по регрессионной модели №4	70

Приложение 6. Матрица расстояний между средними ценами квадратного метра квартиры в зависимости от её типа и класса	73
Приложение 7. Сравнение моделей №5 и №6 с объединенными районами и без объединенных районов.....	74
Приложение 8. Проверка гипотезы о нормальности распределения остатков в модели №9	75
Приложение 9. Проверка гипотезы о нормальности распределения остатков в модели №11	76
Приложение 10. Shapley decomposition для таких характеристик районов как класс, тип квартиры и район, в котором она располагается.....	77
Приложение 11. Проверка нормальности распределения остатков в модели №12	79

ВВЕДЕНИЕ

Гедонистические исследования цен на рынке жилья (как первичного, так и вторичного) становятся достаточно популярными в последнее время. В целом, гедонистический подход является достаточно перспективным при исследовании цен и спроса на различные потребительские блага. Одна из первых работ в этой области была посвящена исследованию цен на овощи (F. Waugh, 1928), отцом же современного гедонистического анализа считается Цви Грилихес (Griliches, 1961). Далее последовало развитие данного вида анализа, так, было проведено первое исследование спроса на компьютеры по собранному массиву о ценах и характеристиках компьютеров (Chow, 1967).

В основе гедонистического подхода заложена теория К.Ланкастера о том, что полезность предмета потребления представляет сумму полезностей отдельных качеств, характеристик данного предмета, таким образом, потребитель, уплачивая определенную цену за товар, платит не за сам товар, а за определенный набор характеристик, которыми он обладает.

Использование гедонистического подхода при исследовании рынка жилья преследует несколько различных целей. Во-первых, это изучение цены как функции отдельных характеристик жилья. Жилье является сильно дифференцированным продуктом и обладает набором сразу нескольких характеристик (Sheppard, 1999), соответственно гедонистический подход дает определенные методы, позволяющие оценить вклад каждой характеристики в цену товара. Существует ряд отечественных исследований, объясняющих структуру цен квартир, к примеру, исследование московского рынка жилья (Магнус Я. Р., Пересецкий А. А., 2010). Так же можно оценить, как и в какой степени влияет наличие или отсутствие определенных факторов, а так же их интенсивность, на цену жилья. Например, экологические факторы (Катышев П., Хакимова Ю., 2014).

Во-вторых, гедонистический подход позволяет строить индексы цен на жилье. Так, жилищная недвижимость является одним из крупнейших активов в

развитых странах и составляет приблизительно треть мирового богатства (Syz, 2008). Существует исследование, согласно которому изменение цен на рынке жилья оказывает большее влияние на потребление домохозяйств, чем изменение цен на фондовых рынках (Quigly and Shiller 2005). Данные факты говорят в пользу того, что индексы цен на жилье – это важная информация, на которую стоит ориентироваться при формировании политики государства. В частности, например, существует вопрос – как применить понятие инфляции к рынку жилья (Cecchetti 2007).

И, наконец, третье, не менее важное направление использования гедонистического подхода – это оценка недвижимости. Так, оценщикам стоит сосредотачиваться на статистической обработке уже существующих данных, которая позволяет выявлять значимые переменные при ценообразовании на рынке (Kelley Pace, 2002).

При использовании гедонистического подхода также возникает определенный ряд проблем, в частности проблема пропущенных переменных, неправильный выбор функциональной формы, смещение выборки, однако данный круг вопросов является типичным при проведении любого статистического исследования. Несколько более детального внимания заслуживают вопросы, связанные со сбором и качеством данных. Так, отечественные исследователи используют цены предложения в качестве альтернативы ценам сделок на квартиры в силу отсутствия информации о последних. Правомерность такого подхода рассматривается в статье, сравнивающей цены предложения и цены сделок (Баталёва А.В., 2005). Также стоит отметить важность и качество сбора такой характеристики о квартире, как ее местоположение. Практически все исследования отмечают исключительную важность данного фактора, например (R. Hill, 2011). Существует несколько видов учета местоположения, например использование почтовых идентификаторов, как дамми-переменных (Hill, Melser and Syed 2009), либо присвоение каждой квартире принадлежности к определенной зоне (Глущенко К. П., Баталева А. В., 1999).

ГЛАВА I. ТЕОРЕТИЧЕСКИЕ АСПЕКТЫ ГЕДОНИСТИЧЕСКОГО ЦЕНООБРАЗОВАНИЯ

1. 1. Особенности рынка жилья и актуальность проведения качественного анализа динамики цен на жилье.

Рынок жилья обладает некоторыми особенностями, которые необходимо учитывать при проведении анализа цен.

Так, прежде всего, стоит отметить, что рынок жилья включает поиск. Это означает, что покупатель во время поиска несет определенные издержки, однако, в то же время, он ожидает найти более подходящие варианты, повысив тем самым уровень полезности, данный процесс будет продолжаться до тех пор, пока ожидаемое увеличение полезности превышает ожидаемый рост затрат.

Рынок жилья является пространственным рынком, это означает, что две абсолютно идентичные квартиры с разным местоположением будут демонстрировать различную цену.

Помимо этого одновременно существуют как новое, первичное жилье так и вторичное, совместно сосуществуют вторичный и первичный рынок жилья, предоставляет возможность покупателям делать еще и выбор между этими рынками.

Для того чтобы доказать необходимость изучения динамики цен на жилье, стоит привести несколько фактов:

1. Жилье составляет примерно треть суммарного мирового богатства (Syz 2008);
2. Изменение цен на рынке жилья оказывает большее влияние на потребление домохозяйств чем изменение цен на фондовых рынках (Quigly and Shiller 2005);
3. Изменение цен внутри страны оказывает существенное влияние на зарубежные экономики.

Цены на рынке жилья влияют на бумы и спады. Так, бумы приводят к перераспределению богатства от молодого поколения к старшему и увеличивают неравенство внутри поколений в силу существования механизма аренды.

В общем говоря, индекс цен на рынке жилья (house price index) – это показатель состояния экономики, на который стоит ориентироваться при формировании политики. Так, например, существует вопрос – как применить понятие инфляции к рынку жилья (Cecchetti 2007).

Большинство стран исключают жилье из ИПЦ, либо применяют подход эквивалента арендной платы (rental equivalence). Однако, случай США (Poole, Ptacek and Verbrugge 2005) показал, что данный подход не вполне корректен. Так, рост цен на жилье не оказал значимого влияния на ИПЦ в США, в то время как в действительности цены на жилье росли быстрее, чем стоимость арендной платы.

Существует два альтернативных подхода – оценка через первоначальную стоимость (acquisitions approach) и через затраты использования (другое название – «прокатная оценка» было использовано Л.В. Конторовичем) (user cost approach).

При оценке через первоначальную стоимость предпринимается попытка рассматривать жилье как любой другой товар длительного пользования. Стоимость оценивается через стоимость строительных материалов, при более утонченном подходе используется эконометрическое разделение цены самого сооружения и земли.

При использовании оценки через затраты использования рассматривается сумма следующих значений – дохода от прироста капитала, процентов, получаемых от использования капитала и обесценивания (с различными знаками).

Иными словами, альтернативные затраты, связанные с использованием принадлежащего капитала, представляют сумму процентного дохода на рыночную стоимость капитала в начале определенного периода и снижения

его рыночной стоимости в течение данного периода. Эту сумму обычно называют затратами использования. Л.В.Канторович использовал в этой связи термин «прокатная оценка». «Мы употребляем термин „прокатная оценка“, так как это есть оценка той платы, которая была бы оправдана, если бы такая машина бралась на некоторый срок напрокат» (Канторович Л.В. Экономический расчет наилучшего использования ресурсов. М., 1960. С. 102).

Чем обуславливается важность индекса цен на рынке жилья?

Развитие рынка деривативов (производных финансовых инструментов) совместно с формированием качественных индексов цен на рынке жилья помогает в управлении рисками. Хозяйства смогут диверсифицировать риски с учетом это дополнительной категории (жилье). В пользу данного рассуждения может говорить тот факт, что на данный момент жилье составляет куда большую долю в инвестиционном портфеле, нежели пенсионные вклады и средства в инвестиционных фондах открытого типа.

Существует такая проблема: жилье является крупнейшим активом в стране (Ian Macfarlane, Governor of Reserve Bank of Australia, 4 June 2004). При этом информации о ценах этого актива значительно меньше, по сравнению с ценами на другие активы (цены на акции, курсы валют и т.д.).

В общем смысле гедонистический метод выражает цену как функцию качественных характеристик и на данный момент существует несколько видов индексов.

1. 2. Гедонистические функции цен

С одной стороны мы можем рассматривать спрос на товары – даже если они является гомогенными не как спрос на товары, как таковые, а спрос на характеристики, которыми этот товар обладает. Полезность домохозяйства, потребляющего благо, зависит от характеристик, которыми оно обладает (Lancaster, 1966).

С другой стороны, рынки гетерогенных товаров (таких как автомобили, жилье, работники) не могут быть описаны стандартными экономическими моделями, и охарактеризованы единственной ценой, они могут быть охарактеризованы неким ценовым диапазоном, который зависит от качества включаемых в товар характеристик. Гедонистический подход предполагает, что гетерогенное благо «раскладывается» на несколько гомогенных частей, каждая из которых вносит свой вклад в общую структуру цен.

Причины, по которым мы оцениваем гедонистические функции цен: 1) построение индексов для оценки изменения качества товаров, 2) для анализа спроса на различные характеристики гетерогенного товара.

В статье (Seppard, 1999) «Hedonic analysis of housing market» рассматривается гедонистический анализ на рынке жилья, который позволяет оценить спрос на жилищные характеристики и различные другие атрибуты. Рассматриваются теоретические основы, непосредственное применение гедонистического метода, трудности, с которыми мы можем сталкиваться и методы борьбы с ними.

Функция полезности домохозяйства представлена функцией $u = u(Z, Y, \alpha)$, где Z – вектор характеристик товара, Y – потребление блага-комплемента, α – личные предпочтения домохозяйства в отношении этого блага.

Функция «bid rent» задается следующим образом: $\beta(Z, M, u, \alpha)$, M – располагаемый доход.

Домохозяйство максимизирует свою функцию полезности, исходя из ограничений:

$$\max_{z, Y} u(Z, Y, \alpha) \text{ при ограничении: } M > P(Z) + Y$$

Условие оптимального выбора:

$$\frac{\partial \beta}{\partial z_i} = \frac{u_i}{u_Y} = P_i$$

Экономический смысл:

$\frac{\partial \beta}{\partial z_i}$ – как домохозяйство готово изменить свои траты в ответ на увеличение i -ой характеристики, оставаясь на прежнем уровне полезности;

$\frac{u_i}{U_Y} = P_i$ – условие первого порядка для обозначенной проблемы потребителя.

Прибыль производителя:

$\pi = P(Z) N - C(Z, N, Y)$, где N – количество построенных домов, C – функция издержек, Y – вектор, характеризующий производителя. Плотностью вероятности $g(Y)$ задается распределение продавцов.

На этом рынке формируется равновесная цена $P(Z)$.

Каким будет равновесие, и от чего оно зависит, рассматривалось в работах Epple (1987), Tinbergen (1959), Sattinger (1980).

1. 3. Оценка спроса

Чтобы проводить гедонистический анализ, необходимо понять структуру спроса. Мы рассмотрим, как функция спроса может быть рассмотрена в гедонистических рамках, и как будет изменяться благосостояние в случае изменения цен или атрибутов согласно (Seppard, 1999).

Так как гедонистическая цена атрибута не является постоянной, возникает неоднозначность в оценке спроса. Если мы знаем предпочтения домохозяйства, мы можем представить функцию спроса, которая будет отражать эти предпочтения при заданных ценах. В качестве альтернативы можно представить домохозяйство, которое максимизирует свою полезность исходя из некой поведенческой функции. Проще говоря, первый подход предполагает функцию спроса как функцию цен и дохода, а второй – функцию параметров функции гедонистической цены и дохода.

Если следовать по пути, предложенному Розеном (Rosen, 1974), и включающем часть других исследований, можно считать оцененные по

гедонистическому уравнению цены фактическими и включить в анализ доход и наблюдаемый выбор домохозяйства.

Спрос на каждый отдельный атрибут может быть представлен в виде:

$$q_i(P_1, \dots, P_j, M, \alpha)$$

Тогда доля, затраченная домохозяйством на покупку данного атрибута в общей структуре расходов:

$$w_i = \frac{q_i(P_1, \dots, P_j, M, \alpha)}{M} + \varepsilon$$

Существует две проблемы, с которыми мы можем столкнуться уже на этом этапе. Первая – отсутствие независимости между переменными P_i и ε . Вторая – возможность того, что домохозяйство сталкивается с нелинейным бюджетным ограничением. Чтобы понять важность этого, обратимся к рисунку:

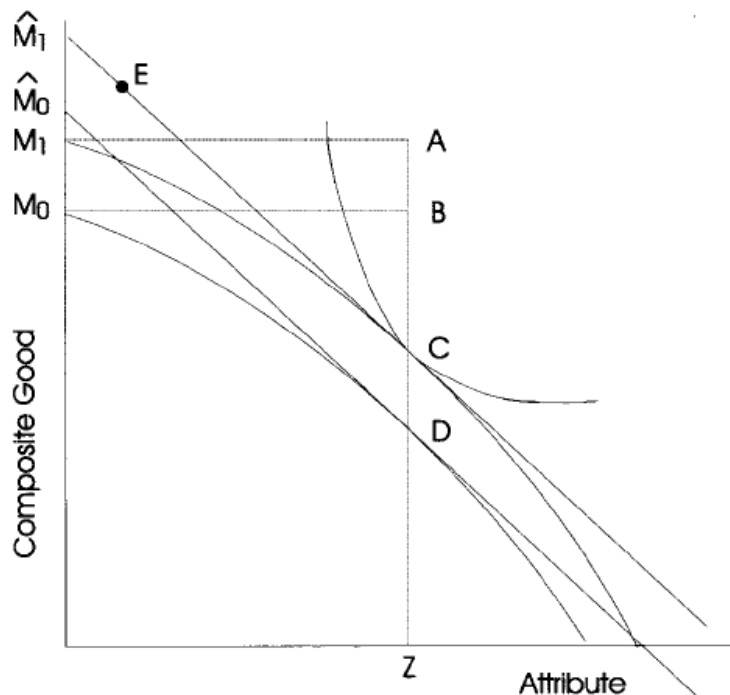


Fig. 1. Hedonic budget lines and consumer choice.

Здесь представлены две бюджетные линии, уравнения которых $Y = M - P(Z)$.

Гедонистическая функция цены независима от доходов домохозяйств (домохозяйства принимают цену) и бюджетные линии является квазилинейной. Выбор домохозяйства с бюджетом M_1 и с бюджетом M_1 с

шапкой совпадают в точке C . При условии, что предпочтения строго выпуклы, спрос будет функцией и линейный спрос будет таким же, как и гедонистический.

Таким образом, если мы хотим оценить структуру предпочтений домохозяйства через спрос на какие-либо характеристики, атрибуты, мы можем использовать следующие данные – гедонистические цены, конкретный выбор домохозяйства, а в качестве дохода мы должны взять $M - P(Z) - \sum_{i=1}^J P_i Z_i$.

Авторы по-разному относятся к оценке спроса и того, как правильно этот спрос называть. Некоторые предпочитают избегать термина «спрос» (McConnell and Phipps, 1987), так как гедонистический (настоящий) спрос не зависит от цен (а зависит от характеристик гедонистической функции). В то же время, линейный спрос, который зависит от цен, на самом деле является надуманным (Murray, 1983), так как покупатель сталкивается с линеаризованными ценами (справедливости ради стоит отметить, что ничего надуманного относительно предпочтений, которые получены из оцененного таким образом спроса, нет).

Существует также важный вопрос относительно линеаризации бюджета: существует мнение, что это приводит к недооценке эластичности замещения между жилищными характеристиками и нежилищными благами (Ohsfeldt and Smith, 1990).

Работа (Day et al., 2004) посвящена оценке уравнений гедонистических функций цен в том случае, когда гедонистическая функция цены имеет нелинейный вид. Недавние теоретические исследования показали, что такой вид функции характерен для функций равновесных гедонистических цен.

Авторы, помимо нелинейной гедонистической функции, используют метод кластеризации; согласно их мнению, гедонистическая функция цен не только не может быть описана линейно, но и более того, оценка на всем пространстве существования различных уровней жилищных атрибутов

приводит к переоценке для одних жилищ, и недооценке для других (кластеризация позволяет выделить области, в рамках которых оценка функций не будет приводить к столь серьезному смещению).

В статье (Quigley, 1982) исследуется, как информации относительно особенностей рынка может быть использована для того, чтобы оценить компенсированный спрос покупателей на жилищный атрибуты.

В работе было показано, что можно наложить несколько ограничений без детального знания распределения дохода и жилищных характеристик. Так для предпочтений потребителя, выраженных функцией Кобба-Дугласа даже необязательно подразумевать вогнутость гедонистической функции. Тем не менее, если гедонистические функции заданы экзогенно, оценить спрос на отдельные характеристики предоставляется возможным. Если предпочтения являются гомотетичными, то можно оценить вид кривой функции полезности и компенсированный спрос на атрибуты в числовом выражении. В противном случае, весьма общих (и негомотетичных) предпочтений могут быть оценены статистически путем наблюдения за выбором потребителя при нелинейном ограничении.

1. 4. Анализ благосостояния

Рассмотрим два вида анализа, согласно (Seppard, 1999), которые могут представлять интерес для применения на рынке жилья: 1) оценка последствий жилищного регулирования на уровень благосостояния, 2) ожидаемые последствия увеличения дохода на жилищный спрос. Полагаем, что гедонистическая функция цен выпукла, и множество доступных домохозяйству наборов так же выпукло.

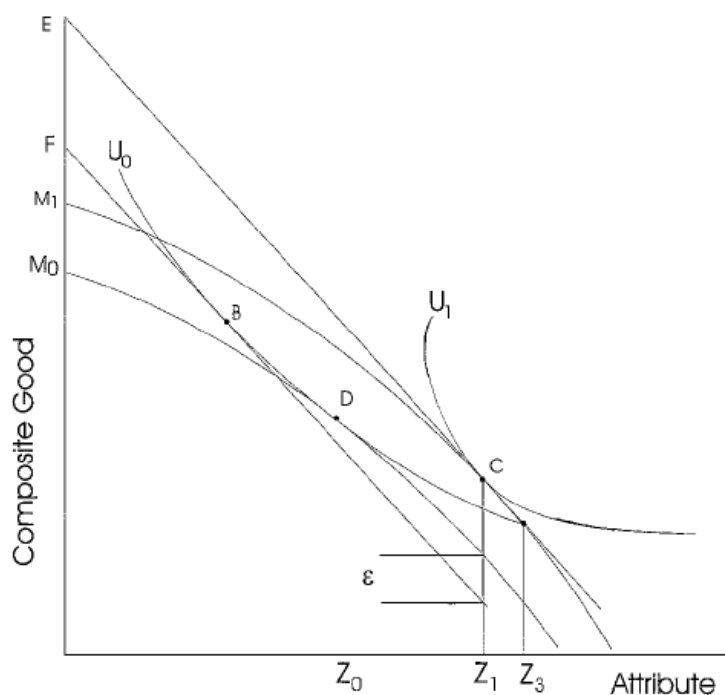


Fig. 2. Estimated attribute demand can overstate welfare costs.

Допустим, домохозяйство достигает максимального уровня полезности U_1 , покупая атрибут в размере Z_1 при доходе M_1 . Если существует некоторое ограничение, которое требует, чтобы было приобретено количество характеристики Z_3 , то домохозяйство опустится на уровень полезности U_0 , и потери общественного благосостояния составят $M_1 - M_0$. Если мы воспользуемся линеаризованной функцией бюджета, то получится, что потери составят $E - F$, что больше, чем истинные потери общества!

Единственным случаем, когда будет наблюдаться нулевой эффект дохода для домохозяйства будет та маловероятная ситуация, когда столкнувшееся с уменьшением дохода домохозяйство не будет изменять уровень потребления атрибута, а общее изменение расходов будет равно изменению дохода.

Нелинейность дохода также сказывается на прогнозировании изменения спроса на атрибут в ответ на изменение дохода домохозяйства. Возвращаясь к рисунку – увеличение спроса на характеристику с Z_0 до Z_1 может быть создано за счет увеличения доходов в размере $E - F$. На самом деле это может быть достигнуто за счет меньшего в абсолютном выражении

изменения $M_1 - M_0$. Очевидно, что точность аппроксимации зависит от структуры предпочтений и кривизны функции гедонистической цены. Аналогичные проблемы возникают, когда мы пытаемся оценить влияние изменения гедонистической функции цены. Предположим изменение в сторону увеличения, когда при каждом уровне потребления атрибута гедонистическая цена увеличивается, а бюджетное ограничение становится круче.

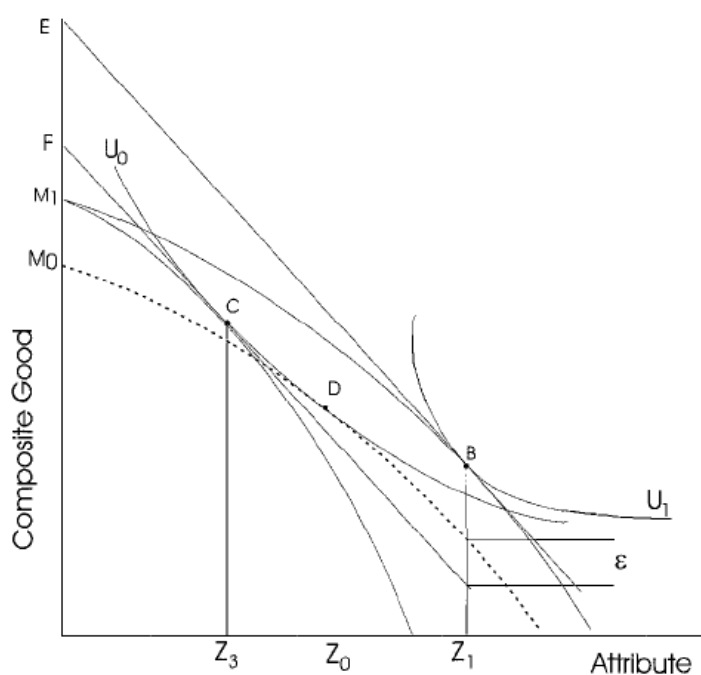


Fig. 3. Measuring welfare effects of a price change.

Домохозяйство изменяет уровень потребления с В на С. Данный рисунок показывает, что если гедонистические цены на характеристики останутся постоянными, в то время как доход сократится на величину $M_1 - M_0$, это будет производить такой же эффект на изменение благосостояния как и увеличение цен на атрибуты.

ГЛАВА II. МЕТОДОЛОГИЯ ПОСТРОЕНИЯ ГЕДОНИСТИЧЕСКИХ УРАВНЕНИЙ И ИНДЕКСОВ

2. 1. Метод фиктивных переменных (time-dummy methods)

Описание метода

Обычно используется логлинейная функциональная форма (semi-log, log-linear)

$$y = Z\beta + D\delta + \varepsilon$$

Обозначение	Размерность	Содержание
y	$N \times 1$	$Y_h = \ln p_h$
Z	$N \times C$	Матрица характеристик (некоторые могут быть дамми-переменными)
β	$C \times 1$	Вектор, определяющий влияние характеристик на цену
D	$N \times T-1$	Матрица дамми-переменных, обозначающих период
δ	$T-1 \times 1$	Вектор, определяющий влияние периода на цену

Удобство такой функциональной формы заключается в том, что индекс цен в периоде t можно найти как $e^{\delta t}$

Для устранения смещенности стоит использовать формулу $e^{\delta t + \frac{1}{2}\sigma_t^2}$, где σ_t^2 – оценка дисперсии δ_t .

Проблема: такая коррекция приводит к систематическому увеличению индекса. Индекс смещается в зависимости от выбора базового периода.

Возможное решение с точки зрения автора: составить T наборов, беря попеременно в качестве базового периода все периоды по очереди и затем в качестве оценки взять среднее геометрическое для всех наборов.

Однако: для данных Sydney, Australia за 2001-2006 разница в индексах, посчитанных обоими способами, присутствовала лишь после 3-его либо 4-ого знака после запятой, что позволяет игнорировать факт, обозначенный в постановке проблемы. Но будет ли она [разница] такой же для других данных – это вопрос.

Один из ключевых факторов в формировании цены жилья – его местонахождение. Включение этого фактора приводит к увеличению R^2 , а гедонистическая модель приобретает следующий вид:

$$y = Z\beta + V\gamma + D\delta + \varepsilon,$$

где V имеет размерность $N \times (M-1)$, $\gamma - (M-1) \times 1$, а M обозначает число идентификаторов местонахождения (почтовых индексов и т.п.)

Сильные и слабые стороны метода

+

Простота. Индексы цен получаются прямо из гедонистического уравнения.

Уравнение дает стандартные ошибки индексов

-

Накладывает ограничения на потенциальное взаимодействие между периодами и характеристиками. Однако можно сделать модель более гибкой, позволив такое взаимодействие. Этот подход требует оценивать больше параметров и делает индекс более сложным.

Метод дамми-переменных использует индекс Ласпейреса, соответственно проблемы:

1) Может возникнуть смещение из-за:

а) того, что характеристики могут по-разному влиять на цену в различных периодах (стабильны ли коэффициенты во времени?)

б) того, что сами характеристики (их качество) могут меняться во времени

в) того, что индексы смещаются по отношению друг к другу (а не по отношению к базовому периоду)

2) С добавлением еще одного периода изменяются все индексы цен

Решение обеих проблем: использовать только смежные периоды (и добавлять/убирать только смежные с уже использованными).

2. 2. Метод подстановок (Imputation method)

Описание метода

Используются индексы Ласпейреса, Пааше, геом. индекс Ласпейреса, геом. Индекс Пааше, Фишера и Торнквиста.

Для пропущенных переменных можно использовать подстановки. Подстановки можно так же использовать и не для пропущенных переменных, тогда

1) Решается проблема смещения коэффициентов

2) Подобное сравнивается с подобным, иначе используются то фактические, то предсказанные значения.

Для данного метода не требуется определенная функциональная форма.

Как и в методе, использующем дамми-переменные, требуется делать корректировку для нахождения предсказанных значений.

Как можно использовать подстановки? Допустим, есть два периода, s и t .

1. Подставлять (чит.: предсказывать) все цены для периода t . Single imputation method

2. Подставлять цены только на то жилье в период t , которое не было продано и в s и в t (здесь используется минимум подстановок)

3. Подставлять все цены для обоих периодов. Double imputation method

4. Подставлять все цены, за исключением цен на то жилье, которое было продано в обоих периодах.

С точки зрения автора предпочтительным является использование метода 3.

Различие между этими двумя методами (звездочкой обозначены предсказанные по регрессии значения):

Single imputation method $p^*_{th}(z_{sh})/p_{sh}$

Double imputation method $p^*_{th}(z_{sh})/p^*_{sh}(z_{sh})$

На взгляд автора предпочтительным является использование формулы Торнквиста. Удобство заключается в том, что общее изменение индекса представляет собой произведение изменений за счет отдельных характеристик.

При использовании данного метода нельзя разделять выбор функциональной формы и формулы для построения индекса:

Логлинейная функциональная форма – формула Торнквиста

Линейная функциональная форма – формула Фишера

Сильные и слабые стороны метода

+

Гибкость. Отображает изменение влияния характеристик на цену во времени.

Использование double imputation method уменьшает проблему смещения, порожденную наличием пропущенных переменных.

Хорошо совмещается с более старыми индексами.

Совместим с непараметрическими оценками гедонистической модели.

-

Необходимо оценивать отдельно уравнения для каждого периода предварительно, до того как исследовать взаимодействие между уравнениями.

Нет стандартных ошибок индексов.

2. 3. Построение индекса в пространстве характеристик (characteristics methods)

Описание метода

Аналогично, модель оценивается для всех отдельных периодов. Также используются стандартные формулы для расчета индексов.

Ключевое отличие состоит в том, что определяется в пространстве характеристик.

Для подсчета индекса определяется цена гипотетического «среднего» жилья (под средним может подразумеваться как среднее, так и медианное значения, помимо этого среднее жилье определяется в пространстве характеристик, то есть обладает средними значениями характеристик). Цена определяется как функция характеристик.

Индекс получается делением цен таких квартир (жилищ) из двух разных периодов.

Для построения используется double imputation method.

Другой подход к построению индекса состоит в том, чтобы дефлировать количественный индекс, используя поправку на качество.

Сильные и слабые стороны метода

Плюсы и минусы в основном те же, что и для double imputation method. Однако, по мнению автора, imputation method является более надежным, так как используется подстановка цен для фактически существующих, а не гипотетических жилищ.

2. 4. Использование непараметрических методов (nonparametric methods)

Использование непараметрических методов становится популярным в последнее время, так как при применении метода отсутствует проблема ошибки спецификации (неправильного выбора функциональной формы).

Однако гибкость модели обходится дорого в том смысле, что в отличие от параметрических моделей, непараметрические не всегда обеспечивают существование теневых цен характеристик. Однако можно получить информацию, аналогичную интересующей нас, с помощью других средств в непараметрической модели.

Использованию непараметрических методов уделено относительно немного внимания, однако, по мнению автора, хорошим использованием данных методов будет совместное их применение с *double imputation method*. Все, что требуется – это чтобы существовала возможность предсказать цену для любого набора характеристик, и затем, чтобы эти предсказанные цены также завесили от времени.

Также непараметрическая гедонистическая модель может быть использована совместно с *characteristics method*.

2. 5. Повторные продажи и гибридная гедонистическая модель (Repeat sales and hedonic hybrid methods)

При использовании метода повторных продаж для построения индекса, выборка формируется только из того жилья, которое между первой и повторными продажами не претерпело изменений своих качественных характеристик.

Оригинальная методика использования метода повторных продаж была предложена (Bailey, Muth and Nourse, 1963). Предполагая, что жилью не подвергается изменениям, можно узнать как изменялась цена, посмотрев на разницу в цене продаж одного и того же дома в разные моменты времени.

Можно сказать, что таким образом появляется возможность решить проблему об изменчивости «состава блага», которую мы имеем при использовании среднего или медианного индекса (где в построении участвуют все проданные дома, которые при этом обладают различными характеристиками, т.е. «состав» блага, для которого строится индекс изменяется). Основная идея заключается в том, что значение индекса оценивается с помощью регрессии цены продажи на период, в который эта продажа была совершена. Таким образом, значения индексов могут быть получены прямо из коэффициентов уравнения.

Модификация оригинальной методики была разработана (Case, Shiller, 1987, 1989)¹. Авторы не были согласны с предположением о постоянстве дисперсии ошибок, утверждая, напротив, присутствие гетероскедастичности. Исследователи предложили использовать взвешенный метод повторных продаж (Weighted Repeat-Sales Method). Авторы указывали на то, что, так как повторные продажи осуществляются через разные интервалы времени, при использовании метода (Bailey, Muth and Nourse, 1963) цены домов, проданных через больший временной промежуток будут оказывать большее влияние на индекс, чем жилья, проданного через более короткий промежуток. Соответственно стоит задать таким наблюдениям меньший вес при построении индекса.

Решением проблемы гетероскедастичности, обозначенной (Case and Shiller, 1987) является построение регрессии с помощью трехступенчатой процедуры. На первом шаге строится регрессия разности логарифмов цен в первом и последующем периоде на набор дамми-переменных, обозначающих

¹ *Индекс Кейса-Шиллера (Case-Shiller home-price index, S&P Case-Shiller Composite-20 Home Price Index). «Взвешенный по стоимости арифметический индекс повторных продаж, позволяющий проследить за продажными ценами одних и тех же домов, и избежать изменений средних цен, которые были бы результатом общей тенденции к увеличению размеров и повышению качества домов с течением времени. Цель заключается в том, чтобы получить оценки цены стандартного, неменяющегося дома, с тем чтобы индекс цен отражал бы результаты инвестиций в дом, аналогично результатам инвестиций в акции, отслеживаемые с помощью индекса фондового рынка, который основан на тех же самых, неменяющихся акциях, которые наблюдаются при покупке и продаже в различные моменты времени» (Мауро П. От провидца к новатору // Финансы и развитие, 2008, №4, с. 4-7.).

время продажи. На втором этапе происходит непосредственно определение весов – для этого строится регрессия квадратов ошибок (полученных на первом шаге) на константу и разницу во времени между повторными продажами. Наконец, строится взвешенная регрессия (методом наименьших квадратов), повторяющая регрессию на первом шаге с учетом весов для наблюдений, полученных на втором шаге.

Два основных вопроса, обозначенных в работе (Meese, Wallace, 1997) связанные с данным методом касаются выборки, принимающая участие в построении модели, (является ли она репрезентативной), и оценок стоимости характеристик (являются ли они устойчивыми во времени)?

Проблема нерепрезентативности выборки связана с тем, что одни дома продаются чаще, чем другие, поэтому они будут представлены в избытке в выборке. А когда различные виды жилья демонстрируют различные изменения в ценах, это неизбежно приводит к смещению. К примеру, если дешевые дома продаются чаще, чем дорогие, при этом дорогие дома поднялись в цене не так сильно, как дешевые, это приведет к завышенной оценке индекса (Naan, Diewert, 2011).

Для решения второй проблемы исследователями были предложены различные варианты. Так, в частности, (Case, Shiller, 1987) указали, что можно использовать лишь данные о тех домах «которые были проданы дважды и для которых не наблюдается видимого изменения качества». К более поздним работам относятся попытки создания гибридной модели, совмещающей характеристики метода повторных продажи и гедонистического метода, к примеру, работа (Case and Quigley, 1991).

Более того, самой главной проблемой, обсуждаемой всеми авторами, при использовании метода повторных продаж является сложность сбора необходимых данных, так как информация по проданным повторно домам является лишь малой частью всей существующей информации о продажах жилья. Здесь же стоит упомянуть о такой проблеме, как отказ от большого

количества данных – информации о том жилье, которое не участвовало в повторных продажах.

Обобщая, можно выделить как преимущества, так и недостатки метода повторных продаж.

Преимущества:

- Не требуется каких-либо специфических характеристик о жилье кроме адреса
- Простота метода построения регрессии и получения индексов цен
- В модели не используются подстановки

Недостатки:

- Метод неэффективен в том плане, что используются не все данные
- Стандартный метод повторных продаж игнорирует амортизацию жилого фонда
- Возможно смещение вследствие нерепрезентативности выборки
- Индекс не позволяет разделить индекс цен на земельные участки и здание в отдельности
- Сложно построить данный вид индекса для отдельных видов собственности, информации, о повторных продажах которых, чрезвычайно мало
- Согласно идее, переоценка индекса должна производиться каждый раз по мере того, как появилась новая информация о повторной продаже.

Предпринимались различные попытки совместить метод повторных продаж и гедонистический подход.

Quigley (1991) и Quigley (1995) использовал выборки единожды проданных и проданных несколько раз жилищ для оценки индекса цен используя метод обобщенного МНК.

Hill, Knight и Sirmans (1997) использовали метод максимального правдоподобия. Их версия модели представляла собой гибридную модель, состоящую из двух совместно оцениваемых уравнений. Первое уравнение содержало гедонистическую модель, второе – модель повторных продаж.

Сторонники гибридной модели утверждают, что этот подход призван взять лучшее от обеих моделей. Так, они утверждают, что использование цен полученных по данным о повторных продажах позволяет лучше отражать изменения качества, чем *double imputation method*. Но метод повторных продаж отбрасывает все данные по жилью, которое было продано единожды. Совмещая же эти два метода, никакие данные не будут отброшены, при этом ключевую роль в построении индекса будет определять все-таки модель повторных продаж.

По мнению автора, нет необходимости использовать гибридную модель вместо *double imputation method* при прочих равных.

2. 6. Существующие исследования в России и выявленные проблемы при использовании гедонистического подхода

В статье Баталевой А.В., Глушенко К.П. (1999) представлена гедонистическая модель цен на жилье на вторичном рынке жилья Новосибирска.

В качестве базовой модели была выбрана мультипликативная форма, что позволяет придать модели прозрачный экономический смысл, а именно: если значения всех объясняющих переменных будут равны 1, то цена будет представлять собой цену некой эталонной квартиры, (для которой значения всех характеристик равны 1), скорректированную с помощью поправочных коэффициентов. Помимо этого модель проста для учета инфляции. Ее не нужно пересчитывать заново, а нужно лишь рассчитать средние значения объясняемой и объясняющих переменных для нового периода, получив обновленную величину вектора цен эталонной квартиры.

В статье было получено несколько выводов. Так, гипотеза об устойчивости структуры во времени должна быть отвергнута. Для ответа на вопрос об устойчивости коэффициентов во времени необходимо изучать динамику цен на более продолжительном интервале времени.

Помимо этого при сравнении модели цен Новосибирске с моделью цен в Москве, поведение цен в отношении изменений характеристик совпадает, если учесть необходимые модификации для адекватного сравнения.

В статье Баталёвой А.В. «Структура цен на квартиры: цены и предложения и цены сделок» рассматривается вопрос о справедливости замены цен сделок ценами предложений при статистическом изучении структуры цен на рынке жилья. В качестве предмета исследования был взят рынок вторичного жилья Новосибирска, объекта исследования – цены на обозначенном рынке.

Как показал статистический анализ, проведенный автором, данная замена является вполне приемлемым приемом при работе с данными. Значимость большинства наиболее важных коэффициентов была подтверждена при тестировании обеих моделей. Однако если отвечать формально на вопрос, поставленный в статье, то ответ будет отрицательным, так тест Чоу на совпадение всей совокупности коэффициентов был отвергнут. Однако, стоит отметить некоторые особенности данных, которые, возможно так же повлияли на результат. Так, выборки по ценам предложения и цена сделок были различны. Возможно, оценки коэффициентов были более близки, если бы имелась информация по ценам на одни и те же квартиры. Помимо этого для этих двух выборок использовалась различная методология при отнесении квартир к какой-либо зоне города. Обе эти проблемы были озвучены автором.

Наконец, стоит отметить один очень важный момент, также обозначенный в статье: из двух построенных моделей лучше объясняет структуру цен квартир та, которая содержит данные о ценах сделок. Как отметил автор статьи, возможно, продавец не вполне корректно оценивает

квартиру, используя какие-либо субъективные соображения. Во время же торга цена квартиры изменяется исходя из каких-либо объективных характеристик квартиры.

В целом данное статистическое исследование является актуальным и содержит ряд ценных выводов и объяснений. При этом стоит отметить, что оно было проведено при использовании данных о вторичном рынке жилья. Это очень важный момент, так как когда мы работаем с рынком первичного жилья, цены предложения и цены сделок либо практически всегда совпадают, либо достаточно близки. Вопрос о том, как изменились бы оценки коэффициентов (и изменились бы они?) в случае, если бы мы рассматривали рынок первичного жилья, остается открытым.

ГЛАВА III. ГЕДОНИСТИЧЕСКИЙ АНАЛИЗ ЦЕН НА РЫНКЕ ПЕРВИЧНОГО ЖИЛЬЯ В САНКТ-ПЕТЕРБУРГЕ

3. 1. Подготовка данных

Для проведения гедонистического анализа цен на жилье мне была предоставлена база цен на первичное жилье в Санкт-Петербурге информационно-аналитическим центром ООО «Норминфо». База содержит 24 параметра (переменных) по каждому наблюдению.

Анализ проводится в программном продукте Stata.

Прежде чем приступать к построению модели необходимо удалить те наблюдения, в которых отсутствует информации о цене квадратного метра квартиры в случае 100% оплаты. Данная стоимость выбрана исходя из тех соображений, что не по всем наблюдениям имеется возможность оплаты в рассрочку.

В итоге было удалено 99 наблюдений из 1091. Стоит также отметить, что в базе присутствует как стоимость одного квадратного метра при покупке квартиры с максимальной площадью, так и с минимальной. Поскольку отдать предпочтение первому или второму показателю не возможно, в данном случае было решено использовать среднюю цену квадратного метра квартиры при 100% оплате.

Прежде чем приступать к построению модели необходимо проверить на нормальность зависимую переменную. Если зависимая переменная будет распределена нормально, то возрастает вероятность получить нормально распределённые остатки по модели, которые необходимы для корректной проверки всех гипотез по модели.

Проведённый графический и статистический анализ (см. приложение 1) говорят нам о том, что на уровне значимости 5% распределение зависимой переменной далеко от нормального, поэтому было решено прологарфимировать зависимую переменную и проверить её на нормальность распределения.

Проведённый графический и статистический анализ (см. приложение 2) всё также на уровне значимости 5% отвергают гипотезу о нормальности распределения зависимой переменной, представленной в виде логарифма первоначальной зависимой переменной. Однако значения расчётных статистик в тестах на нормальность распределения стали значительно меньше, что говорит о приближении к тому, чтобы принять нулевую гипотезу, которая состоит в том, что проверяемая переменная распределена по нормальному закону. Исходя из вышесказанного, было решено проводить дальнейший анализ с использованием в качестве зависимой переменной логарифма цены квадратного метра квартиры в случае 100% оплаты.

3. 2. Построение простейших гедонистических моделей и тестирование базовых гипотез о влиянии отдельных характеристик квартиры на ее цену

3. 2. 1. Влияние классов квартиры на цену квадратного метра

Первая гипотеза, которая кажется наиболее правдоподобной, - эта гипотеза о том, что на среднюю стоимость квадратного метра квартиры влияет её класс: так, например, переход от эконом-класса к комфорт-классу увеличивает цену квадратного метра квартиры и так далее. Чтобы проверить данную гипотезу, следует построить таблицу средних цен квадратного метра квартиры в зависимости от её класса:

Класс квартиры	Summary of Средняя стоимость квадратного метра квартиры при 100% оплате Mean
-----	-----

0		69291.191
1		91804.046
2		120137.5
3		122801.43
4		164588.34
-----+		
Total		92677.52

Табл. 1: Таблица средних цен квадратного метра квартиры в зависимости от её класса

Представленная выше таблица полностью подтверждает выдвинутое предположение, поэтому построим регрессионную модель зависимости логарифма цены квадратного метра квартиры от фиктивных переменных классов квартиры. В качестве базового класса квартиры предлагается использовать «Эконом-класс»:

-----		class4	0.851***
(1)			(0.0321)
model01		_cons	11.12***
-----			(0.0121)
class1	0.277***	N	992
	(0.0164)	R-sq	0.501
class2	0.574***	adj. R-sq	0.499
	(0.0204)	AIC	39.16
class3	0.545***	BIC	63.66
	(0.0405)	-----	
		Standard errors in parentheses	
		* p<0.05, ** p<0.01, *** p<0.001	

Табл. 2: Результаты оценивания регрессионной модели №1

Исходя из выше представленной таблицы, на уровне значимости 5% все классы различимы по степени влияния на цену квадратного метра квартиры. Однако для корректной проверки гипотез о значимости (незначимости) тех или иных коэффициентов модели необходимо, чтобы остатки модели были распределены нормально. В данном случае графический анализ не подтвердил нормальность распределения остатков, однако один из трёх проведённых тестов на уровне значимости 5% подтвердил нормальность распределения остатков модели (см. приложение 3), поэтому данным оценкам можно доверять.

Однако, если внимательно посмотреть на коэффициенты рассматриваемой модели, то можно заметить, что влияние 2 классов квартир («Комфорт-класс (smart-комплекс)» и «Бизнес-класс») на цену квадратного метра квартиры почти одинаково. Для удобства с помощью программного продукта Microsoft Excel представим имеющиеся данные в виде таблицы расстояний между средними ценами квадратного метра квартир различных классов:

Класс квартиры		Эконом-класс	Комфорт-класс	Smart-комплекс	Бизнес-класс	Элит-класс
	Цена	69291	91804	120138	122801	164588
Эконом-класс	69291	0	22513	50846	53510	95297
Комфорт-класс	91804		0	28333	30997	72784
Smart-комплекс	120138			0	2664	44451
Бизнес-класс	122801				0	41787
Элит-класс	164588					0

Табл. 3: Таблица расстояний между средними ценами квадратного метра квартиры в зависимости от её класса

Из представленной выше таблицы видно, что небольшое различие наблюдается только между ценами квадратного метра квартиры комфорт-класса (smart-комплекс) и квартиры бизнес-класса. Это даёт основания предполагать, что данные классы квартир оказывают одинаковое влияние на цену квадратного метра квартиры. Поэтому стоит рассмотреть возможность объединения этих двух классов квартир по степени влияния на цену квадратного метра квартиры. Сравним имеющуюся модель с моделью, получающейся в случае объединения двух описанных выше классов квартир по степени влияния на цену квадратного метра квартиры:

-----		class23		0.548***
(1)	(2)			(0.0368)
model01	model02			
-----		-----		
class1	0.277*** (0.0164)	0.277*** (0.0164)	_cons	11.12*** (0.0121)
class2	0.574*** (0.0204)		N	992
class3	0.545*** (0.0405)		R-sq	0.501
			adj. R-sq	0.499
			AIC	39.16
			BIC	63.66

			Standard errors in parentheses	

class4	0.851*** (0.0321)	0.851*** (0.0321)	* p<0.05, ** p<0.01, *** p<0.001
--------	----------------------	----------------------	----------------------------------

Табл. 4: Таблица сравнения регрессионных моделей №1 и №2

Приведённая выше таблица показывает, что по информационным критериям Акайке и Шварца лидирует вторая регрессионная модель (37.26 и 39.16 против 56.86 и 63.66, соответственно), которая также не уступает и по значению скорректированного коэффициента детерминации (0.499 против 0.499). Исходя из вышесказанного, вторая модель является предпочтительной. Проверка данной модели на нормальность распределения остатков на уровне значимости 5% подтвердила нормальность их распределения (см. приложение 4). Что касается дальнейшего объединения классов квартиры по степени влияния на цену квадратного метра квартир, то автор считает его бессмысленным, поскольку коэффициенты модели перед фиктивными переменными классов квартир в регрессионной модели очень сильно различаются.

В итоге была получена следующая регрессионная модель цены квадратного метра квартиры от её класса, в которой можно выделить 4 группы квартир по степени влияния на цену квадратного метра:

$$\ln P = 11.12 + 0.277 \cdot group0_1 + 0.548 \cdot group0_2 + 0.851 \cdot group0_3,$$

где P - цена квадратного метра квартиры (в рублях); $group0_1$ - фиктивная переменная «Квартира комфорт-класса» (0 – Нет, 1 – Да), $group0_2$ - фиктивная переменная «Квартира комфорт-класса (smart-комплекс) или бизнес-класса» (0 – Нет, 1 – Да), $group0_3$ - фиктивная переменная «Квартира элит-класса» (0 – Нет, 1 – Да).

Интерпретация коэффициентов полученной модели позволяет среднюю цену квадратного метра в каждой группе квартир:

1. Средняя цена квадратного метра квартиры эконом-класса равна 69870 рублей 52 копейки;

2. Средняя цена квадратного метра квартиры комфорт-класса на 31.8655% выше средней цены квадратного метра квартиры эконом-класса и составляет 92135 рублей 11 копеек;
3. Средняя цена квадратного метра квартиры комфорт-класса (smart-комплекс) или бизнес-класса на 72.9873% выше средней цены квадратного метра квартиры эконом-класса и составляет 120867 рублей 11 копеек;
4. Средняя цена квадратного метра квартиры элит-класса на 134.2353% выше средней цены квадратного метра квартиры эконом-класса и составляет 163661 рубль 38 копеек.

3. 2. 2. Влияние типов и классов квартиры на цену квадратного метра

Следующая гипотеза, которая кажется наиболее правдоподобной, состоит в том, что на цену квадратного метра влияет не только класс квартиры, но и тип. Так квадратный метр квартиры-студии состоит, скорее всего, меньше, чем квадратный метр однокомнатной квартиры, и так далее. Проверим наше предположение с помощью построение таблицы средних цен квадратного метра квартиры в зависимости от её типа:

Тип квартиры	Summary of Среднийайа стоимость квадратного метра квартиры при 100% оплате Mean
0	91588.712
1	93751.152
2	89461.41
3	90706.125
4	150065
5	162725

Total | 92677.52

Табл. 5: Таблица средних цен квадратного метра квартиры в зависимости от её типа

Приведённая выше таблица частично опровергает данное предположение, так цена квадратного метра квартиры-студии оказалась ниже, чем цена квадратного метра однокомнатной квартиры, но в целом некоторая зависимость цены квадратного метра квартиры от её класса присутствует, хотя не монотонная. Поэтому можно попробовать построить регрессионную зависимость логарифма цены квадратного метра квартиры от её класса и типа, выбрав в качестве базового типа квартиры-студии, добавив к базовой модели фиктивные переменные всех типов квартир кроме фиктивной переменной «Квартира-студия» и сравнив полученные результаты моделью, включающей только классы:

	(1) model102	(2) model103		
group0_1	0.277*** (0.0164)	0.279*** (0.0159)	flat3	-0.140*** (0.0236)
group0_2	0.548*** (0.0368)	0.551*** (0.0375)	flat4	0.0627 (0.0659)
group0_3	0.851*** (0.0321)	0.856*** (0.0316)	flat5	0.101 (0.124)
flat1		-0.0400 (0.0230)	_cons	11.12*** (0.0121)
flat2		-0.107*** (0.0229)		11.20*** (0.0201)
			N	992
			R-sq	0.501
			adj. R-sq	0.499
			AIC	37.26
			BIC	56.86
			Standard errors in parentheses	
			* p<0.05, ** p<0.01, *** p<0.001	

Табл. 6: Таблица сравнения регрессионных моделей №2 и №3

В данном случае получилось, что коэффициенты перед 3 из 5 фиктивными переменными типов квартиры оказались не значимыми на уровне значимости 5%.

Такая незначимость коэффициентов модели перед фиктивными переменными типов квартиры, скорее всего, связана с влиянием классов квартир, которые мы рассматривали выше: так, например, однокомнатная квартира эконом-класса может стоить как квартира-студия комфорт-класса и так далее. Проверим данное предположение можно с помощью таблицы

средних цен квадратного метра квартиры в зависимости от её типа и группы, к которой она относится:

Means of Средняя стоимость квадратного метра квартиры при 100% оплате

Тип квартиры	Группы квартир в зависимости от их класса				Total
	0	1	2	3	
0	76382.985	98881.325	109618.75	151300	91588.712
1	72335.822	95513.603	120073.68	164907.5	93751.152
2	67055.728	89126.825	119480.95	155738.83	89461.41
3	62185.333	85715.909	126247.92	160701.4	90706.125
4	66650	94110	138870	212887.5	150065
5	.	.	165450	160000	162725
Total	69291.191	91804.046	122528.21	164588.34	92677.52

Табл. 7: Таблица средних цен квадратного метра квартиры в зависимости от её типа и группы, к которой она относится в результате анализа, проведённого в предыдущей главе

В данном случае представленная выше таблица показывает, что в случае, например, с квартирой из групп №0 и №1, то есть в случае квартиры эконом-класса и комфорт-класса, цена квадратного метра квартиры уменьшается с увеличением количества комнат, но она нарушается, когда речь идёт о переходе от трёхкомнатной к четырёхкомнатной квартире. Поэтому с помощью cross-dummy переменных стоит попробовать учесть в модели этот факт, но квартиры не всех типов бывает в каждой из групп квартир. Исходя из этого, некоторые cross-dummy переменные не нужно будет создавать. Для ответа на данный вопрос можно воспользоваться представленной выше таблицей, из которой видно, что квартиры-студии, однокомнатные, двухкомнатные, трёхкомнатные и четырёхкомнатные квартиры встречаются во всех группах квартир, а вот, например, пятикомнатные квартиры бывают только в группах №2 и №3, то есть бизнес-класса или элит-класса, что не удивительно. Исходя из проведённого выше анализа, необходимо создать cross-dummy переменные фиктивных переменных классов квартиры на фиктивные переменные типа квартиры, учитывая тот факт, что квартиры комфорт-класса (smart-комплекс) и бизнес-класса, как было доказано в предыдущей главе, оказывают одинаковое воздействие на среднюю цену

квадратного метра. Далее построим регрессионную модель зависимости цены квадратного метра от только что полученных фиктивных переменных, чтобы не возникло полной мультиколлинеарности откажемся от фиктивных переменных классов квартиры и от фиктивной переменной «Квартира-студия эконом-класса», которая будет базовой:

-----		-----	
(1)			
model04			
-----		-----	
f10_gr0_1	0.258*** (0.0373)	f13_gr0_1	0.109** (0.0362)
f10_gr0_2	0.358*** (0.0899)	f13_gr0_2	0.479*** (0.0704)
f10_gr0_3	0.704*** (0.0509)	f13_gr0_3	0.732*** (0.0585)
f11_gr0_0	-0.0548 (0.0362)	f14_gr0_0	-0.124 (0.115)
f11_gr0_1	0.221*** (0.0346)	f14_gr0_1	0.219** (0.0808)
f11_gr0_2	0.425*** (0.0781)	f14_gr0_2	0.573*** (0.147)
f11_gr0_3	0.757*** (0.0666)	f14_gr0_3	1.007*** (0.107)
f12_gr0_0	-0.129*** (0.0355)	f15_gr0_2	0.797*** (0.0282)
f12_gr0_1	0.148*** (0.0352)	f15_gr0_3	0.763*** (0.0282)
f12_gr0_2	0.429*** (0.0707)	_cons	11.22*** (0.0282)
f12_gr0_3	0.701*** (0.0607)	-----	-----
f13_gr0_0	-0.198*** (0.0353)	N	992
		R-sq	0.535
		adj. R-sq	0.525
		AIC	-0.668
		BIC	97.33
		-----	-----
		Standard errors in parentheses	
		* p<0.05, ** p<0.01, *** p<0.001	

Табл. 8: Таблица результатов оценивания регрессионной модели №4

В представленной выше таблице не приведены результаты оценивания предыдущей регрессионной модели, поскольку набор используемых независимых переменных кардинально поменялся, а именно, в последней регрессионной модели не используется ни одна из независимых переменных предыдущей регрессионной модели. В данном случае из представленной выше

таблицы видно, что два коэффициента модели не значимы на уровне значимости 5%, но, чтобы это однозначно можно было утверждать, надо проверить остатки модели на нормальность, которая на уровне значимости 5% не подтверждается ни одним из тестов (см. приложение 5). Поэтому необходимо провести бутстрапирование этой модели, которое подтверждает не значимость на уровне значимости 5% двух коэффициентов модели, а именно, коэффициентов перед фиктивными переменными «Однокомнатная квартира эконом-класса» и «Четырёхкомнатная квартира эконом-класса» (см. приложение 5). Однако пока не стоит утверждать, что однокомнатная и четырёхкомнатные квартиры эконом-класса оказывают такое же влияние на цену квадратного метра как и квартиры-студии эконом-класса, лучше обратиться к таблице расстояний между средними ценами квадратного метра квартиры в зависимости от её класса и типа (см. приложение 6). Из данной таблицы видно, что некоторые средние цены квадратного метра квартиры в зависимости от её класса и типа очень сильно похожи. Так можно предполагать, что равны средние цены квадратного метра:

1. Двухкомнатной, трёхкомнатной и четырёхкомнатной квартир эконом-класса;
2. Однокомнатной квартиры элит-класса и пятикомнатной квартиры комфорт-класса (smart-комплекс) или бизнес-класса;
3. Однокомнатной и двухкомнатной квартир комфорт-класса (smart-комплекс) или бизнес-класса;
4. Трёхкомнатной и пятикомнатной квартиры элит-класса;
5. Квартиры-студии, однокомнатной и четырёхкомнатной квартир комфорт-класса;
6. Двухкомнатной и трёхкомнатной квартир комфорт-класса;
7. Квартиры-студии и однокомнатной квартиры эконом-класса;
8. Квартиры-студии и двухкомнатной квартиры элит-класса.

Проверяя данные предположения с помощью формального теста на равенство коэффициентов внутри групп, тест подтверждает выдвинутые гипотезы который на уровне значимости 5% ($p\text{-value} = 0.3217$).

Исходя из выше изложенного, можно объединить указанные фиктивные переменные квартир по их классу и типу и сравнить результаты:

	(1) model104	(2) model105			
f10_gr0_1	0.258*** (0.0373)		f14_gr0_0	-0.124 (0.115)	
f10_gr0_2	0.358*** (0.0899)	0.392*** (0.0868)	f14_gr0_1	0.219** (0.0808)	
f10_gr0_3	0.704*** (0.0509)		f14_gr0_2	0.573*** (0.147)	0.606*** (0.144)
f11_gr0_0	-0.0548 (0.0362)		f14_gr0_3	1.007*** (0.107)	1.040*** (0.104)
f11_gr0_1	0.221*** (0.0346)		f15_gr0_2	0.797*** (0.0282)	
f11_gr0_2	0.425*** (0.0781)		f15_gr0_3	0.763*** (0.0282)	
f11_gr0_3	0.757*** (0.0666)		f1014_gr0_1		0.268*** (0.0234)
f12_gr0_0	-0.129*** (0.0355)		f102_gr0_3		0.735*** (0.0500)
f12_gr0_1	0.148*** (0.0352)		f112_gr0_2		0.460*** (0.0515)
f12_gr0_2	0.429*** (0.0707)		f11_gr0_3~2		0.792*** (0.0599)
f12_gr0_3	0.701*** (0.0607)		f1234_gr0_0		-0.124*** (0.0235)
f13_gr0_0	-0.198*** (0.0353)		f123_gr0_1		0.163*** (0.0235)
f13_gr0_1	0.109** (0.0362)		f135_gr0_3		0.767*** (0.0524)
f13_gr0_2	0.479*** (0.0704)	0.512*** (0.0666)	_cons	11.22*** (0.0282)	11.19*** (0.0177)
f13_gr0_3	0.732*** (0.0585)		N	992	992
			R-sq	0.535	0.530
			adj. R-sq	0.525	0.525
			AIC	-0.668	-7.739
			BIC	97.33	51.06
			Standard errors in parentheses		
			* p<0.05, ** p<0.01, *** p<0.001		

Табл. 3.5: Таблица сравнения регрессионных моделей №4 и №5

Из представленной выше таблицы видно, что у регрессионной модели №4 объясняющая способность (скорректированный коэффициент детерминации) такая же как и у регрессионной модели №5, но по

информационным критериям лидирует регрессионная модель №5, к тому же в ней все коэффициенты модели значимы на уровне значимости 5%.

В итоге получена регрессионная модель цены квадратного метра квартиры от её класса и типа, в которой чётко выделяются 12 групп квартир по степени влияния на цену квадратного метра:

$$\ln P = 11.19 + 0.268 \cdot group1_1 + 0.392 \cdot group1_2 + 0.735 \cdot group1_3 + \\ + 0.46 \cdot group1_4 + 0.792 \cdot group1_5 - 0.124 \cdot group1_6 + 0.163 \cdot group1_7 + \\ + 0.512 \cdot group1_8 + 0.767 \cdot group1_9 + 0.606 \cdot group1_10 + 1.04 \cdot group1_11$$

, где P - цена квадратного метра квартиры (в рублях); $group1_1$ - фиктивная переменная «Квартира-студия, однокомнатная или четырёхкомнатная квартира комфорт-класса» (0 – Нет, 1 – Да), $group1_2$ - фиктивная переменная «Квартира-студия комфорт-класса (smart-комплекс) или бизнес-класса» (0 – Нет, 1 – Да), $group1_3$ - фиктивная переменная «Квартира-студия или двухкомнатная квартира элит-класса» (0 – Нет, 1 – Да), $group1_4$ - фиктивная переменная «Однокомнатная или двухкомнатная квартира комфорт-класса (smart-комплекс) или бизнес-класса» (0 – Нет, 1 – Да), $group1_5$ - фиктивная переменная «Однокомнатная квартира элит-класса или пятикомнатная квартира комфорт-класса (smart-комплекс) или бизнес-класса» (0 – Нет, 1 – Да), $group1_6$ - фиктивная переменная «Двухкомнатная, трёхкомнатная или четырёхкомнатная квартира эконом-класса» (0 – Нет, 1 – Да), $group1_7$ - фиктивная переменная «Двухкомнатная или трёхкомнатная квартира комфорт-класса» (0 – Нет, 1 – Да), $group1_8$ - фиктивная переменная «Трёхкомнатная квартира комфорт-класса (smart-комплекс) или бизнес-класса» (0 – Нет, 1 – Да), $group1_9$ - фиктивная переменная «Трёхкомнатная или пятикомнатная квартира элит-класса» (0 – Нет, 1 – Да), $group1_10$ - фиктивная переменная «Четырёхкомнатная квартира комфорт-класса (smart-комплекс) или бизнес-класса» (0 – Нет, 1 – Да), $group1_11$ - фиктивная переменная «Четырёхкомнатная квартира элит-класса» (0 – Нет, 1 – Да).

Давайте проинтерпретируем коэффициенты полученной модели и посмотрим на средние цены квадратного метра в каждой группе квартир:

1. Средняя цена квадратного метра квартиры-студии или однокомнатной квартиры эконом-класса равна 74386 рублей 03 копейки;
2. Средняя цена квадратного метра квартиры-студии, однокомнатной или четырёхкомнатной квартиры комфорт-класса на 30.7436% выше средней цены квадратного метра квартиры-студии или однокомнатной квартиры эконом-класса и составляет 97254 рубля 96 копеек;
3. Средняя цена квадратного метра квартиры-студии комфорт-класса (smart-комплекс) или бизнес-класса на 47.9554% выше средней цены квадратного метра квартиры-студии или однокомнатной квартиры эконом-класса и составляет 110058 рублей 11 копеек;
4. Средняя цена квадратного метра квартиры-студии или двухкомнатной квартиры элит-класса на 108.4774% выше средней цены квадратного метра квартиры-студии или однокомнатной квартиры эконом-класса и составляет 155078 рублей 08 копеек;
5. Средняя цена квадратного метра однокомнатной или двухкомнатной квартиры комфорт-класса (smart-комплекс) или бизнес-класса на 58.4754% выше средней цены квадратного метра квартиры-студии или однокомнатной квартиры эконом-класса и составляет 117883 рубля 54 копейки;
6. Средняя цена квадратного метра однокомнатной элит-класса или пятикомнатной квартиры комфорт-класса (smart-комплекс) или бизнес-класса на 120.8483% выше средней цены квадратного метра квартиры-студии или однокомнатной квартиры эконом-класса и составляет 164280 рублей 25 копеек;
7. Средняя цена квадратного метра двухкомнатной, трёхкомнатной или четырёхкомнатная квартиры эконом-класса на 11.7016% ниже средней цены квадратного метра квартиры-студии или однокомнатной квартиры эконом-класса и составляет 65681 рубль 64 копейки;

8. Средняя цена квадратного метра двухкомнатной или трёхкомнатной квартиры комфорт-класса на 17.7129% выше средней цены квадратного метра квартиры-студии или однокомнатной квартиры эконом-класса и составляет 87561 рубль 97 копеек;
9. Средняя цена квадратного метра трёхкомнатной квартиры комфорт-класса (smart-комплекс) или бизнес-класса на 66.8427% выше средней цены квадратного метра квартиры-студии или однокомнатной квартиры эконом-класса и составляет 124107 рублей 68 копеек;
10. Средняя цена квадратного метра трёхкомнатной или пятикомнатной квартиры элит-класса на 115.2323% выше средней цены квадратного метра квартиры-студии или однокомнатной квартиры эконом-класса и составляет 160102 рубля 77 копеек;
11. Средняя цена квадратного метра четырёхкомнатной квартиры комфорт-класса (smart-комплекс) или бизнес-класса на 83.3585% выше средней цены квадратного метра квартиры-студии или однокомнатной квартиры эконом-класса и составляет 136393 рубля 11 копеек;
12. Средняя цена квадратного метра четырёхкомнатной квартиры элит-класса на 182.9307% выше средней цены квадратного метра квартиры-студии или однокомнатной квартиры эконом-класса и составляет 210460 рублей 87 копеек.

3. 3. 3. Влияние района на цену квадратного метра

Следующим шагом будет проверка влияния района, в котором располагается квартира, на ее цену.

Построим модель, где будут выделена принадлежность к одному из районов Санкт-Петербурга или Ленинградской области, где в качестве базовых районов взяты Адмиралтейский район (в Санкт-Петербурге) и Всеволожский район (в Ленинградской области):

F(26, 965) = 195.74
 Prob > F = 0.0000
 R-squared = 0.8582
 Root MSE = .23732

ln_price100	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
square	.0138192	.0005141	26.88	0.000	.0128103	.014828
spb	.5615917	.0502276	11.18	0.000	.4630238	.6601596
district02	-.00928	.0870573	-0.11	0.915	-.1801235	.1615634
district03	-.396245	.0547592	-7.24	0.000	-.5037059	-.2887841
district04	-.2993165	.0518119	-5.78	0.000	-.4009935	-.1976395
district05	-.2433986	.0551764	-4.41	0.000	-.3516781	-.1351191
district06	-.6090162	.0743392	-8.19	0.000	-.7549013	-.4631311
district07	-.2230336	.0704881	-3.16	0.002	-.3613611	-.084706
district08	-.2894248	.0523541	-5.53	0.000	-.3921658	-.1866838
district09	-.405856	.0643007	-6.31	0.000	-.5320413	-.2796708
district10	-.4125732	.0789674	-5.22	0.000	-.5675408	-.2576055
district11	-.1356675	.0535497	-2.53	0.011	-.2407547	-.0305802
district12	-.2894928	.0532271	-5.44	0.000	-.3939471	-.1850386
district13	.2263085	.0689108	3.28	0.001	.0910761	.3615409
district14	-.3867767	.0663406	-5.83	0.000	-.5169651	-.2565883
district15	-.1001091	.0567155	-1.77	0.078	-.211409	.0111908
district16	-.4480635	.0804703	-5.57	0.000	-.6059805	-.2901465
district17	-.3135827	.0609569	-5.14	0.000	-.4332061	-.1939592
district18	.230302	.0761123	3.03	0.003	.0809373	.3796666
districtlr2	-.2615733	.0492143	-5.31	0.000	-.3581526	-.164994
districtlr3	-.240208	.0645729	-3.72	0.000	-.3669276	-.1134885
districtlr4	-.30491	.0475473	-6.41	0.000	-.3982181	-.211602
districtlr5	-.3497435	.0900363	-3.88	0.000	-.526433	-.173054
districtlr6	-.2209704	.0331755	-6.66	0.000	-.2860749	-.1558659
districtlr7	-.5719814	.0224452	-25.48	0.000	-.6160283	-.5279344
districtlr8	-.3737636	.0675892	-5.53	0.000	-.5064024	-.2411248
_cons	14.29338	.0294173	485.88	0.000	14.23565	14.3511

Стоит также обратить внимание на тот факт, что в модель №6 включена дамми-переменная *spb*, которая означает принадлежность квартиры к какому-либо району Санкт-Петербурга (то есть расположение квартиры в городе, а не области).

В данном случае получилось, что на уровне значимости 5% цены на квартиры Василеостровском и Приморском районах Санкт-Петербурга формируются так же, как и в Адмиралтейском районе Санкт-Петербурга. Проверить это можно с помощью формального теста на одновременное равенство 0 коэффициентов перед фиктивными переменными данных районов:

- (1) *district02* = 0
- (2) *district15* = 0

$$\text{chi2}(2) = 3.97$$

Prob > chi2 = 0.1372

В данном случае на уровне значимости 5% подтвердилась нулевая гипотеза об одновременном равенстве 0 упомянутых выше коэффициентов модели. Посмотрим, как изменится качество модели при объединении этих районов по влиянию на цену квартиры. В данном случае модель №7 улучшилась (см. приложение 7).

Проводя аналогичным образом объединение районов, в итоге получилась следующая модель №8:

```
Linear regression                               Number of obs   =      992
                                                Replications    =      428
                                                Wald chi2(9)    =  3893.27
                                                Prob > chi2     =    0.0000
                                                R-squared       =    0.8562
                                                Adj R-squared   =    0.8548
                                                Root MSE       =    0.2369
```

ln_price100	Observed Coef.	Bootstrap Std. Err.	z	P> z	Normal-based [95% Conf. Interval]	
square	.0138733	.0005358	25.89	0.000	.0128231	.0149235
spb	.463375	.0235439	19.68	0.000	.4172299	.5095201
district0309101416	-.3096405	.0275026	-11.26	0.000	-.3635447	-.2557363
district040507081217	-.178465	.0218013	-8.19	0.000	-.2211947	-.1357352
district06	-.5112753	.0603454	-8.47	0.000	-.62955	-.3930006
district1318	.3236459	.0382967	8.45	0.000	.2485858	.398706
districtlr236	-.2319376	.0328793	-7.05	0.000	-.2963798	-.1674954
districtlr458	-.3501074	.0478441	-7.32	0.000	-.4438801	-.2563346
districtlr7	-.5718293	.0233344	-24.51	0.000	-.6175639	-.5260947
_cons	14.29062	.030722	465.16	0.000	14.23041	14.35084

Note: one or more parameters could not be estimated in 72 bootstrap replicates; standard-error estimates include only complete replications.

Дальнейшие попытки объединения районов по степени влияния на формирования цены квартиры можно считать бесполезными.

Если попытаться проинтерпретировать полученные результаты, то можно отметить следующие:

В данном случае константу, включённую в модель, трудно проинтерпретировать, но если всё же попробовать её проинтерпретировать в данной модели, то – это стоимость самого факта приобретения квартиры во Всеволожском районе Ленинградской области, которая составляет 1 186 348 рублей 40 копеек;

1. Каждый дополнительный метр площади квартиры во Всеволожском районе Ленинградской области увеличивает её цену на 1.397%;
2. Если квартира располагается в Адмиралтейском районе Санкт-Петербурга, то это увеличивает её цену на 58.9429% по сравнению с квартирой во Всеволожском районе Ленинградской области;
3. Если квартира располагается в Выборгском, Кронштадтском, Курортном, Петродворцовом или Пушкинском районе Санкт-Петербурга, то это увеличивает её цену на 16.8121% по сравнению с квартирой во Всеволожском районе Ленинградской области и уменьшает – на 26.6289% по сравнению с квартирой в Адмиралтейском районе Санкт-Петербурга;
4. Если квартира располагается в Выборгском, Кронштадтском, Курортном, Петродворцовом или Пушкинском районе Санкт-Петербурга, то это увеличивает её цену на 32.9642% по сравнению с квартирой во Всеволожском районе Ленинградской области и уменьшает – на 16.3447% по сравнению с квартирой в Адмиралтейском районе Санкт-Петербурга;
5. Если квартира располагается в Колпинском районе Санкт-Петербурга, то это уменьшает её цену на 4.6771% по сравнению с квартирой во Всеволожском районе Ленинградской области и – на 40.027% по сравнению с квартирой в Адмиралтейском районе Санкт-Петербурга;
6. Если квартира располагается в Петроградском или Центральном районе Санкт-Петербурга, то это увеличивает её цену на 119.6842% по сравнению с квартирой во Всеволожском районе Ленинградской области и – на 38.2158% по сравнению с квартирой в Адмиралтейском районе Санкт-Петербурга;
7. Если квартира располагается в Выборгском, Гатчинском или Ломоносовском районе Ленинградской области, то это уменьшает её цену на 20.7004% по сравнению с квартирой во Всеволожском районе Ленинградской области;

8. Если квартира располагается в Кингисепском, Кировском или Тосненском районе Ленинградской области, то это уменьшает её цену на 29.5388% по сравнению с квартирой во Всеволожском районе Ленинградской области;
9. Если квартира располагается в Приозерском районе Ленинградской области, то это уменьшает её цену на 43.5508% по сравнению с квартирой во Всеволожском районе Ленинградской области.

Перейдём к более удобному виду модели: будем рассчитывать не цену квартиры, а цену 1 квадратного метра, и проведём похожий анализ, чтобы посмотреть, как изменятся результаты. К тому же в рассмотренной выше модели константа была неинтерпретируема.

Linear regression	Number of obs	=	992
	Replications	=	500
	Wald chi2(1)	=	499.53
	Prob > chi2	=	0.0000
	R-squared	=	0.2964
	Adj R-squared	=	0.2957
	Root MSE	=	0.2921

ln_price10~e	Observed Coef.	Bootstrap Std. Err.	z	P> z	Normal-based [95% Conf. Interval]
spb	.4062305	.0181758	22.35	0.000	.3706066 .4418544
_cons	11.09575	.0132323	838.53	0.000	11.06982 11.12169

На уровне значимости 5% на цену квадратного метра влияет тот факт, что квартира расположена в Санкт-Петербурге, а не в Ленинградской области.

Давайте проинтерпретируем полученные результаты:

1. Цена квадратного метра в Ленинградской области составляет 69 220 рублей 34 копейки;
2. Если квартира располагается в Санкт-Петербурга, то это увеличивает цену квадратного метра на 50.1149% по сравнению с квартирой в Ленинградской области.

В данном случае модель №8 может объяснить только 29.64% вариации логарифма цены за квадратный метр, что может быть связано с тем, что в разных районах Санкт-Петербурга и Ленинградской области цена квадратного метра разная давайте попробуем проверить этот факт. В качестве базовых

районов опять же будут выбраны Адмиралтейский район в случае Санкт-Петербурга и Всеволожский район в случае Ленинградской области:

Linear regression	Number of obs =	992
	F(24, 966) =	.
	Prob > F =	.
	R-squared =	0.5897
	Root MSE =	.22584

ln_pricel0~e	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
spb	.534639	.0517597	10.33	0.000	.4330645	.6362135
district02	.0263022	.0818354	0.32	0.748	-.1342935	.1868979
district03	-.3824087	.0555766	-6.88	0.000	-.4914734	-.2733439
district04	-.2845701	.0539122	-5.28	0.000	-.3903687	-.1787716
district05	-.2508693	.0550586	-4.56	0.000	-.3589177	-.142821
district06	-.6364454	.0752754	-8.45	0.000	-.7841675	-.4887233
district07	-.2238117	.0703043	-3.18	0.002	-.3617784	-.085845
district08	-.2808871	.0544179	-5.16	0.000	-.3876781	-.1740962
district09	-.360355	.0668089	-5.39	0.000	-.4914623	-.2292478
district10	-.3796061	.0853652	-4.45	0.000	-.5471288	-.2120834
district11	-.1314341	.0544909	-2.41	0.016	-.2383684	-.0244999
district12	-.3062889	.0563281	-5.44	0.000	-.4168284	-.1957494
district13	.3332924	.0664752	5.01	0.000	.20284	.4637449
district14	-.4208543	.0680369	-6.19	0.000	-.5543715	-.2873372
district15	-.1060027	.0582568	-1.82	0.069	-.2203272	.0083218
district16	-.427137	.0709793	-6.02	0.000	-.5664283	-.2878456
district17	-.3510669	.0567759	-6.18	0.000	-.4624852	-.2396487
district18	.2805367	.0706968	3.97	0.000	.1417997	.4192736
districtlr2	-.3202032	.0480998	-6.66	0.000	-.4145954	-.225811
districtlr3	-.227842	.0472648	-4.82	0.000	-.3205954	-.1350885
districtlr4	-.3650138	.0132074	-27.64	0.000	-.3909324	-.3390953
districtlr5	-.3085962	.0805952	-3.83	0.000	-.466758	-.1504344
districtlr6	-.240378	.0368446	-6.52	0.000	-.3126827	-.1680732
districtlr7	-.612675	.0315729	-19.41	0.000	-.6746343	-.5507156
districtlr8	-.321214	.0463444	-6.93	0.000	-.4121613	-.2302666
_cons	11.14397	.0132074	843.76	0.000	11.11805	11.16989

На уровне значимости 5% ни один из 3 тестов не отвергает нулевую гипотезу о нормальности распределения остатков (см. приложение 8), поэтому мы можем доверять рассчитанным при построении регрессионной модели №9 статистикам.

В ходе объединения некоторых районов, которые оказывают одинаковое влияние на цену, была получена следующая модель №10:

Linear regression	Number of obs	=	992
	Replications	=	423
	Wald chi2(9)	=	1898.52
	Prob > chi2	=	0.0000
	R-squared	=	0.5870
	Adj R-squared	=	0.5832
	Root MSE	=	0.2247

ln_pricel0square	Observed Coef.	Bootstrap Std. Err.	z	P> z	Normal-based [95% Conf. Interval]	
spb	.5509213	.0427125	12.90	0.000	.4672064	.6346362
district030910141617	-.4084386	.0457279	-8.93	0.000	-.4980637	-.3188135

district0405070812	-.2882459	.0428564	-6.73	0.000	-.3722429	-.2042489
district06	-.6527277	.0771893	-8.46	0.000	-.804016	-.5014394
district1115	-.1347358	.0443297	-3.04	0.002	-.2216203	-.0478513
district1318	.2934089	.0515968	5.69	0.000	.1922809	.3945368
districtlr2458	-.3241396	.0345412	-9.38	0.000	-.3918391	-.25644
districtlr36	-.2350054	.0303182	-7.75	0.000	-.2944279	-.1755828
districtlr7	-.612675	.033168	-18.47	0.000	-.6776829	-.547667
_cons	11.14397	.0128145	869.64	0.000	11.11885	11.16909

Note: one or more parameters could not be estimated in 77 bootstrap replicates;
standard-error estimates include only complete replications.

В рассмотренной выше модели у нас было жёсткое разделение между районами Санкт-Петербурга и районами Ленинградской области, однако получилось, что в Колпинском районе Санкт-Петербурга цена квадратного метра оказалась ниже, чем цена квадратного метра во Всеволожском районе Ленинградской области, хотя мы первоначально предполагали, что цена квадратного метра в Санкт-Петербурге выше, чем цена квадратного метра в Ленинградской области. Исходя из выше сказанного, можно отказаться от жёсткого разделения между Санкт-Петербургом и Ленинградской областью и ещё раз провести подобный. В качестве базового района в данном случае будет выбран Адмиралтейский район Санкт-Петербурга. Также следует провести объединение районов (аналогичным образом), для того, чтобы получить более компактную модель:

Linear regression	Number of obs	=	992
	Replications	=	440
	Wald chi2(8)	=	1973.55
	Prob > chi2	=	0.0000
	R-squared	=	0.5861
	Adj R-squared	=	0.5828
	Root MSE	=	0.2249

ln_price100square	Observed Coef.	Bootstrap Std. Err.	z	P> z	Normal-based [95% Conf. Interval]	
district030910141617	-.4084386	.0490546	-8.33	0.000	-.5045838	-.3122933
district0405070812	-.2882459	.0479084	-6.02	0.000	-.3821446	-.1943472
district06lr1	-.5546642	.0473144	-11.72	0.000	-.6473987	-.4619296
district1115	-.1347358	.04934	-2.73	0.006	-.2314405	-.0380311
district1318	.2934089	.0593391	4.94	0.000	.1771064	.4097113
districtlr2458	-.8750609	.0559126	-15.65	0.000	-.9846476	-.7654742
districtlr36	-.7859267	.053813	-14.60	0.000	-.8913981	-.6804552
districtlr7	-1.163596	.0542633	-21.44	0.000	-1.26995	-1.057242
_cons	11.69489	.0454398	257.37	0.000	11.60583	11.78395

Note: one or more parameters could not be estimated in 60 bootstrap replicates;
standard-error estimates include only complete replications.

Дальнейшее объединение районов по степени влияния на цену квадратного метра квартиры можно считать бесполезным. Что же касается получившейся модели №11, то у неё остатки распределены не нормально, однако автор пока не будет возвращаться к предыдущей модели, поскольку существует ещё много неучтённых характеристик квартиры, включение которых в модель может вернуть нормальность распределения остатков модели. Давайте проинтерпретируем полученные в модели коэффициенты:

1. Цена квадратного метра в Адмиралтейском районе Санкт-Петербурга составляет 123 578 рублей 84 копейки;

2. Если квартира располагается в Выборгском, Кронштадтском, Курортном, Петродворцовом, Пушкинском или Фрунзенском районе Санкт-Петербурга, то это уменьшает цену квадратного метра на 33.5313% по сравнению с квартирой в Адмиралтейском районе Санкт-Петербурга;

3. Если квартира располагается в Калининском, Кировском, Красногвардейском, Красносельском или Невском районе Санкт-Петербурга, то это уменьшает цену квадратного метра на 25.0423% по сравнению с квартирой в Адмиралтейском районе Санкт-Петербурга;

4. Если квартира располагается в Колпинском районе Санкт-Петербурга или во Всеволожском районе Ленинградской области, то это уменьшает цену квадратного метра на 42.5735% по сравнению с квартирой в Адмиралтейском районе Санкт-Петербурга;

5. Если квартира располагается в Московском или Приморском районе Санкт-Петербурга, то это уменьшает цену квадратного метра на 12.6053% по сравнению с квартирой в Адмиралтейском районе Санкт-Петербурга;

6. Если квартира располагается в Петроградском или Центральном районе Санкт-Петербурга, то это увеличивает цену квадратного метра на 34.0991% по сравнению с квартирой в Адмиралтейском районе Санкт-Петербурга;

7. Если квартира располагается в Выборгском, Кингисепском, Кировском или Тосненском районе Ленинградской области, то это уменьшает цену квадратного метра на 58.3163% по сравнению с квартирой в Адмиралтейском районе Санкт-Петербурга;

8. Если квартира располагается в Гатчинском или Ломоносовском районе Ленинградской области, то это уменьшает цену квадратного метра на 54.4303% по сравнению с квартирой в Адмиралтейском районе Санкт-Петербурга;

9. Если квартира располагается в Приозерском районе Ленинградской области, то это уменьшает цену квадратного метра на 68.7639% по сравнению с квартирой в Адмиралтейском районе Санкт-Петербурга.

3. 3. Построение полной гедонистической модели цены на рынке первичного жилья в Санкт-Петербурге

В предыдущем разделе, посвященном построению простейших гедонистических моделей, были протестированы основные гипотезы относительно влияния на цену таких жилищных характеристик как класс и тип квартиры, а также ее расположение.

Можно утверждать, что данные характеристики вносят решающий вклад в формирование цены квартиры. Такой вывод можно сделать на основе проведенного анализа Shapley decomposition (см. приложение 10).

Однако не стоит забывать о том, что квартира обладает еще определенным рядом характеристик, которые, возможно, не вносят столь существенного вклада по сравнению с классом квартиры или ее расположением, однако, тем не менее, оказывают значимое влияние на цену, и являются важными с точки зрения покупателя.

С одной стороны, можно было попробовать создать группы квартир (аналогично модели, построенной в разделе 3 .2. 2.), принадлежность к которым приводит к одинаковому влиянию на цену. Но данная идея тяжело

применима с точки зрения практического подхода – создавая группы, придется «пересекать» все возможные комбинации признаков, и количество групп станет очень большим. Одно лишь пересечение признаков «класс» и «тип квартиры» дает 30 групп (хотя некоторые из них и окажутся пустыми, все равно количество групп останется значительным), а дальнейшее пересечение делает анализ практически бессмысленным.

Поэтому было принято решение включить в модель дамми-переменные, обозначающие обладание каким-либо конкретным признаком, и не составлять так называемые «группы характеристик», исключение было сделано для районов. Результаты оценивания модели представлены ниже:

	Model12		
-----	-----		
d~0309101416 (0.0238)	-0.231***	permission (0.0484)	0.263***
dist~7081217 (0.0206)	-0.107***	material1 (0.0298)	-0.0951**
district06 (0.0696)	-0.320***	material2 (0.129)	-0.491***
district1318 (0.0309)	0.169***	material3 (0.0359)	-0.103**
district~236 (0.0393)	-0.167***	material5 (0.0518)	-0.262***
distric~r458 (0.0568)	-0.298***	material7 (0.0344)	-0.160***
districtlr7 (0.149)	-0.463**	concierge (0.0164)	0.0567***
square (0.000240)	0.0131***	parking (0.0166)	0.0143
spb (0.0243)	0.303***	_cons (0.0577)	14.11***
class1 (0.0176)	0.104***	N	992
class2 (0.0771)	0.203**	R-sq	0.890
class3 (0.0319)	0.258***	adj. R-sq	0.888
		AIC	-278.1
		BIC	-170.3
		Standard errors in parentheses	
		* p<0.05, ** p<0.01, *** p<0.001	

```
class4          0.437*** |
(0.0328)
```

Стоит отметить, что по информационным критериям данную модель можно признать наилучшей. Графические и формальные тесты на нормальность распределения остатков (см. приложение 11) не подтверждают гипотезы о том, что остатки распределены нормально.

Попытаемся проанализировать полученные коэффициенты. Обращая внимание на тип зависимой переменной (цена квартиры при стопроцентной оплате), можно узнать, чему будет равна цена квартиры с нулевой площадью в Адмиралтейском, Приморском или Московском районах квартиры-студии эконом-класса. Данная квартира расположена в кирпичном или монолитно-панельном доме без консьержа и паркинга при условии, что строительной организацией не было получено разрешение на строительство. Цена такой квартиры составит 101742 рубля. Принимая во внимание тот факт, что квартира имеет нулевую площадь, можно сказать, что вышеупомянутой стоимостью обладает сам факт покупки. Чтобы попытаться объяснить, что это может означать в реальной практике, стоит вспомнить, как происходит процесс покупки. В случае если заключается договор долевого участия (ДДУ) происходит оплата выбранной квартиры в полном объеме. Однако если заключается предварительный договор долевого участия, что происходит как раз в том случае, если застройщиком еще не получено разрешение на строительство, происходит не полная оплата квартиры, внесение некоторого взноса, который гарантирует намерение заключить ДДУ в дальнейшем со стороны как покупателя, так и застройщика.

Интерпретация остальных коэффициентов регрессионной модели дает следующие результаты:

1. Если квартира располагается в Выборгском, Кронштадтском, Курортном, Петродворцовом, Пушкинском или Фрунзенском районе Санкт-Петербурга, то это уменьшает цену на 20.603595% по сравнению с квартирой в Адмиралтейском районе Санкт-Петербурга;

2. Если квартира располагается в Калининском, Кировском, Красногвардейском, Красносельском или Невском районе Санкт-Петербурга, то это уменьшает цену на 10.152626% по сравнению с квартирой в Адмиралтейском районе Санкт-Петербурга;

3. Если квартира располагается в Колпинском районе Санкт-Петербурга, то это уменьшает цену на 27.336863% по сравнению с квартирой в Адмиралтейском районе Санкт-Петербурга;

4. Если квартира располагается в Петроградском или Центральном районе Санкт-Петербурга, то это увеличивает цену на 18.310835% по сравнению с квартирой в Адмиралтейском районе Санкт-Петербурга;

5. Если квартира располагается в Выборгском, Гатчинском или Ломоносовском районе Ленинградской области, то это уменьшает цену на 15.4573% по сравнению с квартирой в Адмиралтейском районе Санкт-Петербурга;

6. Если квартира располагается в Кингисепском, Кировском или Тосненском районе Ленинградской области, то это уменьшает цену на 25.861497% по сравнению с квартирой в Адмиралтейском районе Санкт-Петербурга;

7. Если квартира располагается в Приозерском районе Ленинградской области, то это уменьшает цену на 37.101405% по сравнению с квартирой в Адмиралтейском районе Санкт-Петербурга.

8. Увеличение площади квартиры на 1 квадратный метр увеличивает цену квартиры на 1.3235%;

9. Факт расположения квартиры в Санкт-Петербурге, а не Ленинградской области увеличивает цену квартиры на 35.222114%;

10. Принадлежность квартиры к комфорт-классу увеличивает цену квартиры на 11.077375%;

11. Принадлежность квартиры к комфорт-классу смарт (smart) комплекса увеличивает цену квартиры на 22.707218%;

12. Принадлежность квартиры к бизнес-классу увеличивает цену квартиры на 29.608887%;

13. Принадлежность квартиры к элит-классу увеличивает цену квартиры на 54.978929%;

14. Факт получения застройщиком разрешения на строительство увеличивает цену квартиры на 30.06409%;

15. Если квартира находится в кирпично-монолитном, монолит-газобетонном, монолитном, монолитно-каркасном или панельном доме, то это уменьшает цену квартиры на 9.5319046%, 39.118583%, 10.303579%, 23.419099% и 15.210037% соответственно;

16. Наличие консьержа в доме, в котором располагается квартира, увеличивает ее цену на 5.7542084%;

17. Наличие паркинга в доме, в котором располагается квартира, увеличивает ее цену на 1.4001166%.

Как уже было сказано ранее, то, что константа объясняет цену квартиры с нулевой площадью, может быть отнесено к недостаткам модели, приводит к необходимости рассмотреть модель №13, объясняющую цену одного квадратного метра квартиры:

-----		district3	-0.196***
(1)			(0.0368)
model13		district4	-0.200***
-----			(0.0400)
class1	0.116***	district5	-0.401***
	(0.0154)		(0.0816)
class2	0.285***	district6	-0.175***
	(0.0352)		(0.0477)
class3	0.296***	district7	-0.228***
	(0.0336)		(0.0358)
class4	0.468***	district8	-0.226***
	(0.0298)		(0.0492)
concierge	0.0437**	district9	-0.331***
	(0.0160)		(0.0655)
district2	-0.318***	district10	-0.122**
	(0.0413)		

	(0.0393)		(0.0174)
district11	-0.206*** (0.0378)	flat2	-0.118*** (0.0171)
district12	0.122* (0.0524)	flat3	-0.178*** (0.0183)
district13	-0.317*** (0.0651)	lift	0.0675* (0.0309)
district14	-0.130*** (0.0386)	material1	-0.0734*** (0.0186)
district15	-0.276*** (0.0606)	material2	-0.367*** (0.0640)
district16	-0.0925 (0.0478)	material3	-0.0735** (0.0245)
district17	0.115* (0.0479)	material5	-0.212*** (0.0506)
district18	-0.360*** (0.0367)	material7	-0.151*** (0.0221)
district19	-0.598*** (0.0653)	parking	0.0274 (0.0141)
district20	-0.461*** (0.0585)	permission	0.269*** (0.0582)
district21	-0.688*** (0.0512)	turn	0.0338* (0.0137)
district22	-0.587*** (0.0470)	_cons	11.26*** (0.0757)
district23	-0.559*** (0.0475)	N	992
district24	-0.828*** (0.0444)	R-sq	0.740
district25	-0.721*** (0.0488)	adj. R-sq	0.729
flat1	-0.0535**	AIC	-534.6
		BIC	-328.8
		Standard errors in parentheses	
		* p<0.05, ** p<0.01, *** p<0.001	

Согласно данной модели, цена квадратного метра квартиры эконом-класса, расположенной в Адмиралтейском районе в кирпичном или монолитно-панельном доме без консьержа и паркинга при условии, что строительной организацией не было получено разрешение на строительство составляет 79946.335 рубля.

Так же наличие автостоянки в жилом комплексе не влияет на цену квадратного метра квартиры в нём. Помимо этого, Адмиралтейский и Василеостровский районы оказывают одинаковое влияние на цену квадратного метра квартиры, как оказывают одинаковое влияние на цену квадратного метра квартиры фиктивные переменные "Квартиры-студии", "Четырёхкомнатной квартиры" и "Пятикомнатной квартиры". Стоит также отметить, что цены квадратного метра квартиры в кирпичном, монолитно-газобетонном и монолитно-панельном доме равны.

Интерпретация следующих коэффициентов дает следующие результаты:

1. Если квартира располагается в Выборгском районе, это уменьшает ее цену на 27.241201%;
2. Если квартира располагается в Калининском районе, это уменьшает ее цену на 17.830232%;
3. Если квартира располагается в Кировском районе, это уменьшает ее цену на 18.134571%;
4. Если квартира располагается в Колпинском районе, это уменьшает ее цену на 33.049267%;
5. Если квартира располагается в Красногвардейском районе, это уменьшает ее цену на 16.065341%;
6. Если квартира располагается в Красносельском районе, это уменьшает ее цену на 20.361946%;
7. Если квартира располагается в Кронштадтском районе, это уменьшает ее цену на 20.243677%;
8. Если квартира располагается в Курортном районе, это уменьшает ее цену на 28.207885%;
9. Если квартира располагается в Московском районе, это уменьшает ее цену на 11.492249%;
10. Если квартира располагается в Невском районе, это уменьшает ее цену на 18.598603%;

11. Если квартира располагается в Петроградском районе, это увеличивает ее цену на 13.022766%;
12. Если квартира располагается в Петродворцовом районе, это уменьшает ее цену на 27.193393%;
13. Если квартира располагается в Приморском районе, это уменьшает ее цену на 12.184503%;
14. Если квартира располагается в Пушкинском районе, это уменьшает ее цену на 24.082118%;
15. Если квартира располагается во Фрунзенском районе, это уменьшает ее цену на 8.8385209%;
16. Если квартира располагается в Центральном районе, это увеличивает ее цену на 12.193124%;
17. Если квартира располагается во Всеволожском районе Ленинградской области (ЛО), это увеличивает ее цену на 30.205114%;
18. Если квартира располагается в Выборгском районе ЛО, это увеличивает ее цену на 45.020216%;
19. Если квартира располагается в Гатчинском районе ЛО, это увеличивает ее цену на 36.956698%;
20. Если квартира располагается в Кингисепском районе ЛО, это увеличивает ее цену на 49.758412%;
21. Если квартира располагается в Кировском районе ЛО, это увеличивает ее цену на 44.395443%;
22. Если квартира располагается в Ломоносовском районе ЛО, это увеличивает ее цену на 42.829903%;
23. Если квартира располагается в Приозерском районе ЛО, это увеличивает ее цену на 56.325562%;
24. Если квартира располагается в Тосненском районе ЛО, это увеличивает ее цену на 51.36411%;

25. Квадратный метр однокомнатной, двухкомнатной и трехкомнатной квартиры на 5.2141433%, 11.164215% и 16.343758% дешевле квадратного метра квартиры-студии, четырех- или пятикомнатной квартир;

26. Если квартира находится в кирпично-монолитном, монолит-газобетонном, монолитном, монолитно-каркасном или панельном доме, то это уменьшает цену квадратного метра на 7.0732467%, 30.748648%, 7.0885892%, 19.142316% и 13.974291% соответственно;

27. Факт получения застройщиком разрешения на строительство увеличивает цену квадратного метра на 30.909542%;

28. Если квартира находится не в ближайшей очереди сдачи, это увеличивает цену одного квадратного метра квартиры на 3.4398618%.

Согласно информационным критериям данная модель лучше предыдущей, несмотря на то, что она демонстрирует более низкое значение R^2 .

ЗАКЛЮЧЕНИЕ

В данной работе был представлен гедонистический анализ на рынке первичного жилья. Работа посвящена как теоретическому изучению данного вопроса, касающемуся особенностей рынка жилья, построения гедонистической ценовой функции, анализа спроса, так и эмпирическому исследованию данной темы.

В ходе гедонистического изучения цен на рынке первичного жилья в Санкт-Петербурге были построены базовые модели, тестирующие влияние отдельных переменных на цену квартиры, либо одного квадратного метра квартиры, а также построены полные гедонистические модели для обоих видов зависимой переменной.

Результатом работы было представление цены как функции отдельных характеристик квартиры.

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

- 1) Bartik T. J. The estimation of demand parameters in hedonic price models //The Journal of Political Economy. – 1987. – С. 81-88.
- 2) Blomquist N. S. Comparative statics for utility maximization models with nonlinear budget constraints //International Economic Review. – 1989. – С. 275-296.
- 3) Case K. E., Shiller R. J. Prices of single family homes since 1970: New indexes for four cities. – 1987.
- 4) Cebula R. J. The hedonic pricing model applied to the housing market of the City of Savannah and its Savannah Historic Landmark District //The Review of Regional Studies. – 2009. – Т. 39. – №. 1. – С. 9-22.
- 5) Day B., Bateman I., Lake I. Nonlinearity in hedonic price equation: An estimation strategy using model-based clustering. – CSERGE Working Paper EDM, 2004. – №. 04-02.
- 6) Epple D. Hedonic prices and implicit markets: estimating demand and supply functions for differentiated products //The Journal of Political Economy. – 1987. – С. 59-80.
- 7) Goodman A. C. Hedonic prices, price indices and housing markets //Journal of Urban Economics. – 1978. – Т. 5. – №. 4. – С. 471-484.
- 8) Harrison Jr D., Rubinfeld D. L. Hedonic housing prices and the demand for clean air //Journal of environmental economics and management. – 1978. – Т. 5. – №. 1. – С. 81-102.
- 9) Hill R. Hedonic price indexes for housing. – OECD Publishing, 2011. – №. 2011/1.
- 10) Magnus J., Peresetsky A. The price of Moscow apartments. – Litres, 2014.

- 11) Mayo S. K. Theory and estimation in the economics of housing demand //Journal of Urban Economics. – 1981. – Т. 10. – №. 1. – С. 95-116.
- 12) Nagaraja C. H., Brown L. D., Wachter S. M. Repeat sales house price index methodology //J. Real Estate Lit. Forthcoming. – 2014.
- 13) Ohsfeldt R. L., Smith B. A. Estimating the demand for heterogeneous goods //The Review of Economics and Statistics. – 1985. – С. 165-171.
- 14) Quigley J. M. Nonlinear budget constraints and consumer demand: An application to public programs for residential housing //Journal of Urban Economics. – 1982. – Т. 12. – №. 2. – С. 177-201.
- 15) Rosen S. Hedonic prices and implicit markets: product differentiation in pure competition //The journal of political economy. – 1974. – С. 34-55.
- 16) Sheppard S. Hedonic analysis of housing markets //Handbook of regional and urban economics. – 1999. – Т. 3. – С. 1595-1635.
- 17) Sunding D. L., Swoboda A. M. Hedonic analysis with locally weighted regression: An application to the shadow cost of housing regulation in Southern California //Regional Science and Urban Economics. – 2010. – Т. 40. – №. 6. – С. 550-573.
- 18) Wallace N. E. Hedonic-based price indexes for housing: theory, estimation, and index construction //Economic Review. – 1996. – С. 34-48.
- 19) Wallace N. E., Meese R. A. The construction of residential housing price indices: a comparison of repeat-sales, hedonic-regression, and hybrid approaches //The Journal of Real Estate Finance and Economics. – 1997. – Т. 14. – №. 1-2. – С. 51-73.
- 20) Баталева А.В. Структура цен на квартиры: цены и предложения и цены сделок. Новосибирск, Труды НГУ. – вып.22. – 2005. – с.67-69.
- 21) Глущенко К. П., Баталева А. В. Структура цен на вторичном рынке жилья Новосибирска //Регион: экономика и социология. – 1999. – №. 4. – С. 111-129.

22) Матвеев В. Д. Свойства функций полезности, зависящих от характеристик благ // Вестник УрФУ. Серия экономика и управление. – 2012. – №. 2. – С. 16-27.

23) Цыплаков А. Экскурс в мир инструментальных переменных //квантиль. – 2007. – Т. 2. – С. 21-47.

ПРИЛОЖЕНИЕ

Приложение 1. Проверка нормальности распределения цены квадратного метра

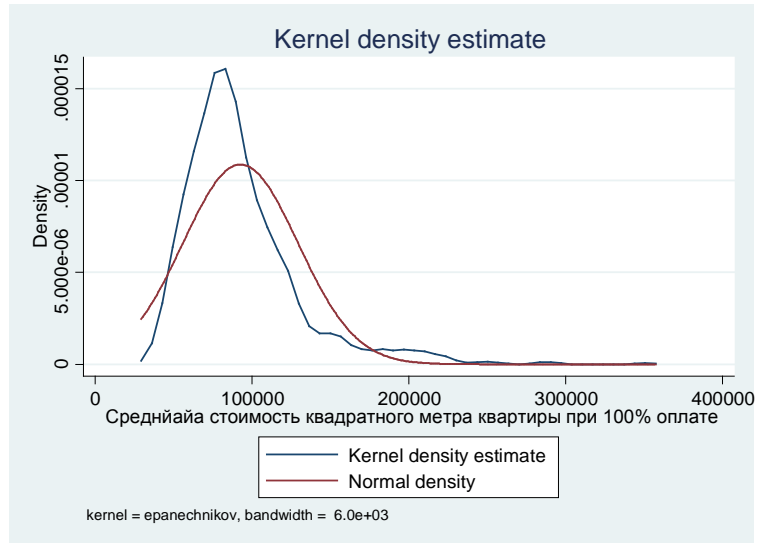
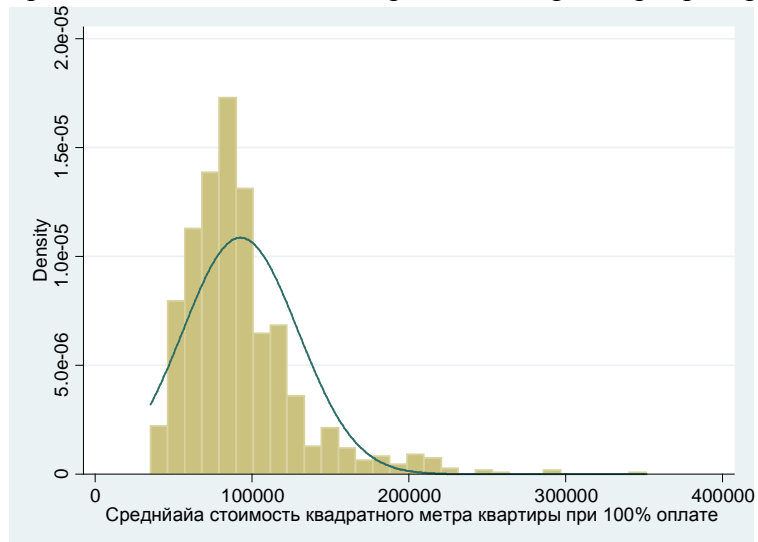


Рис.1.1: График ядерной плотности цены квадратного метра квартиры при 100% оплате



(bin=29, start=35350, width=10901.724)

Рис.1.2: Гистограмма распределения цены квадратного метра квартиры при 100% оплате

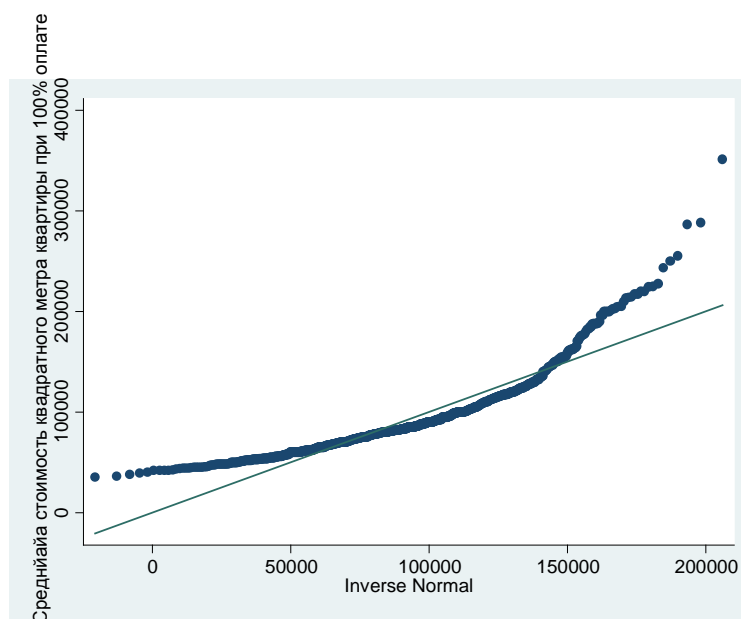


Рис.1.3: График квантилей распределения цены квадратного метра квартиры при 100% оплате

Shapiro-Francia W' test for normal data

Variable	Obs	W'	V'	z	Prob>z
price100sq~e	992	0.85362	97.321	10.490	0.00001

Табл.1.1: Результаты теста Шапиро-Франция на нормальность распределения цены квадратного метра квартиры при 100% оплате

Shapiro-Wilk W test for normal data

Variable	Obs	W	V	z	Prob>z
price100sq~e	992	0.85472	90.957	11.166	0.00000

Табл.1.2: Результаты теста Шапиро-Вилка на нормальность распределения цены квадратного метра квартиры при 100% оплате

Skewness/Kurtosis tests for Normality

Variable	Obs	Pr(Skewness)	Pr(Kurtosis)	adj chi2(2)	joint Prob>chi2
price100sq~e	992	0.0000	0.0000	.	0.0000

Табл.1.3: Результаты теста на нормальность распределения цены квадратного метра квартиры при 100% оплате, основанного на значениях показателей асимметрии и эксцесса

Приложение 2. Проверка нормальности распределения логарифма цены квадратного метра

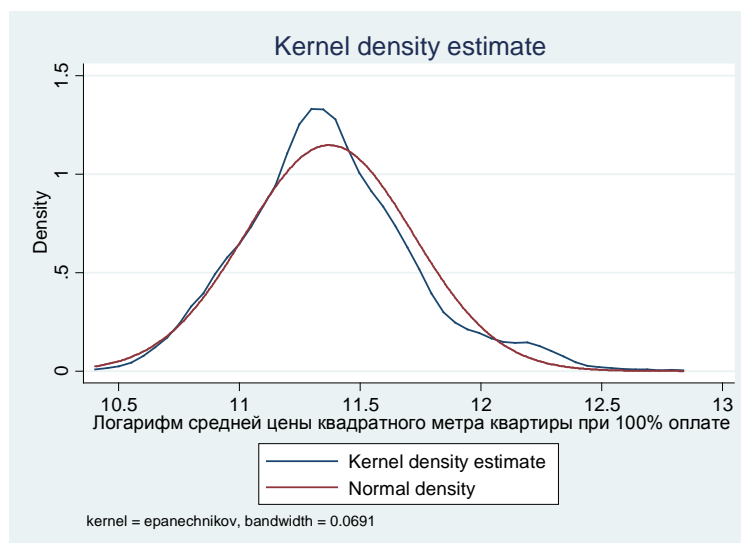
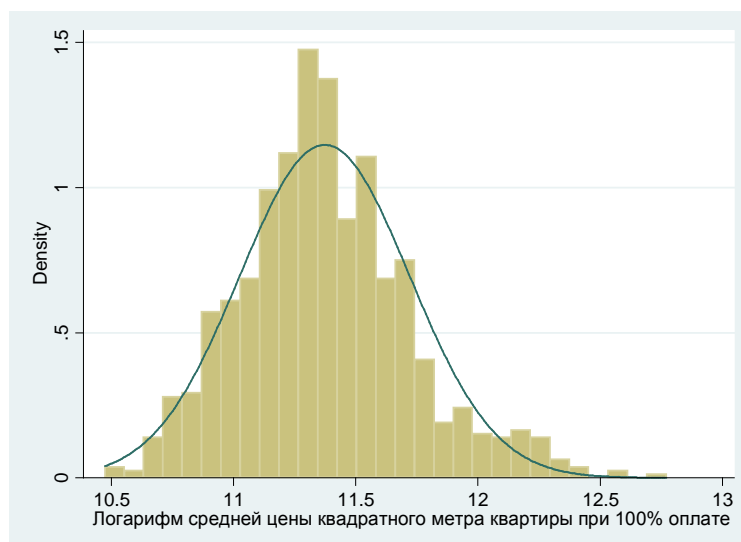


Рис.2.1: График ядерной плотности логарифма цены квадратного метра квартиры при 100% оплате



(bin=29, start=10.473054, width=.07920384)

Рис.2.2: Гистограмма распределения логарифма цены квадратного метра квартиры при 100% оплате

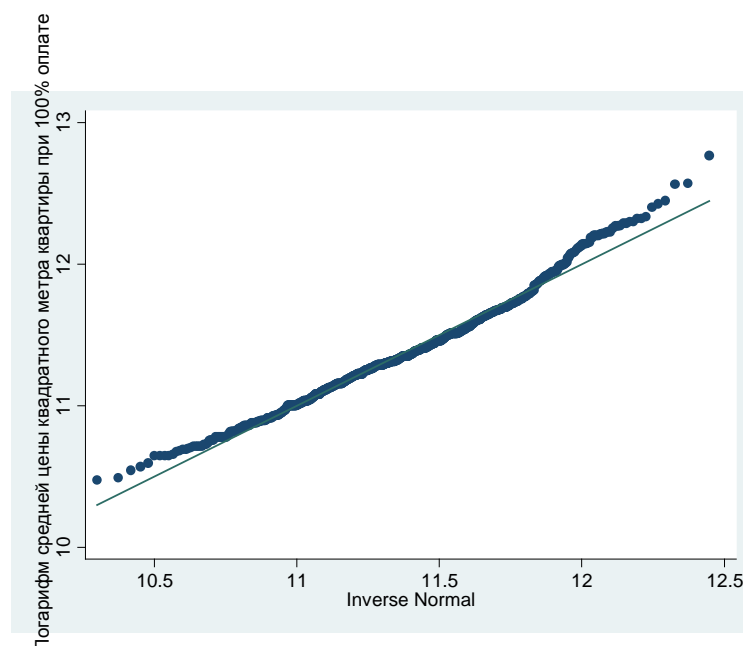


Рис. 2.3: График квантилей распределения логарифма цены квадратного метра квартиры при 100% оплате

Shapiro-Francia W' test for normal data

Variable	Obs	W'	V'	z	Prob>z
ln_price10~e	992	0.98390	10.705	5.432	0.00001

Табл. 2.1: Результаты теста Шапиро-Франция на нормальность распределения логарифма цены квадратного метра квартиры при 100% оплате

Shapiro-Wilk W test for normal data

Variable	Obs	W	V	z	Prob>z
ln_price10~e	992	0.98390	10.080	5.720	0.00000

Табл. 2.2: Результаты теста Шапиро-Вилка на нормальность распределения логарифма цены квадратного метра квартиры при 100% оплате

Skewness/Kurtosis tests for Normality

Variable	Obs	Pr(Skewness)	Pr(Kurtosis)	adj chi2(2)	joint Prob>chi2
ln_price10~e	992	0.0000	0.0027	38.02	0.0000

Табл. 2.3: Результаты теста на нормальность распределения логарифма цены квадратного метра квартиры при 100% оплате, основанного на значениях показателей асимметрии и эксцесса

Приложение 3. Проверка гипотез по регрессионной модели №1

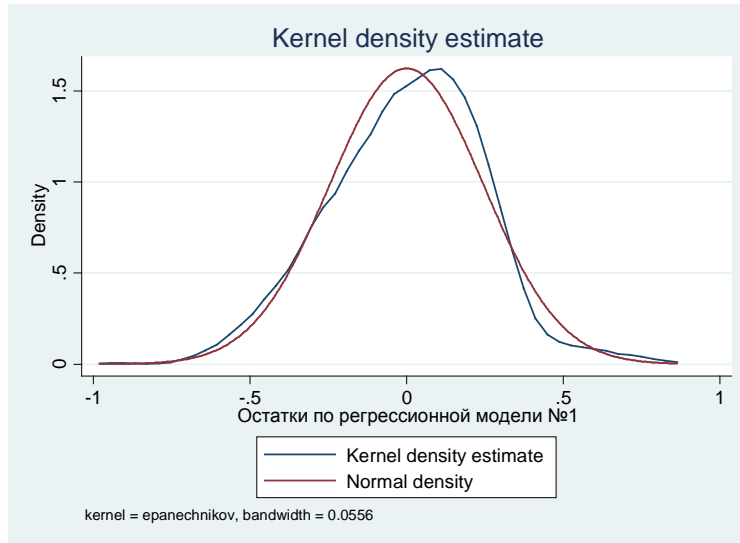
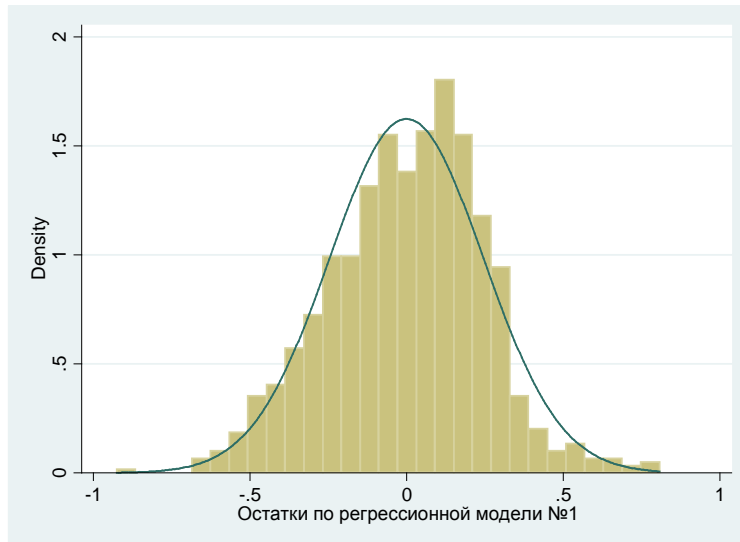


Рис. 3.1: График ядерной плотности остатков регрессионной модели №1



(bin=29, start=-.92498994, width=.05975927)

Рис. 3.2: Гистограмма распределения остатков регрессионной модели №1

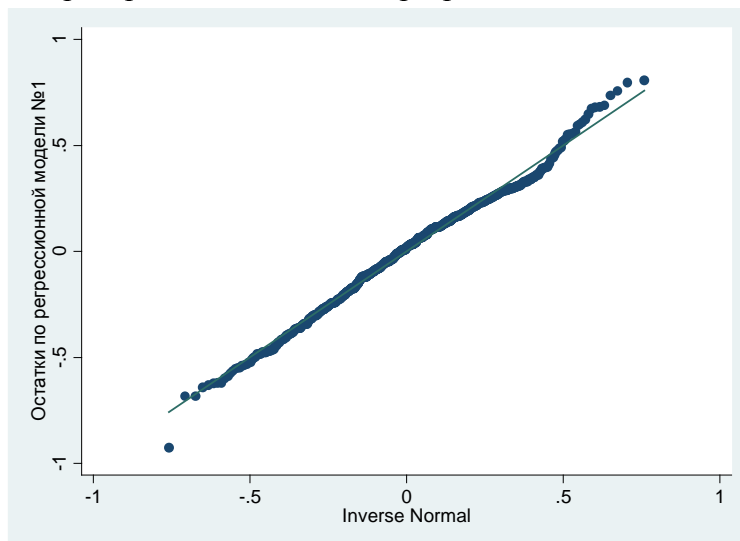


Рис. 3.3: График квантилей распределения остатков регрессионной модели №1
Shapiro-Francia W' test for normal data

Variable	Obs	W'	V'	z	Prob>z
res_model01	992	0.99329	4.461	3.427	0.00031

Табл. Д1.1: Результаты теста Шапиро-Франца на нормальность распределения остатков регрессионной модели №1

Shapiro-Wilk W test for normal data

Variable	Obs	W	V	z	Prob>z
res_model01	992	0.99346	4.095	3.490	0.00024

Табл. Д1.2: Результаты теста Шапиро-Вилка на нормальность распределения остатков регрессионной модели №1

Skewness/Kurtosis tests for Normality

Variable	Obs	Pr(Skewness)	Pr(Kurtosis)	adj chi2(2)	joint Prob>chi2
res_model01	992	0.1348	0.1458	4.34	0.1141

Табл. Д1.3: Результаты теста на нормальность распределения остатков регрессионной модели №1

(1) class2 - class3 = 0

$$F(1, 987) = 0.47$$

$$\text{Prob} > F = 0.4930$$

Табл. 3.5: Результаты теста Фишера на равенство коэффициентов перед фиктивными переменными «Бизнес-класс» и «Комфорт-класс (smart-комплекс)» друг другу

Приложение 4. Проверка гипотез по регрессионной модели №2

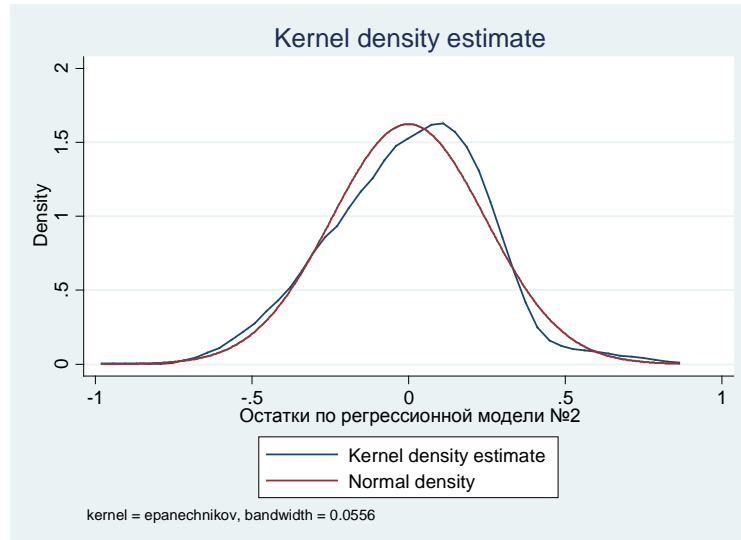
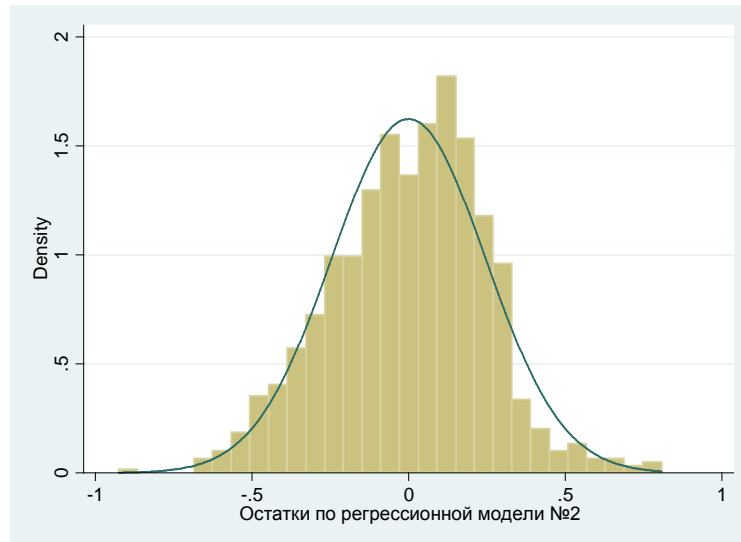


Рис. 4.1: График ядерной плотности остатков регрессионной модели №2



(bin=29, start=-.92498994, width=.05975927)

Рис. 4.2: Гистограмма распределения остатков регрессионной модели №2

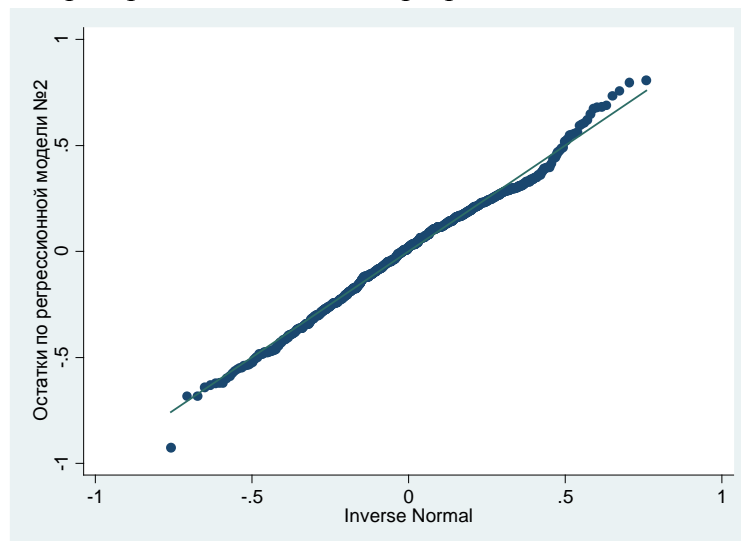


Рис. 4.3: График квантилей распределения остатков регрессионной модели №2
Shapiro-Francia W' test for normal data

Variable	Obs	W'	V'	z	Prob>z
res_model02	992	0.99316	4.547	3.470	0.00026

Табл. 4.1: Результаты теста Шапиро-Франца на нормальность распределения остатков регрессионной модели №2

Shapiro-Wilk W test for normal data

Variable	Obs	W	V	z	Prob>z
res_model02	992	0.99333	4.176	3.539	0.00020

Табл. 4.2: Результаты теста Шапиро-Вилка на нормальность распределения остатков регрессионной модели №2

Skewness/Kurtosis tests for Normality

Variable	Obs	Pr(Skewness)	Pr(Kurtosis)	adj chi2(2)	joint Prob>chi2
res_model02	992	0.1211	0.1469	4.50	0.1056

Табл. 4.3: Результаты теста на нормальность распределения остатков регрессионной модели №2

Приложение 5. Проверка гипотез по регрессионной модели №4

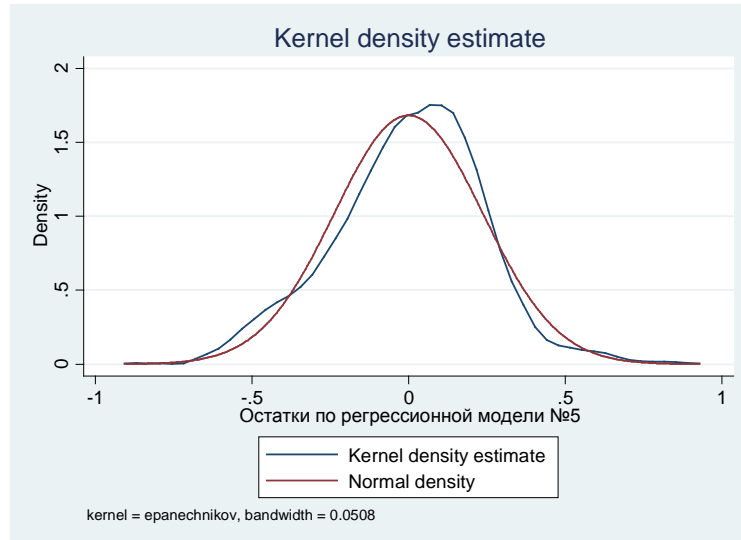
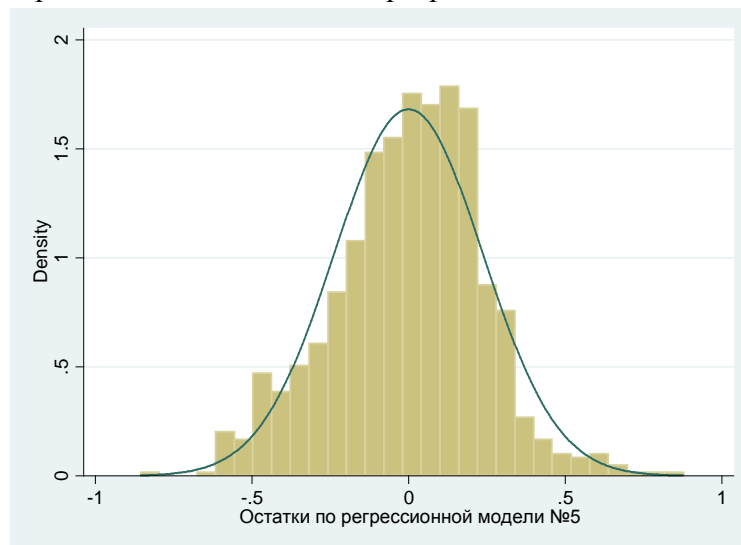


Рис. 5.1: График ядерной плотности остатков регрессионной модели №4



(bin=29, start=-.85527545, width=.05975927)

Рис. 5.2: Гистограмма распределения остатков регрессионной модели №4

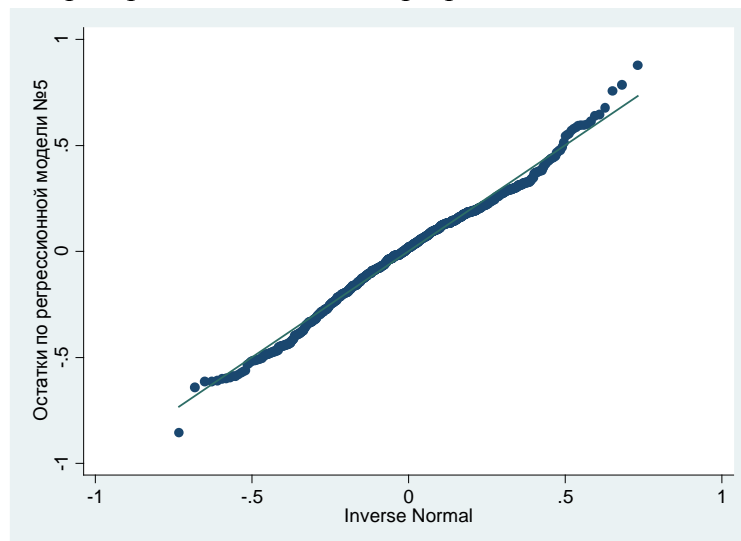


Рис. 5.3: График квантилей распределения остатков регрессионной модели №4
Shapiro-Francia W' test for normal data

Variable	Obs	W'	V'	z	Prob>z
res_model05	992	0.99089	6.060	4.128	0.00002

Табл. 5.1: Результаты теста Шапиро-Франция на нормальность распределения остатков регрессионной модели №4

Shapiro-Wilk W test for normal data

Variable	Obs	W	V	z	Prob>z
res_model05	992	0.99119	5.518	4.228	0.00001

Табл. 5.2: Результаты теста Шапиро-Вилка на нормальность распределения остатков регрессионной модели №4

Skewness/Kurtosis tests for Normality

Variable	Obs	Pr(Skewness)	Pr(Kurtosis)	adj chi2(2)	joint Prob>chi2
res_model05	992	0.0415	0.0361	8.36	0.0153

Табл. 5.3: Результаты теста на нормальность распределения остатков регрессионной модели №4

```

Bootstrap replications (500)
-----+----- 1 -----+----- 2 -----+----- 3 -----+----- 4 -----+----- 5
xxxx.xxxx.x..x.x..x...xxxxx.xxx...x.xxxx...x..x.xx 50
x..xxxx.xxx.xxxx..xx.xx.xx..xx.x..xxx.x..xxxxx...x 100
xxxxx.xxxxxxx.x.x..xx..xxxxxx..xxx...x.x..xx.xxxxx 150
xxx.x..xx.xx.xxxx..x.x..xxx.xx...xxxxx..xx..xxxxxx 200
..xx.xx..xxxxx.xx..x.xxx.x.xx..x...xxx.xxxxxxx..x. 250
..xxxxxxx.xx.xxx.xxxx..xx.xx.xxxxx.xx...x.xx.xx.xx 300
x.xx..xx..xxxxxxx..xxxx.xxx.xxx.xxxxxxxxxxxx..xx..x. 350
..xxx..x...x.xxxxxx...x..x.xxxxxx.x.xxxxx.x.x.x..x 400
..xxxxxxxxxxxxx.x.xxx.xx.xxxxxxxxxx..xx.x..xxxxxxx..x.x 450
x..xxx.xxxxxxxxxxxx.x..xxxxxxxxx.xxxxxxxx.....xxxxxxx 500

```

```

Linear regression                               Number of obs   =       992
Replications                                   =           174
Wald chi2(20)                                  =           .
Prob > chi2                                     =           .
R-squared                                       =       0.5346
Adj R-squared                                  =       0.5246
Root MSE                                       =       0.2397

```

ln_pricel0~e	Observed Coef.	Bootstrap Std. Err.	z	P> z	Normal-based [95% Conf. Interval]	
f10_gr0_1	.2579529	.0356943	7.23	0.000	.1879934	.3279124
f10_gr0_2	.3584136	.0999682	3.59	0.000	.1624796	.5543476
f10_gr0_3	.7040729	.0552881	12.73	0.000	.5957101	.8124356
f11_gr0_0	-.0547748	.0357275	-1.53	0.125	-.1247993	.0152498
f11_gr0_1	.2211526	.0355042	6.23	0.000	.1515657	.2907396
f11_gr0_2	.4252403	.0831085	5.12	0.000	.2623505	.58813
f11_gr0_3	.7570816	.0700683	10.80	0.000	.6197502	.894413
f12_gr0_0	-.1288599	.0324452	-3.97	0.000	-.1924513	-.0652686
f12_gr0_1	.148231	.0359529	4.12	0.000	.0777647	.2186974
f12_gr0_2	.4287869	.0661857	6.48	0.000	.2990654	.5585085
f12_gr0_3	.7009124	.0596572	11.75	0.000	.5839864	.8178384
f13_gr0_0	-.1983737	.0347243	-5.71	0.000	-.2664321	-.1303153
f13_gr0_1	.108829	.036604	2.97	0.003	.0370865	.1805715
f13_gr0_2	.4785546	.0731302	6.54	0.000	.3352221	.6218872
f13_gr0_3	.7321022	.058945	12.42	0.000	.6165721	.8476323
f14_gr0_0	-.1244949	.1175273	-1.06	0.289	-.3548441	.1058544
f14_gr0_1	.2190273	.0943004	2.32	0.020	.0342019	.4038526

f14_gr0_2		.5729463	.1698942	3.37	0.001	.2399598	.9059328
f14_gr0_3		1.006705	.1076398	9.35	0.000	.7957348	1.217675
f15_gr0_2		.7969238	.0294903	27.02	0.000	.7391238	.8547238
f15_gr0_3		.7634288	.0294903	25.89	0.000	.7056289	.8212288
_cons		11.2195	.0294903	380.45	0.000	11.1617	11.2773

Note: one or more parameters could not be estimated in 326 bootstrap replicates; standard-error estimates include only complete replications.

Табл. 5.4: Результаты бутстрапирования регрессионной модели №4

Приложение 6. Матрица расстояний между средними ценами квадратного метра квартиры в зависимости от её типа и класса

Тип квартиры	Группа	Цена	Квартира-студия				Однокомнатная квартира				Двухкомнатная квартира			
			0	1	2	3	0	1	2	3	0	1	2	3
		Цена	76147	98881	115707	127283	72165	95514	120847	150356	66919	89127	120087	147487
Квартира-студия	0	76147	0	22735	39560	51137	3982	19367	44700	74209	9228	12980	43940	71340
	1	98881		0	16826	28402	26716	3368	21966	51475	31963	9755	21206	48606
	2	115707			0	11576	43542	20194	5140	34649	48788	26580	4380	31780
	3	127283				0	55118	31770	6436	23073	60365	38157	7196	20204
Однокомнатная квартира	0	72165					0	23349	48682	78191	5246	16962	47922	75322
	1	95514						0	25333	54842	28595	6387	24573	51973
	2	120847							0	29509	53928	31720	760	26640
	3	150356								0	83437	61229	30269	2869
Двухкомнатная квартира	0	66919									0	22208	53168	80568
	1	89127										0	30960	58360
	2	120087											0	27400
	3	147487												0
...

Табл. 6.1: Таблица расстояний между средними ценами квадратного метра квартиры в зависимости от её типа и группы

Приложение 7. Сравнение моделей №5 и №6 с объединенными районами и без объединенных районов

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
with_di~0215	992	-935.7702	32.93883	27	-11.87766	120.4149
without~0215	992	-935.7702	30.55034	25	-11.10067	111.3924

Note: N=Obs used in calculating BIC; see [R] BIC note

Приложение 8. Проверка гипотезы о нормальности распределения остатков в модели №9

Shapiro-Francia W' test for normal data

Variable	Obs	W'	V'	z	Prob>z
residuals	992	0.99731	1.791	1.336	0.09085

. swilk residuals

Shapiro-Wilk W test for normal data

Variable	Obs	W	V	z	Prob>z
residuals	992	0.99749	1.569	1.116	0.13225

. sktest residuals

Skewness/Kurtosis tests for Normality

Variable	Obs	Pr(Skewness)	Pr(Kurtosis)	adj chi2(2)	joint Prob>chi2
residuals	992	0.2617	0.0586	4.82	0.0897

Приложение 9. Проверка гипотезы о нормальности распределения остатков в модели №11

. sfrancia residuals

Shapiro-Francia W' test for normal data

Variable	Obs	W'	V'	z	Prob>z
residuals	992	0.99654	2.301	1.909	0.02812

. swilk residuals

Shapiro-Wilk W test for normal data

Variable	Obs	W	V	z	Prob>z
residuals	992	0.99683	1.986	1.698	0.04475

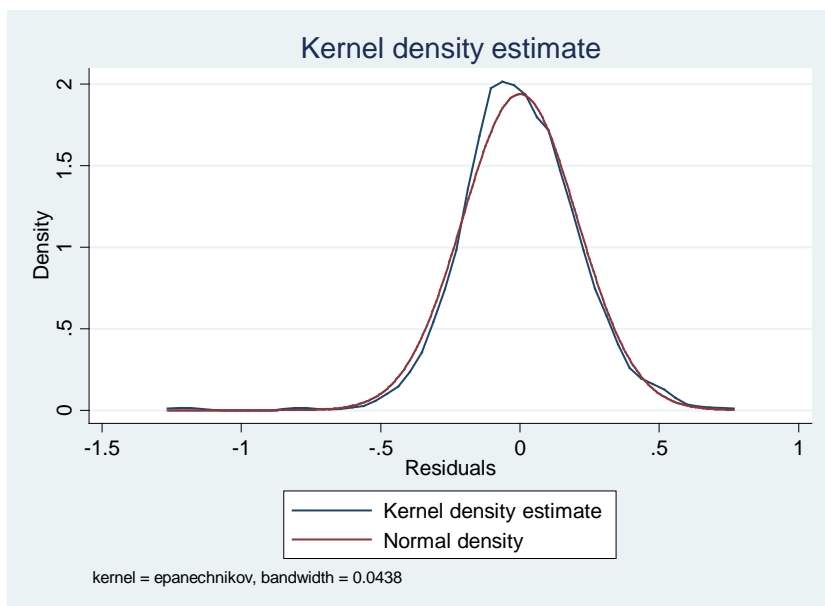
. sktest residuals

Skewness/Kurtosis tests for Normality

Variable	Obs	Pr(Skewness)	Pr(Kurtosis)	adj chi2(2)	joint Prob>chi2
residuals	992	0.0857	0.0427	7.04	0.0297

	Регион, в котором располагается квартира																									
	100.0000%																									
	10.6160%																									
	16.5116%																									
Sum Shapley values	Район города Санкт-Петербурга, в котором располагается квартира												Район Ленинградской области, в котором располагается квартира													
Shapley value	100.0003%												99.9997%													
Shapley value adjusted	66.3410%												33.6590%													
Shapley value total	66.3410%												33.6590%													
	10.9539%												5.5576%													
	Адмиралтейский район города Санкт-Петербурга	Василеостровский район города Санкт-Петербурга	Выборгский район города Санкт-Петербурга	Калининский район города Санкт-Петербурга	Кировский район города Санкт-Петербурга	Колпинский район города Санкт-Петербурга	Красногвардейский район города Санкт-Петербурга	Красносельский район города Санкт-Петербурга	Кронштадтский район города Санкт-Петербурга	Курортный район города Санкт-Петербурга	Московский район города Санкт-Петербурга	Невский район города Санкт-Петербурга	Петроградский район города Санкт-Петербурга	Петродворцовый район города Санкт-Петербурга	Приморский район города Санкт-Петербурга	Пушкинский район города Санкт-Петербурга	Фрунзенский район города Санкт-Петербурга	Центральный район города Санкт-Петербурга	Всеволожский район Ленинградской области	Выборгский район Ленинградской области	Гатчинский район Ленинградской области	Кингисеппский район Ленинградской области	Кировский район Ленинградской области	Ломоносовский район Ленинградской области	Приозерский район Ленинградской области	Тосненский район Ленинградской области
	1.8020%	3.5025%	10.8530%	2.6590%	0.7750%	6.4246%	0.9835%	2.6490%	0.4961%	1.1993%	1.8080%	2.2167%	31.7220%	2.7233%	2.6573%	5.7155%	1.0055%	20.8080%	46.4820%	3.6104%	8.0786%	4.9602%	6.4211%	11.5050%	11.8640%	7.0784%
	0.1974%	0.3837%	1.1888%	0.2913%	0.0849%	0.7037%	0.1077%	0.2902%	0.0543%	0.1314%	0.1980%	0.2428%	3.4748%	0.2983%	0.2911%	0.6261%	0.1101%	2.2793%	2.5833%	0.2007%	0.4490%	0.2757%	0.3569%	0.6394%	0.6594%	0.3934%

Приложение 11. Проверка нормальности распределения остатков в модели №12



```
. sfrancia res_model12
```

Shapiro-Francia W' test for normal data

Variable	Obs	W'	V'	z	Prob>z
res_model12	992	0.97684	15.400	6.266	0.00001

```
. swilk res_model12
```

Shapiro-Wilk W test for normal data

Variable	Obs	W	V	z	Prob>z
res_model12	992	0.97834	13.561	6.454	0.00000

```
. sktest res_model12
```

Skewness/Kurtosis tests for Normality

Variable	Obs	Pr(Skewness)	Pr(Kurtosis)	adj chi2(2)
res_model12	992	0.0011	0.0000	59.89

joint Prob>chi2 0.0000