

И.Н. Анисимова, канд. физ.-мат. наук;
Н.П. Баринов, канд. техн. наук;
С.В. Грибовский, д-р экон. наук,
г. Санкт-Петербург

О требованиях к числу сопоставимых объектов при оценке недвижимости сравнительным подходом

С развитием отечественного рынка недвижимости все большее значение приобретают методы сравнительного подхода к оценке типовых объектов, каждый из которых можно рассматривать как сопоставимый с однородной группой объектов, представленных на рынке. В оценке недвижимости уже не является новинкой моделирование рынка методами многомерного корреляционно-регрессионного анализа. Примеры использования этого мощного математического аппарата при оценке недвижимости вошли в учебники по эконометрике [1, 2] и оценке [3], им посвящены публикации в периодических изданиях [4, 5, 6], появилась и первая монография на русском языке [7]. Наконец, более пяти лет в Санкт-Петербурге успешно применяется методика определения арендной платы за нежилые помещения, построенная по результатам статистического моделирования городского рынка арендных ставок методами множественной регрессии.

Тем не менее, при решении практических задач индивидуальной оценки объектов недвижимости хорошо разработанные теоретически, легко реализуемые на современных компьютерах и достаточно объективные модели множественной регрессии используются лишь узким кругом российских оценщиков. На наш взгляд, это связано не столько с молодостью оценочной деятельности в современной России и относительной неразвитостью рынка, сколько с отсутствием каких-либо обоснованных требований к объему выборки рыночных данных, необходимому для решения задач оценки методами регрессионного анализа. Или, другими словами, остается неясным, сколько аналогов (сопоставимых объектов) необходимо использовать при расчете рыночной стоимости объекта недвижимости методом сравнительного анализа рыночных данных.

Действительно, сегодня для построения надежной эконометрической модели рекомендуют [1, 7, 10] использовать объем выборки рыночных данных, равный 5–7-кратному числу используемых моделью независимых факторных переменных (ценообразующих факторов). Применительно к наиболее распространенным значениям числа основных влияющих факторов (4–7), позволяющих строить адекватные модели для отдельных сегментов рынка недвижимости, необходимый объем выборки однородных рыночных данных должен составлять в этом случае порядка 25–50 сделок или предложений к ним. Опыт показывает, что это слишком жесткое требование для пас-

сивных рынков российских поселений, за исключением, может быть, таких городов, как Москва, Петербург, Екатеринбург, Нижний Новгород и некоторые другие.

Но можно ли рассматривать данную рекомендацию и вытекающие из нее требования к объему выборки как *минимально достаточные* для решения практической задачи индивидуальной оценки объекта недвижимости? Мы полагаем, что нет, так как реально требуемый объем выборки для решения данной задачи может быть существенно меньше. Покажем это.

Требование 5–7-кратного превышения объемом выборки числа факторных переменных призвано обеспечить значимость не только регрессионного уравнения в целом, но также и каждого из коэффициентов регрессии. Практическая ценность значимости коэффициентов регрессии заключается в выявлении характера влияния каждого из факторов на исследуемую величину (результатирующий признак). Вскрытие закономерностей, т.е. характера взаимосвязей между результирующим признаком и влияющими факторами в классической эконометрической постановке задачи статистического моделирования зачастую является главной целью исследования, так как призвано обеспечить возможность формирования количественно обоснованного управляющего воздействия на тот или иной влияющий фактор для придания результирующему признаку заданных значений.

На наш взгляд, выявление тонкой структуры указанных связей применительно к задаче индивидуальной оценки объектов недвижимости является избыточным. Действительно, оценщик, проводя расчет рыночной стоимости (или ставки аренды) конкретного здания, помещения или земельного участка, не решает задачи иной, чем количественное определение *суммарного результирующего* влияния основных ценообразующих факторов на формирование цены (арендной ставки) данного объекта. При этом на практике почти всегда наблюдается частичная мультиколлинеарность¹ влияющих факторов, которая порой сводит к минимуму значимость влияния каждого из факторов на выходную величину. Напомним, что наличие частичной мультиколлинеарности факторных переменных при выполнении остальных предпосылок линейной регрессии не смещает оценку среднего, а лишь ухудшает качество интервальной оценки (см., например, [1, 2, 8]).

Таким образом, для решения задачи индивидуальной оценки объекта недвижимости сравнительным подходом достаточным, на наш взгляд, условием адекватности полученной модели множественной регрессии может являться *статистическая значимость уравнения в целом* (при этом предполагается выполнение условий, обеспечивающих получение несмещенных оценок, обсуждение которых мы откладываем до следующих публикаций).

Статистическая значимость линейного регрессионного уравнения в целом проверяется с помощью статистики Фишера, критические значения $F_{кр}$, которой зависят от числа используемых факторов-регрессоров k , объема выборки n , а также заданного уровня значимости. Процедура сводится к проверке нуль-гипотезы о том, что ни один из факторов, включенных в модель, не оказывает существенного влияния на результирующий показатель. Если расчетное значение F -критерия превышает критическое, то гипотеза о несоответствии используемых в модели связей реально существующим отвергают и считают, что уравнение в целом статистически значимо и хорошо соответствует рыночным данным.

Расчетное значение F -критерия определяется как

$$F_p = R^2(n - k - 1)/k(1 - R^2), \quad (1)$$

где R^2 — коэффициент детерминации (квадрат коэффициента множественной корреляции):

$$R^2 = \frac{Q_R}{Q_R + Q_{ост}}; \quad Q_{ост} = \sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n (y_i - \tilde{y}_i)^2; \quad Q_R = \sum_{i=1}^n (\tilde{y}_i - \bar{y})^2.$$

Здесь y_i — наблюдаемые, а \tilde{y}_i — прогнозируемые моделью значения результирующей переменной.

Критические значения F -критерия определяются по таблицам (публикуемым в литературе или электронным) при заданном уровне значимости (в экономических задачах широко используется значение 0,05) и числе степеней свободы $x_1 = k$, $x_2 = n - k - 1$.

Обратим внимание на связь расчетного значения F -критерия со значением коэффициента детерминации R^2 , который, в свою очередь, часто используется для оценки качества применяемой модели², так как характеризует долю дисперсии, объясняемую моделью, в общей дисперсии результирующего признака. В практических задачах оценка качества модели по показателю R^2 счи-

¹ В классической модели линейной регрессии, когда факторные переменные считаются детерминированными величинами, корректнее говорить об их мультиколлинеарности.

тается весьма высокой, если его значения достигают 0,9 и выше, и вполне достаточной при значениях $\geq 0,7$. Из выражения (1) для F -критерия легко заметить, что его расчетное значение растет с ростом объема выборки и коэффициента детерминации. Однако критическое значение критерия при заданном уровне значимости и числе используемых моделью влияющих факторов падает с ростом объема выборки. Таким образом, в предположении, что будет обеспечено требуемое качество модели, отражаемое значением коэффициента детерминации, всегда можно указать минимальный объем выборки, начиная с которого расчетное значение критерия будет превышать критическое.

Интуитивно понятно, что объем выборки и “качество” модели являются обменными величинами. Так, например, если бы перед оценщиком стояла задача определения рыночной стоимости единицы площади квартиры без отделки в строящемся с привлечением средств инвесторов типовом многоквартирном доме, то, скорее всего, ему достаточно было бы знать цену этой площади, указанную в оферте девелопера. А наличие трех (вполне возможно, очень близких) значений из заключенных договоров инвестирования строительства не вызывало бы никаких вопросов, кроме, быть может, вопроса о скидках на цену квартир последнего этажа. В данном случае все объекты можно считать практически одинаковыми по уровням основных влияющих факторов, и адекватной моделью стоимости является, скорее всего, среднее по выборке цен значение.

При использовании объектов-аналогов, отличающихся от оцениваемого объекта уровнями по нескольким основным ценообразующим факторам, “качество” модели зависит как от степени этих отличий, так и от спецификации модели, т.е. адекватности моделируемого характера зависимости результирующего показателя (стоимости, ставки аренды) от факторов-регрессоров. В отличие от задач массовой оценки, первая составляющая — диапазон отклонений значений влияющих факторов объектов-аналогов от соответствующих значений оцениваемого объекта — является, в определенной степени, управляемой величиной. Другими словами, оценщик формирует контролируемую выборку данных, следя за тем, чтобы в нее попали наиболее близкие к объекту по значениям основных ценообразующих факторов (но не по стоимости!) аналоги. Здесь нет необходимости останавливаться на проблемах такого отбора и методах их решения, так как одни частично освещены (см., например [9]), другие требуют более подробного рассмотрения, которое мы отложим на будущее. Отметим только, что при одном и том же объеме выборки ее однородность, плотность группирования вокруг оцениваемого объекта может быть различной, что во многом определяет и статистические свойства модели.

Важно отметить, что в результате статистического моделирования должно быть получено значимое уравнение множественной регрессии с достаточно высокими показателями достоверности, чтобы результаты расчета по ней могли претендовать на использование в отчете об оценке. Как уже отмечалось, одним из используемых показателей адекватности уравнения является коэффициент детерминации, нижней границей значений которого при решении задачи индивидуальной оценки может быть принято значение 0,7.

Можно задать следующий вопрос: если предположить, что отобранные аналоги обеспечат после проведения моделирования уровень коэффициента детерминации не менее 0,7, какое число таких аналогов необходимо найти на рынке, чтобы обеспечить статистическую значимость регрессионного уравнения? А если однородность выборки и спецификация модели обеспечит значения коэффициента детерминации 0,8 или 0,9?

Ответы на эти вопросы для классической линейной множественной регрессии могут быть легко получены сравнением расчетного $F_{[R^2]}$ и критического F_{kr} значений F -критерия при заданных уровне значимости (0,05), числе используемых моделью факторов-регрессоров и объеме выборки (см. таблицы расчетов). Для наглядности сравнения результатов объем n выборки выражен через дополнение к числу k используемых в модели факторных переменных.

В затемненных областях таблиц представлены расчетные значения F -критерия, достигающие и превышающие соответствующие критические его значения. Анализ полученных результатов для моделей, использующих число факторных переменных k от 3 до 7 включительно, позволяет сделать следующие обобщения.

Для получения представляющих практический интерес результатов моделирования цены (ставки аренды) оцениваемого объекта недвижимости методами множественной линейной регрессии по выборке рыночных данных о ценах (арендных ставках) его аналогов, необходимо иметь в распоряжении, как минимум $n = 2(k + 2)$ аналогов оцениваемого объекта. Такой объем выборки достаточен при требуемом уровне коэффициента детерминации $R^2 = 0,7$.

² Точнее, его скорректированное значение (см., например, список литературы к статье).

$k=3$

n	$k+2$	$k+3$	$k+4$	$k+5$	$k+6$	$k+7$	$k+8$	$k+9$
$n-k-1$	1	2	3	4	5	6	7	8
F_{kr}	216,0	19,20	9,28	6,59	5,41	4,76	4,35	4,07
R_{kr}^2	0,998	0,966	0,903	0,832	0,764	0,704	0,651	0,604
$F_{[\hat{\rho}=0,70]}$	0,79	1,59	2,38	3,17	3,97	4,76	5,55	6,35
$F_{[\hat{\rho}=0,77]}$	1,08	2,16	3,25	4,33	5,41	6,49	7,57	8,66
$F_{[\hat{\rho}=0,83]}$	1,65	3,30	4,94	6,59	8,24	9,89	11,53	13,18
$F_{[\hat{\rho}=0,90]}$	3,09	6,19	9,28	12,37	15,47	18,56	21,65	24,75
$F_{[\hat{\rho}=0,97]}$	9,60	19,20	28,80	38,40	48,00	57,60	67,20	76,80

 $k=4$

n	$k+2$	$k+3$	$k+4$	$k+5$	$k+6$	$k+7$	$k+8$	$k+9$
$n-k-1$	1	2	3	4	5	6	7	8
F_{kr}	225,0	19,20	9,12	6,39	5,19	4,53	4,12	3,84
R_{kr}^2	0,999	0,975	0,924	0,865	0,806	0,751	0,702	0,658
$F_{[\hat{\rho}=0,70]}$	0,59	1,18	1,77	2,35	2,94	3,53	4,12	4,71
$F_{[\hat{\rho}=0,75]}$	0,76	1,51	2,27	3,02	3,78	4,53	5,29	6,04
$F_{[\hat{\rho}=0,81]}$	1,04	2,08	3,11	4,15	5,19	6,23	7,27	8,30
$F_{[\hat{\rho}=0,87]}$	1,60	3,20	4,79	6,39	7,99	9,58	11,18	12,78
$F_{[\hat{\rho}=0,90]}$	3,04	6,08	9,12	12,16	15,20	18,24	21,28	24,32

 $k=5$

n	$k+2$	$k+3$	$k+4$	$k+5$	$k+6$	$k+7$	$k+8$	$k+9$
$n-k-1$	1	2	3	4	5	6	7	8
F_{kr}	230,0	19,30	9,01	6,26	5,05	4,39	3,97	3,69
R_{kr}^2	0,999	0,980	0,938	0,887	0,835	0,785	0,739	0,698
$F_{[\hat{\rho}=0,70]}$	0,46	0,92	1,38	1,85	2,31	2,77	3,23	3,69
$F_{[\hat{\rho}=0,74]}$	0,57	1,13	1,70	2,27	2,84	3,40	3,97	4,54
$F_{[\hat{\rho}=0,79]}$	0,73	1,46	2,20	2,93	3,66	4,39	5,12	5,85
$F_{[\hat{\rho}=0,84]}$	1,01	2,02	3,03	4,04	5,05	6,06	7,07	8,08
$F_{[\hat{\rho}=0,89]}$	1,57	3,13	4,69	6,26	7,82	9,39	10,96	12,52
$F_{[\hat{\rho}=0,94]}$	3,00	6,01	9,01	12,01	15,02	18,02	21,02	24,03

$k = 6$

n	$k + 3$	$k + 4$	$k + 5$	$k + 6$	$k + 7$	$k + 8$	$k + 9$	$k + 10$
$n - k - 1$	2	3	4	5	6	7	8	9
F_{kr}	19,30	8,94	6,16	4,95	4,28	3,87	3,58	3,37
R^2_{kr}	0,983	0,947	0,902	0,856	0,811	0,768	0,729	0,692
$F_{[\hat{\rho} = 0,70]}$	0,75	1,12	1,50	1,87	2,25	2,62	3,00	3,37
$F_{[\hat{\rho} = 0,73]}$	0,90	1,34	1,79	2,24	2,69	3,13	3,58	4,03
$F_{[\hat{\rho} = 0,77]}$	1,11	1,66	2,21	2,76	3,32	3,87	4,42	4,98
$F_{[\hat{\rho} = 0,81]}$	1,43	2,14	2,85	3,57	4,28	4,99	5,71	6,42
$F_{[\hat{\rho} = 0,86]}$	1,98	2,97	3,96	4,95	5,94	6,93	7,92	8,91
$F_{[\hat{\rho} = 0,90]}$	3,08	4,62	6,16	7,70	9,24	10,78	12,32	13,86
$F_{[\hat{\rho} = 0,95]}$	5,96	8,94	11,92	14,90	17,88	20,86	23,84	26,82

$k = 7$

n	$k + 4$	$k + 5$	$k + 6$	$k + 7$	$k + 8$	$k + 9$	$k + 10$	$k + 11$
$n - k - 1$	3	4	5	6	7	8	9	10
F_{kr}	8,89	6,09	4,88	4,21	3,79	3,50	3,29	3,14
R^2_{kr}	0,954	0,914	0,872	0,831	0,791	0,754	0,719	0,687
$F_{[\hat{\rho} = 0,70]}$	0,94	1,26	1,57	1,88	2,20	2,51	2,83	3,14
$F_{[\hat{\rho} = 0,72]}$	1,10	1,46	1,83	2,19	2,56	2,92	3,29	3,66
$F_{[\hat{\rho} = 0,76]}$	1,31	1,75	2,19	2,63	3,06	3,50	3,94	4,38
$F_{[\hat{\rho} = 0,79]}$	1,62	2,17	2,71	3,25	3,79	4,33	4,87	5,41
$F_{[\hat{\rho} = 0,83]}$	2,11	2,81	3,51	4,21	4,91	5,61	6,32	7,02
$F_{[\hat{\rho} = 0,87]}$	2,93	3,90	4,88	5,86	6,83	7,81	8,78	9,76
$F_{[\hat{\rho} = 0,91]}$	4,57	6,09	7,61	9,14	10,66	12,18	13,70	15,23
$F_{[\hat{\rho} = 0,95]}$	8,89	11,85	14,82	17,78	20,74	23,71	26,67	29,63

Если отличия в характеристиках объекта оценки и отобранных аналогах достаточно малы и модель хорошо специфицирована (коэффициент детерминации $R^2 = 0,8$), требуемый объем выборки может быть оценен значением $n = 2(k + 1)$.

При наличии на рынке очень близких аналогов и верной спецификации модели ($R^2 = 0,9$) для статистической значимости регрессионного уравнения достаточно иметь объем выборки $n = k + 5$.

Таблица сравнения требуемых объемов выборки рыночных данных

Качество модели R^2 , %	Число ценообразующих факторов k					Расчетная формула
	3	4	5	6	7	
70	10	12	14	16	18	$n = 2(k + 2)$
80	8	10	12	14	16	$n = 2(k + 1)$
90	8	9	10	11	12	$n = k + 5$
—	21	28	35	42	49	$n = 7k$

В любом из этих случаев требуемый (минимально достаточный) объем выборки существенно ниже, чем определенный выражением $n = (5 - 7)k$ — см. таблицу.

Опыт практического моделирования сегментов рынка коммерческой недвижимости в Санкт-Петербурге при решении задач индивидуальной оценки показывает, что существующее состояние рынка позволяет формировать выборки из объектов-аналогов, обеспечивающие получение коэффициента детерминации не ниже 0,8. Это означает, что для наиболее часто применяемых моделей с числом факторов 4–5 статистическая значимость уравнения обеспечивается уже при объеме выборки, содержащем 10–12 аналогов соответственно. При удовлетворении существующих рекомендаций потребовалось бы формировать выборку из 25–30 аналогов, что для большинства оцениваемых объектов пока невыполнимо. Отметим попутно, что с повышением однородности выборки (близости аналогов к объекту оценки) растет и адекватность применения собственно линейной модели регрессии.

Полученные соотношения, определяющие минимально достаточный объем выборки, не носят абсолютного характера, так как базируются на классической параметрической модели линейной регрессии с ее основными предпосылками. Тем не менее, эти оценки подтверждаются практикой моделирования и существенно расширяют область применения методов аппарата множественной регрессии при решении задач индивидуальной оценки объектов недвижимости на отечественном рынке.

ВЫВОДЫ

1. Существующие рекомендации по определению объема выборки рыночных данных, необходимого для эконометрического моделирования и равного 5–7-кратному числу используемых в модели факторных признаков (ценообразующих факторов), применительно к задачам индивидуальной оценки объектов недвижимости являются в значительной степени избыточными.

2. Количество аналогов n , минимально необходимых для формирования адекватных моделей множественной линейной регрессии при индивидуальной оценке недвижимости, может быть определено значительно (в 2–3 раза) меньшим числом, связанным с числом используемых факторов k как $n = 2(k + 2)$ или даже $n = 2(k + 1)$.

3. При современном уровне развития отечественных региональных рынков недвижимости существенное снижение требуемого объема выборки рыночных данных расширяет возможности решения задач индивидуальной оценки недвижимости методами регрессионного анализа, в значительной степени лишенными недостатков субъективной оценки, присущих альтернативным методам экспертного оценивания.

Литература

1. Эконометрика: Учеб. / Под ред. И.И. Елисеевой. М.: Финансы и статистика, 2001.
2. Магнус Я.Р., Катышев П.К., Пересецкий А.А. Эконометрика. Начальный курс: Учеб. 5-е изд., испр. М.: Дело, 2001.
3. Грибовский С.В. Оценка доходной недвижимости. СПб: Питер, 2001.
4. Сивец С.А. Построение и практическое применение многофакторной гибридной модели оценки доходной недвижимости // Вопросы оценки. 2001. №4.
5. Грибовский С.В., Сивец С.А., Левыкина И.А. Новые возможности сравнительного подхода при решении старых проблем // Вопросы оценки. 2002. №4.
6. Трифонов Н.Ю., Шимановский С.А. Эконометрическая модель рынка квартир // Вопросы оценки. 2002. №4.
7. Сивец С.А., Левыкина И.А. Эконометрическое моделирование в оценке недвижимости: Учебно-практическое пособие для оценщиков. Запорожье: Полиграф, 2003.
8. Доугерти К. Введение в эконометрику: Пер. с англ. М.: ИНФРА-М, 2001. (Сер. “Университетский учебник”).
9. Грибовский С.В., Баринов Н.П., Анисимова И.Н. О повышении достоверности оценки рыночной стоимости методом сравнительного анализа // Вопросы оценки. 2002. №1.
10. Теория статистики / Под ред. Р.А. Шмойловой. М.: Финансы и статистика, 1998.