

ЭКОНОМЕТРИЧЕСКОЕ МОДЕЛИРОВАНИЕ СЕЗОННОСТИ НА ОСНОВЕ РЕГРЕССИОННОЙ МОДЕЛИ С ФИКТИВНЫМИ ПЕРЕМЕННЫМИ

Цыпина Ю.С., Цыпин А.П.

**Оренбургский филиал ФГБОУ ВПО «РЭУ им. Г.В. Плеханова»,
ФГБОУ ВПО «Оренбургский государственный университет»,
г. Оренбург**

В статье рассматривается проблема построения эффективных эконометрических моделей учитывающих наряду с основной тенденцией, сезонную составляющую, для решения поставленной цели используется регрессионная модель с фиктивными переменными учитывающая выявленные составляющие временного ряда. В качестве тестовых совокупностей используются временные ряды ввода в действие жилых домов в РФ и инвестиций в основной капитал России за период 1993-2012 гг. На основе модели проводится прогнозирование уровней анализируемого ряда на 2013 год, прогнозные значения сопоставляются с фактическими на основе коэффициента Тейла.

Ключевые слова: динамика, сезонность, тренд, ввод в действие жилых домов, модель, фиктивные переменные, прогноз

Значительные колебания уровней временных рядов макроэкономических показателей России, приводит к несостоятельности «классических» методов описания тенденций. Подели полученные на основе аналитического выравнивания, как правило, оказываются статистически не значимыми, соответственно прогнозы по ним отличаются от фактических данных. По видимому, в случае наличия высокой флуктуации уровней временного ряда, под влиянием долговременной, сезонной и случайной составляющих необходимо использовать более гибкие методы эконометрического моделирования, по нашему мнению, выходом из подобной ситуации может стать включение в модель фиктивных переменных.

Для иллюстрации возможностей регрессионных моделей с фиктивными переменными, в рамках настоящей статьи, рассмотрим поквартальную динамику ввода жилья за период 1993-2012 года (рисунок 1) [1, 3, 4].

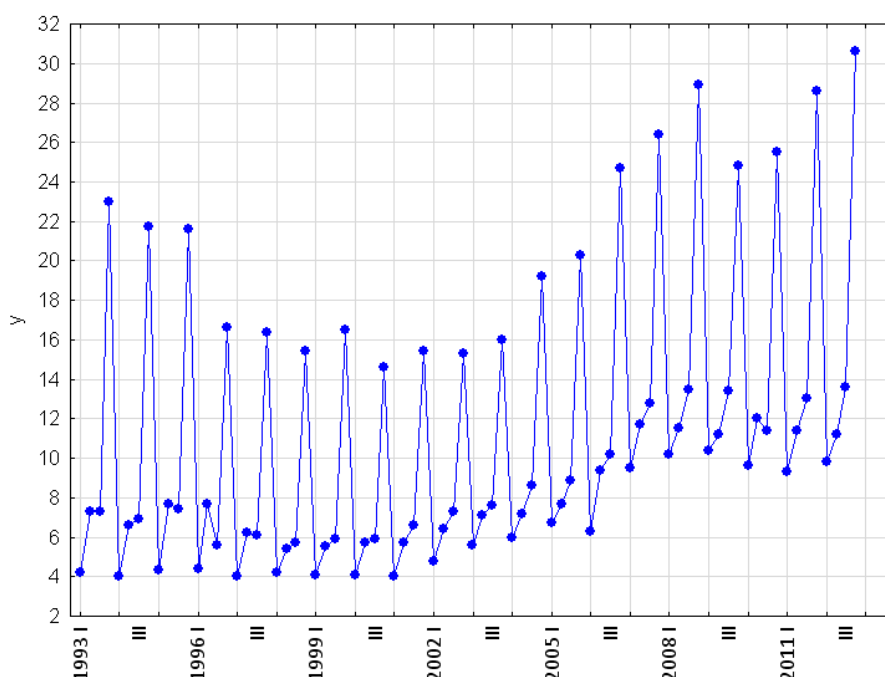


Рисунок 1 - Динамика ввода в действие жилых домов, млн. кв. м

Согласно приведенным на рисунке 1 данным, прослеживается сезонная волна с минимумом в первом квартале каждого года и максимумом в четвертом. Также можно указать на присутствие долговременной составляющей нелинейного характера.

Для учета всех составляющих используем регрессионную модель с фиктивными переменными, что позволит нам [5]:

- 1) протестировать гипотезу о наличии сезонности с помощью t -статистики Стьюдента;
- 2) учесть (смоделировать) в одной модели долговременную и сезонную составляющие временного ряда;
- 3) построить прогноз развития показателя на краткосрочную перспективу с учетом выделенных компонент временного ряда.

В общем виде множественная регрессионная модель с фиктивными переменными будет иметь следующий вид:

$$\hat{y}_t = a_0 + a_1 t_t + \alpha_1 D_{1t} + \alpha_2 D_{2t} + \alpha_3 D_{3t} + \varepsilon_t$$

Фиктивные переменные вводятся согласно следующего правила:

$$D_{1t} = \begin{cases} 1, & \text{если рассматривается II квартал} \\ 0, & \text{в противном случае} \end{cases}$$

$$D_{2t} = \begin{cases} 1, & \text{если рассматривается III квартал} \\ 0, & \text{в противном случае} \end{cases}$$

$$D_{3t} = \begin{cases} 1, & \text{если рассматривается IV квартал} \\ 0, & \text{в противном случае} \end{cases}$$

Коэффициент α_i в приведенной модели иногда называется дифференциальным коэффициентом свободного члена, так как он показывает, на какую величину отличается свободный член модели при значении фиктивной переменной, равном единице, от свободного члена модели при базовом значении фиктивной переменной.

Заметим, что число кварталов равно четырем, следовательно, число фиктивных переменных должно быть равно трем. В нашем примере в качестве базы выбран I квартал. Если значения y существенно различаются по кварталам (сезонам), то в уравнении коэффициенты при фиктивных переменных окажутся статистически значимыми. Тогда ожидаемое значение y по кварталам определяется следующими соотношениями:

$$\text{I квартал} \quad f(y|D_{1t} = 0, D_{2t} = 0, D_{3t} = 0) = a_0 + a_1 t_t$$

$$\text{II квартал} \quad f(y|D_{1t} = 1, D_{2t} = 0, D_{3t} = 0) = (a_0 + a_1) + a_1 t_t$$

$$\text{III квартал} \quad f(y|D_{1t} = 0, D_{2t} = 1, D_{3t} = 0) = (a_0 + a_2) + a_1 t_t$$

$$\text{IV квартал} \quad f(y|D_{1t} = 0, D_{2t} = 0, D_{3t} = 1) = (a_0 + a_3) + a_1 t_t$$

Обратимся к пакету программ STATISTICA (модуль «Множественная регрессия») и получим результаты представленные в таблице 1.

Все параметры получены статистически значимыми, по t -критерию Стьюдента, что указывает на наличие сезонности (значимость a_0, a_1, a_2, a_3) и тренда (значимость параметра a_1). Коэффициент детерминации указывает на высокую подгонку модели (описывается 88% колеблемости зависимой переменной).

Таблица 1 - Результаты оценки коэффициентов регрессионной модели описывающей тренд-сезонность во временном ряду ввода в действие жилых домов (долговременная составляющая в форме прямой)

	β	Стандартная ошибка β	Параметры уравнения	Стандартная ошибка параметров	$t(75)$	p -уровень значимости
Свободный член уравнения			2,11	0,70	3,01	0,00
t	0,37	0,04	0,11	0,01	9,17	0,00
D_1	0,12	0,05	1,85	0,76	2,43	0,02
D_2	0,15	0,05	2,40	0,76	3,16	0,00
D_3	0,93	0,05	14,48	0,76	19,05	0,00
Примечание: $R=0,93$; $R^2= 0,88$; скорректированный $R^2= 0,87$; $F(4,75)=138,97$; $p<0,00$; стандартная ошибка оценки: 2,40						

Повторим расчет, добавив в уравнение переменную t^2 , тем самым учтем параболический тренд.

Таблица 2 - Результаты оценки коэффициентов регрессионной модели описывающей тренд-сезонность во временном ряду ввода в действие жилых домов (долговременная составляющая в форме параболы второго порядка)

	β	Стандартная ошибка β	Параметры уравнения	Стандартная ошибка параметров	$t(74)$	p -уровень значимости
Свободный член уравнения			5,55	0,73	7,63	0,00
t	-0,50	0,12	-0,15	0,04	-3,98	0,00
t^2	0,89	0,12	0,003	0,00	7,13	0,00

D_1	0,12	0,04	1,85	0,59	3,15	0,00
D_2	0,15	0,04	2,40	0,59	4,08	0,00
D_3	0,93	0,04	14,48	0,59	24,59	0,00
Примечание: $R=0,96$; $R^2=0,93$; скорректированный $R^2=0,92$; $F(5,74)=195,29$; $p<0,00$; стандартная ошибка оценки: 1,86						

Как видим модель в целом, и параметры в частности, также получены статистически значимыми, но стоит обратить внимание на более высокий коэффициент детерминации - 93% против 88% в предыдущей модели, а также на стандартную ошибку, в случае введения в модель параболической составляющей она ниже. Отсюда следует вывод о большей успешности второй модели.

Обратившись к стандартизированным коэффициентам регрессии можно сделать вывод о пике в 4 квартале каждого года, так как β -коэффициент при переменной D_3 получен наибольшим - 0,93.

Если обратиться к графическим возможностям пакета, то можно представить результаты моделирования в виде рисунка 2. Теоретические (предсказанные) уровни достаточно хорошо согласуются с фактическими, различия наблюдаются лишь после 2008 года, что объясняется структурными изменениями в механизме формирования уровней связанными с внешним шоком под влиянием финансового кризиса.

Если перейти от общего уравнения к частным уравнениям, то получаем 4 параболических тренда, приведенных на рисунке 3. В целом использование параболы в качестве модели описывающей долговременную составляющую позволило уловить изменения уровней, но падение в 2008 году объемов повлияло на форму частных кривых (более медленное ускорение, нежели до точки перелома) и как следствие наблюдаются расхождения.

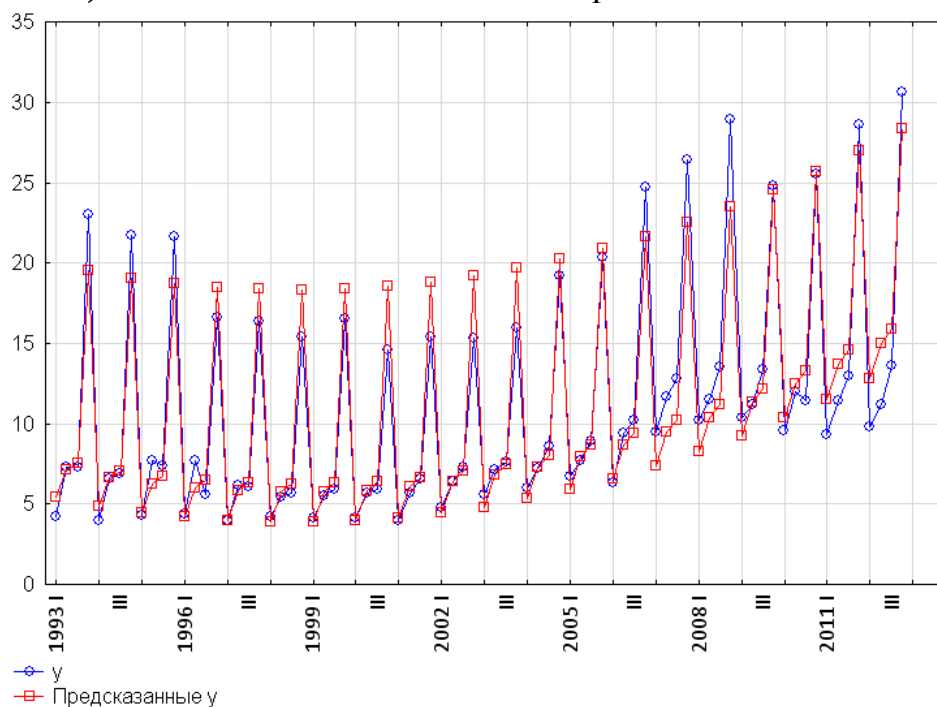


Рисунок 2 - Фактические и теоретические уровни временного ряда ввода в действие жилых домов, млн. кв. м.

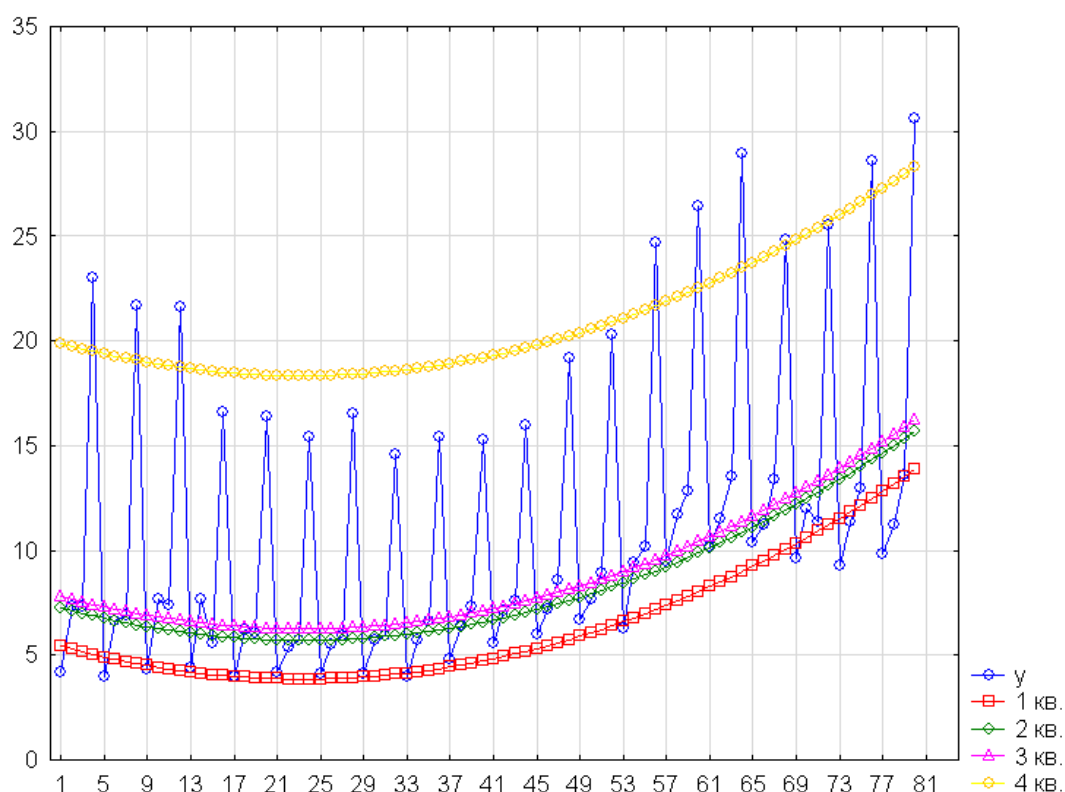


Рисунок 3 - Частные модели, характеризующие квартальную динамику ввода в действие жилых домов

Так как рассмотренное регрессионное уравнение статистически значимо, то его можно использовать для прогнозирования, при этом необходимо подставить в уравнение соответствующие значения (таблица 3).

Таблица 3 - Прогнозные значения ввода в действие жилых домов

Период	t	t^2	D_1	D_2	D_3	Уравнение	$\hat{y}_{t+1}$
2013/I	81	6561	0	0	0	$\hat{y}_t = 5,55 - 0,15 \times 81 + 0,003 \times 6561$	14,20
2013/II	82	6724	1	0	0	$\hat{y}_t = 5,55 - 0,15 \times 82 + 0,003 \times 6724 + 1,85 \times 1$	15,27
2013/III	83	6889	0	1	0	$\hat{y}_t = 5,55 - 0,15 \times 83 + 0,003 \times 6889 + 2,40 \times 1$	16,17
2013/IV	84	7056	0	0	1	$\hat{y}_t = 5,55 - 0,15 \times 84 + 0,003 \times 7056 + 14,48 \times 1$	28,60

В пакете STATISTICA имеется возможность молочения точечных прогнозных значений дополненных 95% доверительными границами, приведем полученные значения и фактические уровни в таблице 4.

Таблица 4 - Прогнозные и фактические уровни ввода в действие жилых домов

Период	Фактические уровни	Точечный прогноз	Нижняя доверительная граница	Верхняя доверительная граница
2013/I	10,4	14,20	12,72	15,68
2013/II	12,2	16,41	14,89	17,94
2013/III	16,1	17,33	15,76	18,91
2013/IV	30,7	29,78	28,15	31,41

Получаем, некоторое расхождение между прогнозами и фактическими уровнями (прогнозы занижены). Для объективной оценки величины расхождения (совпадения) обратимся к коэффициенту Тейла:

$$U = \frac{\sqrt{\sum (\hat{y}_t - y_t)^2}}{\sqrt{\sum y_t^2}}$$

где: $\hat{y}_t$ и y_t – соответственно предсказанное и фактическое (реализованное) изменения переменной.

В случае равенства коэффициента нулю имеем совершенный прогноз. Соответственно при $U > 1$ прогноз дает худшие результаты, чем предположения о неизменности исследуемого явления.

В нашем случае значение коэффициента равно 0,26, что указывает на высокую согласованность прогнозов и фактических данных.

В приведенном примере рассматривается ситуация в которой квартальные различия отражаются лишь в расхождении свободных членов моделей. И такой подход справедлив для моделирования временных рядов, в которых сезонная составляющая входит аддитивно, но что если рассматриваемая сезонная волна развивается мультипликативно. В качестве примера можно рассмотреть динамику инвестиций в основной капитал России за период 2003-2012 года [2].

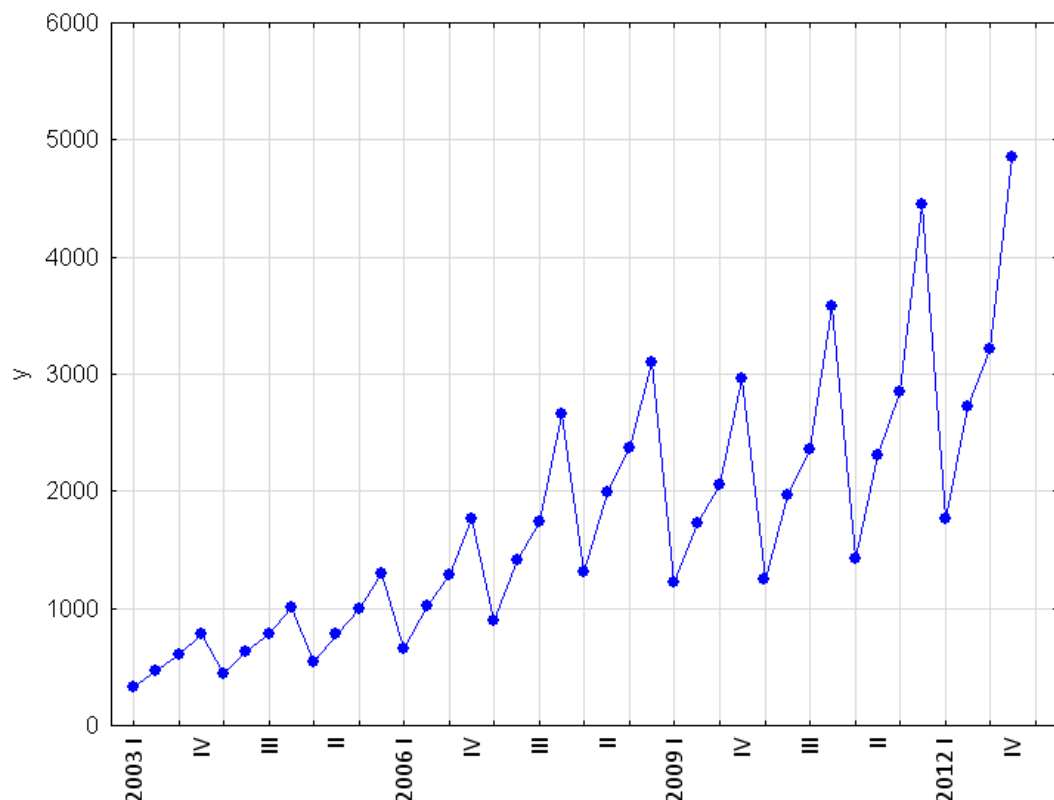


Рисунок 4 - Динамика инвестиций в основной капитал России, млрд. руб.

Согласно приведенному рисунку, четко проявляется сезонность с пиком в каждом 4 квартале, при этом по мере удаления от начала точки отсчета размах колебания увеличивается, что указывает на мультипликативность. Наряду с сезонностью в рассматриваемом временном ряду прослеживается

долговременная составляющая в форме линейного тренда. Очевидно, что введение в регрессионное уравнение фиктивных переменных классическим образом не принесет наилучших результатов (если вообще параметры будут статистически значимыми), поэтому необходимо использовать такие регрессоры, параметры которых будут корректировать угол наклона регрессии, т.е. параметр a_1 . В таком случае получаем следующую модель:

$$\hat{y}_t = a_0 + a_1 t_t + \alpha_1 D_{1t} t_t + \alpha_2 D_{2t} t_t + \alpha_3 D_{3t} t_t + \varepsilon_t$$

Фиктивные переменные вводятся согласно следующего правила:

$$D_{1t} t_t = \begin{cases} t_t, & \text{если рассматривается II квартал} \\ 0, & \text{в противном случае} \end{cases}$$

$$D_{2t} t_t = \begin{cases} t_t, & \text{если рассматривается III квартал} \\ 0, & \text{в противном случае} \end{cases}$$

$$D_{3t} t_t = \begin{cases} t_t, & \text{если рассматривается IV квартал} \\ 0, & \text{в противном случае} \end{cases}$$

Опираясь на общее уравнение можно получить следующие частные уравнения, для каждого квартала:

$$\begin{aligned} \text{I квартал} \quad \hat{y}_t &= a_0 + a_1 t_t \\ \text{II квартал} \quad \hat{y}_t &= a_0 + (a_1 + \alpha_1) t_t \\ \text{III квартал} \quad \hat{y}_t &= a_0 + (a_1 + \alpha_2) t_t \\ \text{IV квартал} \quad \hat{y}_t &= a_0 + (a_1 + \alpha_3) t_t \end{aligned}$$

Оценка параметров общего уравнения приводит нас к результатам представленным в таблице 5.

Таблица 5 - Результаты оценки коэффициентов регрессионной модели описывающей тренд-сезонность во временном ряду инвестиций в основной капитал (долговременная составляющая в форме прямой)

	β	Стандартная ошибка β	Параметры уравнения	Стандартная ошибка параметров	$t(35)$	p -уровень значимости
Свободный член уравнения			237,33	58,12	4,08	0,00
t	0,42	0,04	39,10	3,41	11,48	0,00
$D_1 t$	0,23	0,03	23,73	3,56	6,66	0,00
$D_2 t$	0,36	0,03	35,88	3,50	10,25	0,00
$D_3 t$	0,73	0,04	71,75	3,45	20,82	0,00
Примечание: $R=0,99$; $R^2=0,98$; скорректированный $R^2=0,97$; $F(4,35)=354,99$; $p<0,00$; стандартная ошибка оценки: 180,11						

На основе полученной статистически значимой модели построим частные модели (рисунок 5).

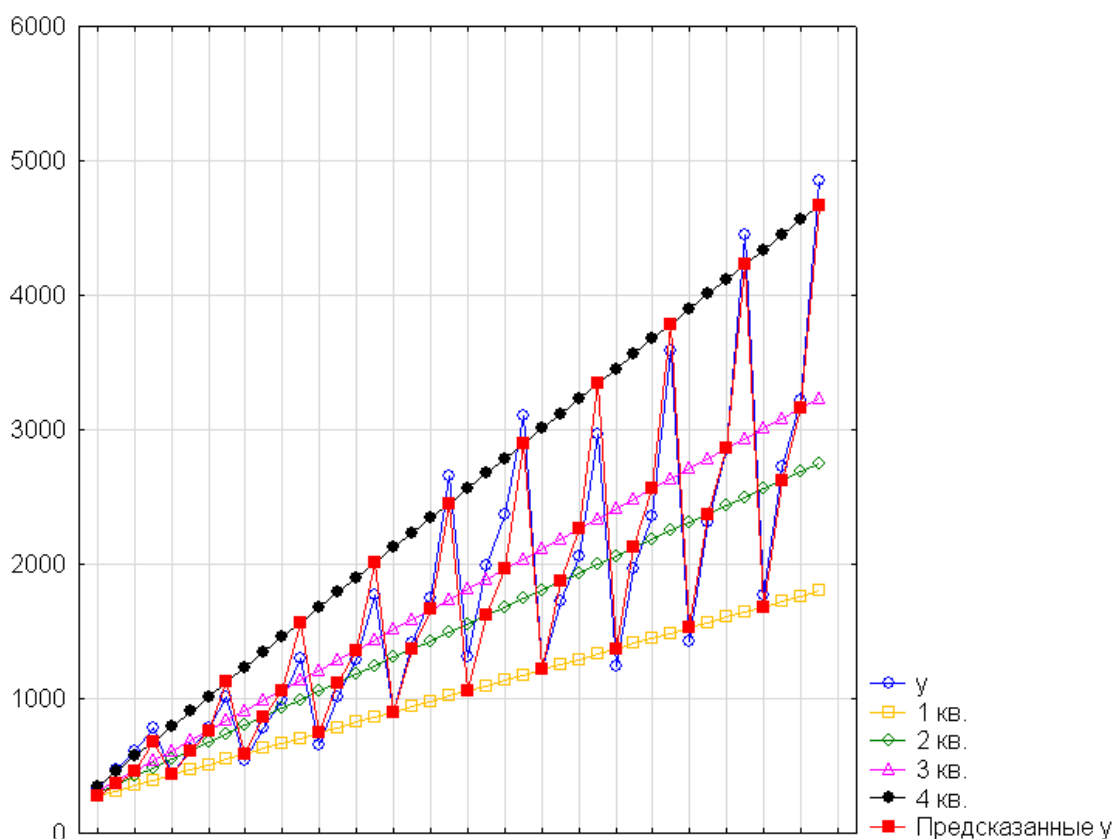


Рисунок 5 - Частные модели, характеризующие квартальную динамику инвестиций в основной капитал

Согласно полученным данным предсказанные (прогнозные) уровни и частные модели достаточно хорошо согласуются фактическими точками, что указывает на высокую подгонку модели.

Подводя итог проведенного исследования, можно констатировать наличие мощного инструмента позволяющего описывать реальные процессы и явления с высокой точностью.

Список литературы

1. Цыпин А.П. Конкурентоспособность в банковской сфере России / А.П. Цыпин // Вестник магистратуры. 2012. № 5. С. 90-91.
 2. Цыпин А.П. Оценка доли иностранного капитала в промышленности России / А.П. Цыпин, В.А. Овсянников // Молодой ученый. 2014. № 12 (71). С. 195-198.
 3. Цыпина Ю.С. Моделирование динамики объемов ипотечного жилищного кредитования в РФ / Ю.С. Цыпина // Научное обозрение. 2012. № 3. С. 464-469.
 4. Цыпина Ю.С. Статистические методы в изучении ипотечного жилищного кредитования России / Цыпина Ю.С., Цыпин А.П. // Новый университет. Серия: Экономика и право. 2012. № 6. С. 10-13.
- Цыпина Ю.С. Статистическое моделирование влияния социально-экономических факторов на объем ипотечного жилищного кредитования на основе нейронных сетей / Ю.С. Цыпина // Интеллект. Инновации. Инвестиции. 2012. № 4. С. 104-106.