

К.В. Мотричкин, А.В. Степанов
**ПРОГНОЗИРОВАНИЕ РАЗВИТИЯ ЭКОНОМИКИ РЕГИОНА
В УСЛОВИЯХ НЕПОЛНОТЫ ИНФОРМАЦИИ***

В статье рассмотрен вопрос построения системно-динамических экономических моделей в условиях, когда доступный объём статистических данных мал. Приведён анализ подходов к решению проблемы малого количества данных. Показаны подходы к построению динамических моделей прогнозирования в условиях, когда данных недостаточно для полноценной эконометрической оценки. Предлагается вводить дополнительные предпосылки, благодаря которым можно обойти оценивание одного или нескольких коэффициентов уравнений. Описаны подходы к построению модели прогнозирования, которая может быть адекватно оценена на малом объёме данных.

Ключевые слова: агентное моделирование, выборка ограниченного объёма, малая выборка, модель прогнозирования, системная динамика, экономика региона.

Форм. 12. Лит. 12.

К.В. Мотрічкін, А.В. Степанов
**ПРОГНОЗУВАННЯ РОЗВИТКУ ЕКОНОМІКИ РЕГІОНУ
В УМОВАХ НЕПОВНОТИ ІНФОРМАЦІЇ**

У статті розглянуто питання побудови системно-динамічних економічних моделей в умовах, коли доступний об'єм статистичних даних є малим. Наведено аналіз підходів до вирішення проблеми малої кількості даних. Показано підходи до побудови динамічних моделей прогнозування в умовах, коли даних недостатньо для повноцінної економічної оцінки. Запропоновано ввести додаткові передумови, завдяки яким можна обійти оцінювання одного або декількох коефіцієнтів рівнянь. Описано підходи до побудови моделі прогнозування, яка може бути адекватно оцінена на малому обсязі даних.

Ключові слова: агентне моделювання, вибірка обмеженого обсягу, мала вибірка, модель прогнозування, системна динаміка, економіка регіону.

К.В. Motrichkin¹, А.В. Stepanov²
**FORECASTING REGION'S ECONOMY DEVELOPMENT UNDER
THE CONDITIONS OF INCOMPLETE INFORMATION**

The article considers the issues of constructing system dynamics economic models under the conditions when the available statistical data is small in volume. The analysis of various approaches to solving the problem of insufficient data is carried out. The approaches to construction of dynamic models for forecasting are demonstrated, especially for those conditions when there is a lack of sufficient information for a full econometric estimation. It is suggested to introduce additional prerequisites which would enable overlooking several indices saving the estimation as a whole. Several approaches to construction of a forecasting model are described; such a model can perform adequate estimation on a insignificant data volume.

Keywords: agent-based modeling; limited sample; small sample; forecasting model; system dynamics; region's economy.

Постановка проблемы. При прогнозировании экономических показателей региона, особенно отраслей региона, может возникнуть проблема нехватки данных. Например, данные по некоторым показателям видов экономической деятельности (ВЭД) доступны только за шестилетний период — с 2005 г. по

* Статья подготовлена при финансовой поддержке Министерства образования и науки Российской Федерации (Постановление Правительства РФ от 09.04.2010 № 218) в рамках выполнения Договора № 13.G25.31.0065.

¹ Junior Consultant, "IBS Soft" Ltd. (Moscow, Russian Federation).

² Junior Consultant, "IBS Soft" Ltd. (Moscow, Russian Federation).

2010 год. Это обусловлено тем, что общероссийский классификатор видов экономической деятельности (ОКВЭД) введён в действие в 2003 году. При наличии всего 6 наблюдений оценка эконометрических уравнений не представляется возможной. В общем случае, при доверительной вероятности равной 0,8 адекватная погрешность (меньше 0,5) может быть получена при выборке, содержащей более 15 наблюдений [4]. При условиях выборки ограниченного объёма необходимо строить модель таким образом, чтобы оценка её коэффициентов была возможной. Это, в свою очередь, требует особого применения методов моделирования.

Анализ последних исследований и публикаций. Задача статистической оценки по малой выборке или при малом количестве данных возникает в разных областях. Хотя проблема эта имеет одинаковую природу, в различных областях она решается по-разному – с учётом специфики рассматриваемых систем [12]. Решить проблему малой выборки данных иногда можно с помощью перехода к другим методам моделирования. Например, метод нейросетевого прогнозирования позволяет прогнозировать, имея небольшой массив данных [8]. При применении моделирования временных рядов существуют приёмы, которые позволяют достичь высокой точности моделирования на сравнительно малых объёмах выборки [9]. Иногда есть возможность использовать специальные критерии [1; 10] или приёмы [2; 3; 6] для малых выборок.

Нерешённые части проблемы. Проблема оценивания моделей на выборках относительно малого объёма не решена для случая построения системно-динамических моделей. В случае системной динамики можно использовать некоторые свойства этого метода моделирования для решения проблемы малой выборки.

Цель исследования – разработать и обосновать методы построения системно-динамических моделей, оцениваемых на выборках относительно малого объёма.

Основные результаты исследования. В ходе построения модели прогнозирования кадровой потребности для отрасли региона автор выделил 3 метода изменения модели, которые позволяют обойти проблему нехватки данных, вызванную использованием выборки малого объёма:

1. Введение и проверка гипотез о значениях коэффициентов уравнений.
2. Обобщение конкретного свойства системы.
3. Применение агентного моделирования.

Эти приёмы можно использовать, когда нет возможности оценить эконометрическое уравнение, но есть необходимость как-то описать взаимосвязь между переменными. Применение таких подходов зачастую сопровождается усилением предпосылок модели. Рассмотрим эти методы отдельно.

1. Введение и проверка гипотез о значениях коэффициентов уравнений. Для этого необходимо, чтобы уже было составлено эконометрическое уравнение. Если оценить некоторый коэффициент не получается, но есть некоторое предположение о том, чему он может быть равен, то можно составить и проверить гипотезу о равенстве коэффициента некоторому значению. Важно, чтобы гипотеза была теоретически обоснована.

Рассмотрим уравнение для основных средств:

$$K_t = \gamma_1 \times K_{t-1} + \gamma_2 \times I_{t-1}, \quad (1)$$

где K_t – стоимость основных средств в год t ; I_t – объём инвестиций в основные средства в год t ; γ_1, γ_2 – коэффициенты регрессии. Коэффициент γ_1 отражает амортизацию основных средств. То есть $1-\gamma_1$ – норма амортизации. Коэффициент γ_2 показывает, какая доля инвестиций идёт на непосредственное наращение основных средств.

Естественно можно предположить, что коэффициент $\gamma_2 = 1$, что значит, что все инвестиции переходят в основные средства. Если данных для оценки уравнения мало, то всё равно можно оценить верность этой гипотезы. Оценку можно провести, например, статистическим тестом Вальда [5] или использовать специальные критерии для малых выборок [1]. Если тест не отвергнет гипотезу, то можно не оценивать коэффициент, а использовать значение из гипотезы.

Важно, чтобы гипотеза была верной в смысле соблюдения логической и экономической когерентности. Не стоит пытаться подобрать коэффициент. При применении подхода усиливаются предпосылки модели. Они дополняются предпосылкой о конкретном значении коэффициента.

Иногда значение коэффициента для гипотезы нужно рассчитать. Обратим внимание на уравнение для основных средств ещё раз. Выше уже говорилось, что $1-\gamma_1$ – это, по сути, норма амортизации. Несмотря на то, что существуют разные методы списания амортизации, норма амортизации мало меняется во времени. Значение $1-\gamma_1$ можно оценить как среднее отношение амортизации основных средств к основным средствам. Получившееся таким образом значение коэффициента можно проверить тем же тестом Вальда.

2. Обобщение конкретного свойства системы. Этот метод применяется в случае, когда некоторое свойство моделируемой системы справедливо и для других систем. Тогда это свойство можно оценивать не только по данным одной системы, но и по данным всех систем, которым присуще данное свойство. То есть можно отбросить некоторые особенности, если это не противоречит целям исследования. Получившиеся коэффициенты можно проверить уже на данных исходной системы каким-нибудь статистическим критерием, например, критерием Вальда.

Например, есть некоторое уравнение для параметра отрасли региона. Если есть причины полагать, что взаимосвязи между этим параметром и объясняющими параметрами одинаковы для всех регионов страны, то можно оценить коэффициенты уравнения на данных отрасли страны. Далее следует проверить получившиеся коэффициенты на данных отрасли региона. Если проверка успешна, можно использовать получившиеся коэффициенты.

Следует осторожно применять описанный выше способ оценки коэффициентов. Так, при построении модели кадровой потребности оказалось, что обобщение конкретного свойства системы подходит для коэффициента γ_1 вышеописанного уравнения, но не подходит для γ_2 .

При применении подхода происходит усиление предпосылок – появляется предпосылка о том, что поведение элемента системы одинаково в некотором классе систем.

3. Применение агентного моделирования. При отсутствии возможности оценить взаимосвязь между параметрами системы на статистических данных можно описать их исходя из того, чем эти взаимосвязи порождаются. То есть исходя из мотивов экономических агентов, которые могут влиять на переменную.

Для этого выбирается агент, составляется функция, которую агент максимизирует, составляется оптимизационная задача так, чтобы из её решения можно было бы записать уравнение для искомой переменной. Соответственно, при этом предполагается рациональное поведение агентов и, может быть, вводятся предположения об их ожиданиях.

При применении агентного моделирования может оказаться, что отдельно оценивать уравнение для параметра не нужно. Однако бывает нужно оценить параметры функции, которую максимизируют.

Например, пусть надо составить модель для потребности экономики отрасли региона в кадрах. Составим ее с помощью агентного моделирования, как это делается в [7]. Агентом здесь является работодатель, который порождает эту потребность. Поскольку все работодатели отрасли действуют одинаково, их всех можно считать одним агентом. Цель работодателей как экономических фирм – максимизировать свою прибыль. Записываем выражение для прибыли:

$$\Pi_t = Y_t - C_t, \tag{2}$$

где Π – прибыль; Y – выручка; C – издержки; t – индекс времени. Далее определяем входящие в это уравнение переменные. Пусть выручка определяется производственной функцией Кобба-Дугласа:

$$Y_t = P_t \times A_t \times K_t^a \times L_t^b, \tag{3}$$

где a, b – оцениваемые параметры ($a + b = 1, 0 < a < 1, 0 < b < 1$); P – уровень цен на продукцию отрасли; A – параметр технологического уровня и прочих необъяснённых факторов отрасли; K – стоимость основных средств отрасли; L – численность занятых в отрасли.

Заметим, что необязательно оценивать производственную функцию эконометрическими методами. Существует способ оценки производственной функции по долям затрат на факторы производства [11].

Издержки удобно разделить на две группы: зависящие от численности сотрудников и не зависящие от неё. Пусть издержки – это сумма затрат на оплату труда работников $w_t \times L_t$, где w – среднегодовая зарплата в отрасли, и прочих расходов, выраженных функцией Φ :

$$C_t = \Phi_t + w_t \times L_t. \tag{4}$$

Тогда задачу, решаемую агентом, можно записать так:

$$P_t \times A_t \times K_t^a \times L_t^b - \Phi_t - w_t \times L_t \rightarrow \max_{L_t} \tag{5}$$

Решение:

$$(P_t \times A_t \times K_t^a \times L_t^b - \Phi_t - w_t \times L_t)'_{L_t} = 0; \tag{6}$$

$$b \times A_t \times P_t \times K_t^a \times L_t^{b-1} - w_t = 0. \tag{7}$$

При условии $a + b = 1$ получаем:

$$b \times A_t \times P_t \times \frac{K_t^{1-b}}{L_t^{1-b}} - w = 0. \quad (8)$$

Помножив обе части дроби на L_t^b :

$$A_t \times L_t^b \times K_t^a \times \frac{b}{L_t} - w_t = 0. \quad (9)$$

Выделив Y_t , получим условие для оптимального количества сотрудников:

$$L_t = \frac{b \times Y_t}{w_t}. \quad (10)$$

То есть в оптимуме работодатель тратит на оплату труда всегда одинаковую долю своего дохода, равную b .

Таким образом, для достижения некоторого уровня выпуска работодателю нужно нанять определённое количество работников. Однако из получившегося равенства можно понять, что работники должны наниматься для удовлетворения этой потребности до того, как становится известен соответствующий ей уровень выпуска, так как Y_t — выпуск, обеспечиваемый численностью сотрудников L_t . То есть работодатель руководствуется при принятии решения о найме сотрудников не реально достигнутым уровнем выпуска, а некоторым представлением о том, каким будет выпуск. Необходимо математически описать ожидания работодателя. Для этого рассматриваются ожидаемый выпуск отрасли и ожидаемая среднегодовая заработная плата в отрасли.

При предположении о существовании описанной выше производственной функции самым оптимистичным для работодателя будет получить постоянную отдачу от масштаба. То есть получить рост производства пропорционален росту стоимости основных средств. Ожидаемый уровень выпуска отрасли Y_{t+1}^* будет вычисляться:

$$Y_{t+1}^* = Y_t \times \frac{K_{t+1}}{K_t}. \quad (11)$$

Пусть у работодателя рациональные ожидания среднегодовой заработной платы, то есть его ожидания совпадают с модельным значением заработной платы. Тогда формулу расчёта потребности в кадрах N можно записать так:

$$N_t = b \times \frac{Y_{t-1} \times \frac{K_t}{K_{t-1}}}{w_t}. \quad (12)$$

Заметим, что использование агентного моделирования возможно даже без каких-либо идей относительно вида эконометрического уравнения для рассматриваемого параметра.

При использовании этого подхода вместо эконометрических уравнений происходит изменение предпосылок. Неизбежно появляется предпосылка о некоторой рациональности экономических агентов.

Выводы. В статье были рассмотрены способы построения системно-динамических моделей, оцениваемых на малых объёмах данных. Рассмотрены два способа оценки коэффициентов эконометрических уравнений при отсутст-

вии достаточно большого массива данных. Был рассмотрен способ применения агентного моделирования в целях избежания оценки эконометрических уравнений. При этом агент рассматривался как элемент системы, то есть модель рассматривалась как системно-динамическая. Описанные подходы были апробированы авторами и дали адекватный результат. Но нужно понимать, что при применении описанных подходов происходит усиление предположений — прогноз, получаемый на основе модели, становится более условным.

1. *Балицкая Е.О., Буяк А.Н., Золотухина Л.А.* Специальные критерии согласия для малых выборок // Прикладная и вычислительная математика в судостроении: Научн. труды (Ленинградский кораблестроительный институт). — Ленинград, 1981. — С. 14–21.

2. *Долгов Ю.А., Ваняшкин М.М., Деткова А.В.* Определение границ и ширины ядра эквивалентных выборок в методе точечных распределений. — К., 2010. — 280 с.

3. *Капур К., Ламберсон Л.* Надежность и проектирование систем / Пер. с англ. — М.: Мир, 1980. — 604 с.

4. *Кольхан Н.В., Тюрлев В.С.* Информационные технологии статистической обработки данных выборок ограниченного объема // Инженерный вестник Дона. — 2007. — №2. — С. 37–50.

5. *Леман Э.* Проверка статистических гипотез. — М.: Наука, 1979. — 408 с.

6. *Литвинов С.* Использование понятия относительной погрешности оценивания для расчета выборки из генеральной совокупности с малой долей признака // Социология: теория, методы, маркетинг. — К., 2008. — С. 141–150.

7. *Мотричкин К.В.* Оценка кадровой потребности в экономике для решения задач стратегического планирования // Информационные бизнес системы: Труды 54-й научной конференции МФТИ «Проблемы фундаментальных и прикладных естественных и технических наук в современном информационном обществе». — М., 2011. — С. 45–47.

8. *Осипов В.Ю.* Нейросетевое прогнозирование социально-экономических процессов // Материалы Всероссийской конференции, посвященной 75-летию со дня рождения первого директора СПб ЭМИ РАН, зам. председателя Президиума СПб НЦ РАН, проф. Б.Л. Овсевича. — Санкт-Петербург, 2011. — С. 149–150.

9. *Семёнычев В.К., Павлов В.Д., Семёнычев В.В.* Моделирование и прогнозирование временного ряда суммой логистической, линейной и гармонической компонент на основе ARMA-модели. — Самара, 2008. — 14 с.

10. Специальные критерии согласия для малой выборки при оценке надежности нефтепромыслового оборудования / И.Е. Ишемгузин, А.Р. Атнагулов, А.Н. Зотов, Е.И. Ишемгузин. — Уфа: Нефтегазовое дело, 2008. — 44 с.

11. Факторы экономического роста в регионах РФ / С.В. Дробышевский, О. Луговой, Е. Астафьева, Д. Полевой, А. Козловская, П. Трунин, Л. Ледерман. — М.: Агентство СІР РГБ, 2005. — 278 с.

12. *Фокин В.А.* Статистическое моделирование данных при оценке состояния биологических систем // Известия Томского политехнического университета. — 2007. — Т. 311, №5. — С. 132–135.

Стаття надійшла до редакції 20.06.2012.