

## МОДЕЛИ ARIMA В КРАТКОСРОЧНОМ ПРОГНОЗИРОВАНИИ ВНУТРЕННЕЙ МИГРАЦИИ В РОССИИ

Е.В. Павловский

На основе данных о прибытиях и выбытиях по федеральным округам и типам населенных пунктов Российской Федерации за 2000–2015 гг. проведена статистическая проверка гипотезы об инерционности внутренней миграции. Определены коэффициенты автокорреляции и выявлена автокорреляционная зависимость в 243 из 256 временных рядов, что позволило сделать вывод о наличии инерционности во внутренней миграции России. На примере анализа 28 временных рядов данных о внешней миграции в страны Европейского союза и трех - в государства - члены Североамериканского соглашения о свободной торговле за 2000–2015 гг. были рассчитаны коэффициенты автокорреляции и оценена их статистическая значимость, что позволило проверить и подтвердить потенциальную применимость метода ARIMA для краткосрочного прогнозирования международной миграции.

Обобщены результаты подбора параметров моделей типа ARIMA для временных рядов показателей внутренней миграции в России. На основе ретроспективных прогнозов внутренней миграции в России для периода 2013–2015 гг. проведено сравнение средних относительных ошибок прогнозирования при использовании моделей ARIMA и метода экспоненциального сглаживания. Сделан вывод о том, что модели ARIMA обеспечивают большую по сравнению с методом экспоненциального сглаживания точность краткосрочных адаптивных прогнозов внутренней миграции в России.

На основе расчета доверительных интервалов для средней определено, что модели семейства ARIMA с одинаковой точностью прогнозируют въездную и выездную миграцию для двух типов населенных пунктов (города и сельской местности). На примере данных об объемах внутренней миграции Республики Беларусь за 2000–2016 гг. доказано, что модели семейства ARIMA могут быть использованы для краткосрочного прогнозирования внутренней миграции других государств.

*Ключевые слова:* миграция, инерционность миграции, автокорреляция, внутренняя миграция, прогнозирование, экспоненциальное сглаживание, интегрированная модель авторегрессии и скользящего среднего ARIMA.

JEL: C53, J62.

### 1. Введение и теоретический обзор

Миграция оказывает значительное влияние не только на демографическое, но и на экономическое развитие регионов. В регионах-реципиентах мигранты увеличивают объемы спроса и предложения на всех рынках - труда (трудова́я миграция), образовательных услуг (учебная миграция) и т. д., что способствует экономическому росту и в то же время приводит к возникновению диспропорций в социально-экономической сфере (нехватка мест в дошкольных учреждениях и школах, проблемы с медицинским обслуживанием прибывших и т. д.). В регионах-донорах, если они не являются трудоизбыточными, миграционная убыль ведет к снижению объемов экономической деятельности. В России в связи с ростом объемов

внутренней миграции (число внутрироссийских перемещений в 2015 г. по сравнению с 2011 г. возросло на 1077386 единиц, или на 35,2%, и составило 4135906)<sup>1</sup> вопрос ее прогнозирования на общефедеральном и региональном уровнях становится все более актуальным.

Большинство теоретических работ, а также работ в области прогнозирования миграции посвящено исследованию внешних перемещений. В то же время ряд российских ученых, в том числе В.М. Моисеенко [1], Ж.А. Зайончковская [2] и др., уделяют большое внимание вопросам внутренней миграции. Так, исследуются причины внутренней миграции в России, в первую очередь значительная дифференциация социально-экономического развития регионов, а также культурные и религиозные различия между территориальными

Павловский Егор Витальевич (pavlovskij.egor@gmail.com) - аспирант, Санкт-Петербургский государственный экономический университет (г. Санкт-Петербург, Россия).

<sup>1</sup> Росстат. Внутрироссийская миграция по территориям прибытия и выбытия («шахматка» по федеральным округам). URL: [www.gks.ru/free\\_doc/new\\_site/population/demo/migr3.xls](http://www.gks.ru/free_doc/new_site/population/demo/migr3.xls).

единицами, которые часто становятся факторами, обуславливающими смену места жительства. Как правило, при изучении и прогнозировании внутренней миграции в России использовались гравитационные модели [3, 4, 5], основоположником которых стал американский астроном Дж. Стюарт. Он утверждал: «Демографическая сила притяжения между регионами прямо пропорциональна населению в регионе выбытия и регионе прибытия и обратно пропорциональна квадрату расстояния между ними» [6, с. 445]. В дальнейшем приверженцы данного подхода всячески модифицировали его, вводя в качестве объясняющих переменных экономические, социальные и иные индикаторы, но детерминантой в большинстве случаев оставался фактор «расстояние». Несмотря на высокую оценку точности прогнозирования как внутренней, так и внешней миграции на основе гравитационных моделей в некоторых странах Европы, например в Румынии [7], для России данный подход имеет ограничения. Как показывают некоторые исследования, из-за высокой степени дифференциации регионов страны по ряду признаков (в том числе экономических, климатических и культурных), а также больших расстояний между субъектами Российской Федерации, измеряемых в некоторых случаях десятками тысяч километров, статистические модели имеют довольно низкие показатели качества [3]. Основной причиной, по нашему мнению, является то, что данный подход предполагает симметричность противотоков, а этого вряд ли можно ожидать в условиях критически дальних расстояний между регионами (при достижении некоторой точки максимума расстояние перестает быть ключевым фактором решения о переезде) и большого количества территориальных единиц. Даже группировки регионов, используемые для преодоления данных ограничений, не обеспечивают должного качества получаемых статистических моделей [3].

Несмотря на высокую степень актуальности изучения проблемы внутренней миграции, методология и методика ее прогнозирования развиты явно недостаточно. Это во многом обусловлено тем, что теоретические модели внутренней миграции и эмпирические данные сосуществуют в некотором отрыве друг от друга. В результате наблюдается нехватка теоретических гипотез, прошедших статистическую проверку на данных по внутренней миграции. Кроме того, существу-

ющие теории уделяют больше внимания внешней миграции, чем внутренней. Имеющиеся пробелы требуют адаптации теорий, посвященных в большей степени внешней миграции, к объяснению и прогнозированию внутренних перемещений. Это особенно важно для России, для экономики которой, в силу неоднородности территорий, внутренние миграции имеют особое значение.

Во многих работах, посвященных внешней миграции, отмечается ее инерционность. Так, Дж. Тейлор [8] указывает, что весьма важным фактором притяжения мигранта в конкретный регион является наличие так называемой межличностной мигрантской сети в принимающей стране. Такая сеть, состоящая из родственников, друзей и сограждан, помогает каждому следующему поколению мигрантов приспособиться ко многим аспектам повседневной жизни, что и обуславливает инерционность миграции. Главная роль сетей, по мнению Д. Мэсси, являющегося последователем теории Тейлора, состоит в уменьшении различного рода издержек (не только денежных, но и психологических, а также других) и рисков, связанных с миграцией [9, р. 448]. Статистическая проверка гипотезы о влиянии мигрантских сетей на объемы миграции была подтверждена Т. Хаттоном и Дж. Уильямсоном на основе данных о выбытиях из Европы за 1850-1914 гг. [10] и получила подтверждение в последующих работах [11, 12]. Вероятно, удовлетворившись результатами исследований конца прошлого столетия, ученые, публикующие свои работы в XXI веке, не проводят статистическую проверку данной гипотезы и свои изыскания основывают на результатах своих предшественников, сосредоточиваясь в большей степени на теоретическом обосновании [13] и подкрепляя его данными социологических исследований [14, 15].

По мнению Т. Фэйста, разработавшего понятие «транснациональные социальные пространства», весьма близкого понятию «мигрантские сети», транснациональные социальные пространства могут рассматриваться как мосты между странами-донорами и странами-реципиентами международной миграции [11]. В 1993 г. Д. Мэсси дополнил теорию мигрантских сетей, указав на значимость коммерческих и некоммерческих организаций, облегчающих переезд мигрантов, например компаний, организующих поиск рабочей силы, в том числе за рубежом [9]. Ряд социологических исследований, посвященных

миграции из Филиппин в США, также подтверждают данную концепцию: мигранты с большей частотой обращаются за подробной информацией о визах и по прочим организационным вопросам не в официальные представительства США, а к успешно социализированным в новой стране бывшим соотечественникам [13, с. 57-58; 16]. На наш взгляд, данный подход тесно связан с концепцией социального капитала в случае рассмотрения его в качестве ресурса, помогающего людям достичь своих целей с помощью личностных связей и активов.

Можно утверждать, что в настоящее время гипотеза об инерционности внешней миграции имеет не только теоретическое обоснование, но и прошла статистическую проверку. В отличие от этого, проверка данной гипотезы в отношении внутренней миграции имеет некоторые пробелы как в теоретическом, так и в статистическом плане. Указанное обстоятельство и обусловило цели настоящего исследования: во-первых, проверить гипотезу об инерционности внутренней миграции на примере России и определить возможность ее использования в прогнозировании; во-вторых, рассмотреть возможности краткосрочного адаптивного прогнозирования внутренней миграции с помощью моделей типа ARIMA и выявить на этой основе место данного подхода в системе методов прогнозирования миграции.

## 2. Данные и методика прогноза

Прогнозирование внутренней миграции, несмотря на отсутствие ряда факторов, затрудняющих прогноз внешней миграции (международных отношений, визовой политики и т. д.), также представляет собой весьма сложный процесс. В некоторых известных работах справедливо отмечается, что многие исследователи при построении прогнозных моделей внутренней миграции необоснованно пренебрегают географическими факторами (в том числе типом поселения, расположением региона) и выстраивают свои концепции для всей страны в целом [17, р. 6-7; 18, р. 122-124]. Если для небольших стран такой подход еще может оказаться адекватным, то для России, где степень различий социально-экономического развития между регионами весьма высока, подобный подход неприемлем.

Представляется целесообразным осуществлять прогнозирование на уровне страны «снизу вверх», позволяющее учесть для каждого временного ряда (каждой территориальной единицы) его динамические особенности.

Исходя из сказанного выше, нами был выбран следующий подход к анализу и прогнозированию рядов по федеральным округам Российской Федерации:

1. Прогноз выездов и въездов по типам поселений (городское или сельское);
2. Прогноз миграционного оборота как суммы количества перемещений по типам поселений;
3. Прогноз объемов внутренней миграции на уровне страны как сумма прогнозов по всем федеральным округам.

Исследование проводилось на основе годовых данных о межрегиональной миграции за 2000-2015 гг. по территориям выбытия и прибытия<sup>2</sup> с распределением мигрантов по федеральным округам Российской Федерации и типу поселения (городское и сельское). Таким образом, было получено 2<sup>8</sup>, то есть 256 временных рядов длиной по 16 лет. Данные по Крыму и Севастополю были исключены из анализа в связи с критично малой длиной рядов.

Гипотеза об инерционности внутренней миграции проверялась на основе анализа наличия автокорреляции для каждого динамического ряда. Отметим, что в 2011 г. в России была изменена методика статистического учета прибывших и выбывших. Прибывший (согласно новой методике) - лицо, регистрирующееся по месту жительства либо пребывания на срок не менее девяти месяцев; выбывшими же считаются лица, регистрация которых по месту региона проживания истекла. Такое изменение привело к проблемам, вызванным несопоставимостью данных временных рядов для двух периодов - 2000-2010 и 2011-2015 гг. Для периода с 2011 по 2015 г. длина ряда составляет всего пять лет, что не позволяет использовать классические статистические методы факторного прогнозирования. Даже введение фиктивных переменных, позволяющих включить в модель фактор изменения методики учета, удлинит ряд до 16 лет, не дает возможности провести многофакторный регрессионный анализ, так как будет нарушено одно из основных его условий: для получения надежных оценок коэффициен-

<sup>2</sup> Росстат. Внутрироссийская миграция по территориям прибытия и выбытия («шахматка» по федеральным округам).

тов регрессии число объясняющих переменных должно быть не менее чем в шесть раз меньше числа наблюдений [19, с. 98].

В сложившихся условиях целесообразно, по нашему мнению, применять методы бесфакторного адаптивного прогнозирования, основывающиеся на исторических данных. Наиболее распространенными в социально-экономических исследованиях методами прогнозирования, позволяющими учесть не только прошлые значения, но и ошибки при прогнозе, являются метод экспоненциального сглаживания и интегрированная модель авторегрессии и скользящего среднего ARIMA [20]. *Метод экспоненциального сглаживания* характеризуется использованием для прогноза всех предыдущих наблюдений с разным экспоненциально убывающим весом. *Метод ARIMA* основан на использовании в прогнозе некоторого количества прошлых значений, определяемого исходя из автокорреляционных зависимостей ряда, и прошлых значений величины ошибки (остатков).

Модели типа ARIMA широко используются при прогнозировании внешней миграции в странах Европы [21, р. 776-777]. Впервые такой подход был использован в Голландии в 1997 г. с параметром  $AR = 1$  без применения параметра скользящего среднего. В 1998 г. в Финляндии впервые прогноз внешней миграции был осуществлен на основе модели ARIMA (0;1;1)<sup>3</sup> с логарифмированием, так как использовались помесечные данные для построения модели [22]. С 2001 г. прогноз прибытий в Норвегию осуществляется на основе модели ARIMA (1;0;1), в большинстве случаев используемой в проведенном нами исследовании, но с логарифмированием, так как для временного ряда характерна месячная сезонность [23, 24]. К 2004 г. такие модели применялись уже в 18 странах Европы, и их популярность постоянно растет [22, р. 38]. Прогнозирование международных прибытий с учетом автокорреляции используется также в Австралии и Новой Зеландии [25]. Нами было проанализировано 28 рядов данных о внешней миграции в страны Европейского союза<sup>4</sup> и три ряда - в государства - члены Североамериканского соглашения о свободной торговле (НАФТА) длиной по 15 лет и определены значимые коэффициенты автокорреляции в 26 рядах при условии

стационарности всех рядов (в широком смысле), что еще раз подтверждает потенциальную применимость метода ARIMA для прогнозирования внешних прибытий и выбытий. Однако временные ряды по международной миграции стран Северной Америки, а именно США и Канады, не обладают автокорреляционными связями, что противоречит результатам исследования по крупным странам-реципиентам из Европейского союза (Франция, Великобритания, Германия).

Модели ARIMA широко используются не только при прогнозировании миграции, но и при построении прогнозных моделей других демографических индикаторов. Так, например, данный метод прошел апробацию и получил высокую оценку при краткосрочном прогнозировании смертности в Соединенных Штатах Америки [26], при исследовании фертильности в Норвегии [27]. Основное преимущество подхода к прогнозированию на основе ARIMA, по мнению специалистов, состоит в возможности быстрого реагирования на изменение / уточнение данных и статистические выбросы (их учет возможен в некоторых пакетах обработки статистических данных, например SPSS). С учетом всего сказанного, в качестве основного метода исследования был избран метод прогнозирования на основе модели ARIMA.

### 3. Результаты

Для проверки гипотезы об инерционности внутренней миграции и определения параметров моделей ARIMA, используемых для каждого временного ряда, нами были рассчитаны коэффициенты автокорреляции первого и второго порядка по всем 256 рядам с 2000 по 2015 г. Отметим, что изменение учета миграции в 2011 г. могло несколько исказить значения коэффициентов автокорреляции в сторону занижения по причине их расчета, в том числе и на переходном периоде 2010-2011 гг., но все же не привело к отрицанию автокорреляционных связей, свойственных ряду, что объясняется сопоставимостью остальных составляющих временных рядов.

Результаты расчетов коэффициентов автокорреляции первого и второго порядка представлены в таблицах 1 и 2.

<sup>3</sup> Здесь и далее первая цифра в скобках означает порядок автокорреляции  $AR(p)$ , вторая - коэффициент учета наличия тренда в случае нестационарности ряда ( $d$ ), третья - порядок скользящей средней  $MA(q)$ .

<sup>4</sup> Immigration by age and sex. URL: [http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=migr\\_imm8&lang=en](http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=migr_imm8&lang=en).

Таблица 1

**Значения коэффициентов автокорреляций 1-го порядка, значимых на уровне  $\alpha = 0,05$ , по рядам количества прибытий и выбытий в федеральные округа Российской Федерации**

Регион выбытия  Регион прибытия - ФО		Федеральный округ (ФО)															
		Центральный		Северо-Западный		Южный		Северо-Кавказский		Приволжский		Уральский		Сибирский		Дальневосточный	
		город	село	город	село	город	село	город	село	город	село	город	село	город	село	город	село
Центральный	город	0,79	0,79	0,78	0,77	0,79	0,77	0,74	0,73	0,76	0,76	0,77	0,78	0,75	0,74	0,79	0,74
	село	0,83	0,82	0,83	0,75	0,85	0,76	0,79	0,73	0,80	0,78	0,81	0,64	0,81	0,69	0,80	0,54
Северо-Западный	город	0,80	0,80	0,82	0,77	0,82	0,80	0,76	0,75	0,78	0,79	0,77	0,77	0,75	0,75	0,77	0,75
	село	0,82	0,79	0,83	0,81	0,84	0,82	0,80	0,78	0,80	0,77	0,78	0,75	0,79	0,80	0,77	0,44
Южный	город	0,80	0,74	0,81	0,76	0,81	0,73	0,75	0,79	0,80	0,82	0,81	0,81	0,80	0,75	0,81	0,79
	село	0,81	0,75	0,82	0,74	0,81	0,79	0,70	0,77	0,76	0,74	0,77	0,71	0,77	0,71	0,78	0,73
Северо-Кавказский	город	0,79	0,64	0,82	0,77	0,83	0,77	0,74	*	0,77	0,75	0,79	0,50	0,78	0,52	0,76	0,66
	село	0,80	0,74	0,84	0,80	0,83	0,80	0,72	0,84	0,83	0,71	0,84	*	0,83	0,65	0,79	0,80
Приволжский	город	0,80	0,76	0,80	0,73	0,82	0,78	0,60	0,70	0,80	0,76	0,79	0,66	0,79	0,52	0,78	0,60
	село	0,82	0,75	0,83	0,77	0,84	0,55	0,72	0,73	0,84	0,81	0,82	0,66	0,83	0,50	0,78	*
Уральский	город	0,80	0,81	0,77	0,78	0,82	0,78	0,70	0,81	0,82	0,83	0,82	0,79	0,83	0,81	0,77	0,60
	село	0,81	0,50	0,79	0,69	0,80	0,72	0,72	0,69	0,75	0,68	0,82	0,78	0,79	0,66	0,70	*
Сибирский	город	0,79	0,75	0,78	0,68	0,82	0,80	0,56	0,83	0,78	0,79	0,79	0,73	0,81	0,70	0,77	0,70
	село	0,82	0,48	0,83	0,49	0,80	0,71	*	*	0,68	0,60	0,80	0,59	0,83	0,75	0,76	*
Дальневосточный	город	0,78	0,59	0,81	0,76	0,84	0,83	0,77	0,76	0,70	0,79	0,60	*	0,82	0,80	0,78	0,74
	село	0,67	*	0,85	*	0,83	0,80	0,64	*	0,45	0,45	0,63	*	0,82	*	0,76	0,81

Примечание: «\*» - коэффициент автокорреляции не является значимым на уровне  $\alpha = 0,05$ .

Источник: расчеты автора на основе данных Росстата: URL: [www.gks.ru/free\\_doc/new\\_site/population/demo/migr3.xls](http://www.gks.ru/free_doc/new_site/population/demo/migr3.xls).

Таблица 2

**Значения коэффициентов автокорреляций 2-го порядка, значимых на уровне  $\alpha = 0,05$ , по рядам количества прибытий и выбытий в федеральные округа Российской Федерации**

Регион выбытия  Регион прибытия - ФО		Федеральный округ (ФО)															
		Центральный		Северо-Западный		Южный		Северо-Кавказский		Приволжский		Уральский		Сибирский		Дальневосточный	
		город	село	город	село	город	село	город	село	город	село	город	село	город	село	город	село
Центральный	город	0,55	0,54	0,53	0,49	0,54	0,49	0,47	0,47	0,48	0,50	0,48	0,52	0,46	0,47	0,49	0,44
	село	0,56	0,54	0,57	0,44	0,57	0,44	0,50	0,51	0,49	0,45	0,48	*	0,54	*	0,51	*
Северо-Западный	город	0,55	0,52	0,55	0,53	0,57	0,53	0,45	0,47	0,49	0,51	0,45	0,50	0,45	0,44	0,46	0,47
	село	0,55	0,50	0,56	0,55	0,58	0,56	0,57	0,51	0,50	0,48	0,49	*	0,52	0,53	0,46	*
Южный	город	0,57	0,44	0,56	0,45	0,52	*	*	*	0,49	0,50	0,47	0,46	0,51	0,46	0,50	0,54
	село	0,57	*	0,57	*	0,53	0,48	*	*	0,44	*	0,44	*	0,50	*	0,49	*
Северо-Кавказский	город	0,59	*	0,59	0,46	0,57	0,43	*	*	*	*	0,44	*	0,44	*	*	*
	село	0,59	*	0,61	0,49	0,57	*	*	0,60	0,52	*	0,56	*	0,58	*	0,52	0,50
Приволжский	город	0,58	0,50	0,56	0,44	0,56	0,49	*	0,50	0,44	0,45	0,45	*	0,48	*	0,46	*
	село	0,59	0,45	0,59	*	0,58	*	0,47	0,54	0,54	0,52	0,53	*	0,54	*	0,48	*
Уральский	город	0,57	0,51	0,58	0,48	0,59	0,52	*	0,55	0,51	0,53	0,50	0,46	0,56	0,48	0,45	*
	село	0,55	*	0,56	*	0,53	*	0,48	0,48	*	*	0,48	*	0,47	*	*	*
Сибирский	город	0,55	*	0,54	*	0,58	0,51	*	0,52	*	*	0,47	*	0,51	*	0,44	*
	село	0,56	*	0,54	*	0,58	*	*	*	*	*	0,47	*	0,55	*	0,45	*
Дальневосточный	город	0,46	*	0,56	0,48	0,58	0,52	0,44	0,50	*	0,50	*	*	0,53	0,47	0,44	0,44
	село	*	*	0,55	*	0,62	0,52	*	*	*	*	*	*	0,55	*	0,44	0,51

Примечание: «\*» - коэффициент автокорреляции не является значимым на уровне  $\alpha = 0,05$ .

Источник: расчеты автора на основе данных Росстата: URL: [www.gks.ru/free\\_doc/new\\_site/population/demo/migr3.xls](http://www.gks.ru/free_doc/new_site/population/demo/migr3.xls).

Таким образом, только в 13 (причем в 12 из этих 13 случаев местом прибытия стали сельские поселения) из 256 временных рядов было выявлено одновременное отсутствие значимых на уровне 0,05 автокорреляций первого и второго порядка. Для рядов «село-село» и «город-село» значимые автокорреляционные связи часто отмечаются лишь на первом лаге, тогда как для 71,2% всех анализируемых в исследовании рядов число мигрантов зависит от двух предыдущих лет. Для межгородской миграции во всех 64 случаях коэффициенты автокорреляций оказались значимыми на уровне 0,05; при этом в 54 случаях - при лагах в один и два года, а в остальных лишь при лаге в один год. Если для миграции в город вне зависимости от типа поселения исхода мигрантов в большинстве регионов отмечаются значимые коэффициенты корреляции на первых двух лагах, то для миграции в городские поселения Северного Кавказа в 50% рядов (в 8 из 16) они отмечаются только на первом лаге, а в одном случае не наблюдаются вовсе. Таким образом, Северо-Кавказский федеральный округ стал некоторым исключением и в изучении автокорреляционных связей внутри федерального округа: значимые коэффициенты наблюдаются лишь в трех случаях из четырех; при этом в двух из них только при лаге в один год.

Все вышесказанное говорит о том, что гипотеза об инерционности находит статистическое подтверждение не только для внешней, но и для внутренней миграции.

Определение параметров для моделей ARIMA производилось на основе анализа графиков автокорреляций и частных автокорреляций, а также их сравнения с теоретическими значениями коэффициентов, предложенных Дж. Боксом и Г. Дженкинсом - разработчиками этого адаптивного метода прогнозирования [20, 28].

Данные полученных графиков свидетельствуют о возможности использования для большинства моделей параметра  $AR(p)$ , равного 1 или 2. Указанный параметр отражает порядок автокорреляции, являющейся значимой в каждом конкретном случае. Решение об окончательном выборе параметров проводилось по результатам сопоставления качества двух моделей. Для рядов со значимой автокорреляцией на первом лаге применим параметр  $AR(p) = 1$ , а для рядов без зависимости от прошлых значений -  $AR(p) = 0$ .

Выбор параметра  $MA(q)$ , определяющего количество прошлых значений, используемых для

расчета скользящей средней, включаемой в прогноз, показал, что для большинства рядов (84,4%) параметр  $MA(q)$  должен составить 1; в 16 случаях (6,3% от общего числа) он составил 0; для 24 рядов окончательное решение о значении коэффициента было принято после сопоставления качества моделей при  $MA(q) = 1$  и  $MA(q) = 2$ . Параметр  $MA(q)$  позволяет учитывать ошибки прошлых периодов, что дает возможность «сглаживать» предсказываемые значения. Анализ графиков автокорреляции показал, что ни в одном из 256 рядов не отмечается наличие линейного или нелинейного тренда, что говорит о стационарности рядов. Таким образом, значение параметра  $d$  для всех рядов было принято равным 0.

В целях оценки приемлемости использования моделей типа ARIMA для прогнозирования внутренней миграции России было проведено ретроспективное прогнозирование. Сводная характеристика значимых спецификаций моделей ARIMA приведена в таблице 3.

Таблица 3

Распределение моделей ARIMA по типу спецификаций

Параметры модели ARIMA	Число рядов, для которых данная модель значима, ед.	Доля рядов, для которых данная модель значима, в %
$p = 1, d = 0, q = 1 (1;0;1)$	172	67,2
$p = 2, d = 0, q = 1 (2;0;1)$	66	25,8
$p = 0, d = 0, q = 1 (0;0;1)$	5	1,9
$p = 0, d = 0, q = 2 (0;0;2)$	1	0,4
Модель типа ARIMA незначима	12	4,7
Итого	256	100,0

Источник: расчеты автора на основе данных Росстата: URL: [www.gks.ru/free\\_doc/new\\_site/population/demo/migr3.xls](http://www.gks.ru/free_doc/new_site/population/demo/migr3.xls).

Таким образом, в большинстве случаев (67,2% от общего числа) наиболее распространенным набором параметров значимых моделей был ( $p = 1; d = 0; q = 1$ ). Во всех случаях, когда модель была значима, коэффициент  $d$  был равен 0. Это говорит о том, что все значимые модели основываются на одновременном учете автокорреляции и скользящего среднего (ARMA-модели) или только скользящего среднего (MA-модели) без тренда. Для 12 рядов (то есть 4,7% от их общего числа) модель типа ARIMA оказалась незначимой, что привело к необходимости поиска другого метода прогнозирования для этих данных. Оценка адекватности каждой модели производилась на

основе анализа остатков и  $R^2$ . Модели экспоненциального сглаживания для данных рядов также оказались незначимыми, и прогнозирование осуществлялось по средней арифметической за 2011-2015 гг. Средние относительные ошибки прогнозирования (МАРЕ) за 2013-2015 гг. для различных территориальных уровней приведены в таблице 4. Расчет среднего значения ошибки для каждого временного ряда за 2011 и 2012 гг. является нецелесообразным, так как прогнозирование производится на основе данных прошлых периодов (2010 и 2011 гг.), а учет перемещений в 2010 г. осуществлялся по другой методике.

Таблица 4

**Средние относительные ошибки прогнозирования (МАРЕ) внутрироссийской миграции на 2013-2015 гг. (в процентах)**

Федеральный округ	Тип перемещения	Тип поселения	2013	2014	2015
Центральный	Выезд	Город	2,15	0,49	3,66
		Село	2,21	4,91	1,20
		Население в целом	2,17	1,09	3,13
	Въезд	Город	1,65	1,73	4,05
		Село	6,71	2,95	5,95
		Население в целом	0,40	0,59	4,50
Северо-Западный	Выезд	Город	5,75	0,90	7,73
		Село	5,48	5,84	1,81
		Население в целом	5,68	2,04	5,18
	Въезд	Город	9,57	0,17	5,72
		Село	9,76	1,75	10,33
		Население в целом	9,61	0,21	6,66
Южный	Выезд	Город	5,58	3,79	3,58
		Село	9,56	2,01	3,38
		Население в целом	7,22	3,09	5,41
	Въезд	Город	14,42	6,76	7,50
		Село	11,77	5,18	1,39
		Население в целом	13,51	2,26	5,19
Северо-Кавказский	Выезд	Город	5,74	1,34	4,00
		Село	6,26	3,88	5,87
		Население в целом	6,01	2,61	7,59
	Въезд	Город	9,37	6,00	0,87
		Село	3,32	1,67	2,29
		Население в целом	6,63	4,08	1,51

Окончание таблицы 4

Федеральный округ	Тип перемещения	Тип поселения	2013	2014	2015
Приволжский	Выезд	Город	4,03	3,68	4,47
		Село	9,48	0,55	0,90
		Население в целом	6,41	1,88	3,87
	Въезд	Город	6,11	4,79	3,06
		Село	4,26	3,25	5,38
		Население в целом	5,40	4,21	0,26
Уральский	Выезд	Город	6,44	1,01	2,55
		Село	9,16	4,40	3,35
		Население в целом	7,20	0,42	4,10
	Въезд	Город	7,14	1,20	0,20
		Село	6,63	1,77	5,66
		Население в целом	7,02	0,54	1,45
Сибирский	Выезд	Город	6,42	1,62	7,53
		Село	8,79	2,49	3,36
		Население в целом	7,33	1,95	3,67
	Въезд	Город	5,50	1,03	3,46
		Село	4,98	1,12	10,05
		Население в целом	5,33	1,06	5,62
Дальневосточный	Выезд	Город	7,02	0,76	5,67
		Село	9,61	0,96	5,12
		Население в целом	7,81	0,82	4,20
	Въезд	Город	9,75	1,41	7,39
		Село	3,19	5,51	4,99
		Население в целом	8,12	0,38	6,75

Источник: расчеты автора на основе данных Росстата: URL: [www.gks.ru/free\\_doc/new\\_site/population/demo/migr3.xls](http://www.gks.ru/free_doc/new_site/population/demo/migr3.xls).

Средние относительные ошибки, как известно, зависят от масштаба абсолютных величин, именно поэтому зачастую максимальные значения таких ошибок отмечаются в миграции из сельской или в сельскую местность, минимальный абсолютный показатель (число выбытий) для которых в 2015 г. составил 104, тогда как для города - почти в девять раз больше - 903. Различия относительных ошибок прогноза для въездной и выездной миграции по городской и сельской местности в 2015 г. отображены на рис. 1 и 2.

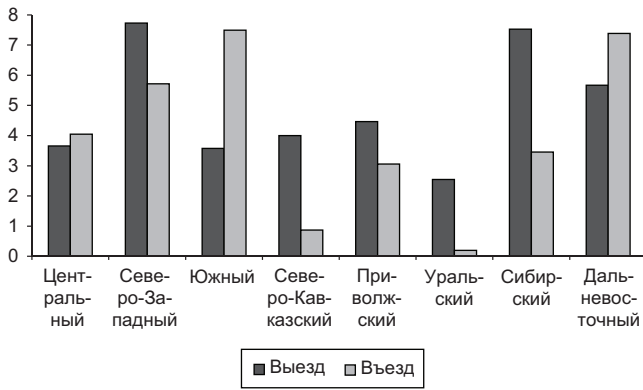


Рис. 1. Средние относительные ошибки прогноза (MAPE) для городской миграции России за 2015 г. по типу перемещения (в процентах)

Источник: расчеты автора на основе данных Росстата: URL: [www.gks.ru/free\\_doc/new\\_site/population/demo/migr3.xls](http://www.gks.ru/free_doc/new_site/population/demo/migr3.xls).

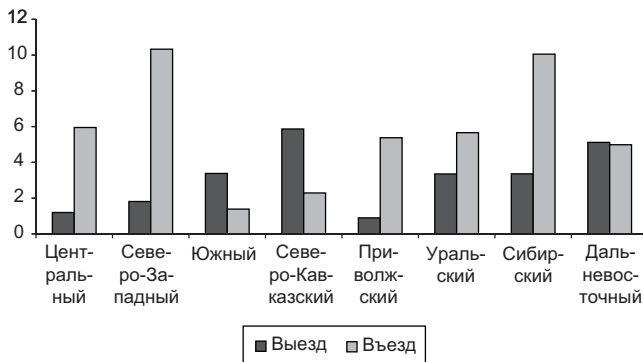


Рис. 2. Средние относительные ошибки прогноза (MAPE) для сельской миграции России за 2015 г. по типу перемещения (в процентах)

Источник: расчеты автора на основе данных Росстата: URL: [www.gks.ru/free\\_doc/new\\_site/population/demo/migr3.xls](http://www.gks.ru/free_doc/new_site/population/demo/migr3.xls).

Данные таблицы 4 и рис. 2 дают основания полагать, что ошибки прогноза въездов в сельские поселения выше, чем выездов из них. Однако на основе сравнения доверительных интервалов средней можно утверждать, что статистически значимых различий между ошибками нет. Нами было также проведено аналогичное сравнение средних относительных ошибок прогноза для городских и сельских поселений и выявлено, что различия статистически незначимы на уровне 0,05. Это говорит о том, что модели ARIMA в равной степени качественно прогнозируют и въездную, и выездную миграцию, независимо от типа поселения.

Ошибки прогнозирования на уровне федеральных округов, как видно из данных таблицы 4, варьируют от 0,26% по въездной миграции (Приволжский федеральный округ) до 7,59% по выездной миграции (Северо-Кавказский ФО). Таким образом, миграционная ситуация Северо-Кавказского федерального округа во многом уникальна, что было выявлено еще при анализе автокорреляций.

На основе полученных данных была рассчитана интенсивность миграции на уровне Российской Федерации как сумма предсказанных значений перемещений по федеральным округам. Ошибка прогноза для 2013 г. составила 5,8%, для 2014 г. - 0,6, для 2015 г. - 3,9%.

Для сравнительной оценки эффективности методов бесфакторного адаптивного прогнозирования был построен ретроспективный прогноз внутренней миграции в России методом экспоненциального сглаживания. Анализ значений  $R^2$  и графиков остатков каждой модели показал пригодность данного метода для прогнозирования лишь для 171 временного ряда из 256, то есть в 66,8% случаев. Для остальных рядов прогноз осуществлялся по средней арифметической, рассчитываемой на основе данных второй части каждого временного ряда за 2011-2015 гг. После этого были вычислены ретроспективные прогнозные значения миграционного оборота Российской Федерации и определены средние относительные ошибки прогнозирования (MAPE) в 2013-2015 гг. В 2013 г. ошибка прогноза оборота внутренней миграции составила 7,8%, в 2014 г. - 6,2, в 2015 г. - 7,7%. Таким образом, сравнение величины ошибок прогнозирования свидетельствует о том, что при прогнозировании внутренней миграции в России более целесообразно использование метода ARIMA, чем метода экспоненциального сглаживания.

Целесообразность использования метода ARIMA для прогнозирования внутренней миграции была оценена также на основе данных Республики Беларусь за 2000-2016 гг.<sup>5</sup> Ввиду малой длины рядов по областям проверка осуществлялась по данным внутренних перемещений на уровне всей страны в целом. Средняя относительная ошибка (MAPE) модели составила 2,1%, а средняя абсолютная ошибка (MAE) - 4716 человек. Параметры  $p$ ,  $d$ ,  $q$  модели ARIMA для

<sup>5</sup> Белстат. Общие итоги миграции населения. URL: [http://www.belstat.gov.by/ofitsialnaya-statistika/socialnaya-sfera/demografiya\\_2/g/obschie-itogi-migratsii-naseleniya-respubliki-belarus/](http://www.belstat.gov.by/ofitsialnaya-statistika/socialnaya-sfera/demografiya_2/g/obschie-itogi-migratsii-naseleniya-respubliki-belarus/).



прогноза внутренних перемещений в Республике Беларусь составили 2; 0; 2 соответственно, что говорит о значимой автокорреляции второго порядка ( $p$ ), отсутствии тренда ( $d$ ) и количестве прошлых значений, используемых для расчета скользящей средней ( $q$ ), равном двум. Низкая ошибка прогноза обеспечена также и тем, что в стране, в отличие от России, методика учета на протяжении 2000–2016 гг. не менялась.

#### 4. Заключение

Проведенное исследование свидетельствует об инерционности миграционных потоков между субъектами Российской Федерации. Такая инерционность, на наш взгляд, обусловлена факторами, действующими как на уровне индивида, так и на мезо- либо макроуровнях. На индивидуальном уровне решение о перемещении места жительства внутри страны, как и в случае внешней миграции, во многом зависит от имеющегося у потенциального мигранта социального капитала. Любое перемещение связано с большим числом рисков, однако наличие в регионе въезда знакомых, друзей или родственников позволяет прибывшему легче найти работу и жилье. Это в свою очередь снижает его опасения возможных неудач и обеспечивает более оптимистическую оценку перспектив. На мезо- и макроуровнях инерционность может определяться как постоянством влияния привлекающих и выталкивающих факторов, незначительными темпами изменения численности населения регионов выбытия и прибытия, так и рядом других причин. Точное определение силы их влияния представляет собой отдельную задачу, требующую дополнительных расчетов, выходящих за рамки данной статьи.

Полученные результаты свидетельствуют о целесообразности использования моделей ARIMA для краткосрочного прогнозирования внутренней миграции в России. В настоящее время использование метода ARIMA для перспективного прогнозирования внутрироссийской миграции возможно лишь на 2–3 года вперед, что обусловлено малой длиной временных рядов, вызванных относительно недавним изменением методики учета. Однако по мере удлинения сопоставимых отрезков временных рядов параметры модели могут пересматриваться, а горизонт прогнозирования несколько увеличиваться.

#### Литература

1. **Моисеенко В.М.** Внутренняя миграция населения. М.: ТЭИС, 2004. 285 с.
2. **Зайончковская Ж.А.** Внутренняя миграция в России и в СССР в XX веке как отражение социальной модернизации // Мир России. 1999. № 4. С. 22–34.
3. **Вакуленко Е.С., Мкртчян Н.В., Фурманов К.К.** Моделирование регистрируемых миграционных потоков между регионами Российской Федерации // Прикладная эконометрика. 2011. Т. 21. № 1. С. 35–55.
4. **Василенко П.В.** Применение гравитационной модели для анализа внутриобластных миграций на примере Новгородской и Псковской областей // Псковский регионологический журнал. 2013. № 15. С. 83–90.
5. **Соболева С.В.** Демографические процессы в региональном социально-экономическом развитии. Новосибирск, 1988.
6. **Изард У.** Методы регионального анализа: введение в науку о регионах. М.: Прогресс, 1966. 660 с.
7. **Balan M., Vasile V., Stanescu M.** Gravity model application to estimate Romanian migration size // Knowledge horizons. 2013. No. 2. P. 74–78.
8. **Taylor J.E.** Differential migration, networks, information and risk // O. Stark (ed.). Migration, human capital, and development. Greenwich, CT, JAI Press, 1986. P. 147–171.
9. **Massey D.S.** et al. Theories of international migration: Review and appraisal // Population and Development Review. 1993. Vol. 19. P. 431–466.
10. **Hatton T.J., Williamson J.G.** The age of mass migration: Causes and economic impact. Oxford, Oxford University Press, 1998.
11. **Faist T.** The volume and dynamics of international migration and transnational social spaces. Oxford, Oxford University Press, 2000.
12. **Pries L.** Migration and transnational social spaces. Ashgate, Aldershot, 1999.
13. **Сталкер П.** Международная миграция / пер. с англ. А. Захарова. М.: Книжный Клуб Книговек, 2015. 192 с.
14. IMF/World Bank. Market access for developing countries' exports. Paper prepared by the staffs of the IMF and the World Bank, 2001.
15. Virtual champions: India's IT stars are still rising fast // The Economist, 1 June, 2006.
16. **Freeman R., Oostendorp R.** Wages around the world: Pay across occupations and countries. Working Paper 8058. Cambridge, MA, National Bureau of Economic Research, 2000.
17. **Öberg S., Wils A.B.** East-West migration in Europe. Can migration theories help estimate the numbers // Popnet. 1992. Vol. 22. P. 1–7.
18. **Kupiszewski M.** Modelowanie dynamiki przemian ludności w warunkach wzrostu znaczenia migracji międzynarodowych. Warszawa, IGiPZ PAN, 2002.

19. Статистика: учеб. / под ред. И.И. Елисеевой. М.: ООО «Прспект», 2010. 448 с.
20. Бокс Дж., Дженкинс Г. Анализ временных рядов / пер. с англ. А. Левшина. М.: Мир, 1974. 406 с.
21. Greene W.H. Econometric analysis. Upper Saddle River, N. J., Prentice Hall, 2000.
22. Alho J. A stochastic forecast of the population of Finland. Reviews 1998/4. Helsinki, Statistics Finland, 1998.
23. Keilman N., Pham D.Q. Empirical errors and predicted errors in fertility, mortality and migration forecasts in the European Economic Area. Discussion Paper No. 386. Social and Demographic Research. Oslo, Statistics Norway, 2004.
24. Keilman N., Pham D.Q., Hetland A. Norway's uncertain demographic future. Oslo, Statistics Norway, 2001.
25. Gorbey S., James D., Poot J. Population forecasting with endogenous migration: An application to trans-Tasman migration // International Regional Science Review. 1999. Vol. 22. P. 69-101.
26. Lawrence R.C. Forecasting U.S. mortality: A comparison of Box-Jenkins ARIMA and structural time series models // The Sociological Quarterly. 1996. Vol. 37. P. 127-144.
27. Keilman N., Pham D.Q., Hetland A. Why population forecasts should be probabilistic - illustrated by the case of Norway // Demographic Research. 2002. Vol. 36. P. 409-454.
28. Ханк Дж., Уичерн Д., Райтс Дж. Бизнес-прогнозирование / пер. с англ. М.: Издательский дом «Вильямс», 2003. 656 с.

## ARIMA MODELS IN THE SHORT-TERM FORECASTING OF INTERNAL MIGRATION IN RUSSIA

Egor V. Pavlovskij

Author affiliation: Saint-Petersburg State University of Economics (Saint-Petersburg, Russia).

E-mail: pavlovskij.egor@gmail.com.

Based on the data on inflows and outflows for the federal districts and types of settlements in the Russian Federation in 2000-2015 a statistical test of the hypothesis of the inertia of internal migration was carried out. The coefficients of autocorrelation were determined and autocorrelation dependence was found in 243 out of 256 time series, which made it possible to conclude that there is inertia in the internal migration of Russia. On the example of analysis of 28 time series of data on external migration to the countries of the European Union and 3 member states of the North American Free Trade Agreement for 2000-2015, the coefficients of autocorrelation were calculated and their statistical significance was estimated. It allowed to verify and confirm the potential applicability of the ARIMA method for short-term forecasting of international migration.

The results of the selection of ARIMA models parameters for time series of internal migration indicators in Russia are summarized. Based on the retrospective forecasts of internal migration in Russia for the period 2013-2015, the forecasting errors using ARIMA models and the method of exponential smoothing are compared. It is concluded that ARIMA models provide more accuracy of short-term adaptive forecasts of internal migration in Russia in comparison with the method of exponential smoothing.

Based on the calculation of confidence intervals for the mean, it has been demonstrated that the ARIMA models with the same accuracy predict inflows and outflows for two types of settlements (urban and rural). It was proved that ARIMA models can be used for short-term forecasting of internal migration of other states (on the example of data on the volumes of internal migration of the Republic of Belarus for 2000-2016).

*Keywords:* migration, inertia of migration, autocorrelation, internal migration, forecasting, exponential smoothing, autoregressive integrated moving average ARIMA.

*JEL:* C53, J62.

## References

1. Moiseenko V.M. *Vnutrennyaya migratsiya naseleniya* [Internal migration of the population]. Moscow, TEIS Publ., 2004. 285 p. (In Russ.).
2. Zaionchkovskaya Zh.A. *Vnutrennyaya migratsiya v Rossii i v SSSR v XX veke kak otrazhenie sotsial'noi modernizatsii* [Internal migration in Russia and the USSR in the XXth century as a reflection of social modernization]. *Universe of Russia*, 1999, no. 4, pp. 22-34. (In Russ.).
3. Vakulenko E.S., Mkrtchyan N.V., Furmanov K.K. *Modelirovanie registriruemyykh migratsionnykh potokov mezhdru regionami Rossiiskoi Federatsii* [Modeling registered migration flows between regions of the Russian Federation]. *Applied Econometrics*, 2011, vol. 21, iss. 1, pp. 35-55. (In Russ.).
4. Vasilenko P.V. *Primenenie gravitatsionnoi modeli dlya analiza vnutrioblastnykh migratsii na primere Novgorodskoi i Pskovskoi oblasti* [Applying gravity model to interregional migration analysis in relation to Novgorod and Pskov region]. *Pskovskii regionologicheskii zhurnal*, 2013, no. 15, pp. 83-90. (In Russ.).
5. Soboleva S.V. *Demograficheskie protsessy v regional'nom sotsial'no-ekonomicheskom razviti* [Demographic processes in the regional socio-economic development]. Novosibirsk, 1988. (In Russ.).

6. **Isard W.** *Methods of regional analysis: An introduction to regional science*. Cambridge, 1960 (Russ. ed. Izard U. *Metody regional'nogo analiza: vvedenie v nauku o regionakh*. Moscow, Progress Publ., 1966. 660 p.).
7. **Balan M., Vasile V., Stanescu M.** Gravity model application to estimate Romanian migration size. *Knowledge horizons*, 2013, no. 2, pp. 74-78. (In Russ.).
8. **Taylor J.E.** Differential migration, networks, information and risk. In: O. Stark (ed.). *Migration, human capital, and development*. Greenwich, CT, JAI Press, 1986. P. 147-171.
9. **Massey D.S.** et al. Theories of international migration: Review and appraisal. *Population and Development Review*, 1993, vol. 19, pp. 431-466.
10. **Hatton T.J., Williamson J.G.** *The age of mass migration: Causes and economic impact*. Oxford, Oxford University Press, 1998.
11. **Faist T.** *The volume and dynamics of international migration and transnational social spaces*. Oxford, Oxford University Press, 2000.
12. **Pries L.** *Migration and transnational social spaces*. Ashgate, Aldershot, 1999.
13. **Stalker P.** *The no-nonsense guide to international migration*. Oxford, New Internationalist, 2008 (Russ. ed. Stalker P. *Mezhdunarodnaya migratsiya*. Moscow, Knizhnyi Klub Knigovek Publ., 2015. 192 p.).
14. IMF/World Bank. *Market access for developing countries' exports*. Paper prepared by the staffs of the IMF and the World Bank, 2001.
15. Virtual champions: India's IT stars are still rising fast. *The Economist*, 1 June, 2006.
16. **Freeman R., Oostendorp R.** *Wages around the world: Pay across occupations and countries*. Working Paper 8058. Cambridge, MA, National Bureau of Economic Research, 2000.
17. **Öberg S., Wils A.B.** East-West migration in Europe. Can migration theories help estimate the numbers. *Popnet*, 1992, vol. 22, pp. 1-7.
18. **Kupiszewski M.** *Modelowanie dynamiki przemian ludności w warunkach wzrostu znaczenia migracji międzynarodowych* [Population dynamics modeling under the increasing importance of international migration]. Warsaw, IG&SO PAS, 2002. (In Polish).
19. **I.I. Eliseeva (ed.)**. *Statistika: ucheb. [Statistics: Textbook]*. Moscow, Prospect LLC Publ., 2010. 448 p.
20. **Box G.E.P., Jenkins G.M.** *Time series analysis. Forecasting and control*. San Francisco: Holden Day., 1970. (Russ. ed. Boks Dzh., Dzhenkins G. *Analiz vremennykh ryadov*. Moscow, Mir Publ., 1974. 406 p.).
21. **Greene W.H.** *Econometric analysis*. Upper Saddle River, N. J., Prentice Hall, 2000.
22. **Alho J.** *A stochastic forecast of the population of Finland*. Reviews 1998/4. Helsinki, Statistics Finland, 1998.
23. **Keilman N., Pham D.Q.** *Empirical errors and predicted errors in fertility, mortality and migration forecasts in the European Economic Area*. Discussion Paper No. 386. Social and Demographic Research. Oslo, Statistics Norway, 2004.
24. **Keilman N., Pham D.Q., Hetland A.** *Norway's uncertain demographic future*. Oslo, Statistics Norway, 2001.
25. **Gorbey S., James D., Poot J.** Population forecasting with endogenous migration: an application to trans-Tasman migration. *International Regional Science Review*, 1999, vol. 22, pp. 69-101.
26. **Lawrence R.C.** Forecasting U.S. mortality: a comparison of Box-Jenkins ARIMA and structural time series models. *The Sociological Quarterly*, 1996, vol. 37, pp. 127-144.
27. **Keilman N., Pham D.Q., Hetland A.** Why population forecasts should be probabilistic - illustrated by the case of Norway. *Demographic Research*, 2002, vol. 36, pp. 409-454.
28. **Hanke J.E., Wichern D.W., Reitsch A.G.** *Business forecasting*. Upper Saddle River, N. J., Prentice Hall, 2000. (Russ. ed. Khank Dzh., Uichern D., Raits Dzh. *Biznes-prognozirovanie*. Moscow, Williams Publ. House, 2003. 656 p.).