

ЭКОНОМИЧЕСКАЯ СОЦИОЛОГИЯ

УДК 314.18 (477)

ВОССТАНОВЛЕНИЕ ВОЗРАСТНЫХ РАСПРЕДЕЛЕНИЙ ОСНОВНЫХ ДЕМОГРАФИЧЕСКИХ ПРОЦЕССОВ

© 2014 г. Сучков А. Н.

Московский государственный университет

Важнейшим условием построения модуля демографических процессов является использование статистической информации, которая нуждается в определенной предварительной обработке. В частности, необходимы половозрастные данные рождаемости, смертности и миграционного прироста, с возможностью улавливать их динамику при изменении ожидаемой продолжительности жизни и среднего возраста матери. Необходимо знать возрастное распределение мигрантов для оценки их влияния на рождаемость в стране. Поэтому необходимо проанализировать возрастные распределения демографических процессов и восстановить их годовые возрастные коэффициенты с использованием как общеизвестных, так и новых авторских методов.

Ключевые слова: рождаемость; смертность; миграция; гамма-распределение; модель Гомперца; средний возраст матери; сплайновая интерполяция.

The most important condition to build a module of demographic processes is using statistical information, which needs some preprocessing. In particular, the required data includes fertility, mortality and migration rates by age and sex, with the ability to capture their dynamics with changes in life expectancy and the average mothers' age. We also have to know the age distribution of migrants to assess their influence in the birth rate in the country. Therefore it is necessary to analyze the age distributions of demographic processes and restore their annual age coefficients, using both well-known and new author's methods.

Key words: fertility; mortality; migration; gamma distribution; the Gompertz model; the average age of the mother; spline interpolation.

Демографическое прогнозирование — крайне актуальная задача в условиях современного демографического состояния России. Прогнозирование и упреждение неблагоприятных демографических тенденций необходимо для развития экономики и страны в целом. Для прогнозирования численности населения используется модель передвижки по возрастам, в которую закладывают динамические меняющиеся показатели рождаемости, смертности и миграции. Все эти процессы характеризуются определенными

возрастными распределениями и меняют возрастную структуру населения.

Статистическими ведомствами данные рождаемости и смертности зачастую приводятся группировками по пять соседних возрастов, однако для построения современных моделей демографического анализа и прогнозирования зачастую оказывается недостаточно, этой слишком общей информации. Особенно критична ситуация для рождаемости, так как для анализа приводятся данные 7–9 чисел. Данные по возрастному распределению миг-

рационального прироста еще более разрозненны. Поэтому, в качестве отдельной части исследования, видится смысл в восстановлении годовых данных для каждого процесса, используя математический аппарат, целью является выработать алгоритм восстановления годовых возрастных распределений и продемонстрировать его на современных российских данных определенного календарного года.

Данные вычисления дадут на два основных важных вывода. Во-первых, будут восстановлены недостающие данные модели. Во-вторых, полученную модель можно использовать в аналитическом виде для подстановки в аналитическую модель демографического прогноза. Для прогнозов на срок до 2100 года следует задуматься о том, как ведут себя повозрастные коэффициенты демографических процессов при увеличении ожидаемой продолжительности жизни и среднего возраста матери при рождении ребенка.

Возрастное распределение смертности

Рассмотрим половозрастные показатели смертности на предмет выявления общей зависимости коэффициентов в разных возрастах для обоих полов. Для описания динамики

смертности удобно использовать модель Гомперца [8]. В возрасте старше 20 лет возрастной показатель смертности $d(a)$ в промилле определяется следующим образом:

$$d(a) = 1000(1 - e^{1 - e^{k_1 + k_2 a}}),$$

где a — возраст человека в годах; k_1 и k_2 — параметры модели Гомперца.

Проанализируем показатели смертности на предмет линейности следующего отношения $DL(a)$:

$$\begin{aligned} DL(a) &= \text{Ln}(-(\text{Ln}(1 - 0.001 * d(a))) = \\ &= k_1 + k_2 a. \end{aligned}$$

Данная модель имеет небольшое упрощение — дело в том, что $\text{Ln}(1 - 0.001 * d(a))$ приблизительно равен $-0.001 * d(a)$, т. е. показатель пересчитывается до $\text{Ln}(0.001 * d(a))$. То есть зависимость смертности от возраста приближена к экспоненциальной. Данный факт имеет социологическое обоснование: чем старше становится человек, тем больше у него вероятность смертности, причем с течением времени появляются новые болезни, поэтому вероятность смерти возрастает мультипликативно, что и отражает данная экспоненциальная зависимость.

Таблица 1

Апробация модели Гомперца для возрастных коэффициентов смертности женщин в России за 2010 год

Номер возрастной группы	Нижняя граница	Верхняя граница	Смертность $d(a)$	Соотношение $DL(a)$
0	20	24	0.78	-7.1558
1	25	29	1.3	-6.6447
2	30	34	1.93	-6.2493
3	35	39	2.41	-6.0269
4	40	44	3.15	-5.7588
5	45	49	4.18	-5.4754
6	50	54	5.93	-5.1248
7	55	59	9.13	-4.6916
8	60	64	13.1	-4.3286
9	65	69	20.11	-3.8964
10	70	74	32.15	-3.4210
11	75	79	57.81	-2.8210
12	80	84	98.83	-2.2628
13	85	+	188.21	-1.5678

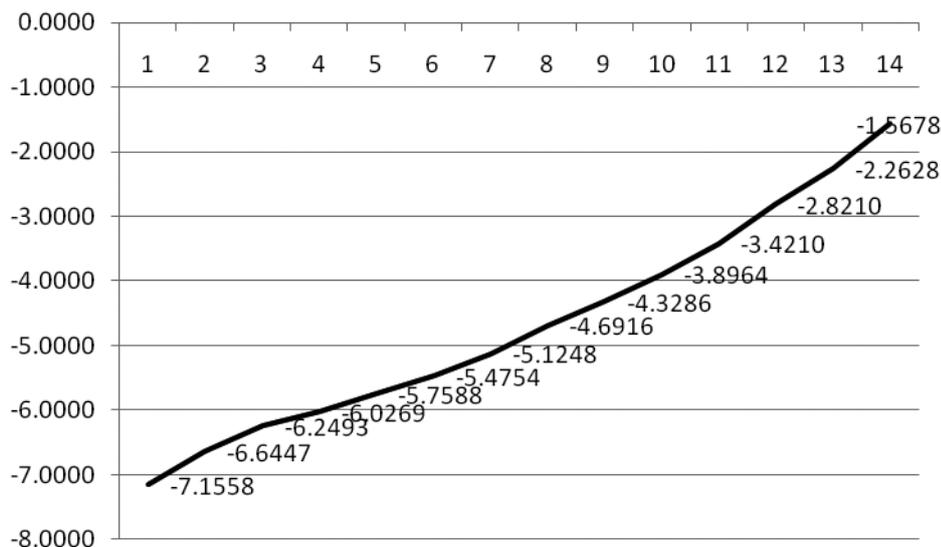


Рис. 1. Значение функции $DL(a)$ в каждой возрастной группе

Коэффициенты модели вычисляются исходя из экспериментальных данных.

Существуют уточнение модели Гомперца, например модель Гомперца — Мекхама

$$d(a) = K + 1000(1 - e^{1 - e^{k_1 + k_2 * a}}),$$

которая вводит промежуточный дополнительную константу K в модель, которая отражает в основном смертность от несчастных случаев, тогда как k_1 , k_2 — от заболеваний, накапливаемых в течение жизни. В модели Гомперца — Мейкхама константу K можно выделить из статистики внешних причин смерти. В России в 2012 году смертности 1,35 промилле.

Для модели возможны и другие уточнения, обусловленные тем, что смертность, как от заболеваний, так и от несчастных случаев, имеет крайне неравномерное распределение по возрасту, зависящее от конкретного заболевания. Можно заметить: по смертности в разные годы половозрастные коэффициенты смертности изменяли по-разному, существуют неоднородности и флуктуации смертности между возрастными группами. Также соотношение между половозрастными коэффициентами различно в разных странах мира. На этот случай предложены типовые модели смертности Брасса [2], региональные модельные таблицы дожития Коула — Демени [9], которые более детально отражают динамику процесса смертности в конкретной стране [4; 11].

Однако модель смертности Гомперца в России в целом хорошо отражает динамику изменения возрастных коэффициентов внутри возрастной группы, достаточно для вывода годичных коэффициентов смертности из пятилетних.

Восстановление проводим по следующему алгоритму сглаживания.

1. Проставление среднего значения смертности по возрастной группе на эту середину возрастной группы (для группы 85+ середину следует взять в возрасте 88 лет с учетом среднего количества лиц данного возраста, а в возрастах 88–99 лет использовать другой коэффициент роста смертности, основанный на дожитии из таблиц численности населения).

2. Интерполяция экспонентой внутри полученных интервалов, а также экстраполяция экспонентой на крайние возраста (1 год, 89–99 лет).

3. Нормирование полученных возрастных коэффициентов с учетом данных по пятилетним (умножение на нужный коэффициент для возрастов каждой возрастной группы для равенства средним значениям).

Возрастное распределение рождаемости

Повозрастная динамика процесса рождаемости весьма специфична. Во-первых, рождаемость в подростковых возрастах (15–19) очень мала, также как рождаемость в поздних возрастах (40–45), однако сам процесс роста рождаемости с ростом в первых и снижения

Таблица 2

**Увеличение смертности при переходе к следующей возрастной группе
в разные годы у мужчин и женщин**

	Мужчины					Женщины				
	1990	1995	2000	2005	2010	1990	1995	2000	2005	2010
15–19	1.61	1.80	2.33	2.39	2.14	1.14	1.22	1.41	1.53	1.47
20–24	1.27	1.27	1.21	1.66	1.76	1.08	1.23	1.18	1.55	1.67
25–29	1.30	1.34	1.17	1.27	1.52	1.36	1.38	1.29	1.37	1.48
30–34	1.30	1.36	1.30	1.26	1.17	1.47	1.44	1.35	1.34	1.25
35–39	1.37	1.41	1.39	1.39	1.23	1.53	1.52	1.45	1.37	1.31
40–44	1.53	1.37	1.41	1.35	1.37	1.59	1.50	1.51	1.37	1.33
45–49	1.38	1.41	1.37	1.38	1.39	1.43	1.47	1.49	1.45	1.42
50–54	1.46	1.25	1.37	1.27	1.41	1.60	1.35	1.50	1.44	1.54
55–59	1.46	1.35	1.33	1.37	1.41	1.57	1.50	1.39	1.41	1.43
60–64	1.36	1.31	1.34	1.25	1.34	1.63	1.51	1.62	1.43	1.54
65–69	1.45	1.28	1.36	1.38	1.37	1.69	1.59	1.61	1.67	1.60
70–74	1.48	1.41	1.31	1.37	1.42	1.68	1.66	1.64	1.67	1.80
75–79	1.46	1.43	1.44	1.31	1.41	1.70	1.68	1.71	1.63	1.71
80–84	1.57	1.54	1.57	1.63	1.44	1.86	1.88	1.92	1.95	1.90
85+						85+				

рождаемости в последних заметно отличается — рост гораздо более активный, а снижение более медленное.

Для анализа рождаемости я использую ГКС и возрастные коэффициенты рождаемости за разные годы, как в городской, так и в сельской среде.

Есть большое количество аналитических моделей для описания рождаемости. Их точ-

ность колеблется в разных диапазонах, для каждого распределения рождаемости.

1) Экспоненциальная модель рождаемости

$$b(t) = A * e^{-b*(t-t_0)}.$$

t_0 обычно берется около 20 лет. Она довольно хорошо аппроксимирует падение рождаемости в поздних возрастах, однако в рождаемость в подростковых возрастах в совре-

Таблица 3

Восстановление возрастных коэффициентов смертности для женщин в 2010 году

0	6.88	10	0.23	20	0.65	30	1.68	40	2.82	50	5.07	60	11.18	70	25.98	80	78.84	90	202.83
1	0.61	11	0.24	21	0.71	31	1.81	41	2.97	51	5.44	61	12.01	71	28.54	81	87.76	91	215.57
2	0.51	12	0.24	22	0.76	32	1.96	42	3.13	52	5.84	62	12.91	72	31.35	82	97.70	92	228.99
3	0.43	13	0.28	23	0.84	33	2.05	43	3.32	53	6.36	63	14.07	73	35.25	83	108.77	93	243.09
4	0.36	14	0.32	24	0.94	34	2.15	44	3.51	54	6.94	64	15.33	74	39.64	84	121.09	94	257.90
5	0.30	15	0.41	25	1.07	35	2.19	45	3.69	55	7.70	65	16.72	75	45.42	85	129.90	95	273.43
6	0.25	16	0.47	26	1.18	36	2.29	46	3.90	56	8.39	66	18.22	76	51.28	86	144.61	96	289.68
7	0.21	17	0.54	27	1.31	37	2.39	47	4.13	57	9.15	67	19.85	77	57.44	87	160.99	97	306.67
8	0.22	18	0.59	28	1.41	38	2.52	48	4.43	58	9.84	68	21.80	78	63.94	88	179.28	98	324.39
9	0.22	19	0.64	29	1.53	39	2.66	49	4.75	59	10.57	69	23.95	79	71.18	89	190.74	99	342.84

Таблица 4

Восстановленные возрастные данные по суммарному коэффициенту рождаемости

Возраст	Рождаемость	Возраст	Рождаемость	Возраст	Рождаемость
15	7.27	25	107.74	35	38.34
16	14.44	26	107.54	36	31.24
17	24.36	27	104.48	37	25.08
18	36.57	28	99.01	38	19.84
19	50.23	29	91.68	39	15.48
20	64.3	30	83.09	40	11.92
21	77.69	31	73.8	41	9.05
22	89.38	32	64.31	42	6.79
23	98.58	33	55.05	43	5.03
24	104.77	34	46.31	44	3.69

менных условиях данная модель не описывает. Она может быть применима в странах, где насторого запрещены ранние рождения или уточняются другой компонентой. Существуют различные вариации экспоненциальной модели, часто рассматривается сумма двух экспонент с разными множителями

$$b(t) = A_0 * e^{-b_0*(t-t_0)} + A_1 * e^{-b_1*(t-t_1)}.$$

2) Модель рождаемости Брасса

$$b(t) = A * (t - t_0)^a * (t_1 - t)^b.$$

Данная модель рассматривает как увеличение рождаемости в ранних возрастах, так и снижение в поздних. В этом и заключается основное преимущество модели перед предыдущей, однако проведенные мной исследования, показали существенные неточности в оценке возрастного распределения при ее использовании.

Таблица 5

Сравнение точности полученных оценок

Возраст	По ГКС	По модели
15–19	27	26.58
20–24	87.5	86.94
25–29	99.2	102.09
30–34	67.3	64.51
35–39	30	25.99
40–44	5.9	7.3

3) Гамма-распределение рождаемости [1; 3]:

$$b(t) = t^{a-1} * \frac{e^{-t/b}}{b^a * \Gamma(a)},$$

где $\Gamma(a)$ — гамма-функция Эйлера от фиксированного параметра a .

Фактически данная модель основана на отношении степенной функции с некоторым коэффициентом к экспоненциальной функции с другим коэффициентом. Это и есть два параметра: параметр формы a и параметр масштаба b . В ранних возрастах рост рождаемости описывается степенной функцией, основанной на показателе масштаба. В поздних возрастах рождаемость убывает по экспоненциальному закону в соответствии с параметром формы.

Именно гамма-распределение лучше всего описывает изменение рождаемости с возрастом, однако и здесь придется сделать несколько уточнений. Во-первых, в качестве аргумента к плотности гамма-распределения следует брать не сам возраст, а некоторую функцию от него. В простом случае можно взять линейную функцию, однако точность такой модели невысокая, и она не описывает возрастное распределение рождаемости от 15 до 45 лет. Во-вторых, сам результат вычисления плотности следует отнормировать согласно исходным возрастным показателям.

В своей модели я формирую аргумент гамма-функции следующим образом: $x = (0.1 * (age - d))^u$, где d — параметр началь-

Таблица 6

Данные сальдо миграции в России за 2009–2011 годы по возрастным группам

Мужчины				Женщины			
Возраст	2009	2010	2011	Возраст	2009	2010	2011
0–5	3454	2459	4130	0–5	3159	2251	3879
6–13	5790	3977	6037	6–13	5209	3665	5427
14–15	2578	1887	2210	14–15	2202	1571	1836
16–17	2630	1881	2832	16–17	2366	1612	2146
18–19	3819	2940	8780	18–19	3550	2291	4245
20–24	19584	13498	34278	20–24	14750	8729	16500
25–29	25020	16781	35408	25–29	16024	9632	17294
30–39	35436	22793	48534	30–39	24854	14634	25666
40–49	20327	12977	32793	40–49	15174	8721	15578
50–54	6462	4353	10593	50–54	7108	4454	7819
55–59	4344	2822	6442	55–59	6220	3954	6967
60–64	2145	1609	3660	60–64	3466	2172	4393
65+	3337	2012	3918	65+	8441	4403	8396
Всего	134926	89989	199615	Всего	112523	68089	120146

ного возраста, u — параметр растяжения. Необходимость введения параметра растяжения обусловлена тем, что при использовании обычной линейной функции в возрастах 30–44 рождаемость по модели убывает слишком медленно, а при выборе u в пределах от 1 до 2 модель серьезно уточняется. Саму модель я строю согласно приведенной формуле

$$\frac{x^a}{e^{bx}}.$$

Для оценивания параметров используется подбор необходимых коэффициентов с помощью компьютерной программы, которая на основании исходных пятилетних данных подбирает данные модели. Параметр d лучше всего приближается 12 годами, $u = 1,65$; $a = 2,18$; $b = 1,19$.

В таблице 5 приведены различия восстановленных коэффициентов рождаемости.

Как видно данные по модели достаточно точны, поэтому данную модель можно использовать в расчетах. Также для уточнения можно провести процедуру, описанную в пункте 3 как при восстановлении смертности — умножить коэффициент внутри возрастной группы на соответствующее отношение

между модельными и исходными коэффициентами возрастной группы.

Половозрастное распределение сальдо миграции

Для процесса миграции половозрастные данные публикуются редко, однако они для прогнозирования они очень необходимы. В демографическом ежегоднике приводятся данные по миграции за последние годы по некоторым возрастным группам, более разрозненные, нежели данные по рождаемости и по смертности. Количество возрастных категорий, как правило, меньше, и распределение обладает меньшей устойчивостью во времени [7].

Проведем восстановление половозрастного распределения мигрантов. Данные получены из демографического ежегодника России 2012 года [6].

Рассмотрим соответствующую таблицу за 2009–2011 годы, проведем анализ по возрастам для мужского и женского населения. Динамика каждого года отличается своей спецификой, поэтому для демографического прогноза будем использовать данные за три года (2009, 2010 и 2011). Доля мужчин среди

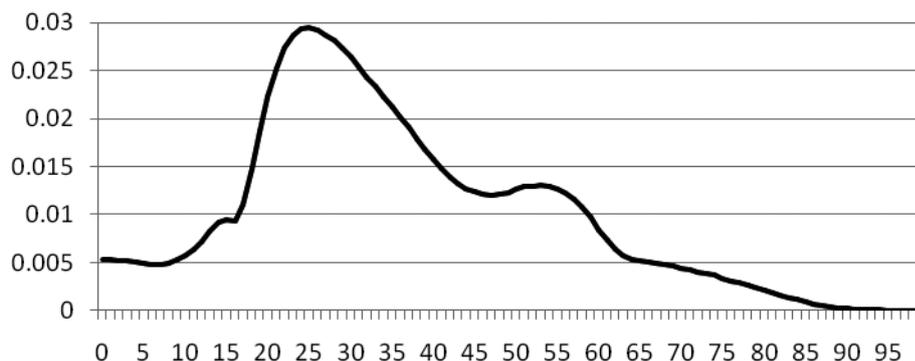


Рис. 2. Плотность возрастного распределения сальдо миграции у женщин

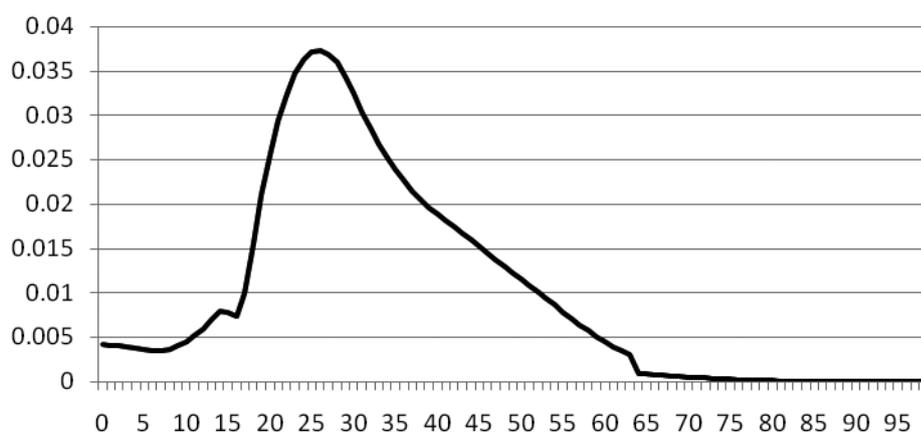


Рис. 3. Плотность возрастного распределения сальдо миграции у мужчин

мигрантов равна 58,5%, доля женщин среди мигрантов соответственно 41,5%.

Возрастные группы распределены следующим образом, в результате чего удалось получить данные, представленные в таблице 6.

Далее, восстановление происходит на основе интерполяции сплайнами удельных долей миграции. В возрастах до 64, соответствующие возрастные группы разделены довольно точно, и данные на этом гладко аппроксимируются. После 65 лет, оказывается необходима экстраполяция более точным способом.

Так как данные по миграции в поздних возрастах хоть и не так точны, используем другую систему. Миграционная подвижность населения в старших возрастах мала и уменьшается с возрастом. Само количество людей в старших возрастах уменьшается с возрастом. Поэтому предполагаем, что с возрастом возрастные коэффициенты миграции пропорциональны функции дожития в некоторой степени.

На рис. 2 приведен график восстановленного возрастного распределения миграционного сальдо среди женщин.

На рис. 3 приведены данные по сальдо миграционного прироста мужчин.

Выводы

Рассмотрена методология восстановления годичных возрастных данных по основным компонентам численности населения с использованием математических методов. В качестве примера восстановлены данные по смертности среди женщин в 2010 году и рождаемости в 2010 году, и сальдо миграции за 2009–2011 годы. При анализе каждого демографического процесса были предположены соответствующие возрастные распределения и восстановлены годичные данные о возрастном распределении каждого процесса.

1. Возрастное распределение смертности в России хорошо описывается моделью Гомперца внутри каждой возрастной груп-

пы со специфическими для России скачками смертности между возрастными группами.

2. Возрастное распределение рождаемости в России с хорошей степенью точности описываются гамма-распределением, в качестве аргумента которого используется линейное степенное преобразование возраста.

3. Возрастное распределение миграционного сальдо восстанавливается, используя сплайновую интерполяцию.

Литература

1. Колмаков И.Б. Система моделей прогнозирования демографических показателей. // ИСЭПН РАН. Народонаселение. — 2005. — № 3.

2. Брасс У. Об одном способе выражения закономерностей смертности. // Изучение продолжительности жизни. Сборник статей под ред. и с предисл. Е.М. Андреева, А.Г. Волкова. — М.: Статистика, 1977.

3. Система моделей долгосрочного прогноза (2070–2095) демографических показателей. [Электронный ресурс] / РЭУ им. Г.В. Плеханова. — Режим доступа: http://www.rea.ru/Main.aspx?page=Sistema_modellej_dolgosrochnogo_proгноза_2070_2095_demograficheskikh_pokazatelej, свободный. — Загл. с экрана.

4. Тебуев Дж.Б. Комбинированный метод оценивания показателей таблиц дожи-

тия: Дисс. ... канд. соц. наук. — Ставрополь, 2008.

5. Баркалов Н.Б. Моделирование демографического перехода. — М.: Изд-во МГУ, 1984.

6. Демографический ежегодник России 2012: статистический сборник. — М.: Федеральная служба государственной статистики (Росстат), 2012.

7. Алейковский И.А. Моделирование миграционных процессов. // Экономические исследования молодых ученых. — М.: ТЕИС, 2005.

8. Демографические процессы и их закономерности. / Под ред. А.Г. Волкова. — М.: Мысль, 1986.

9. Васин С.А. Региональные модельные таблицы смертности. // Социологические исследования. — 1988. — № 4.

10. Единая межведомственная информационно-статистическая система [Электронный ресурс] / Росстат. — Режим доступа: <http://www.fedstat.ru/indicators/start.do>, свободный. — Загл. с экрана.

11. Эдиев Д.М., Тебуев Дж.Б. О выборе стандартной таблицы дожития в модели Брасса при оценивании показателей дожития малочисленного населения. [Электронный ресурс] / Электронный журнал «Исследовано в России». — Режим доступа: <http://zhurnal.apc.relarn.ru/articles/2005/155.pdf>, свободный. — Загл. с экрана.

Поступила в редакцию

25 декабря 2013 г.



Александр Николаевич Сучков — аспирант кафедры социологии управления Московского государственного университета.

Aleksander Nikolayevich Suchkov — postgraduate student of Moscow State University's Sociology of Management department.

119992, Москва, Ломоносовский проспект, д. 27, корп. 4
27 Lomonosovskiy ln., bld. 4, 119992, Moscow, Russia
Тел.: +7 (495) 939-27-83; e-mail: Socio@spa.msu.ru