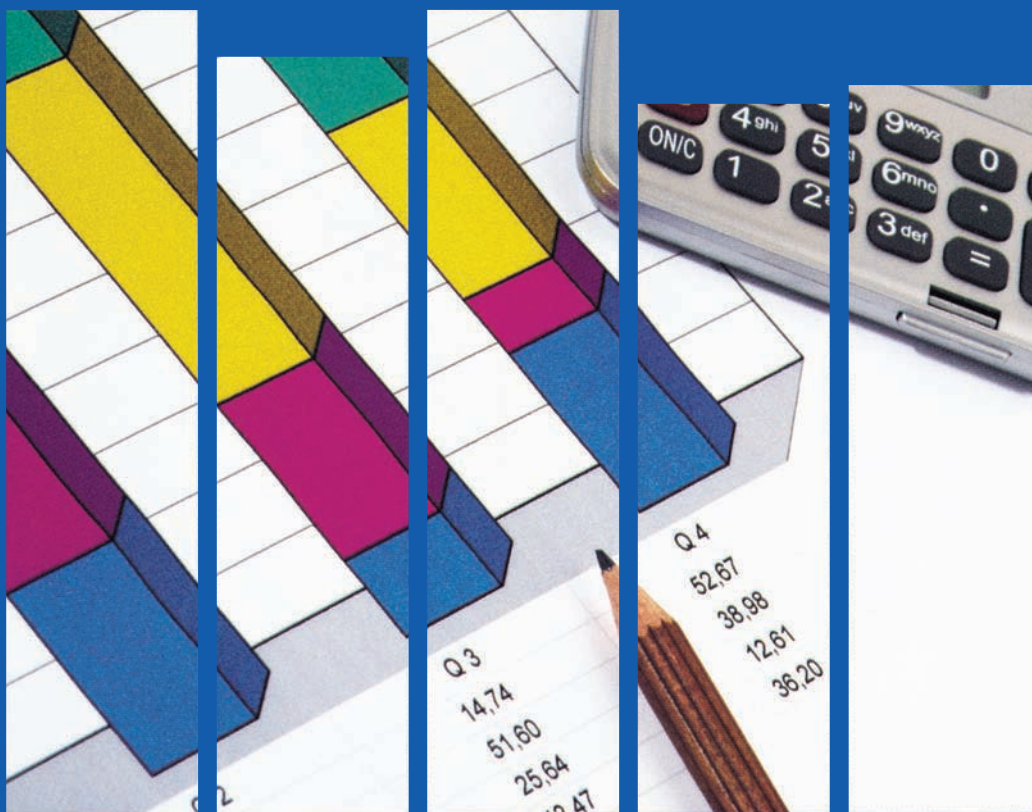
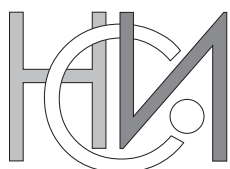


СТАТИСТИКА STATISTICS

3/2014



РЕПУБЛИКА БЪЛГАРИЯ
REPUBLIC OF BULGARIA



НАЦИОНАЛЕН СТАТИСТИЧЕСКИ ИНСТИТУТ
NATIONAL STATISTICAL INSTITUTE

СТАТИСТИКА

STATISTICS

3/2014

СОФИЯ, 2015
SOFIA, 2015

Decorative blue wavy lines at the bottom of the page, consisting of several overlapping curved lines that create a sense of movement and flow.

РЕДАКЦИОННА КОЛЕГИЯ

Главен редактор: д-р Богдан Богданов

Заместник главен редактор: проф. Васил Цанов

Членове:

Теория и практика на статистиката: д-р Б. Богданов, проф. В. Цанов,

И. Балеv, С. Костов, д-р Е. Атанасова, проф. Й. Христосков, д-р Л. Томев,

д-р А. Цветков, доц. Е. Тошева

Факти и събития от вековната история на държавната статистика:

А. Парушева, С. Златанова

Регионални анализи: А. Ахмедов, С. Цонев, Н. Нанев, А. Настева,

Г. Ангелов, К. Казанджиева, Д. Славов

Международни проекти, изяви, анализи и материали: Г. Статева,

Д. Георгиева

Социално-икономически анализи: д-р Д. Димитрова, Е. Йорданова,

Л. Янева, д-р Т. АлексПреводач на английски език: Камен Караджов

Отговорен редактор: Лидия Александрова

Стилов редактор: Мила Трифонова

Адрес на редакцията:

София, 1038, ул. „П. Волов” № 2, ет. V

e-mail: bbogdanov@nsi.bg

EDITORIAL PANEL

Editor in Chief: Bogdan Bogdanov, Ph.D.

Vice Editor in Chief: Prof. Vasil Tsanov, Ph.D.

Members:

Theory and practice of statistics: B. Bogdanov Ph.D., Prof. V. Tsanov

I. Balev, S. Kostov, E. Atanasova Ph.D., J. Hristoskov L. Tomev Ph.D., A. Tsvetkov

Ph.D., Prof. E. Tosheva

Facts and events from the centuries of history of government statistics:

A. Parusheva, S. Zlatanova

Regional Analyses: A. Ahmedov, S. Tsonev, N. Nanev, A. Nasteva,

G. Angelov, K. Kazandjieva, D. Slavov

International projects, events, analysis and materials G. Stateva, D. Georgieva

Socio-Economic Analysis: D. Dimitrova Ph.D., E. Yordanova, L. Ianeva, T. Aleksandrova Ph.D.

Russian Translation: Jaklina Tzvetkova

English Translation: Kamen Karadjov

Responsible editor: Lidia Aleksandrova

Stylistic editor: Mila Trifonova

Editorial address:

2, P. Volov St., Sofia 1038, Bulgaria

Vth floor

e-mail: bbogdanov@nsi.bg



СЪДЪРЖАНИЕ

Стр.

ТЕОРИЯ И МЕТОДОЛОГИЯ НА СТАТИСТИЧЕСКИТЕ ИЗУЧАВАНИЯ

Венелин Бошнаков Екатерина Тошева Ива Тасева	Данъчно-трансферен симулационен модел за България: възможности за сравнителни социално-икономически анализи чрез EUROMOD	9
Александър Найденев	Подходи за месечна оценка на равнището на безработица въз основа на данни от Наблюдението на работната сила	37

СТАТИСТИЧЕСКИ ИЗСЛЕДВАНИЯ И АНАЛИЗИ

Стоян Баев	Статистически поглед към образованието	59
Галя Статева	Големите данни - възможност, предизвикателство или заплаха пред официалната статистика	71

ИСТОРИЧЕСКО РАЗВИТИЕ НА СТАТИСТИЧЕСКАТА ТЕОРИЯ И ПРАКТИКА

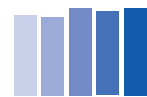
Димитър Аркадиев	Изменения в броя на населението по българските земи в състава на Османската империя	91
---------------------	---	----

ИНФОРМАЦИИ, РЕЦЕНЗИИ, КОНСУЛТАЦИИ

Галя Статева Костадин Георгиев	Информация за проект „Прилагане на технически и статистически SDMX стандарти в Националната статистическа система на България”	125
--------------------------------------	--	-----

СОДЕРЖАНИЕ

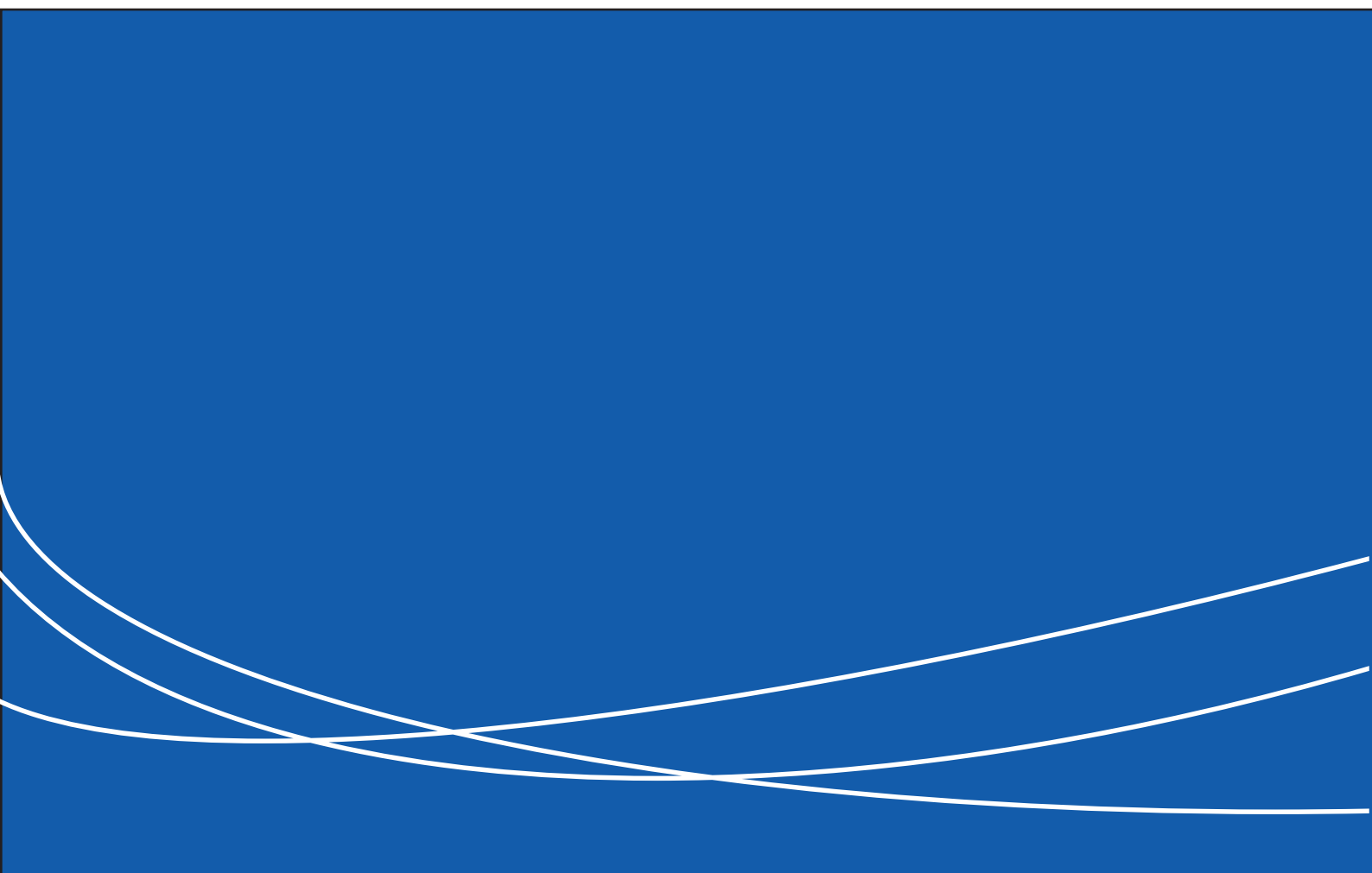
		Стр.
ТЕОРИЯ И МЕТОДОЛОГИЯ СТАТИСТИЧЕСКИХ ИССЛЕДОВАНИЙ		
Венелин Бошнаков Екатерина Тошева Ива Тасева	Налогово-трансферная симуляционная модель для Болгарии: возможности для сопоставительного социально-экономического анализа с помощью EUROMOD	9
Александр Найденов	Подходы для месячной оценки уровня безработицы на основе данных с Исследования рабочей силы	37
СТАТИСТИЧЕСКИЕ ИССЛЕДОВАНИЯ И АНАЛИЗЫ		
Стоян Баев	Статистический взгляд на образование	59
Галя Статева	„Большие данные” - возможность, вызов или угроза для официальной статистики	71
ИСТОРИЧЕСКОЕ РАЗВИТИЕ СТАТИСТИЧЕСКОЙ ТЕОРИИ И ПРАКТИКИ		
Димитр Аркадиев	Изменения в численности населения в болгарских землях в составе Османской империи.....	91
ИНФОРМАЦИИ, РЕЦЕНЗИИ, КОНСУЛЬТАЦИИ		
Галя Статева Костадин Георгиев	Информация о проекте „Применение технических и статистических SDMX стандартов в национальной статистической системе Болгарии”	125



CONTENTS

	Page
THEORY AND METHODOLOGY OF STATISTICAL RESEARCH	
Venelin Boshnakov Ekaterina Tosheva Iva Taseva	Tax-transfer simulation model for Bulgaria: opportunities for comparative socio-economic analyses using EUROMOD 9
Aleksander Naidenov	Approaches for monthly assessment of unemployment level based on data from the Labour Force Survey 37
STATISTICAL SURVEYS AND ANALYZES	
Stoyan Baev	Statistical insight into education 59
Galia Stateva	Big Data - an opportunity, a challenge or threat to official statistics 71
HISTORICAL DEVELOPMENT OF STATISTICAL THEORY AND PRACTICE	
Dimitar Arkadiev	Changes in population numbers in the Bulgarian lands included in the Ottoman Empire 91
INFORMATION, REVIEWS, CONSULTATIONS	
Galia Stateva Kostadin Georgiev	Information about the 'Implementation of SDMX technical and statistical standards in the National Statistical System of Bulgaria' Project 125

**ТЕОРИЯ И МЕТОДОЛОГИЯ НА
СТАТИСТИЧЕСКИТЕ ИЗУЧАВАНИЯ**



ДАНЪЧНО-ТРАНСФЕРЕН СИМУЛАЦИОНЕН МОДЕЛ ЗА БЪЛГАРИЯ: ВЪЗМОЖНОСТИ ЗА СРАВНИТЕЛНИ СОЦИАЛНО-ИКОНОМИЧЕСКИ АНАЛИЗИ ЧРЕЗ EUROMOD

Венелин Бошнаков^{}, Екатерина Тошева^{**}, Ива Тасева^{***}*



1. Въведение

Микросимулационните модели за анализ на ефектите от данъчно-трансферната интервенция върху доходите на лицата и домакинствата имат вече своята значима история. Разработването и усъвършенстването на различни микросимулационни модели за анализ на икономически и социални явления и процеси са с близо петдесетгодишна история в западноевропейските страни и САЩ след пионерната публикация на W. A. Steger (1961). Националните модели са изградени при системно придържане към традициите и акцентирание върху специфичните особености на данъчно-трансферната система в съответната страна. Тяхното създаване и развитие често е иницирано от стремежа да се получи аналитична информация за отделни аспекти, представляващи (в една или друга степен) конюнктурен интерес (Wagenhals, 2004; Sutherland, Figari, 2013).

Процесите на самостоятелно и изолирано развитие на националните данъчно-трансферни симулационни модели довежда до значително ниво на несъпоставимост на получаваните аналитични резултати между отделните страни (Callan, Sutherland, 1997). Във връзка с осигуряването на възможности за

^{*} Доцент д-р в катедра „Статистика и иконометрия”, УНСС - София; e-mail: venelinb@unwe.bg.

^{**} Доцент д-р в катедра „Статистика и иконометрия”, УНСС - София; e-mail: ekaterina_tosheva@yahoo.com.

^{***} Институт по социални и икономически изследвания, Университет Есекс, Обединено кралство; e-mail: itasheva@essex.ac.uk.

моделиране на широк диапазон от схеми както за данъчно облагане, така и за социално подпомагане възниква необходимостта от синхронизация на тези процеси чрез възприемане на интегриран подход за моделиране на данъчно-трансферните системи. В основата на тази интеграция е поставено изискването за придържане към базови стандартизирани дефиниции за елементите на дохода и разходите. Тази стандартизация трябва да осигурява максимално широки възможности за **симулиране на отделни предложения за специфична политика**, насочена към реформиране на един или друг елемент на подходното облагане (включително социалноосигурителните компоненти) или специфичен инструмент на социалната политика.

От тази гледна точка подходът за развитие на интегриран мултинационален симулационен модел притежава безспорното предимство да възприема решения за различните проблеми на информационното осигуряване още на етапа на неговото изграждане. Контролът върху валидността на резултатите за отделната страна естествено предполага съпоставянето на набор от основни статистически показатели, получавани едновременно от националния и мултинационалния (общоевропейски) интегриран модел. Крайната цел на микросимулационния анализ е да предостави количествени оценки за ефектите от промяната в една или повече подходни политики (облагане, социално осигуряване или социално подпомагане) - въпроси с висока обществена и политическа значимост, търсещи „печелещите и губещите“ от дадена промяна (Atkinson et al., 2000). Тези задачи са безспорно актуални и в светлината на поставените стратегически цели по неравенството и бедността в рамката на стратегия „Европа 2020“ интелигентен, устойчив и „включващ“ растеж на ЕС.

Паневропейският модел EUROMOD (Tax-Benefit Microsimulation Model for the EU) е единственият данъчно-трансферен микросимулационен модел за Европейския съюз (Sutherland, Figari, 2013)¹. Той може да бъде разглеждан като уникален ресурс за международни изследвания, предоставящ сравними симулационни резултати между държавите - членки на ЕС. Най-общо симулационните резултати, получавани от EUROMOD, дават възможност за извличане на оценки за ефектите от минали и потенциални промени в

¹ Информация за EUROMOD може да се намери на адрес: <https://www.iser.essex.ac.uk/euromod/>.



социалните, осигурителните и данъчните политики върху държавните приходи и разходи, както и върху доходите и жизнения стандарт на населението. EUROMOD може да се използва за анализ на социално-икономически ефекти от типа „в полза на“ и „за сметка на“ отделни социално-демографски групи (например едночленни домакинства, семейства с две деца, лица в работоспособна възраст, домакинства с пенсионери и други). Други приложения на модела включват изготвянето на оценки за очакваните промени в тежестта (маржиналните ефективни ставки) на подоходното облагане, „доходната субституция“ при безработица, пенсиониране и др., както и в различни показатели за икономическо неравенство и бедност.

Целта на настоящата статия е да запознае специалистите от науката и практиката с възможностите на модела EUROMOD - към момента единственият симулационен данъчно-трансферен модел за България. В резултат на популяризирането на модела се очаква да се окуражи използването му от страна на заинтересованите институции и лица в научни и приложни проекти с оглед оценка на ефекта от различни политики в социалното подпомагане, социалното осигуряване и данъчното облагане, както и за обосноваване на нуждата от различни реформи в тези области.

2. Разработване на български модул в EUROMOD

EUROMOD е разработен чрез серия от проекти, финансирани от програми на ЕС за развитие на европейската изследователска инфраструктура. Създаването на Консорциума EUROMOD се осъществява през 1997 г. чрез първоначален тригодишен международен изследователски проект, финансиран по програма TSER (Targeted Socio-Economic Research Programme) на Европейската комисия (ЕК). Началното покритие на модела обхваща 15 държави от ЕС (EC-15). Подготвителната работа по инкорпорирането на присъединените през 2004 г. десет страни започва в рамките на Проект I-CUE (Improving the Capacity and Usability of EUROMOD) с финансовата подкрепа на 6-ата рамкова програма на ЕС (FP6 Research Infrastructures Action). По същество планираното разширяване на покритието до ЕС-27 се осъществява чрез изпълнението на Проект *EUROMODupdate* от разширен Консорциум², координиран от Института за социални и икономически изследвания (ISER) при

² Консорциумът включва водещия екип от Университета в Есекс и 27 екипа от държавите - членки на ЕС. Състои се понастоящем от 92 членове.

Университета в Есекс (University of Essex), Обединено кралство, регулиран чрез Рамково споразумение VS/2008/0318 с Генерална дирекция (ГД) „Заетост, социални въпроси и приобщаване“ на ЕК и финансиран по Програма PROGRESS (Community Programme for Employment and Social Solidarity).

Като участник в проекта *EUROMODupdate* българският екип разработва **подготвителен документ „Проучване на изпълнимостта“ за България**, който съдържа оценка на възможностите за създаване на българска секция в EUROMOD и интегриране на страната в общоевропейския данъчно-трансферен симулационен модел (Boshnakov et al., 2010). Разработването на първия пълнофункционален модул в EUROMOD за България се осъществява по същия проект, като за подпомагане на техническата разработка, изпълнявана от експерти на ISER, българският екип подготвя разширен документ: „Доклад за България: данъчно-трансферни системи 2007 - 2010“. Той съдържа детайлизирано представяне на основните параметри и инструменти за пряка интервенция върху доходите на лицата и домакинствата в България през всяка година от посочения период, както и анализ на микроданните и симулационните резултати (Boshnakov et al., 2012).

Следващите актуализации на българския модул в EUROMOD са осъществени по Проект *EUROMODupdate2*, който се изпълнява чрез Рамково споразумение VS/2011/0445 на Университета в Есекс с ГД „Заетост, социални въпроси и приобщаване“ на ЕК. Последователно са добавени модули за данъчно-трансферните системи за 2011, 2012 и 2013 г., като е извършено и пребазиране на първичната информация от EU-SILC 2008 към EU-SILC 2010 с планирана актуализация за 2014 г. и пребазиране към микроданни от EU-SILC 2012 (последният доклад за България е представен в: Tosheva et al., 2014).

3. Основни етапи в алгоритъма на работа на симулационния модел EUROMOD

За реализиране на симулационните процедури в модела EUROMOD, както и за валидацията и анализа на получените резултати, е необходимо да се следва единна методика. Най-общо тя включва следните етапи: подготвителен етап (подетапи 1а и 1б), същински симулационни процедури (подетапи 2, 3, 4, 5а и 5б) и анализ на получените резултати (подетап 6). Тези етапи със съответните подетапи са показани на фиг. 1.

Фиг. 1. EUROMOD структура



На подготвителния етап се решават две основни задачи. Първата е свързана с избора на източник на оригинални микроданни (подетап 1а). Тук се определят обхватът и съдържанието на базата микроданни (първичните статистически сведения) чрез дефиниране на минималния набор от признаци, необходими за осъществяване на симулациите на осигурителните и данъчните политики и политиките за социално осигуряване в дадена страна. Това са данни за социално-икономическите характеристики на лицата и домакинствата и техните парични приходи (този въпрос е разгледан в точка 4). Втората задача (подетап 1б) се състои в набиране на детайлна информация за законовите регулации в областта на данъчното облагане, социалното осигуряване и социалното подпомагане и е необходима за симулационните процедури.

Вторият етап на работа в EUROMOD - същинските симулационни процедури, включва няколко стъпки. Първата стъпка (подетап 2) предвижда осъществяване на техническа подготовка на входящите информационни масиви и включва техническа подготовка на входната база данни, свързана с унифицирано кодиране и „параметризация” на всеки елемент на данъчно-

трансферната система на всяка страна. В следващата стъпка (подетап 3) данъчните и социалните политики се описват в EUROMOD с помощта на специален „данъчно-трансферен” език, базиращ се на функции и параметри, чрез използване на разработения потребителски интерфейс на EUROMOD. На третата стъпка (подетап 4) се извършва същинската симулация чрез специални програмни модули.

Основните съставни компоненти на паричния доход на домакинството, за които се конструират типови (и във възможно най-висока степен унифицирани) променливи, са: доходи от наемен труд (трудови и приравнени към тях платени възнаграждения); доходи от самонаемане и предприемачество; доходи от собственост; други пазарни доходи; пенсии за прослужено време и инвалидност; наследствени пенсии; тук се обхващат и **симулираните социални плащания и данъчно-осигурителни задължения**. В българския модул на EUROMOD симулираните данъчни и социални политики включват:

- социално- и здравноосигурителни вноски за сметка на наетите и самостоятелно заетите лица;
- социално- и здравноосигурителни вноски на работодателя (моделът осигурява възможност да се симулира допълнително и тази променлива с оглед измерването на т.нар. „данъчен клин”);
 - данък общ доход;
 - обезщетение за безработица;
 - обезщетение за отглеждане на дете до 2-годишна възраст;
 - обезщетение за бременност и майчинство;
 - месечна помощ за отглеждане на дете;
 - целева помощ за ученици;
 - еднократна помощ при раждане;
 - месечна помощ за отглеждане на дете до 1-годишна възраст;
 - месечна добавка за отглеждане на дете с трайни увреждания до завършване на средно образование;
 - еднократна помощ за отглеждане на близнаци до 1-годишна възраст;
 - еднократна помощ за отглеждане на дете до 1-годишна възраст за майки, които са редовни студенти;
 - социална пенсия за възраст;

- гарантиран минимален доход;
- целева помощ за отопление³;
- корекция за възможно укриване на доходите на заети лица (корекцията може да бъде „включена” или „изключена” при симулациите⁴).

Илюстрация на тази стъпка е представена на фиг. 2 и 3 по отношение на симулираните политики, и по-точно за симулацията на конкретен инструмент - данък върху дохода на физическите лица в българския модул на EUROMOD.

Фиг. 2. Симулирани политики в българския модул на EUROMOD

Policy	Grp/No	BG_2011	BG_2012	BG_2013	Comment
7	• TCA_bg	switch	switch	switch	SWITCH: correction for tax evasion: 1) UDB SILC - comparing gross and net income from employment; 2) UDB SILC+ national SILC variables - using information on base for calculating SIC
8	• tsцер_bg	on	on	on	SIC: social insurance contributions employer (вноски за социално осигуряване платени от работодателите)
9	• tscee_bg	on	on	on	SIC: social insurance contributions employee (вноски за социално осигуряване на работниците и служителите)
10	• tsce_bg	on	on	on	SIC: social insurance contributions self-employed (вноски за социално осигуряване на самостоятелно заети)
11	• tin_bg	on	on	on	TAX: income tax (данък общ доход)
12	• bunct_bg	on	on	on	BEN: unemployment benefit (обезщетение за безработица)
13	• bmact_bg	on	on	on	BEN: contributory maternity benefit: cash benefit for bringing up child up to age 2 (обезщетение за отглеждане на дете до 2г.)
14	• bchba_bg	on	on	on	BEN: contributory maternity benefit: cash benefit for pregnancy and childbirth (обезщетение за бременност и майчинство)
15	• bsa00_bg	on	on	on	BEN: social assistance - guaranteed minimum income (помощ за социално подпомагане - гарантиран минимален доход)
16	• bsah_bg	on	on	on	BEN: heating benefit (целева помощ за отопление)
17	• bchmt_bg	on	on	on	BEN: means-tested child benefit (месечна помощ за дете)
18	• bched_bg	on	on	on	BEN: child benefit education (целева помощ за ученици)
19	• bchbals_bg	on	on	on	BEN: birth grant (еднократна помощ при раждане)
20	• bmanc_bg	on	on	on	BEN: benefit for raising a child under the age of 1, non-contributory (месечна помощ за отглеждане на дете до 1г.)
21	• bchnm01_bg	on	on	on	BEN: non means-tested child benefit for mothers in tertiary
22	• bchnm02_bg	on	on	on	BEN: non means-tested child benefit for twins
23	• poamt_bg	on	on	on	BEN: social old-age pension

Източник: EUROMOD, версия G2.0.

Забележка: Редът на симулациите зависи от дефинираните взаимодействия между данъчните и социалните политики.

³ Информацията за други данъци или социални обезщетения, които не са симулирани в EUROMOD, е взета от микроданните.

⁴ Корекцията се основава на сравнение между нетните и брутните доходи от заетост, докладвани в анонимизираните микроданни от всяко едно лице. При този подход се приема, че лицата участват в сивата икономика, ако техните нетни и брутни доходи от заетост са равни, т.е. те не плащат данъци и социални осигуровки. В резултат симулираните данъчни и осигурителни задължения са равни на нула, което се отразява и върху симулираните социални помощи. Следва да се отбележи, че методът за корекция е опростен и няма за цел пресъздаване на поведението на лицата, а единствено калибриране на симулационните резултати. Синтезирано описание на метода може да бъде намерено в Boshnakov et al. (2012) и Tosheva et al. (2014).

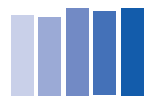
Фиг. 3. Симулация на данък общ доход в българския модул на EUROMOD

Policy	Grp/No	BG_2007	BG_2013	Comment
1	- tin_bg	on	on	TAX: income tax (данък общ доход)
1.1	▸ DefI	on	on	Income that is taxed (before any tax deductions)
1.2	▸ DefTu	on	on	Assessment unit
1.3	▸ BenCalc	on	off	TAX: child deduction (данъчно облекчение за деца)
1.4	▸ ArithOp	on	on	TAX: tax deduction for private pension contributions (данъчно облекчение за частни пенсионни вноски)
1.5	▸ ArithOp	on	on	TAX: tax deduction for income from rent (данъчно облекчение за доходи от наем)
1.6	▸ ArithOp	on	on	Sum of all tax deductions
1.7	▸ ArithOp	on	on	Taxable income
1.8	- SchedCalc	on	off	Progressive tax in 2007
1.8.1	base	tintb_s	n/a	taxable income
1.8.2	band_uplim 1	2400#y	n/a	annual income of less than BGN 2400
1.8.3	band_rate 1	0	n/a	0% tax rate
1.8.4	band_uplim 2	3000#y	n/a	annual income between BGN 2400 and 3000
1.8.5	band_rate 2	0.2	n/a	20% tax rate
1.8.6	band_uplim 3	7200#y	n/a	annual income between BGN 3000 and 7200
1.8.7	band_rate 3	0.22	n/a	22% tax rate
1.8.8	band_rate 4	0.24	n/a	annual income above BGN 7200 - 24% tax rate
1.8.9	output_var	tin_s	n/a	income tax variable
1.8.10	TAX_UNIT	tu_individual_bg	n/a	the assessment unit is the individual
1.9	- SchedCalc	off	on	Flat tax after 2007
1.9.1	base	n/a	tintb_s	taxable income
1.9.2	band_rate 1	n/a	0.1	10% tax rate
1.9.3	output_var	n/a	tin_s	income tax variable
1.9.4	TAX_UNIT	n/a	tu_individual_bg	the assessment unit is the individual

Източник: EUROMOD, версия G2.0.

Четвъртата и петата стъпка (подетапи 5а и 5б) се състоят в производството на изходящи макроданни, включващи симулираните на предходната стъпка променливи, както и тяхната валидация. Оценката на резултатите, получени чрез симулационните процедури в модела EUROMOD, може да се извърши в два аспекта. Единият предполага сравняване на агрегираните показатели за разходите за социални помощи и приходите от данъци и социалноосигурителни вноски с кореспондиращи показатели от външни източници. За целта се използват най-често административни източници на обобщени данни за целевите съвкупности на платци и получатели на доходи. Този подход е известен като „макротвалидация“.

При такъв тип сравнения трябва да се държи сметка за някои особености. **Първо**, административните данни могат да се отнасят за времеви период и/или съвкупност, които не съответстват напълно точно на периода и/или



съвкупността, за които се осигуряват входни данни от представителната извадка. **Второ**, административните данни могат да не съдържат детайлизирани сведения за всички инструменти или източници на доход, като не кореспондират в най-пълен обхват с използваните от EUROMOD. **Трето**, невинаги данните за отделни макроагрегатни величини са публични. А в случаите, когато такива данни се публикуват, това често става с голямо закъснение във времето. **Четвърто**, поради някои проблеми на микроданните от EU-SILC като проблема с неотговорилите респонденти или систематичното деклариране на по-ниски от действителните доходи (Figari et al., 2012), както и поради несъвършенства в симулационните процедури (например грешки при разчитането и симулирането на социалните, осигурителните и данъчните политики), симулационните резултати могат да доведат до подценяване или надценяване на агрегираните показатели.

Вторият аспект предполага съпоставяне на показателите за бедност и неравенство с подобни оценки, получени директно по данни от изследването „Статистика на доходите и условията на живот (EU-SILC)”. И тук съществуват известни различия, дължащи се на различни причини. Като основни могат да се посочат:

Първо, използваните различни дефиниции на базови понятия (например за обхвата на категорията „разполагаем доход“).

Второ, проблемът с единиците, неотговорили по един или повече от ключовите признаци, доколкото те се изключват при изчислителната работа в EUROMOD.

Трето, EUROMOD е симулационен модел, което означава, че на базата на данните за всяка единица се „присвояват“ полагащите ѝ се помощи и дължимите разходи за данъци и осигуровки. В този смисъл резултатите от модела не отразяват реално проявяващи се явления като например фактическо неполучаване на полагащи се социални помощи (Tasseva, 2012).

След валидацията и документацията на резултатите моделът и неговият основен продукт се считат за завършени и могат да бъдат използвани за научни изследвания с теоретична и приложна насоченост (подетап 6). **Основният продукт** на модела представлява база микроданни, които включват информация за различните симулирани и несимулирани социални, осигурителни и данъчни инструменти, както и променливи за пазарните доходи и социално-

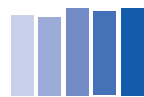
икономическите и демографските характеристики на отделните лица и домакинства. В изходната база микроданни се отразяват промените в първичния (пазарен) и разполагаемия доход на домакинствата като резултат както от проведени реформи в миналото, така и от бъдещи потенциални промени в социалните, осигурителните и данъчните политики. Това дава възможност да се оценят коефициенти на неравенство и бедност и да се анализира степента на доходното преразпределение чрез данъчно-трансферната интервенция (примери за такъв вид анализ ще бъдат разгледани в точка 5).

4. Източници на данни за модела EUROMOD

За осигуряване на симулациите в рамката на EUROMOD хронологично са използвани **различни източници на данни**. В началните фази на неговото разработване информационното осигуряване е базирано на панелното изследване на домакинствата в държавите от ЕС през 90-те години на XX век (ECHP - European Community Household Panel). Стремещът на проектантите на модела е бил да се осигури в максимална степен съпоставимост на входните данни за отделните страни, което налага използване на идентични дефиниции и инструментариум за регистрация на първичните данни за всички ключови променливи. Същевременно ефективното приложение на симулационните механизми, чрез които се генерират основни променливи (компоненти на осигурителни/данъчни задължения и социални трансфери), изисква по-детайлна информация от тази, осигурявана от стандартизирания модел на изследването ECHP. Това налага търсене на допълнителни източници на данни, налични в националните извадкови изследвания на домакинствата.

В тази ситуация данни от две от вълните на ECHP са били използвани при създаването на базите с микроданни на EUROMOD за пет от страните (Австрия, Дания, Гърция, Португалия и Испания) за периода до 2000 година. На този начален етап от развитието на модела са търсени възможности за оползотворяване на информационни масиви, конструирани на базата на национални представителни извадкови изследвания на домакинските бюджети в държавите от ЕС-15. Като национални източници на данни в EUROMOD за останалите страни имат принос изследвания на домакинските бюджети или други източници:

- Ирландия (национален панел): Living in Ireland Survey (LII)



- Италия (извадка): Survey of Households Income and Wealth
- Люксембург (национален панел): PSELL-2
- Белгия (национален панел): Panel Survey of Belgian Households (PSBH)
- Финландия (данни от административни регистри + извадка): Income Distribution Survey
- Германия (национален панел): German Socio-Economic Panel
- Франция (извадка): Budget de Famille (BdF)
- Нидерландия (национален панел): Sociaal-economisch Panelonderzoek
- Швеция (данни от административни регистри + извадка): Income Distribution Survey (IDS)
- Обединено кралство: Family Expenditure Survey (FES).

В основата на базите анонимизирани данни за новите версии на EUROMOD, разработвани след 2005 г., се поставя обновеният и усъвършенстван вариант на европейското домакинско изследване „Статистика на доходите и условията на живот (EU-SILC)”. Тези данни са достъпни на Консорциума съгласно серия от мрежови споразумения на координиращата институция (University of Essex) с Евростат за използване на потребителските бази данни (UDB: User Data Base), предоставяни от Евростат за научноизследователски цели (договори EU-SILC/2009/17, EU-SILC/2011/55). С някои от статистическите институции на държавите - членки на ЕС, е постигнато споразумение за предоставяне на допълнителни променливи с анонимизирани записи, каквито липсват в стандартния обхват на променливите от потребителската база данни на Евростат. В тази насока все още не са преодолени различни затруднения от организационно и административно-правно естество, като продължават усилията за постигане на напредък в отношенията с институциите в някои от държавите - членки на ЕС. При разработване на българския модул в EUROMOD Националният статистически институт предостави националните променливи от проведеното през 2008 г. изследване „Статистика на доходите и условията на живот“ и това даде възможност да се подобрят съществено симулационните резултати и да се обогатят възможностите за анализа им (Boshnakov et al., 2012).

5. Изходните резултати от EUROMOD - потенциал за социално-икономически анализи

Основна цел на EUROMOD като европейски мултинационален микросимулационен модел е да осигури детайлизиран инструментариум за провеждане на сравнителен анализ на различни социално-икономически ефекти върху благосъстоянието на населението в ЕС както за всяка държава членка поотделно, така и общо за целия Европейски съюз. Тези ефекти се оценяват въз основа на **лабораторна симулация** на проектирането и прилагането на специфични мерки за реформиране на политиките по данъчното облагане и социалното подпомагане в отделните държави членки.

Различни икономически показатели могат да бъдат преоценени след експериментиране (симулиране) на различни параметри на политиките по облагането на личния и/или семейния доход, социалното осигуряване и/или социалното подпомагане. Получените резултати могат да се сравнят със „статуквото“ и да се осигурят възможности за анализиране на ефектите от интервенцията върху неравенството, преразпределението на дохода, равнището и рисковете за бедност и т.н. Нещо повече, тези показатели могат да се оценят както общо, така и поотделно за определени подсъвкупности, формирани на базата на различни демографски и други критерии (например по възраст, регион, социално-икономическа страта и други). Тези възможности на модела позволяват да се симулират различни преразпределителни механизми с оглед предлагане на различни политики за ефективно подпомагане на рискови групи от населението.

Най-общо EUROMOD може да се използва за оценка и анализ на ефекта от минали и настоящи социални и данъчни политики. Чрез симулациите могат да се оценяват ад хок промените в неравенството и бедността, настъпващи в резултат на реално проведени реформи (Avram et al., 2014; Hills et al., 2014). От съществено значение е фактът, че моделът дава възможност и да се анализират **потенциални промени** при алтернативни сценарии за реформи в данъчната, социалноосигурителната политика и/или политиката за социално подпомагане (Tasseva, 2012). Чрез EUROMOD могат да се анализират ефектите не само от политики, провеждани на национално ниво, но и от реформи, инициирани на



равнището на целия ЕС. Типичен пример е едно от първите приложения на модела, фокусирано върху потенциалните ефекти от въвеждането на т.нар. „минимална европейска пенсия“ (EMP: European Minimum Pension) (Atkinson et al., 2002). Други по-актуални примери включват въвеждането на детски минимален доход в ЕС (Levy et al., 2013) или обезщетение за безработица в еврозоната (Andor et al., 2014).

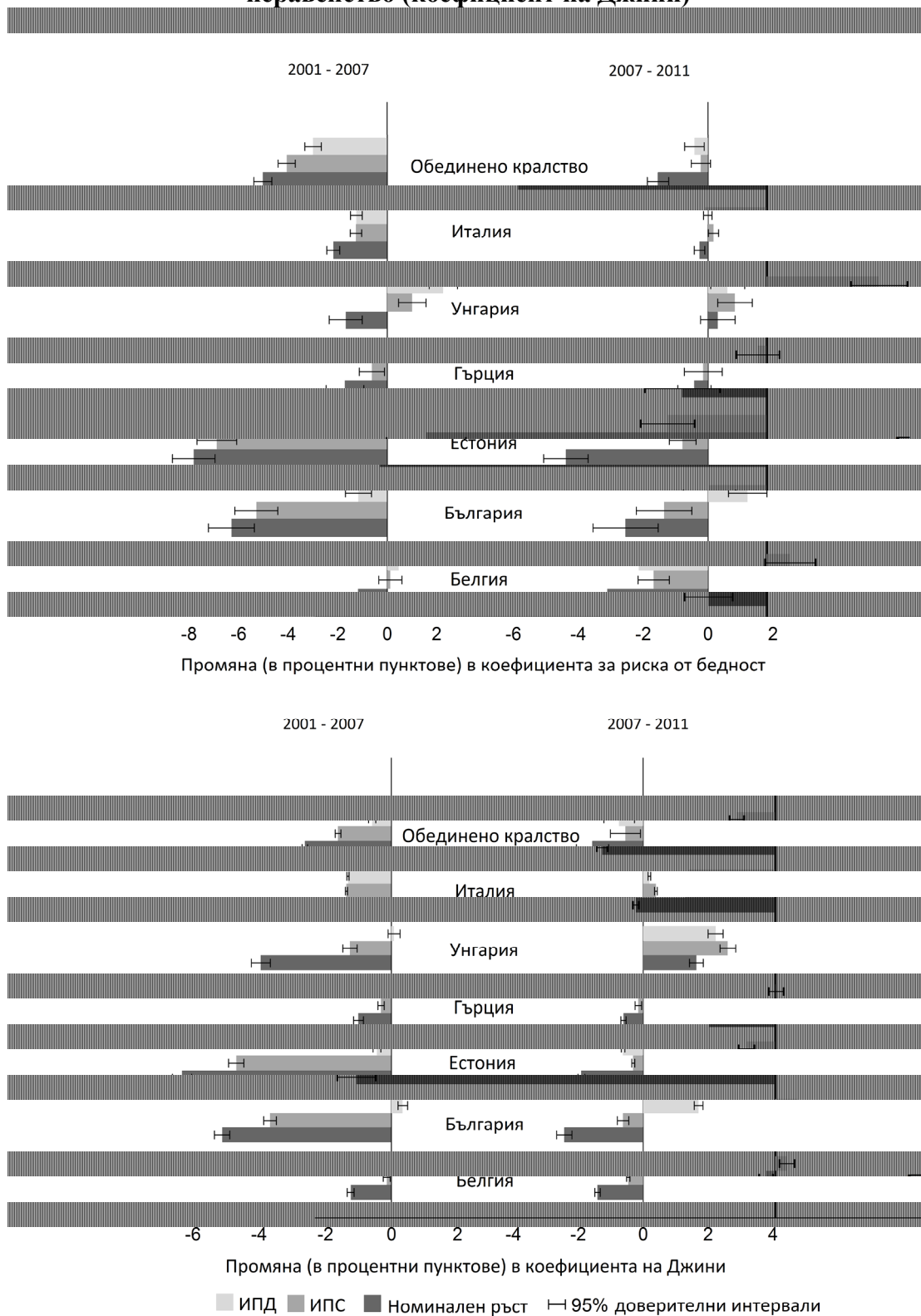
Симулационните резултати дават възможност да се оцени ефектът на различни политики върху държавните приходи и разходи, да се оценят пределните ефективни данъчни ставки (Jara, Tumino, 2013) или да се оцени ефектът от т.нар. „бюджетно неутрални“ реформи (Tasseva, 2012). Чрез участието на всички държави от ЕС в EUROMOD се дава възможност да се симулира размяна между две и повече страни на различни инструменти на политиките по социално осигуряване, данъчно облагане и социално подпомагане (Salanauskaite, Verbist, 2013). Резултатите могат да послужат за оценка на предимствата и слабите страни на националните политики и съответно да се използват за обосноваване на предложения за реформи. В контекста на стратегия „Европа 2020“ EUROMOD може да бъде разглеждан като инструмент, чрез който се оценяват националните и общоевропейските социална и фискална политика с оглед намаляване на бедността, както и за задълбочено изследване на доходното неравенство в държавите членки и между тях.

На базата на няколко примера за проведени анализи на промените в бедността и неравенството, разпределението на доходите и стимулите на работа ще бъдат илюстрирани познавателните възможности на симулационните резултати, получени от модела EUROMOD. Първият пример е основан на резултати от Paulus et al. (2014), където е измерен ефектът на промените в данъчно-осигурителните и социалните политики отделно за периодите на икономически растеж (2001 - 2007 г.) и на икономическа криза (2007 - 2011 г.). Политиките на седем държави - членки на ЕС, включително България, са симулирани в EUROMOD, като симулациите отразяват законовите параметри през 2001, 2007 и 2011 година. Пазарните доходи и демографските и социално-икономическите характеристики на лицата и домакинствата са базирани на микроданни от 2007 година (EU-SILC 2008). Като използват пазарните доходи и характеристиките на лицата и домакинствата през 2007 г., прилагайки

поотделно политиките от 2001, 2007 и 2011 г. върху тях, авторите изолират ефекта от промените в политиките върху доходите на домакинствата⁵ (Paulus et al., 2014). На фиг. 4 е показан ефектът от промените в политиките върху коефициентите на риска от бедност и на неравенството. Ефектът е изчислен в номинална стойност (номинален ръст), в реална стойност (спрямо индекса на покупателната способност: ИПС) и спрямо растежа на средните пазарни доходи в икономиката (спрямо индекса за растежа на пазарните доходи: ИПД).

⁵ Т.е. ефектът е референтен спрямо пазарните доходи и характеристиките на лицата и домакинствата през 2007 година.

Фиг. 4. Ефект от промените в данъчно-осигурителните и социалните политики през 2001 - 2007 и 2007 - 2011 г. върху риска от бедност и неравенство (коэффициент на Джини)



Източник: Paulus et al. (2014), при използване на EUROMOD.

Забележка: Коефициентът за риска от бедност е изчислен въз основа на 60% от медианния еквивалентен располагаем доход на лице от домакинство.

Резултатите от анализа сочат, че в България и другите страни промените в политиките са допринесли повече за намаляване на бедността в първия период в сравнение с втория (изключение е Унгария; фиг. 4, първи панел). Същата тенденция се вижда и в резултатите за коефициента на Джини (фиг. 4, втори панел). Ако тези резултати отразяват по-големите възможности на правителствата да действат, когато икономическите условия са благоприятни, то тогава те трябва да се стремят да постигнат повече през годините на растеж, за да преодолеят негативния ефект от кризисните периоди.

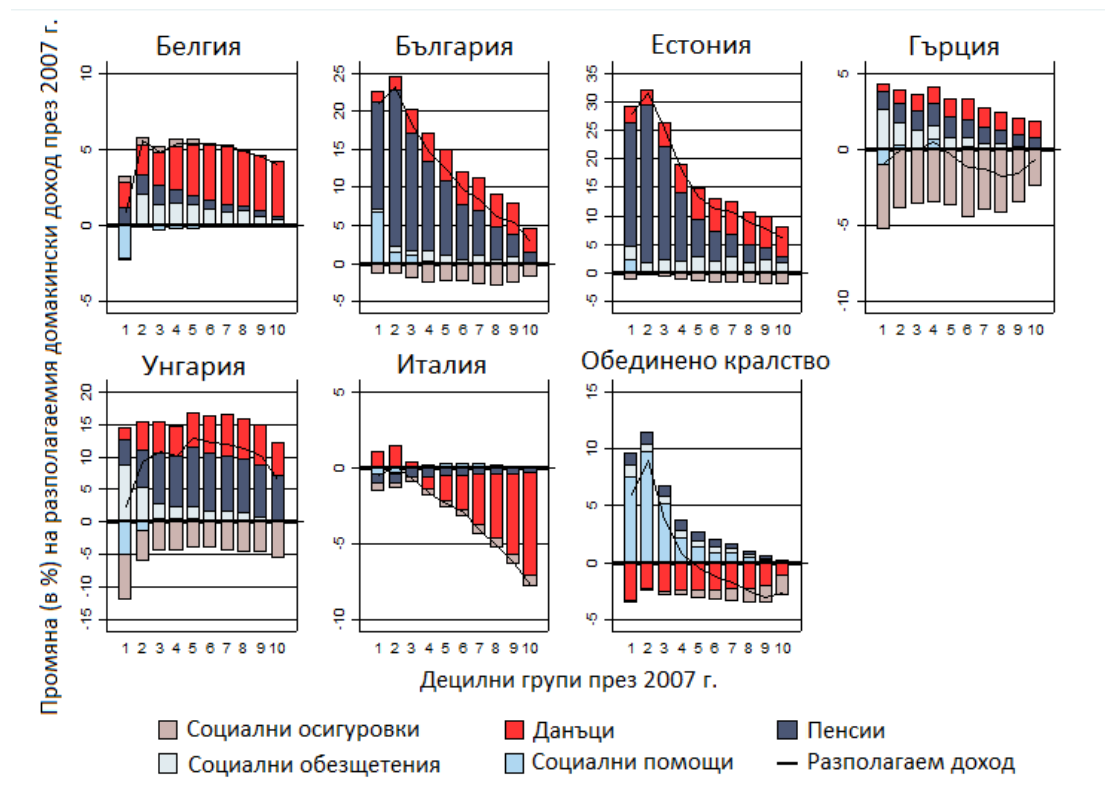
Друга аналитична разработка (Hills et al., 2014) показва, че ефектът на промените в политиките върху доходите на домакинствата във всяка страна е резултат от различни промени в отделните компоненти на данъчните и трансферните системи. На фиг. 5 и 6 са илюстрирани ефектите (в реална стойност) от промените в осигурителните вноски, преките данъци, пенсиите, социалните обезщетения и социалните помощи по децилни групи (по абсцисата) върху доходите на домакинствата (по ординатата). Линията показва общия ефект от всички компоненти. Ако ефектът (поотделно и общо за всички компоненти) е положителен, т.е. измерен над нулата по ординатата, то в резултат на промените доходите на домакинствата в дадена децилна група са се увеличили. Ако ефектът е отрицателен, доходите на домакинствата са се намалили в резултат от промените в политиките.

През периода 2001 - 2007 г. в България поради намаляването на данъчните ставки и реалното увеличение на пенсиите и социалните помощи доходите на домакинствата от всички децилни групи са се увеличили (фиг. 5). От друга страна, увеличението на социалните осигуровки, заплащани от (само)наетите, имат малък, но отрицателен ефект върху доходите. През годините на икономическата криза (2007 - 2011) реалното увеличение на пенсиите има значителен положителен ефект върху разполагаемия доход на домакинствата (фиг. 6). Но от друга страна, този положителен ефект е намален поради замяната на прогресивния данък общ доход с плосък данък, който засяга негативно домакинствата до 8-ата децилна група включително, докато най-богатите две децилни групи печелят от регресивната реформа.

Тук трябва да отбележим, че резултатите на Paulus et al. (2014) и Hills et al. (2014) показват ефекта от промените в данъчните и трансферните политики,

като не се отчитат евентуални промени в пазарните доходи и характеристиките на населението (например застаряването на населението, емиграцията, увеличението на безработицата не са взети под внимание).

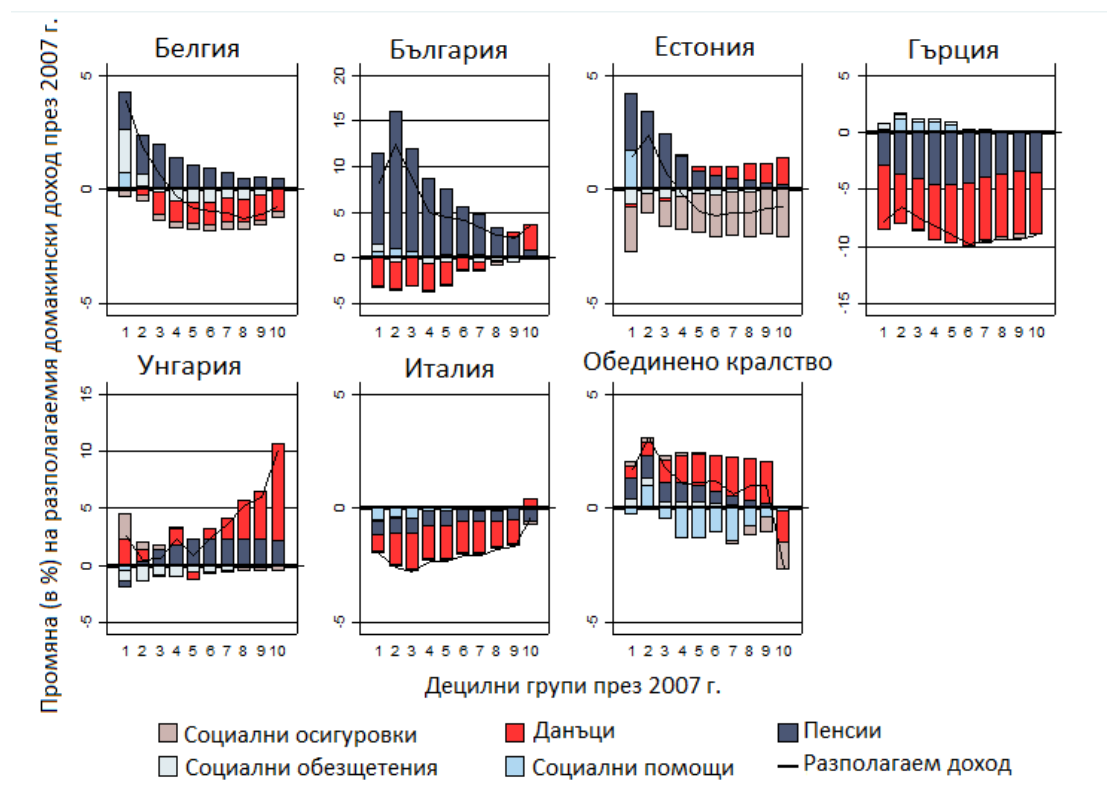
Фиг. 5. Ефект в промените в данъчно-осигурителните и социалните политики през периода 2001 - 2007 г. по отделни компоненти и децилни групи



Източник: Hills et al. (2014), използвайки EUROMOD.

Забележка: Ефектът от промените е изчислен в реална стойност. Децилните групи са изчислени въз основа на еквивалентния разполагаем доход на лице от домакинство; първата децилна група се състои от най-бедните 10%, а десетата децилна група се състои от най-богатите 10% от домакинствата.

Фиг. 6. Ефект от промените в данъчно-осигурителните и социалните политики през периода 2007 - 2011 г. по отделни компоненти и децилни групи



Източник: Hills et al. (2014), използвайки EUROMOD.

1. Ефект на детския минимален доход върху коефициента за риска от бедност при деца до 6 години

	Коефициент за риска от бедност през 2007 г. - %	Коефициент за риска от бедност след реформата - %		Спад в коефициента за риска от бедност - %	
		50 евро	50 евро ппс	50 евро	50 ппс
Белгия	12.2	10.0	9.8	-18	-20
България	26.1	18.1	21.8	-31	-16
Чешка република	9.6	7.2	7.6	-25	-21
Дания	6.9	6.9	6.8	0	-1
Германия	14.2	12.8	12.5	-10	-12
Естония	13.7	9.9	10.9	-27	-20
Ирландия	14.0	13.0	13.0	-7	-7
Гърция	20.4	18.1	18.5	-12	-10
Испания	17.0	15.5	15.5	-9	-9
Франция	16.6	15.4	15.0	-7	-9
Италия	20.2	18.4	18.4	-9	-9
Кипър	13.1	9.9	9.9	-25	-25
Латвия	21.4	17.0	18.6	-20	-13
Литва	15.8	11.6	13.4	-27	-15
Люксембург	9.4	7.5	7.2	-20	-23
Унгария	18.0	11.4	12.9	-37	-28
Малта	18.3	14.9	15.9	-19	-13
Нидерландия	11.7	10.1	9.2	-14	-22
Австрия	13.7	11.0	11.0	-20	-20
Полша	18.5	14.3	15.5	-23	-16
Португалия	15.1	11.7	12.9	-23	-15
Румъния	26.2	16.9	21.3	-35	-19
Словения	11.1	9.4	9.5	-15	-15
Словакия	13.8	9.6	11.2	-30	-19
Финландия	12.4	10.7	10.2	-14	-18
Швеция	12.1	11.9	11.8	-1	-2
Обединено кралство	19.7	17.3	17.3	-12	-12

Източник: Levy et al. (2013), използвайки EUROMOD.

Забележка: Коефициентът за риска от бедност е изчислен въз основа на 60% от медианния еквивалентен разполагаем доход на лице от домакинство. Всички деца на възраст до 6 години получават ДМД.

В друго приложение на EUROMOD Levy et al. (2013) изследват потенциалните последици от въвеждането на една реформа, т.нар. детски минимален доход (ДМД), на равнището на всички държави от ЕС. Симулационните резултати показват, че универсално заплащане от 50 евро на

месец за всяко дете на възраст под 6 години може да помогне на 800 000 деца в ЕС в тази възрастова група да се измъкнат от бедност. Статията също така разглежда ефекта и от реформата с алтернативно заплащане, при което стойността на ДМД отразява покупателната способност в отделните държави членки (50 евро паритет на покупателната способност (ппс)).

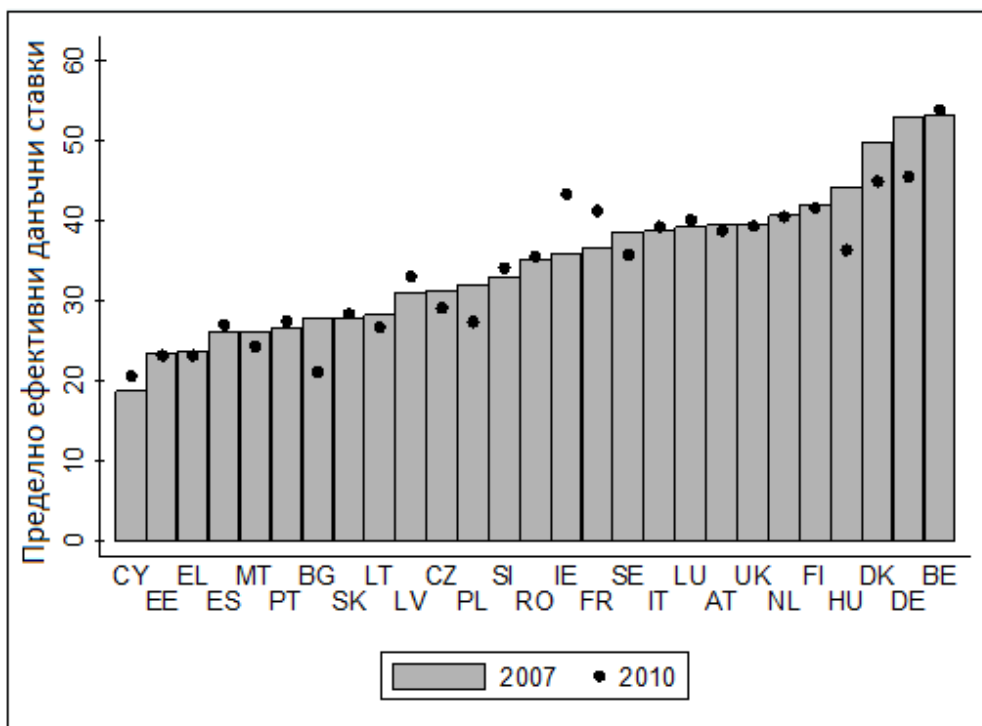
Използвайки микроданни от 2007 г. за всяка държава - членка на ЕС, е изчислен коефициентът на риск от бедност при деца на възраст до 6 години (табл. 1). След симулиране на реформата в EUROMOD с двете алтернативни плащания в третата и четвъртата колона на таблицата са показани преизчислените коефициенти на риска от бедност. Двете най-десни колони показват промяната в коефициентите (в %) при прилагане на двата вида плащания. Намалението в риска от бедност при децата на възраст до 6 години е най-голямо в Унгария (37%), а в Румъния и България надвишава 30%.

Последният пример е основан на разработката на Jaga и Tumino (2013). Фокусът на техния анализ е поставен върху пределните ефективни данъчни ставки (ПЕДС), които са показатели за промяната в облагането на допълнителните единици придобит облагаем доход. Те отразяват финансов стимул за дадено лице да увеличи доходите си от (само)заетост, като повиши броя на работните часове или производителността си, съпътствано обаче от увеличение в данъчно-осигурителните задължения и/или намаление на изплащаните социални обезщетения за домакинството, в което живее. Стойността на ПЕДС варира по правило между 0 и 1 (100%), при което размер на ставката, равен на 0, показва максимален финансов стимул (няма облагане на допълнителната единица доход), докато стойност 1 индикира нулев стимул (пълно изземване на допълнителната единица доход)⁶. Изчисленията се извършват от EUROMOD чрез следните стъпки: първо, изчислява се разполагаемият доход на домакинствата; след това приходите от (само)заетост се увеличават с 3% (това приблизително отговаря на 1 допълнителен час работа)

⁶ При размер на ставката, равен на 0, лицето може да получи 100% от допълнително спечеления доход, т.е. стимулт за работа е много висок. При размер от 50% лицето може да получи само 50% от допълнително спечеления доход, а останалите 50% загубва поради увеличени данъчно-осигурителни задължения и/или намалени изплащания на социални обезщетения. При стойност на ставката от 100% лицето няма никакъв финансов стимул да работи повече часове или при по-висока продуктивност, тъй като целият допълнително заработен доход бива „изгубен“. В зависимост от данъчно-осигурителната и социалната система на страната ПЕДС могат да бъдат отрицателни или да надвишат 100%.

отделно за всеки (само)зает в домакинството; след това се преизчислява отново разполагаемият доход на домакинството. Въз основа на промяната в разполагаемия доход спрямо увеличението на доходите от (само)заетост биват измерени ПЕДС (по-детайлно вж. Jara, Tumino, 2013). На фиг. 7 са изобразени средните стойности на ПЕДС за единиците на всяка държава членка през 2007 и 2010 г., като тяхната промяна се дължи на промените в данъчно-осигурителните и социалните политики (вземайки предвид ръста на доходите през периода) в държавите от ЕС-27 (Jara, Tumino, 2013). Резултатите сочат, че в България стимулите за работа са сравнително високи (ПЕДС са ниски) и биват увеличени през периода 2007 - 2010 г. главно поради въвеждането на плоския данък.

Фиг. 7. Средни за цялата съвкупност стойности (в %) на ПЕДС през 2007 и 2010 година



Източник: Jara, Tumino (2013), използвайки EUROMOD.

6. Бъдещо развитие и предизвикателства

Безспорно симулационният модел EUROMOD е необходим и полезен инструмент за социално-икономически анализ на неравенството, разпределението и преразпределението на дохода. Той притежава значителен потенциал за осигуряване на емпирична основа и обосновки при формулиране на адекватна политика за намаляване на бедността. Подобряването и

усъвършенстването на модела би рефлектирало в повишаване на точността на симулационните резултати, получени на базата на съответните за всяка година параметри на политиките за преразпределение на дохода. EUROMOD е инструмент за научнообосновано изследване на ефекта от различни алтернативни сценарии на реформи и съответстващите промени в статуса на групи от населението, засегнати от тях. Бъдещото развитие на модела се свързва главно със следните насоки:

Първо, да се продължи тенденцията симулационните резултати от EUROMOD да се използват при изготвяне на анализи от страна на Европейската комисия. Такива резултати често се използват в докладите „Annual Report on Employment and Social development” и „EU Employment and Social Situation Quarterly Review”.

Второ, усилията да се насочат към още по-голямо популяризиране на модела сред широката общественост.

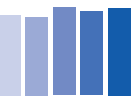
Трето, да се стимулират отговорните институции в държавите членки да използват резултатите при предлагането на реформи в областта на социалното подпомагане, социалното осигуряване и данъчното облагане.

Четвърто, да се активизира работата по използване на модела в различни научни разработки в социално-икономическата област.

За подобряване на работата на модела се очаква усилията да се насочат в следните конкретни направления:

Първо, подобряване на симулационните механизми на базата на направените сравнения с данни от административни източници. В този смисъл се очакват решения на проблеми, свързани с отразяване в модела на неплащани данъци и осигуровки и неполучавани социални помощи. Предвиждат се и изследвания на случаите, в които симулираните величини при конкретни единици по-добре отразяват реалното състояние в сравнение с регистрираните сведения за стойностни величини при извадковите изследвания, които могат да бъдат по-ниски от действителните или изобщо да не се посочват от респондентите.

Второ, важна насока на развитие на модела е и приспособяването му към промените в пазара на труда, настъпили в редица държави от ЕС в последните няколко години, свързани с икономическата криза и последвалото бавно възстановяване на икономиката.



Трето, подобряване на прецизността и равнището на детайлизация на несимулираните доходи за съответната година.

Четвърто, развитието на модела е свързано и с увеличаване на броя на държавите, използващи данни от **собствените национални извадкови изследвания** „Статистика на доходите и условията на живот” вместо ограничените и агрегирани променливи в потребителските бази данни (UDB), предоставяни от Евростат. Както беше споменато, българският модул в EUROMOD използва ограничени данни от националното изследване за 2008 г., което значително спомогна за процеса на интеграция на България в модела и подобри симулационните резултати. Тази насока е особено важна с оглед използването на по-детайлизирана информация за източниците на първичния доход, което би повишило полезността от микросимулационните анализи.

В края на анализа изразяваме нашата **признателност** и висока оценка за приноса на всички бивши и настоящи членове на EUROMOD консорциума. Процесът на разширяване и актуализиране на EUROMOD е подкрепен финансово от Генерална дирекция „Заетост, социални въпроси и приобщаване“ на Европейската комисия (Progress grant no. VS / 2011/0445).

ЦИТИРАНА ЛИТЕРАТУРА:

Andor, L., S. Dullien, X. H. Jara, H. Sutherland, D. Gros (2014). Designing a European unemployment insurance scheme. *Intereconomics* 49(4), 184 - 203.

Atkinson, A. B., F. Bourguignon, C. O'Donoghue, H. Sutherland, F. Utili (2000). Microsimulation and the formulation of policy: A case study of targeting in the European Union. A. B. Atkinson, H. Glennerster, N. Stern, *Putting economics to work: A volume in honour of Michio Morishima*. STICERD Occasional Paper 22, LSE.

Atkinson, T., F. Bourguignon, C. O'Donoghue, H. Sutherland, F. Utili (2002). Microsimulation of social policy in the European Union: Case study of a European Minimum Pension. *Economica* 69, 229 - 243.

Avram, S., H. Levy, H. Sutherland (2014). Income redistribution in the European Union. EUROMOD Working Paper EM8/14.

Avram, S., H. Sutherland, I. Tasseva, A. Tumino (2011). Income protection and poverty risk for the unemployed in Europe. *Social Situation Observatory Research Note* 1/2011.

Boshnakov, V., E. Tosheva, D. Dimitrova, D. Draganov (2010). Feasibility study: Republic of Bulgaria (tax-benefit systems 2007 - 2010, http://www.euro.centre.org/data/1291995455_72301.pdf).

Boshnakov, V., E. Tosheva, D. Dimitrova, D. Draganov, I. Tasseva (2012). EUROMOD Country Report Bulgaria, <https://www.iser.essex.ac.uk/euromod/resources-for-euromod-users/country-reports>.

Callan, T., H. Sutherland (1997). The impact of comparable policies in European countries: Microsimulation approaches. *European Economic Review* 41, 627 - 633.

De Agostini, P., A. Paulus, H. Sutherland, I. V. Tasseva (2014). The effect of tax-benefit changes on the income distribution in EU countries since the beginning of the crisis. EUROMOD Working Paper Series EM9/14.

Figari, F., M. Iacovou, A. J. Skew, H. Sutherland (2012). Approximations to the truth: comparing survey and microsimulation approaches to measuring income for social indicators. *Social Indicators Research* 105(3), 387 - 407.

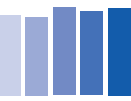
Hills, J., A. Paulus, H. Sutherland, I. Tasseva (2014). A lost decade? Decomposing the effect of 2001 - 11 tax-benefit policy changes on the income distribution in EU countries. *ImPRovE Working paper* 14/03.

Jara, H. X., C. Leventi (2014). Baseline results from the EU27 EUROMOD (2009-2013). EUROMOD Working Paper EM18/14.

Jara, H. X., A. Tumino (2013). Tax-benefit systems, income distribution and work incentives in the European Union. *International Journal of Microsimulation* 6(1), 27 - 62.

Levy, H., M. Matsaganis, H. Sutherland (2013). Towards a European Union Child Basic Income? Within and between country effects. *International Journal of Microsimulation* 6(1), 63 - 85.

Paulus, A., H. Sutherland, I. Tasseva (2014). Indexation matters: The distributional impact of fiscal policy changes in cross-national perspective. Unpublished manuscript presented at the 26th EALE Conference in Ljubljana, September 18 - 20, 2014.



Salauskaite, L., G. Verbist (2013). Is the neighbour's grass greener? Comparing family support in Lithuania and four other New Member States. *Journal of European Social Policy* 23, 315 - 331.

Steger, W. A. (1961). Simulation and tax analysis: A research proposal. *National Tax Journal* XIV(3), 286 - 301.

Sutherland, H., F. Figari (2013). EUROMOD: the European Union tax-benefit microsimulation model. *International Journal of Microsimulation* 6(1), 4 - 26.

Tasseva, I. (2012). Evaluating the performance of means-tested benefits in Bulgaria. ISER Working Paper 2012 - 18.

Tosheva, E., V. Boshnakov, D. Dimitrova, D. Daraganov (2014). EUROMOD Country Report: 2007 - 2013, https://www.iser.essex.ac.uk/files/euromod/country-reports/Year5/BG_CR_year_5_final%28with_annex%2920052014.pdf.

Wagenhals, G. (2004). Tax-benefit microsimulation models for Germany: A survey. IAW Report/Institut fuer Angewandte Wirtschaftsforschung, Tuebingen/32(1), 55 - 74.

ДАНЪЧНО-ТРАНСФЕРЕН СИМУЛАЦИОНЕН МОДЕЛ ЗА БЪЛГАРИЯ: ВЪЗМОЖНОСТИ ЗА СРАВНИТЕЛНИ СОЦИАЛНО-ИКОНОМИЧЕСКИ АНАЛИЗИ ЧРЕЗ EUROMOD

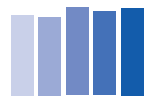
Венелин Бошнаков^{}, Екатерина Тошева^{**}, Ива Тасева^{***}*

РЕЗЮМЕ Целта на статията е да запознае специалистите от науката и практиката с възможностите на европейския модел EUROMOD - симулационен данъчно-трансферен модел за ЕС, в рамките на който има разработен модул за България. Данъчно-трансферните симулации са рядко използвани в социално-икономическите анализи за България до момента поради липсата на надежден симулационен модел. Такива анализи вече са възможни с помощта на EUROMOD, който осигурява уникален ресурс за реализирането на национални и сравнителни изследвания. Представени са различните етапи на конструирането на модела, основните от които са: подготовка на входните данни; симулация на социалните помощи, обезщетения, данъци и социални осигуровки; валидация на получените резултати. Някои възможни приложения на модела са демонстрирани чрез примери от актуални анализи, проведени в сравнителен план на общоевропейско ниво. Оценена е полезността на симулационните анализи в областта на разпределението и преразпределението на дохода, неравенството, бедността и стимулите на работа. Изведени са основни перспективи пред бъдещото развитие и използване на EUROMOD за целите на социално-икономическите анализи и разработването на данъчно-трансферни политики в България.

^{*} Доц. д-р в катедра „Статистика и иконометрия”, УНСС - София; e-mail: venelinb@unwe.bg.

^{**} Доц. д-р в катедра „Статистика и иконометрия”, УНСС - София; e-mail: ekaterina_tosheva@yahoo.com.

^{***} Институт по социални и икономически изследвания, Университет Есекс, Обединено кралство; e-mail: itasseva@essex.ac.uk.



НАЛОГОВО-ТРАНСФЕРНАЯ СИМУЛЯЦИОННАЯ МОДЕЛЬ ДЛЯ БОЛГАРИИ: ВОЗМОЖНОСТИ ДЛЯ СОПОСТАВИТЕЛЬНОГО СОЦИАЛЬНО-ЭКОНОМИЧЕСКОГО АНАЛИЗА С ПОМОЩЬЮ EUROMOD

Венелин Бошнаков^{}, Екатерина Тошева^{**}, Ива Тасева^{***}*

РЕЗЮМЕ Цель статьи состоит в том, чтобы ознакомить специалистов из области науки и практики с возможностями европейской модели EUROMOD – симуляционной налогово-трансферной модели для ЕС, в рамках которой разработан и модуль для Болгарии. К настоящему времени налогово-трансферные симуляции не очень часто используются в социально-экономических исследованиях для Болгарии в связи с отсутствием надежной симуляционной модели. Такие анализы теперь возможны с помощью EUROMOD, которая обеспечивает уникальный ресурс для осуществления национальных и сопоставительных исследований. Представлены различные этапы построения модели, основными из которых являются: подготовка исходных данных; симуляция социальных пособий, компенсаций, налогов и взносов на социальное страхование; проверка полученных результатов. Некоторые возможные приложения модели продемонстрированы на примерах из недавних анализов, проведенных в сопоставительном плане на европейском уровне. Оценивается полезность симуляционных анализов в области распределения и перераспределения дохода, неравенства, бедности и стимулов к труду. Выявлены основные перспективы для дальнейшего развития и использования EUROMOD для целей социально-экономических анализов и разработки налогово-трансферных политик в Болгарии.

* Доцент, д-р на кафедре Статистики и эконометрии Университета национального и мирового хозяйства, г. София; электронная почта: venelinb@unwe.bg.

** Доцент, д-р на кафедре Статистики и эконометрии Университета национального и мирового хозяйства, г. София; электронная почта: ekaterina_tosheva@yahoo.com.

*** Институт социальных и экономических исследований, Университет Есекса, Обединенное королевство; электронная почта: itasseva@essex.ac.uk.

**A TAX-BENEFIT MICROSIMULATION MODEL FOR BULGARIA:
OPPORTUNITIES FOR CROSS-COUNTRY SOCIO-ECONOMIC ANALYSIS
USING EUROMOD**

Venelin Boshnakov, Ekaterina Tosheva**, Iva Taseva****

SUMMARY This paper aims to provide an introduction for experts in the field with benefits of the European model EUROMOD - microsimulation tax-benefit model within the framework of which there is a developed model for Bulgaria. Tax-benefit microsimulation is rarely used for policy analysis and design in Bulgaria because until recently no such reliable model existed. This has now been made possible with the use of EUROMOD which is a unique resource for national and comparative research. The different stages of its construction are described in the model: preparing the input micro-data utilised by the model; social security simulation; compensations tax and social insurance; validation of the results. Some possible applications of the model are demonstrated with examples of analyses at European level in a comparative plan. The usefulness of simulation analyses is evaluated in the field of distribution and redistribution of income inequality, poverty and incentives to work. Outlined are the main prospects for the future development and use of EUROMOD for the purposes of socio-economic analysis and the development of tax-transfer policies in Bulgaria.

* Associate Professor Dr. in 'Statistics and Econometrics' Department, University of National and World Economy - Sofia; e-mail: venelinb@unwe.bg.

** Associate Professor Dr. in 'Statistics and Econometrics' Department, University of National and World Economy - Sofia; e-mail: ekaterina_tosheva@yahoo.com.

*** Institute of Social and Economic Research, University of Essex, United Kingdom; e-mail: itasseva@essex.ac.uk.



ПОДХОДИ ЗА МЕСЕЧНА ОЦЕНКА НА РАВНИЩЕТО НА БЕЗРАБОТИЦА ВЪЗ ОСНОВА НА ДАННИ ОТ НАБЛЮДЕНИЕТО НА РАБОТНАТА СИЛА

*Александър Найденов**



Въведение

Равнището на безработица е един от важните индикатори за състоянието на икономиката на нашата страна и участва в агрегатните индикатори за икономиката на Европейския съюз. Необходимостта от осигуряване на месечни оценки на равнището на безработица в България се детерминира както от нуждата от задълбочен и навременен анализ на социално-икономическите процеси във връзка с осъществяване на контрол върху тяхното развитие, така и от поставените изисквания от Евростат относно „производството“ на месечни оценки на равнището на безработица и хармонизиране на методологията в Общността¹.

1. Подходи при месечната оценка на безработицата в някои европейски страни

В практиката съществува голямо разнообразие от методи за оценка на месечния размер на безработицата. Това се дължи на различни фактори като използвания тип извадка, наличието на достоверна допълнителна информация (напр. от регистри), на методологически добре подготвен екип, който да извърши оценяването, и други.

Някои практики на страни от Европа, и в частност от ЕС², са обобщени в табл. 1.

* Гл. ас. д-р, преподавател към катедра „Статистика и иконометрия“, УНСС - София; e-mail: anaidenov@gmail.bg.

¹ Повече информация може да се намери на сайта на Евростат, в частта, отнасяща се до „Ажустираните оценки на безработицата, получени въз основа на данни от Наблюдението на работната сила“ (Unemployment - LFS adjusted series).

² Пак там.

1. Подходи за месечна оценка на безработицата по страни

Страни	Подход за оценка
Австрия	До края на 2010 г. са използвани методи за моделиране, комбиниращи тримесечни оценки с данни за регистрираните безработни. След 2011 г. се използват месечни директни оценки въз основа на калибрирани данни, като не се извършва сезонно изглаждане.
Германия	До края на 2004 г. се използват годишни данни (наблюдението е част от т.нар. Микросенсус), а от 2005 г. наблюдението става непрекъснато и се тестват месечни оценки, обикновено комбинирани с данни от регистри.
Великобритания (Обединено кралство)	Използват се плъзгащи се средни, като се комбинират данните от три последователни месеца, а резултативната средна се отнася за месеца, който се намира в средата на тримесечието.
Турция	Използват се основно тримесечни плъзгащи се средни.
Португалия	Комбинират се данни от Наблюдението на работна сила (НРС) и регистрираните безработни, като се използва методологията на Чоу-Лин за дезагрегация на данните. Не се извършва сезонно изглаждане.
Италия	От началото на 2004 г. се извършва месечно оценяване въз основа на претегляне чрез калибриране.
Румъния	Използват се директно данните от НРС, като впоследствие се изграждат ARIMA модели от типа TRAMO-SEATS, прилагайки софтуерния продукт Demetra. За разпространение се предоставят само данни за заетите и безработните - общо за възрастните заети и безработни (25+ години). След изтичане на дадено тримесечие данните за предходните месеци се ревизират.

Както става ясно от таблицата, всяка от страните сама определя подхода за решение на проблема по адекватното оценяване на месечните стойности на безработицата. В някои случаи се разчита по-скоро на „директни“ оценки от изследването, свързано с пазара на труда, а в други - на приложението на сложни оценителни процедури.

2. Възможни подходи за получаване на месечни оценки на равнището на безработица в България

По същество месечната оценка на равнището на безработица се получава въз основа на отделното месечно оценяване на броя на безработните и отделната месечна

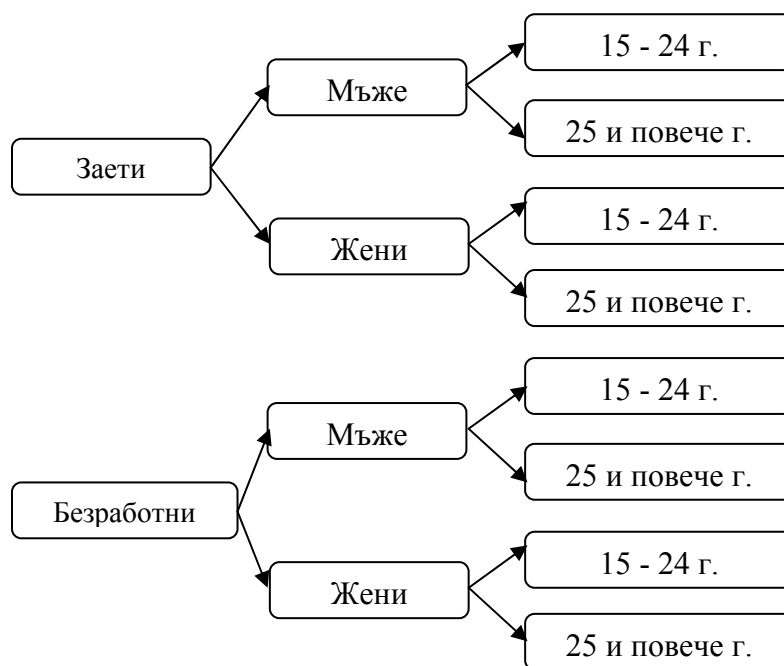
оценка на броя на заетите. Това се дължи на факта, че равнището на безработица е производна величина, включваща посочените два компонента, които, от своя страна, проявяват специфичност по отношение на тяхната същност и динамика.

Преди да се обърне внимание на възможните стратегии за месечно оценяване, е необходимо да се направят следните няколко уточнения:

- Използването на допълнителна информация при комбинирането на данни от Наблюдението на работната сила и регистрираните безработни в процеса на оценяване не е надеждно, тъй като е налице сериозно разминаване между регистрираните безработни в България (от бюрата по труда) и тези, отчетени от НРС на Националния статистически институт. Това вероятно се дължи на силната зависимост на броя на регистрираните от промените в трудовото законодателство на България. Аналогичен е и проблемът при използването на данни от тримесечното изследване на наетите лица, отработеното време, средствата за работна заплата и други разходи за труд (отново на НСИ). Оттук можем да заключим, че единственият възможен вариант е провеждането на анализа **само въз основа на данни** от Наблюдението на работната сила.

- Съобразявайки се със съществуващите практики в страните от ЕС за оценка на безработицата, е желателно анализът да бъде извършен за 8 основни групи (категории), представени на фиг. 1.

Фиг. 1. Групи за анализ и оценка



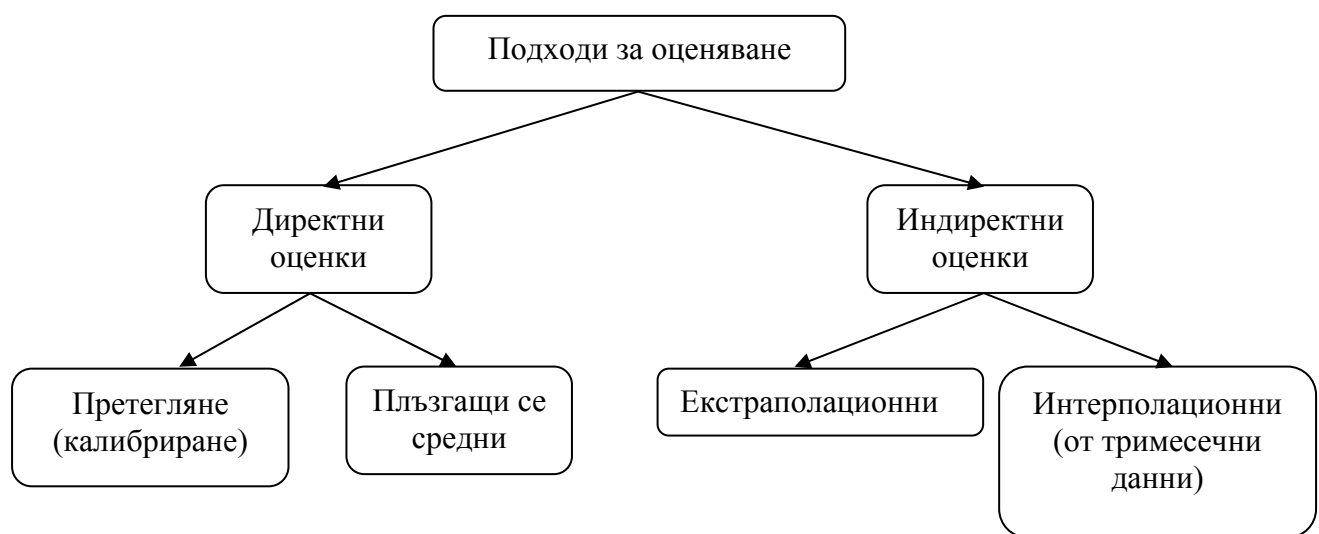
Въз основа на проучената информация, свързана с оценките на равнището на безработица в страните от ЕС и конкретните възможности на НСИ, беше установено, че възможните стратегии за оценка на месечното равнище на безработица (аналогично - брой заети и брой безработни) са най-общо две:

- Като **директни оценки** от данните от Наблюдението на работната сила, използвайки техники за претегляне (калибриране) на данните с цел ажустиране³ на месечните данни към населението по отделни категории - пол, възраст, вид населено място и други.

- Като **индиректни оценки**, комбинирайки данните от НРС и прилагайки методи за моделиране (екстраполиране и интерполиране) на развитието в динамичните редове на заетите и безработните.

Подходите за анализ са представени най-общо на фиг. 2.

Фиг. 2. Подходи за месечна оценка на броя на заетите и броя на безработните



В следващата част от изложението подходите за оценяване са разгледани по-подробно, като са описани както самата процедура, така и софтуерните продукти, с които тя може да бъде реализирана.

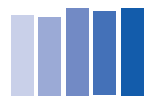
2.1. Директно оценяване

2.1.1. Директни месечни оценки чрез претегляне на данните

През 2007 г. беше извършено за първи път пробно оценяване на месечния брой на заетите и месечния брой на безработните⁴, в процеса на което възникнаха някои проблеми. Те се дължаха на факта, че събраните първични данни от НРС за даден месец се претегляха, използвайки постстратификация по ограничен брой страти - пол, местоживеене (град/село) и четири възрастови интервала. Въпреки че този тип постстратификация е „по-опростен“ в сравнение с прилагания при тримесечните данни, често това води до наличието на празни страти (напр. липса на мъже на възраст 15 - 24

³ Ажустирането (adjustment) е статистическа техника, която цели да „коригира“ резултатите от едно емпирично извадковото изследване съгласно предварително известни данни за генералната съвкупност. Към техниките на ажустиране се включват претеглянето (weighting), калибрирането (calibration) и други. В текста термините са използвани като синоними.

⁴ Проект „Quality improvement in Unemployment Statistics“, ФАР 2004.



години в някои области от NUTS3), което, от своя страна, не дава възможност за коректно възпроизвеждане на генералната съвкупност.

От началото на 2012 г. за претегляне на данните от Наблюдението на работната сила се използва нова методология, която (за разлика от постстратификационния подход) при определяне на тегловните коефициенти включва⁵:

- тегла, представляващи реципрочната стойност на вероятността домакинство от дадена страта да попадне в извадката;
- корекция за неотговорили;
- калибриране (ажустиране на данните от извадката към данните от населението): общо за страната - по възрастови групи, пол и вид на населеното място (град/село), и по области - пол и вид на населеното място (град/село).

За претеглянето на резултатите се използват последните актуални данни за населението (в частност и домакинствата) на Република България, получавани в резултат от дейността на демографската статистика.

Новата методология по претеглянето на данните дава в известна степен възможност да се извърши директна месечна оценка на броя на заетите и броя на безработните лица, като се преодоляват някои недостатъци на постстратификационното претегляне (част от които са празните страти), като същевременно се отчита величината на неотговорилите и се осигурява съпоставимост на месечните с тримесечните оценки. За да бъде потвърдена коректността и адекватността на новата методология обаче, е необходимо минимум експериментиране в продължение на една година. За практическото приложение на оценителните процедури могат да бъдат използвани MS Excel и SPSS (за процедурите по изчисляване на теглата), а за калибрирането - g-Galib.

2.1.2. Месечни оценки въз основа на плъзгащи се тримесечни средни

В някои случаи, за да се избегнат недостатъците на директните месечни оценки, получени в резултат на постстратификационно претегляне, се предлага използването на комбинирана информация, която включва съчетаването на данни в следната последователност:

- „Суровите“ данни от три последователни месеца се обединяват, образувайки „фиктивно тримесечие“ (напр. февруари, март и април вместо стандартните януари, февруари и март).
- Данните се претеглят по стандартната процедура, която се прилага при обикновените тримесечни оценки (постстратификация или калибриране).
- Полученият резултат (съответно за броя на заетите и броя на безработните за описаните 8 групи) се използва като оценка за месеца, намиращ се в средата на съставеното „фиктивно“ тримесечие.

Основен недостатък е „изоставането“ във времето, тъй като резултатите винаги се получават най-малко с един месец закъснение. Въпреки че има страни от ЕС, които използват подобни оценки, трябва да се има предвид, че те се получават с прекалено голямо закъснение, а и водят до прекалено изглаждане на динамичния ред.

⁵ Съгласно методологията на Наблюдението на работната сила от 2012 година.

За приложението на този подход могат да бъдат използвани SPSS и MS Excel, а при евентуално калибриране - и g-Calib.

2.2. Индиректно оценяване

2.2.1. Екстраполиране въз основа на директните месечни оценки

Този подход е неразривно свързан с директното оценяване, тъй като оценките, получени от него, са основа за екстраполация (прогноза) на броя на заетите и броя на безработните за периодите, следващи последния месец, за които се разполага с данни. Използването на подобен подход има едно основно **предимство** - прогнозираният брой за интересувания ни месец е наличен преди неговото изтичане (за разлика от първия директен подход).

Прилагането на методите за екстраполация изисква преминаването през няколко стъпки:

- Извършва се директно оценяване (по описания начин) за определен период от време. Например за периода от началото на 2008 г. до края на 2011 г., като се използват наличните данни за населението от демографската статистика.

- Проверява се „качеството” на динамичния ред, т.е. редът се разглежда по отношение на наличието на: липсващи стойности (missing values), силно отличаващи се стойности (outliers) и съпоставимост на данните във времето от гледна точка на използваната методология.

- Проверява се наличието/отсъствието на стационарност в реда, като се използват автокорелационните коефициенти. Отговаря се на въпроса дали в реда има тенденция на развитие или не.

- Декомпозира се времевият ред, т.е. разграничават се и се оценяват компонентите му: тренд, сезонен компонент и случайни отклонения. В теорията се посочва, че за целите на декомпозицията времевият ред трябва да съдържа поне 4 дължини на сезонния цикъл, т.е. в конкретния случай с месечните оценки: $4 \cdot 12 = 48$ месеца.

- Моделира се трендът - оценяват се параметрите на модел, описващ адекватно тенденцията в изследваното явление. Като най-подходящи в литературата се посочват ARIMA моделите и в частност по-съвременните им варианти: TRAMO-SEATS (разработен в Испания и използван предимно в страните от ЕС) и X-12-ARIMA (използван основно в САЩ). На този етап се решават и въпросите, свързани с определянето на типа модел на връзка между компонентите на времевия ред - адитивен или мултипликативен, и необходимостта от календарни корекции (във връзка с различния брой работни дни в месеца, националните празници и други⁶).

- Прогнозират се броят на заетите и броят на безработните въз основа на оценения в предходния етап модел.

- Коригира се прогнозата със сезонния компонент.

- Изчислява се грешката на прогнозата.

⁶ От предходни изследвания е установено, че върху разглежданите динамични редове на заетите и безработните календарните корекции не са необходими.

За приложението на този тип оценяване основно може да се използва софтуерният продукт Demetra+ ver.1.04, тъй като той комбинира всички необходими процедури за анализ на динамични редове, но могат да бъдат използвани и аналогични продукти като: Winx12, TRAMO-SEATS ver. 1.97, JMultiVM, BV4.1, които обаче осигуряват само част от необходимите процедури на анализа.

2.2.2. Интерполиране въз основа на „оригиналните“ тримесечни оценки

Интерполирането е подход, при който даден времеви ред с по-висока агрегираност (напр. тримесечни оценки) се трансформира във времеви ред с по-голяма дезагрегираност (напр. месечни оценки). Този метод се основава на сложни итеративни процедури като например тези на Чоу-Лин (Chou-Lin), Буут (Boot), Фийбс (Fiebes), Лисман (Lisman) и др., които се осъществяват само с помощта на специализиран статистически софтуер.

Въз основа на получения нов динамичен ред от дезагрегирани месечни оценки се извършва екстраполация по описаната в предходната точка последователност. За интерполацията може да се използва основно ECOTRIM ver.1.01, а за последващата екстраполация - Demetra+.

В процеса на прилагане на изложените подходи е необходимо да се извърши и проверка за „точността“ им с цел да се избере най-подходящият от тях във връзка с евентуалното им практическо приложение. Затова в практиката много често се изчисляват т.нар. грешки на оценката. Те се представят обикновено в абсолютен и относителен вид. Един от най-популярните и лесно разбираеми измерители е т.нар. средна абсолютна грешка, представяна във вида:

- Средна абсолютна грешка (Mean Absolute Error):

$$MAE = \frac{1}{n} \sum_{t=1}^n |Y_{dtr} - Y_{indtr}|,$$

където:

Y_{dtr} – брой на заетите/безработните за съответната група (1...8), получен в резултат на директното оценяване;

Y_{indtr} – брой на заетите/безработните за съответната група (1...8), получен в резултат на индиректното оценяване;

n – брой периоди.

Величината на тази грешка представя средното отклонение на оценките, получени в резултат на моделирането спрямо „действителните“ директни оценки, измерено в брой лица.

- Средна абсолютна грешка, изразена в проценти (средна абсолютна процентна грешка - Mean Absolute Percentage Error):

$$MAPE\% = \frac{1}{n} \sum_{t=1}^n \frac{|Y_{dtr} - Y_{indtr}|}{Y_{dtr} + Y_{indtr}} \cdot 100$$

Този измерител на грешката представя средното отклонение на оценките, получени в резултат на моделирането спрямо „действителните“ директни оценки,

измерено в проценти. Движи се в границите от 0 до 100%, а колкото стойността е по-близка до нула, толкова по-точен е използваният подход за моделиране.

За да се провери предикативната способност на използваните индиректни методи, в практиката обикновено се постъпва по следния начин:

- Прилага се съответният индиректен метод върху разполагаемата информация, като предварително динамичният ред се „скъсява“ изкуствено в края. В конкретния случай могат да бъдат отстранени последните три месеца на 2011 година.

- След моделирането на реда се извършва прогнозиране въз основа на разработения модел за остатъка на реда, който е отстранен предварително - преди моделирането.

- Сравняват се „реалните“ данни от края на реда с прогнозираните⁷.

По този начин може да се провери доколко използваният модел дава адекватни предположения относно бъдещото развитие на разглеждания динамичен ред.

3. Експериментални резултати от приложението на подходите за месечна оценка на безработицата

В процеса на експериментиране се преминава през следните етапи:

Първи етап. Набавяне на необходимата информация

За провеждане на експериментите относно построяването на месечни оценки за броя на заетите/безработните са използвани следните времеви редове:

- Брой заети и брой безработни (по групи: мъже на 15 - 24 и 25+ години, жени на 15 - 24 и 25+ години) по месеци за периода януари 2008 - декември 2011 г.⁸ - непрегледени данни, за да се провери реалният брой лица, които попадат във всяка от 8-те групи.

- Брой заети и брой безработни (по групи: мъже на 15 - 24 и 25+ години, жени на 15 - 24 и 25+ години) по месеци за периода януари 2008 - декември 2011 г., получени въз основа на **директно оценяване** (по новата методология от 2012 година). Трябва да се отбележи, че дължина на динамичния ред от 48 подпериода (месеца) е достатъчна за извършване на моделирането, като същевременно се спазват и изискванията за анализа на динамични редове, налагащи използването на съпоставими във времето данни, т.е. данните за времето преди 2012 г. трябва да бъдат преизчислени по новата методология. Използването на по-дълги времеви редове (напр. включително за периода януари 2003 - декември 2007 г.) е желателно, но е свързано със значително количество допълнителен труд от страна на експертите от Наблюдението на работната сила, тъй като те трябва да преизчислят (препрегледят) данните за всеки месец от всяка година с новата методология поради факта, че до момента данните от предходни периоди са получавани по различна методология.

- Брой заети и брой безработни (по групи мъже на 15 - 24 и 25+ години, жени на 15 - 24 и 25+ години) по тримесечия за периода първо тримесечие 2008 - четвърто

⁷ В статията сравнението се извършва въз основа на средната абсолютна грешка, но в литературата съществуват и други методи като например коефициента на Тейл.

⁸ При експерименталните оценки не са използвани най-актуалните данни, тъй като целта е апробация на методологични подходи, а не получаването на актуални резултати.

тримесечие 2011 г., получени въз основа на новата методология за претегляне на данните от 2012 година.

Втори етап. Форматиране на данните във вид, подходящ за използване в софтуерните продукти

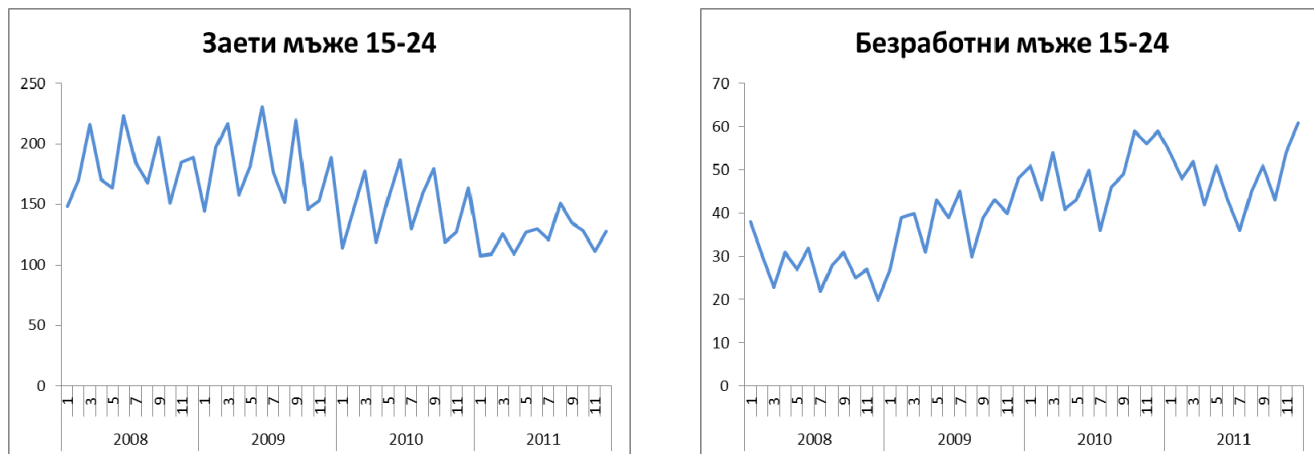
След като експертите от НСИ, провеждащи Наблюдението на работната сила, подготвят необходимите първични данни за анализа, описани в предходната точка, данните се предоставят за тестване на различни подходи за оценка на броя на заетите и броя на безработните лица. След това се преформатират във вид, подходящ за използването им в посочения софтуер за индиректно оценяване, като за целта се използват електронни таблици на MS Excel.

Трети етап. Изчисляване на оценки на броя на заетите и броя на безработните по месеци съгласно изложената методология⁹

3.1. Директни оценки

Директното оценяване е неразривно свързано с реалния брой анкетирани лица, принадлежащи към 8-те групи, тъй като обемът на извадката за съответната група определя и величината на стохастичната грешка на оценката за тази група. Визуално броят на лицата, попаднали в извадката по групи, е представен на фиг. 3.

Фиг. 3. Брой лица в извадката на НРС по месеци и групи



⁹ В настоящата статия не са публикувани подробно конкретните резултати от процеса на оценяване поради тяхната поверителност.



От представената информация става ясно, че проблеми могат да възникнат единствено при оценяването за групите на безработните мъже и жени на възраст 15 - 24 години, тъй като при тях броят на анкетираните лица е изключително малък - за жените достига 9 единици през септември 2008 година. От гледна точка на тенденцията в броя на единиците, попаднали в посочените две групи, се забелязва положителна тенденция (от статистическа гледна точка) към увеличаването им към края на периода.

3.2. Индиректно оценяване

3.2.1. Чрез екстраполация

Въз основа на описаната методология и предоставените данни, използвайки програмния продукт Demetra, е извършена екстраполация (вкл. и сезонно изглаждане) на времевите редове за 8-те разглеждани групи. Трябва да се отбележи, че поради значителното усъвършенстване на този програмен продукт в него са автоматизирани всички етапи, разгледани в т. 2.2.1.

При прилагането на метода „X12“ не са установени отклонения от минималните изисквания¹⁰, поради което получените оценки могат да се окачествят като „добри“. Това може да се види и от обобщаващия прозорец, който се генерира от софтуерния продукт при провеждането на анализа (фиг. 4).

Фиг. 4. Екранна снимка от програмния продукт Demetra относно приложението на методиката „X12“

Series	Method	Estimation	Proces...	Priority	Quality
em1524f	X12	Concurrent	Valid		Good
em1524m	X12	Concurrent	Valid		Good
em25f	X12	Concurrent	Valid		Good
em25m	X12	Concurrent	Valid		Good
un1524f	X12	Concurrent	Valid		Good
un1524m	X12	Concurrent	Valid		Uncertain
un25f	X12	Concurrent	Valid		Good
un25m	X12	Concurrent	Valid		Good

summary
Uncertain
basic checks
definition: Good (0,000)
annual totals: Good (0,000)
visual spectral analysis
spectral seas peaks: Good
spectral td peaks: Good
regarima residuals
normality: Good (0,142)
independence: Good (0,922)
spectral td peaks: Good (0,174)
spectral seas peaks: Bad (0,008)
residual seasonality
on sa: Good (0,989)
on sa (last 3 years): Good (0,970)
on irregular: Good (0,887)
outliers
number of outliers: Good (0,000)

Единствено изключение прави оценката на броя на безработните мъже на възраст 15 - 24 години, където едно от условията не е изпълнено, но въпреки това оценяването е приемливо.

Приложението на метода „TRAMO-SEATS“ също преминава без особени проблеми, включително и за посочената рискова група. Това може да се види от екранната снимка, показана на фиг. 5.

¹⁰ В настоящата статия не се разглеждат подробности относно изискванията за приложение на методите TRAMO-SEATS и X12, тъй като това излиза извън нейното предназначение. Повече информация за изисквания може да бъде открита на адрес: http://www.cros-portal.eu/sites/default/files//Demetra%2B%20User%20Manual%20November%202012_0.pdf.

Фиг. 5. Екранна снимка от програмния продукт Demetra относно приложението на методиката „TRAMO-SEATS“

Series	Method	Estim...	Proces...	Quality
em1524m	TS[RSA3]	Concur...	Valid	Good
un1524m	TS[RSA3]	Concur...	Valid	Good
em1524f	TS[RSA3]	Concur...	Valid	Good
un1524f	TS[RSA3]	Concur...	Valid	Good
em25m	TS[RSA3]	Concur...	Valid	Good
un25m	TS[RSA3]	Concur...	Valid	Good
em25f	TS[RSA3]	Concur...	Valid	Good
un25f	TS[RSA3]	Concur...	Valid	Good

summary
Good
basic checks
definition: Good (0,000)
annual totals: Good (0,002)
visual spectral analysis
spectral seas peaks: Good
spectral td peaks: Good
regarima residuals
normality: Good (0,493)
independence: Good (0,477)
spectral td peaks: Good (0,301)
spectral seas peaks: Uncertain (0,028)
residual seasonality
on sa: Good (0,817)
on sa (last 3 years): Good (0,693)
on irregular: Good (0,976)
outliers
number of outliers: Good (0,022)

При необходимост софтуерният продукт дава възможност анализът на отделния времеви ред да се задълбочи значително. Това се отнася както за графичното представяне на отделните компоненти на времевия ред - тренд, сезонна вълна, остатъци, така и за подробен анализ на сезонността в реда; детайли относно процеса на ARIMA моделирането; спектрален анализ на данните; анализ на разпределението на остатъците и много други.

3.2.2. Интерполиране

Въз основа на предоставените тримесечни данни от Наблюдението на работната сила и софтуерния продукт Escotrim е извършено дезагрегиране на данните до месечна основа, без използването на допълнителна информация (т.е. чрез методите на Boot, Feibes и Lisman).

След интерполирането е извършено екстраполиране на резултатите аналогично на направеното в предходната точка. Тук обаче се появяват някои проблеми при оценката на броя на заетите млади жени и на броя на безработните мъже на възраст 24 и повече години, използвайки методологията „TRAMO-SEATS“, което, от своя страна, води до невъзможност да се оценят и свързаните с тях коефициенти на безработица. Проблемите се дължат основно на нарушаване на изискванията на метода, свързани с независимост на остатъците.

3.3. Сравнение на получените резултати

За да се извърши сравнение на резултатите, получени от различните индиректни методи, са използвани формулите, изложени в т. 2.2.2, отнасящи се до средните абсолютни грешки. В обобщение може да се каже, че оценките, получени чрез екстраполация, дават най-добри резултати, но не трябва да се забравя, че те са

възможни едва когато съответното тримесечие е изтекло, т.е. със значително закъснение.

Съобразявайки се с тази особеност, можем да посочим най-добрите методи за съответната група от 8-те разглеждани групи (табл. 2).

2. Методи по групи и величина на грешката

Групи	Най-добър метод	Величина на грешката - %
Заети мъже (15 - 24 години)	Екстраполиране с „TRAMO-SEATS”	3.1
Безработни мъже (15 - 24 години)	Екстраполиране с „X12“	3.7
Заети жени (15 - 24 години)	Екстраполиране с „TRAMO-SEATS“	3.0
Безработни жени (15 - 24 години)	Екстраполиране с „TRAMO-SEATS“ и „X12“	6.8
Заети мъже (25 и повече години)	Екстраполиране с „X12“	0.7
Безработни мъже (25 и повече години)	Екстраполиране с „TRAMO-SEATS“	3.5
Заети жени (25 и повече години)	Екстраполиране с „TRAMO-SEATS“	0.8
Безработни жени (25 и повече години)	Интерполиране с Ecotrim и екстраполиране с „X12“	3.1

При прогнозните стойности резултатите са аналогични с изключение на:

- Заети жени (15 - 24 години): екстраполиране с „TRAMO-SEATS“ (2.5% грешка)
- Заети мъже (25 и повече години): екстраполиране с „TRAMO-SEATS“ (0.4% грешка)
- Заети жени (25 и повече години): екстраполиране с „X12“ (1.2% грешка).

4. „Качество“ на получените резултати от приложените подходи

За да се оцени „качеството“ на получените резултати, е необходимо да се обърне внимание на грешката, с която са обременени различните методи (фиг. 6).

Фиг. 6. Грешки при различните методи за оценяване

Тип оценка	<i>Директна оценка</i>		Стохастична = грешка (данни от извадка)	Грешка от ажустирание (калибриране)				
	<i>Индиректни оценки</i>	X12	Стохастична = грешка (данни от извадка)	Грешка от ажустирание (калибриране)	Грешка от моделиране (екстраполиране)			
		TRAMO-SEATS	Стохастична = грешка (данни от извадка)	Грешка от ажустирание (калибриране)	Грешка от моделиране (екстраполиране)			
		Ecotrim	Стохастична = грешка (данни от извадка)	Грешка от ажустирание (калибриране)	Грешка от моделиране (интерполиране)			
		Ecotrim-X12	Стохастична = грешка (данни от извадка)	Грешка от ажустирание (калибриране)	Грешка от моделиране (интерполиране)	Грешка от моделиране (екстраполиране)		
		Ecotrim-TRAMO-SEATS	Стохастична = грешка (данни от извадка)	Грешка от ажустирание (калибриране)	Грешка от моделиране (интерполиране)	Грешка от моделиране (екстраполиране)		

От гледна точка на точността на оценяващата процедура става ясно, че най-добрият подход за оценяване на броя на заетите и броя на безработните е директното оценяване, тъй като теоретично при него размерът на грешката е най-малък. Всеки един от индиректните методи не само че „наследява“ грешките от директния, но дори допълнително „влошава“ точността на резултатите.

5. Проблеми при оценяването

В процеса на оценяване въпреки сложността на прилаганата методология не възникват особени затруднения от техническо естество поради наличието на съответния софтуер, с чиято помощ са извършени оценителните процедури. Въпреки това трябва да се отбележи следното:

- Използването на софтуер за индиректно оценяване (екстраполиране и интерполиране) на броя на заетите и броя на безработните по групи изисква задълбочени познания по статистика и в частност статистически анализ на времеви редове, за да се реализират адекватно и пълнофункционално възможностите, които предоставя този софтуер.

- Директното оценяване е трудоемък процес, който изисква сериозни познания, свързани със съставянето на коректен програмен код, който да бъде изпълнен в условията на калибрацията продукт g-Calib. Този продукт предизвиква и допълнителни затруднения, тъй като при създаването му не са съобразени неминуемите еволюционни промени, настъпващи в компютърните операционни системи, а това налага много строги (почти неизпълними) изисквания относно хардуера и софтуера, на които ще бъде инсталиран g-Calib.

От гледна точка на получените резултати възникнаха и някои проблеми при моделиране на времевите редове, което се дължи основно на корелираност на остатъците в динамичния ред.

Изводи

Въз основа на проведенния анализ могат да се направят следните по-важни изводи:

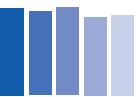
- Осигуряването на месечни оценки на броя на заетите, броя на безработните и произтичащия от тях коефициент на безработица е необходимо и обективно обусловено. Основни причини за това са: необходимостта от вземане на адекватни управленски решения, насочени към намаляване на негативните последици от липсата на препитание, и по-доброто управление на пазара на труда.

- В статистическата практика съществува голямо разнообразие от методи за месечно оценяване на безработицата, но приложението им зависи от възможностите на съответната статистическа институция от гледна точка на наличната информация и методологията за нейното събиране.

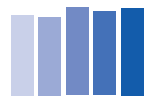
- Възможни за приложение в български условия се оказаха два типа методи: директни (чрез претегляне/калибриране и плъзгащи се средни) и индиректни (чрез екстраполация и интерполация на месечните данни).

- От всички разглеждани методи най-близко до търсения метод за месечно оценяване се оказа директният, използващ калибриране на данните. Въпреки че извадката на Наблюдението на работната сила е проектирана да предоставя предимно тримесечни оценки, тя в известна степен е достатъчна да удовлетвори и нуждата от месечно оценяване. Въведената нова методология по претегляне на данните от Наблюдението през 2012 г. дава нови възможности за по-адекватно месечно претегляне на данните, като същевременно данните при нужда могат да бъдат „произведени“ месечно в срокове, които биха удовлетворили изискванията, поставени от Евростат. Поради посочените причини този метод е препоръчителен за използване в бъдеще при възникване на евентуална необходимост от месечно оценяване на разглежданите тук характеристики, свързани с пазара на труда - заети и безработни. Трябва обаче да се има предвид и това, че от една страна, месечната оценка на безработицата за възрастовия интервал 15 - 24 години е ограничена (най-вече поради малкия брой единици в извадката), а от друга страна, използваната методология за получаване на тези директни оценки предстои да се тества във времето, за да се установи нейната достоверност и адекватност.

- От своя страна и останалите методи (индиректните оценки) имат някои преимущества като например това, че те могат да осигуряват данни за даден месец преди неговото изтичане, т.е. въз основа на динамиката на изследвания динамичен ред в предходни периоди. Основният проблем при тях е усложняването на оценителния процес, затруднения в изчисляването на окончателния размер на стохастичната грешка на оценката и натрупването на значителен размер грешка на оценяването. Въпреки това при необходимост индиректните оценки биха намерили приложение в практиката, съобразявайки се с техните особености. Различието в резултатите, което беше установено в процеса на тестване на различните методики за оценяване, най-вероятно се дължи на „прекаленото“ изглаждане на динамичния ред поради използването на техники, свързани със сезонното изглаждане. Поради недостатъците на тримесечните плъзгащи се средни се предлага те да се използват само в краен случай.



В заключение може да се каже, че оценката на броя на зетите и броя на безработните на месечна основа, използвайки данните от Наблюдението на работната сила и нововъведената методология, с известни ограничения е напълно възможна. Във връзка с намаляване на стохастичната грешка на оценките, получавани от Наблюдението, се препоръчва в бъдеще да бъде увеличен размерът на извадката в съответствие с наличните ресурси на Националния статистически институт.

**ЦИТИРАНА ЛИТЕРАТУРА:**

- Веселинов, Р.** (2011). Кратко ръководство за „Деметра+“.
- Мишев, Г., В. Гоев** (2010). Статистически анализ на времеви редове, С.
- НСИ** (2012). **Методология на Наблюдението на работната сила.**
- Demetra+ User Manual ver. 4** (2011). National Bank of Poland.
- ESS guidelines on Seasonal Adjustment** (2009). Eurostat.
- Infante, E.** (2011). Using DEMETRA+, Paris.
- Kaiser, R., A. Maravall** (2000). An application of TRAMO-SEATS: Changes in seasonality and current trend-cycle assessment.
- Косак, N.** (2011). An Analysis of German Industrial Production with DEMETRA+, Turkish Statistical Institute.
- Rogza, A., A. Secer, J. Arneric** (2008). The Analysis and Forecasting of Unemployment in Croatia.
- Recommendations on Seasonal Adjustment for Employment and Unemployment Data** (2010). ILO.
- Unemployment - LFS adjusted series** (2012). Eurostat.
- Working Instructions for Seasonal Adjustment with Demetra** (2009). UN.
- <http://epp.eurostat.ec.europa.eu/>.
- <http://www.nsi.bg/>.
- <http://www.cros-portal.eu/page/seasonal-adjustment>.
- <http://vesselinov.com/>.

ПОДХОДИ ЗА МЕСЕЧНА ОЦЕНКА НА РАВНИЩЕТО НА БЕЗРАБОТИЦА ВЪЗ ОСНОВА НА ДАННИ ОТ НАБЛЮДЕНИЕТО НА РАБОТНАТА СИЛА

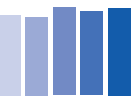
*Александър Найденов**

РЕЗЮМЕ Един от най-важните индикатори за състоянието на икономиката на дадена страна е равнището на безработица. Необходимостта от задълбочен и навременен анализ на социално-икономическите процеси във връзка с осъществяване на контрол върху тяхното развитие, както и съблюдаването на поставените от Евростат изисквания за „производство” на месечни оценки на равнището на безработица и хармонизиране на методологията в Общността, дефинира търсенето на подходи за месечна оценка на равнището на безработица.

В статистическата теория и практика съществува голямо многообразие от методи и подходи за месечна оценка на безработицата, но тяхното използване „наготово“ е невъзможно в българските условия. Причина за това са специфичните особености и наличните ресурси в нашата страна не само от гледна точка на достъпната информация, но и по отношение на степента на подготовка на методологическите екипи. В статията са разгледани основните подходи съобразно конкретните възможности за тяхното приложение. Възможни подходи за решаването на „задачата“ по изготвянето на месечни оценки на безработицата са директните оценки, използвайки техники за претегляне (калибриране) на данните с цел ажустиране на месечните данни към населението по отделни категории, и индиректните оценки, комбинирайки данните от НРС и прилагайки методи за моделиране (екстраполиране и интерполиране) на развитието в динамичните редове на заетите и безработните. Направено е кратко описание на етапите на приложение на тези методи и са посочени възможните софтуерни продукти, които да подпомогнат имплементацията им.

Въз основа на проведеното експериментално изследване с данни от Наблюдението на работната сила се установи, че месечната оценка на равнището на безработицата е напълно възможна, а най-адекватен метод за оценяване е директният, тъй като при него предполагаемата грешка е най-малка.

* Гл. ас. д-р, преподавател в катедра „Статистика и иконометрия“, УНСС - София; e-mail: anaidenov@nsi.bg.



ПОДХОДЫ ДЛЯ МЕСЯЧНОЙ ОЦЕНКИ УРОВНЯ БЕЗРАБОТИЦЫ НА ОСНОВЕ ДАННЫХ С ИССЛЕДОВАНИЯ РАБОЧЕЙ СИЛЫ

*Александър Найденов**

РЕЗЮМЕ Одним из наиболее важных показателей о состоянии экономики данной страны является уровень безработицы. Необходимость в тщательном и своевременном анализе социально-экономических процессов в связи с осуществлением контроля за их развитием, как и соблюдением требований, установленных Евростатом в отношении „производства” месячных оценок уровня безработицы и гармонизации методологии в Сообществе, определяет поиск подходов для месячной оценки уровня безработицы.

В статистической теории и практике существует большое разнообразие методов и подходов для месячной оценки безработицы, но их использование „в готовом виде“ невозможно в болгарских условиях. Причиной этого являются специфические особенности и наличные ресурсы в нашей стране, не только с точки зрения доступной информации, но и в отношении степени подготовки методологических групп. В статье рассматриваются основные подходы в соответствии с конкретными возможностями для их применения. Возможные подходы для решения „задачи“ по изготовлению месячных оценок безработицы это прямые оценки, с использованием приемов взвешивания (калибровки) данных в целях „ажустирования” месячных данных к населению по отдельным категориям, а также и косвенные оценки, получаемые путем комбинирования данных с исследования рабочей силы и с применением методов моделирования (экстраполяции и интерполяции) развития в динамических рядах о занятых и безработных лицах. Коротко описаны этапы в использовании этих методов и упомянуты возможные софтуерные продукты, способствующие их применению.

На основе осуществленного экспериментального исследования с помощью данных с Исследования рабочей силы было установлено, что месячная оценка уровня безработицы вполне возможна, а наиболее адекватным методом оценки является прямая оценка, поскольку она связывается с наименьшей предполагаемой ошибкой.

* Главный ассистент. д-р, преподаватель на кафедре „Статистика и эконометрия“, Университет национального и мирового хозяйства (УНСС) - София; e-mail: anaidenov@nsi.bg.

UNEMPLOYMENT RATE MONTHLY EVALUATION APPROACHES BASED ON DATA FROM THE LABOUR FORCE SURVEY

*Alexander Naydenov**

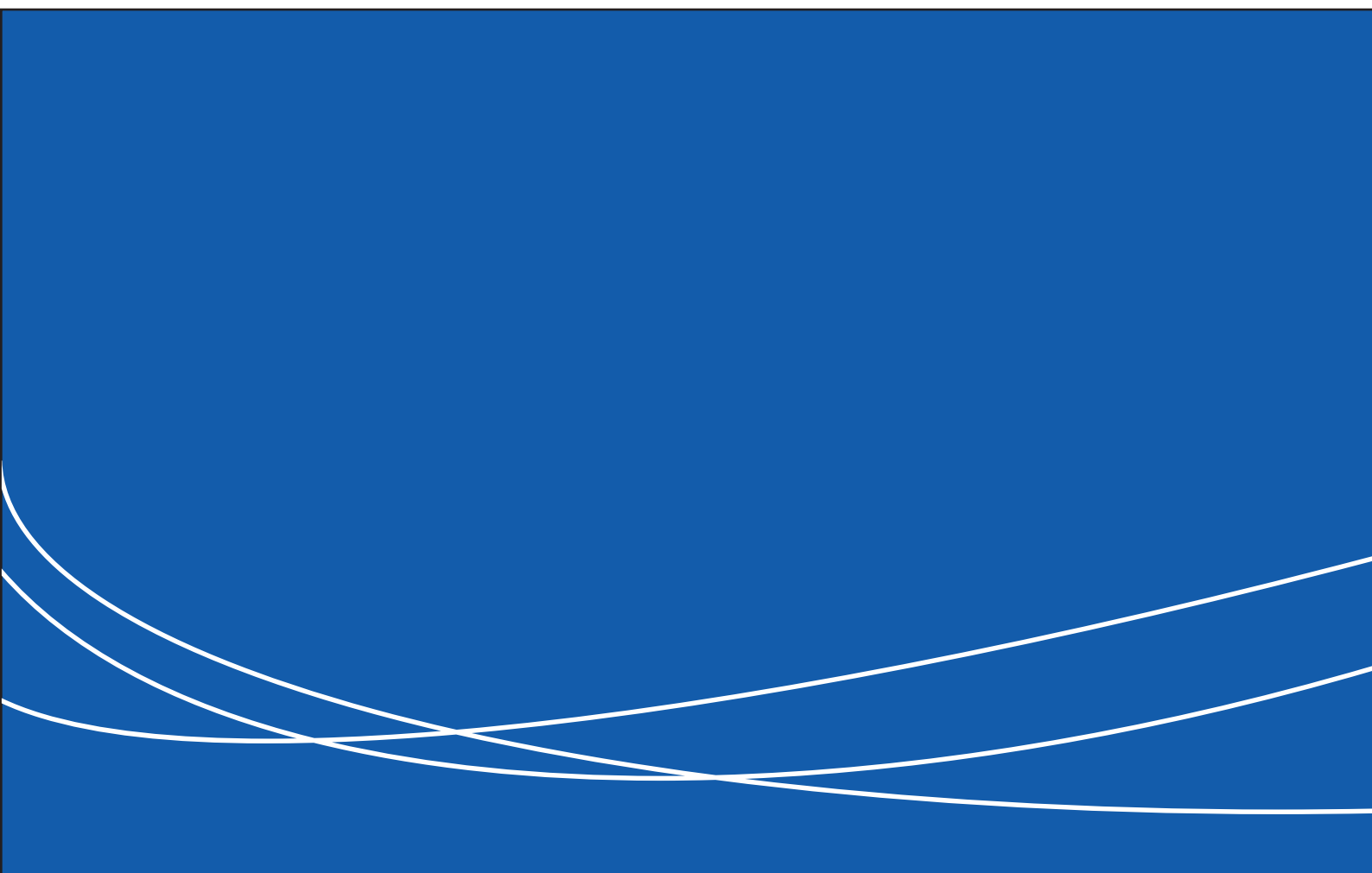
SUMMARY One of the most important indicators of the state of the economy of a given country is the unemployment rate. The need for thorough and timely analysis of the socio-economic processes in relation to the implementation of control over their development, as well as compliance with the requirements set by Eurostat for the ‘production’ of monthly estimates of the unemployment rate and harmonization of the methodology in the Community defines the search of approaches for monthly assessment of the unemployment level.

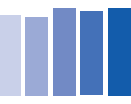
In statistical theory and practice there is a wide variety of methods and approaches for monthly assessment of unemployment, but their ‘ready-made’ use is impossible in Bulgarian conditions. The reasons for this are the specific characteristics and resources available in our country not only in terms of available data, but also in relation to the level of preparation of the methodological teams. The article presents the main approaches according to the specific opportunities for their application. Possible approaches to solving the ‘issue’ in the preparation of monthly estimates of unemployment are the direct assessments using techniques of weighting (calibration) of data in order to adjust the monthly data on population in different categories, and indirect estimates combining data from the NFS and applying modeling techniques (extrapolation and interpolation) of development in the series of the employed and the unemployed. A brief description of the steps of the application of these methods is made, and possible software to assist their implementation are indicated.

Based on experimental study with data from the Labour Force Survey it is established that the monthly assessment of the level of unemployment is entirely possible, and the most adequate method of evaluation is the direct one, since with the alleged error is smallest.

* Ch. Assist. PhD, professor in the ‘Statistics and Econometrics’ Department, University of National and World Economy - Sofia; is-mail: anaydenov@nsi.bg.

**СТАТИСТИЧЕСКИ ИЗСЛЕДВАНИЯ И
АНАЛИЗИ**





СТАТИСТИЧЕСКИ ПОГЛЕД КЪМ ОБРАЗОВАНИЕТО

ИЛИ НЯКОИ ПРОБЛЕМИ НА ОБРАЗОВАТЕЛНАТА СИСТЕМА И НА ИНФОРМАЦИОННОТО ОСИГУРЯВАНЕ НА НЕЙНОТО УПРАВЛЕНИЕ

*Стоян Баев**



Проблемите на образователната система на България и нейното информационно осигуряване не са от вчера, не са напълно идентифицирани и анализирани и негативните последици от това се задълбочават. Тези проблеми, разбира се, са свързани с цялостното икономическо и социално развитие на страната през последните години.

Негативните последици от дълго натрупваните проблеми се изразяват главно в незадоволителните резултати от функционирането на формалната образователна система, чието основно предназначение е да организира и осигурява качествено образование на децата и младите хора от предучилищно до висше образование. Проблемите са многобройни, поради което в анализа се разглеждат само една част от обществено значимите и актуалните сред тях, за които има налични статистически данни. Изложението на проблематиката е структурирано в две части - степени на образование и общосистемни въпроси на информационното осигуряване на образованието.

I. Проблеми на формалната образователна система

1. Предучилищно възпитание и подготовка

1.1. Предучилищното възпитание и подготовка в своята задължителна част **не обхваща всички подлежащи на задължително обучение деца**. Обхватът на 6-годишните деца в образователната система към 31.12.2013 г. е 96.2%. Това на практика е делът на записаните в предучилищна подготовка или в първи клас 6-годишни деца спрямо общия брой на населението на 6 години. Незаписаните 6-годишни деца са около 2 500. Една част от тези деца не са записани поради здравни проблеми и главно поради заминаване заедно с родителите си в чужбина през 2013 г. и предходните години, без родителите да са обявили промяна на настоящия си адрес от България в чужбина. Броят

* Консултант по статистика на образованието и ученето през целия живот; e-mail: stoyan.baev@gmail.com.

на тези деца по експертна оценка на автора е около 1 500, включително около 50 деца, незаписани поради трайна инвалидизация или други тежки заболявания. По този начин броят на необхванатите 6-годишни деца общо за страната се оценява на около 1 000 към края на 2013 година. Тези оценки до голяма степен са релевантни на данните от преброяването на населението към 1.02.2011 г. които показват, че общо 5 464 деца на възраст от 7 до 14 навършени години никога не са посещавали училище. Необходимо е да се отбележи, че данните от домакинствата при преброяването на населението се вписват на декларативен принцип, поради което е вероятно фактическият брой на никога непосещавалите училище деца да е по-голям, т.е. около 6 - 7 хиляди. Следователно причините за неграмотността на младите хора, пък и на по-възрастните, се коренят в неучастието им в предучилищната подготовка и в началния етап на основното образование, включително и поради напускане на училище. Резултатът е красноречив - броят на самоопределилите се като неграмотни (не знаят да пишат и четат) лица на възраст 9 - 29 години към 1.02.2011 г. е 37 155 души! От тях 7 968 се самоопределят като българи, 5 060 - като турци, и 16 707 - като роми.

Независимо от тези обезпокоителни данни следва да се отбележи, че броят на напусналите основното образование деца (I - VIII клас) поради нежелание и семейни причини съществено и систематично намалява през последните десет години - от 13 784 през учебната 2002/2003 година на 6 749 през учебната 2012/2013 година, или с 51%, по данни на НСИ. Важен фактор за тази тенденция са усилията на училищата да предпазват учениците от напускане в условията на делегираните бюджети. Друг е въпросът, че при тези условия се допуска неграмотни ученици да преминават в по-горен клас и дори да завършват четвърти клас.

Основните причини за наличието на необхванати в образователната система деца са главно организационни. От една страна, местните власти не успяват да се справят с проблема, защото не се прилагат строго изискванията за задължително участие на децата в предучилищната подготовка. От друга страна, липсва комуникация между Информационната система за гражданска регистрация ЕСГРАОН и Информационната система за предучилищно и училищно образование на Министерството на науката и образованието (МОН) и по този начин не могат да се идентифицират необхванатите деца. Проблем е също така липсата на пълна актуална адресна регистрация на населението в ЕСГРАОН, недостатъчната строгост на правилата, липсата на контрол и нежеланието за прилагане на тези правила от страна на компетентните органи. Впрочем проблемът за актуалността на адресната регистрация има и други аспекти на държавно и местно равнище, свързани с вътрешната и външната миграция, престъпността и т.н.

Необходимите мерки и дейности за решаване на описаните проблеми се свеждат главно до:

- Спешно проучване и предприемане на мерки и действия от страна на правителството на Република България за решаване на проблема за адресната регистрация. Необходимо е да се предложат и въведат промени в законовата и подзаконовата нормативна уредба. Заедно с това е необходимо да се реши въпросът за мотивиране на гражданите, заминаващи на работа или за цялостно обучение, да

обявяват настоящ адрес в чужбина. Необходими са и съответни стимули и санкции. Опитът на някои от държавите на ЕС показва, че може да бъде създадена административна процедура за своевременна промяна на настоящия адрес от страната в чужбина, мотивирана например от данъчни облекчения за дългосрочно заминаващите (над 1 година) лица за чужбина.

- Определяне на точна възраст за постъпване на децата в задължителната предучилищна подготовка - например в годината, през която детето навършва 6 години, а не както досега - на 5- или 6-годишна възраст по преценка на родителите. Същото се отнася и за възрастта за постъпване на децата в първи клас - да постъпват в годината, през която детето навършва 7 години.

1.2. Вторият важен проблем на предучилищната подготовка е, че през „изхода“ на задължителната предучилищна подготовка преминават деца, които не са подготвени да постъпят в първи клас на началния етап на училищното образование. Макар че липсват статистически данни за това явление, известно е, че се допуска в първи клас да постъпват деца, които не владеят достатъчно говоримия български език. Тези деца още през първата година поради нарушената комуникация между тях и учителя са поставени в неблагоприятна и често стресова среда, изостават в учебния процес, демотивират се и загубват желание за активно участие в ученето. Впоследствие тези негативни последици стават основна причина за напускане и/или отпадане от училище.

Установяването на потребност и полагането на специални допълнителни грижи към тази категория деца следва да се уреди в подзаконовите нормативни документи за предучилищната подготовка и училищното образование.

2. Основно образование - начален и прогимназиален етап

2.1. Главните проблеми на основното образование са известни - остарелите и неадекватни учебни програми, ненужната натовареност на учениците (изисквания за наизустяване на факти), демотивация и нежелание за активно участие в учебния процес, недостатъчна насоченост към практическо използване на знанията и уменията, незадоволителна училищна дисциплина и като краен резултат - ниско качество на образованието. Проблемите на качеството се виждат от слабите резултати на 15-годишните ученици по функционална грамотност по четене, математика и природни науки според международното изследване PISA на Организацията за икономическо сътрудничество и развитие (ОИСР). Резултатите от последното изследване относно равнището на функционална грамотност на 15-годишните ученици у нас показват, че в две от трите познавателни области (четене и математика) България е с най-слабо представяне (на последно място) сред държавите от ЕС¹. Съществената разлика между дела на учениците с резултати под критичното ниво в България и средния за страните от ЕС от приблизително 21.7 процентни пункта по четене и 20.3 процентни пункта по природни науки през 2012 г. потвърждава негативната и тревожна картина на качеството на образованието в България и потребността от промени в учебните планове и програми.

¹ OECD, PISA 2012. Result in focus, page 5, Tables I.2.1a, I.2.1b, I.2.3a, I.2.3b, I.4.3a, I.4.3b, I.5.3a and I.5.3b „Snapshot of performance in mathematics, reading and science”.

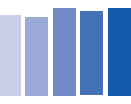
2.2. Нормативното решение учениците да не повтарят от първи до четвърти клас поражда като резултат известен (неуточнен) брой деца, главно от ромски произход, да достигат до пети клас неграмотни. Опитът на други страни показва, че решението не е само в допускането на възможност да се повтаря класът, а в системната работа с учениците и родителите, включително осигуряване на допълнителни ресурси за изоставащите ученици. Проблемът у нас остава нерешен и крайните ефекти за децата от тази група са отпадане от училище, ниска фактическа и функционална грамотност, пораждащи впоследствие безработица и бедност.

2.3. Към признаците за ниско качество на основното образование трябва да се добави и недостатъчната чуждоезикова подготовка на завършилите основно образование с изключение на учениците в езиковите училища и паралелки.

2.4. Друг проблем е преминаването към програми на средното (гимназиалното) образование след седми клас, преди да бъде завършено основното образование. В резултат на това се стигна до парадокса основното образование в профилираните и професионалните гимназии да се завършва „в движение”, някъде след осми клас, а някъде и след девети клас. На този фон, на фона на резултатите от PISA и на фона на структурите на образователните системи на държавите от ЕС² изглежда напълно нелогична идеята за 7-годишна продължителност на основното образование. Такава продължителност, ако се приеме с новия Закон за училищното образование, би била уникална за ЕС. Най-често срещаната продължителност на основното образование е 9 години - в 18 страни от ЕС, в 7 страни (България, Белгия, Хърватия, Италия, Унгария, Малта и Австрия) продължителността е 8 години, а в 3 страни (Германия, Дания и Испания) продължителността е 10 години. Следователно европейските измерения за продължителност на основното образование, което трябва да осигури на учениците определен обем от знания, умения и ключови компетентности, са в рамките от 8 до 10 години. В държавите от ЕС със съответните нормативни актове компетентните органи са преценили, че това е необходимата продължителност на основното образование, за да могат учениците да постигнат международно възприетите критерии за тази степен, препоръчани в Европейската квалификационна рамка. Същите критерии са заложили и в Националната квалификационна рамка на България. И още нещо - при разрастващата се мобилност на българските граждани към други страни от ЕС изглежда малко вероятно 7-годишното основно образование да бъде признато в страни с 9- и 10-годишно основно образование.

2.5. Нерешен остава проблемът за ограмотяването на неграмотните лица, които са над 37 000, между които има и много млади хора. След един по-задълбочен анализ на разпространението на неграмотността би могло директно да се пристъпи към изготвяне, утвърждаване и изпълнение на програма за свеждане до минимум на неграмотността сред населението в трудоспособна възраст чрез различни форми за учене на възрастни. По-нататък логически следва създаването и изпълнението на план за включване на ограмотените лица в програми за възрастни за придобиване на начално образование (четвърти клас) и възможност за придобиване на професионална квалификация.

² The structure of the European education systems 2013/14: schematic diagrams.



3. Средно образование

3.1. Средно общо образование

През последните 10 - 15 години голяма част от средното общо образование беше профилирано, без това да доведе до подобряване на качеството. Изключение от тази констатация до известна степен са езиковите, хуманитарните и природо-математическите гимназии и паралелки и Националната гимназия за древни езици и култури, които „обират“ по-добрите ученици след завършването на седми клас. В тези три вида профилирани гимназии и паралелки през настоящата учебна година учат 66.1% от учениците в общообразователните гимназии. В непрофилирани училища и паралелки учат само 25% от учениците, а останалите 9% завършват средно образование с никаква и непризната от никого „квалификация“ по измислени профили като мениджмънт, вид изкуства и т.н.

3.2. Средно професионално образование

Основните недостатъци на средното професионално образование са:

- Несъответствие на структурата на учениците по професионални направления и професии с реалните потребности на страната и съответните региони от персонал с определени знания, умения и компетентности. Измеренията на това „несъответствие“ не са ясни поради липса на методика за определяне на потребностите от професионално подготвени хора със съответните умения. Но наличието на такова несъответствие е несъмнено, ако се вземат предвид:

- незначителните промени в структурата на учениците по професионални направления и професии за последните 10 - 15 години;

- данните за регистрираните безработни млади хора със завършено средно професионално образование;

- публично изразените мнения на официални представители на работодателски организации.

От друга страна, традиционно пасивната позиция на бизнеса относно заявяването на бъдещите потребности от персонал с определена квалификация не допринася за решаването на този проблем.

- Вторият недостатък, който е свързан с първия, е прекалено раздробеният Списък на професиите за професионално образование и обучение. Броят на професиите в Списъка към април 2014 г. е 230, а на специалностите към тези професии - 538. Същевременно наличието на тесни специалности води до парадокса продължителността на обучението по редица специалности да е с пъти по-голяма от необходимата. Например специалностите „Маникюр, педикюр и ноктопластика“ и „Обслужване на заведения в общественото хранене“ се изучават 4 години и осигуряват на завършилите успешно II степен на професионална квалификация. А такива умения могат да се придобият само за 3 - 4 месеца. Тясната специализация на завършилите не дава възможност за по-лесен преход от училище към работа, тъй като несъответствията с изискванията на конкретното работно място са обикновено значителни. Една по-широка професионална подготовка на учениците в училище би улеснила прехода към работа чрез кратко обучение на работното място или чрез кратък курс.

- Третият недостатък е липсата на гъвкавост в професионалното образование. Т.нар. „традиции” в професионалното образование водят до ежегодно възпроизвеждане на държавния план-прием, слабо съобразен с реалните потребности на икономиката от специалисти, като се толерира запазването на съществуващите професии и специалности с цел запазването на преподавателския състав.

4. Висше образование

4.1. Подобно на средното професионално образование висшето образование традиционно не успява напълно да отговори на потребностите на икономиката от специалисти с определена квалификация и „произвежда” специалисти по многобройни, в някои случаи странни и ненужни специалности. Тясната специализация на студентите по част от специалностите често е пречка да не могат след завършване да намерят работа по специалността си или изобщо да започнат работа. Статистическите данни за структурата на студентите по професионални направления сочат, че тази структура през последните десетина години не се променя съществено, което показва липса на гъвкавост в системата на висшето образование. От друга страна, според редица проучвания практическата подготовка на голяма част от студентите не е достатъчна и те завършват, без да са готови за преход към работа.

4.2. Друг проблем, или по-скоро парадокс, е, че във висшите училища постъпват млади хора, които не пишат граматически правилно на български език и не могат да изразят ясно своята теза писмено. Поради това не е странно, че някои университети вече са въвели задължително обучение по български език на студентите български граждани. Оттук следва изводът, че качеството на средното образование, и в частност подготовката на учениците по български език и литература, не отговаря на изискванията и че външното независимо оценяване следва да бъде решително засилено с цел вземане на мерки за промяна на ситуацията.

4.3. Класификаторът на областите на висше образование и професионалните направления още от създаването си през 2002 г. не съответства на Международната стандартна класификация на образованието (ISCED) на ЮНЕСКО и досега не е приведен в съответствие с нея. Класификаторът представлява списък на професионални направления без каквито и да било методически указания за отнасянето (класифицирането) на специалностите към едно или друго професионално направление. За сравнение - интересуващите се могат да видят какво представлява съвременният вариант на Международната класификация на областите на образованието (ISCED-F-2013) на ЮНЕСКО - документ, който беше утвърден в края на 2013 г., в интернет на адрес: <http://www.uis.unesco.org/Education/Documents/isced-37c-fos-review-222729e.pdf>.

II. Проблеми на информационното осигуряване на управлението на образователната система

Информационните системи на МОН изпълняват незаменяема и важна функция за информационното осигуряване на управлението на образователната система, в т.ч. и главно за финансовото планиране на публичното образование. Наред с положителните си страни функционирането на информационните системи на МОН е съпроводено с редица проблеми и недостатъци, по-важните от които са:

1. Данните от информационните системи не се използват ефективно от широк кръг потребители, тъй като не се разпространяват официално. Макар че тези системи се финансират с бюджетни средства, данните се ползват главно като вътрешноведомствени и ограничено от някои държавни ведомства като НСИ, Министерството на труда и социалната политика и други.

2. Референтната дата за първия кръг събиране на данни в системата „АдминМ” - 15 септември като първи учебен ден, не е подходящо избрана, тъй като към тази дата положението на част от учениците и учителите не е установено и това води до неточност на данните за броя на записаните ученици като цяло, по класове и т.н.

3. В регистъра на учениците липсват някои характеристики като например дали ученикът за първи път се записва в дадената степен на образование, дали за първи път се дипломира в тази степен и други характеристики, изисквани по Регламент (ЕС) № 912/2013 на Комисията относно изготвянето и развитието на статистиката в областта на образованието и обучението през целия живот по отношение на статистиката на образователните системи и системите за обучение.

4. Данните за завършилите основно образование са непълни, тъй като на приетите след седми клас не се издават дипломи.

5. Категориите педагогически персонал, използвани в Информационната система „АдминМ”, се нуждаят от актуализация и прецизиране. Например за ръководни кадри се считат както директори или помощник-директори, така и учители, ако имат преподавателска заетост, т.е. допуска се двойно броене на едни и същи физически лица при извеждане на общия брой на педагогическия персонал.

6. Данните в „АдминМ” за участието на педагогическия персонал в продължаващо обучение са непълни и контролът по въвеждането им е занижен от ниво директор до ниво Инспекторат на МОН.

7. В Министерството на образованието и науката липсва отдел за статистически изследвания и анализи. Същевременно Центърът за информационно осигуряване на образованието (ЦИОО) е със статут на обслужващо звено, като по този начин работещите в него служители са демотивирани и с по-ниски работни заплати в сравнение със служителите например в дирекция „Информационни и комуникационни технологии” на МОН.

Съдържанието и функционирането на информационните системи на МОН се нуждаят от задълбочен анализ и адекватно обновяване в тясно сътрудничество с НСИ. Ако ръководството на МОН счита решаването на тези проблеми за приоритет през

следващите 2 - 3 години, препоръчвам следните основни стъпки и дейности за решаването им:

Като първа стъпка е необходимо да се подготви и проведе среща на министъра на образованието и науката с председателя на НСИ за установяване на трайно сътрудничество между двете ведомства в областта на статистиката на образованието и обучението на основата на използване на информационните системи на МОН. Ключов въпрос е обсъждането на включването на МОН в Националната статистическа система на основание на Закона за статистиката. На тази среща е важно да се обсъди и предложение за създаването на работна група на МОН и НСИ със следните задачи:

1. Изготвяне на проект на споразумение за сътрудничество и обмен на информация между НСИ и МОН в условията на членство на МОН в Националната статистическа система. Основна задача на споразумението е да се реши процедурно въпросът за обмен на индивидуални данни по смисъла на Закона за статистиката.

2. Да се направи преглед, анализ и съпоставка на наличните информационни източници в МОН и НСИ.

3. Да се предложат мерки, свързани с модернизация и актуализация на съществуващите информационни системи на МОН, с оглед привеждането им в съответствие с европейските статистически стандарти.

4. Изготвяне на проект за интегрирана информационна система на МОН и НСИ на основата на съществуващите информационни системи на МОН при спазване на изискванията на Регламент (ЕС) № 912/2013 на Комисията относно изготвянето и развитието на статистиката в областта на образованието и обучението през целия живот по отношение на статистиката на образователните системи и системите за обучение.

5. Проучване на необходимостта от законодателни промени с оглед изграждането и функционирането на интегрираната информационна система.

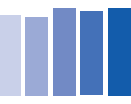
6. Изготвяне на предложение относно задължителното разпространение на данни от провеждани от МОН статистически изследвания и от съвместни изследвания с НСИ, базирани на данни от интегрирана информационна система на МОН и НСИ.

Изграждането и внедряването на интегрирана информационна система на МОН и НСИ на основата на съществуващите информационни системи на МОН при спазване на изискванията на Регламент (ЕС) № 912/2013 на Комисията в условията на членство на МОН в Националната статистическа система се очаква да доведе до постигането на следните положителни резултати:

- Привеждане на статистическите изследвания, провеждани от МОН, в съответствие с Регламент (ЕС) № 912/2013 на Комисията относно изготвянето и развитието на статистиката на образованието и ученето през целия живот по отношение на статистиката на образователните системи и системите за обучение и другите европейски статистически стандарти.

- Осигуряване на МОН с официална актуална статистическа информация за анализ и формиране на политики в областта на образованието и ученето през целия живот, както и за мониторинг на приетите стратегии за периода до 2020 година.

- Постигане на методологично единство между МОН и НСИ при подготовката, производството и разпространението на статистически данни за образованието и



ученето през целия живот на територията на страната, основано на Закона за статистиката.

- Обновяване на информационните системи и регистри на МОН с оглед на тяхното развитие в съответствие с европейските статистически стандарти и използването им съвместно от МОН и НСИ за статистически цели при условията на конфиденциалност на индивидуалните статистически данни.

- Осигуряване на възможност провежданите понастоящем от МОН ведомствени статистически изследвания чрез регистрите на МОН и други източници, при които се събират, обработват и разпространяват статистически данни за училищното и висшето образование, включително рейтинговата система на висшите училища, да бъдат включени в Националната статистическа програма като съвместни изследвания на НСИ и МОН или като изследвания, провеждани от МОН.

- Създаване на условия за бъдещи съвместни дейности на МОН и НСИ в областта на статистиката на образованието и ученето през целия живот като подготовка и провеждане на нови изследвания; представяне на резултати от съвместни статистически изследвания едновременно на сайтовете на МОН и НСИ; издаване на съвместна електронна публикация с данни за образованието в Република България.

- Изпълнение на препоръките на Европейската комисия за намаляване на административната натовареност на респондентите, които са всички видове учебни институции, предоставящи понастоящем дублиращи се данни на НСИ и МОН, и създаване на реална възможност за по-ефективно използване на бюджетните средства за осигуряване на статистически данни за образованието.

В заключение - необходимо е Министерството на образованието и науката да прецени потребността от създаването на отдел по статистически изследвания и анализи в МОН и да набележи мерки за подобряване на мястото и ролята на Центъра за информационно осигуряване на образованието.

СТАТИСТИЧЕСКИ ПОГЛЕД КЪМ ОБРАЗОВАНИЕТО

ИЛИ НЯКОИ ПРОБЛЕМИ НА ОБРАЗОВАТЕЛНАТА СИСТЕМА И НА ИНФОРМАЦИОННОТО ОСИГУРЯВАНЕ НА НЕЙНОТО УПРАВЛЕНИЕ

*Стоян Баев**

РЕЗЮМЕ Проблемите на образователната система на България и нейното информационно осигуряване не са от вчера, не са напълно идентифицирани и анализирани и негативните последици от това се задълбочават. Те се изразяват главно в незадоволителните резултати от функционирането на формалната образователна система. В статията се разглеждат само една част от общественозначимите проблеми, за които има налични статистически данни.

Предучилищното възпитание и подготовка не успява да обхване напълно всички подлежащи на задължително обучение деца. Друг важен проблем е, че през „изхода” на задължителната предучилищна подготовка преминават деца, които не са подготвени да постъпят в първи клас. Основното образование страда от ниско качество - около 40% от 15-годишните ученици са функционално неграмотни. На този фон и на фона на структурите на образователните системи на държавите от ЕС, в които продължителността на основното образование е най-често 9 години, идеята за 7-годишна продължителност изглежда напълно нелогична. Традициите в професионалното образование водят до ежегодно възпроизвеждане на държавния план-прием, който е слабо съобразен с реалните потребности на икономиката от специалисти. Висшето образование традиционно не успява напълно да отговори на потребностите на икономиката и „произвежда” специалисти по многобройни, в някои случаи странни и ненужни специалности.

Информационните системи на МОН изпълняват незаменима и важна функция за информационното осигуряване на управлението на образователната система. Наред с положителните си страни функционирането на информационните системи на МОН е съпроводено с редица проблеми и недостатъци, които са систематизирани в статията. Възможните мерки и действия за преодоляване на тези недостатъци са свързани главно с изграждането и внедряването на интегрирана информационна система на МОН и НСИ, от която се очакват множество положителни резултати, посочени в заключителната част на статията.

* Консултант по статистика на образованието и ученето през целия живот; e-mail: stoyan.baev@gmail.com.

СТАТИСТИЧЕСКИЙ ВЗГЛЯД НА ОБРАЗОВАНИЕ ИЛИ НЕКОТОРЫЕ ПРОБЛЕМЫ СИСТЕМЫ ОБРАЗОВАНИЯ И ИНФОРМАЦИОННОГО ОБЕСПЕЧЕНИЯ ЕЕ УПРАВЛЕНИЯ

*Стоян Баев**

РЕЗЮМЕ Проблемы системы образования Болгарии и ее информационного обеспечения существуют не со вчерашнего дня, а также не выявлены и проанализированы, что приводит к углублению негативных последствий с этого. Эти последствия выражаются в основном в неудовлетворительных результатах функционирования системы формального образования. В статье рассматривается только одна часть общественно значимых проблем, о которых имеются статистические данные.

Дошкольное образование и дошкольная подготовка не в полной мере успевают охватить всех детей, подлежащих обязательному обучению. Второй по важности проблемой является то, что через „выход” обязательной дошкольной подготовки проходят дети, которые не подготовлены к вступлению в первый класс. Основное образование страдает от низкого качества - около 40% 15-летних учеников являются функционально неграмотными. На этом фоне и на фоне структур систем образования стран ЕС, в которых продолжительность основного образования чаще всего девять лет, идея о семилетней продолжительности основного образования кажется совершенно нелогичной. Традиции в области профессионального образования приводят к ежегодному воспроизводству государственного плана приема, соответствующего в очень малой степени реальным потребностям экономики в специалистах. Высшее образование традиционно не в состоянии удовлетворить в полной мере потребности экономики и „производит” специалистов в области многочисленных, иногда странных и ненужных специальностей.

Информационные системы Министерства образования и науки (МОН) выполняют незаменимую и важную функцию по информационному обеспечению управления системой образования. Наряду со своими положительными сторонами, функционирование информационных систем МОН сопровождается рядом проблем и недостатков, перечисленных в статье. Возможные меры и действия по устранению этих недостатков, в основном, связаны с построением и реализацией интегрированной информационной системы МОН и НСИ, от которой ожидается множество положительных результатов, которые упоминаются в заключительной части статьи.

* Консультант по статистике образования и пожизненного обучения; электронная почта: stoyan.baev@gmail.com.

STATISTICAL REVIEW OF EDUCATION
OR SOME ISSUES OF THE EDUCATIONAL SYSTEM AND
THE INFORMATION ASSURANCE OF ITS MANAGEMENT

*Stoyan Baev**

SUMMARY The problems of the education system of Bulgaria and its information provision are nothing new, are not fully identified and analysed, and the negative consequences of this are deepening. They are mainly expressed in the unsatisfactory results of the functioning of the formal education system. The article reviews only one part of the socially significant issues for which statistical data are available.

Pre-school education and school-prep fail to fully cover all children subject to mandatory education. Another important problem is that in the ‘output’ of the mandatory pre-school education pass children who are not prepared to enter first grade. Primary education suffers from low quality - about 40% of 15 year olds are functionally illiterate. Against this background and the background of the structures of the education systems of the EU countries in which the duration of basic education is generally nine years the idea of a 7-year period seems completely illogical. Traditions in vocational education lead to annual reproduction of ‘State admission plan’ which is less consistent with the real needs of the economy of experts. Higher education traditionally fails to meet the needs of the economy and ‘produces’ experts of numerous, in some cases strange and unnecessary specialties.

MES Information systems carry irreplaceable and important function of information assurance of the management of the education system. Along with the positive sides, the functioning of the MES information systems is accompanied by a number of problems and shortcomings, which are summarized in the article. Possible measures and actions to address these deficiencies are mainly related to the construction and implementation of an integrated information system of MES and NSI and from it is expected to produce a number of positive results, which are referred to in the final part of the article.

* Consultant in education and lifelong learning statistics; e-mail: stoyan.baev@gmail.com.

ГОЛЕМИТЕ ДАННИ - ВЪЗМОЖНОСТ, ПРЕДИЗВИКАТЕЛСТВО ИЛИ ЗАПЛАХА ПРЕД ОФИЦИАЛНАТА СТАТИСТИКА

Галя Статева*



Въведение

В нашия модерен свят все повече данни се генерират от световната интернет мрежа и се произвеждат от електронни сензори и устройства, които са навсякъде около нас. Тези данни се създават в резултат от протичане на процеси в различни сфери на общественно-икономическия живот на национално и международно равнище. Тяхното естество трябва да се разглежда в два основни аспекта: като описание на характеристиките, спецификите и особеностите на всеки процес и като ръководство за управление на всеки процес. Обемът на данните и високата скорост, с която се произвеждат, води до създаване на концепцията „големи данни“ (**Big Data**). Това понятие е добре дефинирано от определението на Gartner¹:

Големите данни са данните, които могат да бъдат описани като „голямо количество от разнообразни данни, осигуряващи ефективността и ефикасността на протичащите процеси чрез повишаване на познанието и вземането на правилни решения“.

Големите данни се характеризират още като набор от данни с нарастващ обем, скорост и разнообразие, или т.нар. 3Vs². Те обикновено са неструктурирани, нямат предварително дефиниран модел и мащаб, като най-често са в текстови формат.

Статистическите организации дефинират големите данни като:

Данни, които трудно се събират, съхраняват или обработват с конвенционалните системи на статистическите организации. Техният обем,

* Държавен експерт в отдел „Обща методология и анализ на статистическите изследвания“, НСИ; e-mail: gstateva@nsi.bg.

¹ Повече информация може да се намери на адрес: <http://www.gartner.com/it-glossary/big-data/>.

² Volume, Velocity, Variety (3Vs).

скорост, структура и разнообразие изискват адаптиране на нов статистически софтуер за обработка и/или нова ИТ инфраструктура, за да бъдат ефективни направените разходи.

Според потребителите с източниците на големи данни е трудно да се работи. Основните причини за това са: скоростта на промените; трудностите с тяхната идентификация за възможно най-кратък срок от време; големите мащаби; непознаването на методите за тяхната обработка и превръщането им в обобщени синтетични показатели, позволяващи използването им за аналитични цели. Примери за големи данни са: търговски трансакции с кредитни и дебитни карти; данни за околната среда от различни сензори за наблюдение на въздуха, водата и почвата; трафик информация от камери за наблюдение; социална информация от различни мрежи като Twitter, Facebook, Google+, LinkedIn и други.

Информационните потоци в обществото се променят динамично. Това обстоятелство поставя нови акценти в общественото развитие. Големите данни създават нови търговски възможности в частния и обществения сектор, но освен това могат да бъдат потенциално интересни и като източник за официалната статистика - за самостоятелно използване или в комбинация с традиционните източници на данни - например извадковите изследвания и/или административните източници. Основателно възникват редица въпроси: как големите данни могат да спомогнат да се измерят точно и навременно икономическите, политическите, социалните и природните феномени в нашия постоянно развиващ се свят?

Големите данни възникват от множество източници, които могат да се групират в три основни категории:

- Население (социални мрежи)
- Данни, генерирани от информационни системи (традиционни бизнес системи и уебсайтове)
- Данни, генерирани от машини/сензорни устройства (автоматизирани системи).

Процесът на „добиване” на информация от големи данни и инкорпорирането им в производствен процес на официалната статистика не е никак лека задача.

Какво се случва, когато официалната статистика срещне големите данни?

Официалната статистика играе ключова роля в модерното общество. До 80-те години на миналия век данните бяха оскъдна стока с висока цена. Преди ерата на големите данни информацията не беше толкова достъпна и трябваше да бъде събирана за определени цели. Статистическите данни се получаваха предимно чрез изчерпателни изследвания. Тези данни се използваха предимно за целите на държавното управление. Такива бяха изследванията за държавните предприятия, селското стопанство и други. През 90-те години изчерпателните изследвания отстъпиха място на извадковите изследвания. Постепенно получаването на статистически данни чрез въпросници беше допълнено и с данни от административни източници. Много статистически служби имат достъп по закон до всички държавни, институционални източници на данни и имат право да събират данни от други източници, без да плащат за това на доставчиците. В

частност, възможността за комбиниране на данни от различни източници прави официалната статистика по-прецизна в своята дейност и разширява нейния обхват.

Във връзка с това и в унисон с високото технологично развитие информацията, осигурявана от статистическите институти, запазва своята уникалност и няма алтернатива, тъй като нейното качество се определя от 15-те принципа на Кодекса на европейската статистическа практика³. Професионалистите, работещи в производството на официална статистика, са държавни служители и имат широки и специфични познания в статистическата наука. Те са гарант, който осигурява доверието на обществото в надеждността на произвежданата статистическа информация. Успоредно с това усилията на Европейската статистическа система (ЕСС) са насочени към повишаване на равнището на стандартизиране, хармонизиране и съчетаване на различни източници на данни, включително и на динамично променящите се големи потоци от данни. Това обстоятелство променя облика на официалните статистики в държавите от Европейския съюз.

В контекста на тези обективно случващи се процеси големите данни все повече разширяват своето присъствие в общественото пространство: огромен обем от дигитална информация, произлизаща от всички видове човешки дейности, служи за производство на статистика, която се използва от частни институции или компании. Възникват следните въпроси: дали компаниите, които събират данни, са собственици на тези данни; могат ли да ги използват за различни цели без съгласието на респондентите; дали тези числа могат да се възприемат като конкурентни на официалната статистика; могат ли да се разработват анализи, съчетаващи информация от големи данни и от официалната статистика.

Източниците на големи данни предлагат огромен обем от данни, които изискват съхранение и обработка, надвишаващи капацитета на традиционните статистически средства при процеса на производство на статистика. Поради тази причина биха могли да се прилагат „нови“ техники за извличане на знания от данни (data mining) и прилагане на алгоритми от областта на машинното обучение (machine learning algorithms), имащи изискваната изчислителна ефикасност (Bondi, 2000).

Друго съображение относно използването на големите данни е свързано с представителността и обхвата на произвежданата от тях статистика. С традиционните извадкови техники се осигурява точност на получените статистически оценки на базата на размера на стохастичните грешки. Това може би е приложимо и за големите данни, които могат да бъдат адаптирани към традиционните извадкови техники, но се изискват достатъчно добри аргументи и прецизен анализ на получените емпирични резултати. Успоредно могат да се разработят и алтернативни методи, които отразяват спецификата на големите данни, тяхната динамичност, обхват и области на приложение.

Теоретично погледнато, големите данни могат да бъдат използвани за производство на официална статистика по различни начини: 1) като заменят изцяло статистическите източници, основани на общи дефиниции, класификации и т.н., което е малко вероятно в обозримото бъдеще; 2) частична замяна на

³ Повече информация може да се намери на адрес:
http://www.nsi.bg/sites/default/files/files/pages/Quality/1.1.%20CoP_ALL_BG.pdf

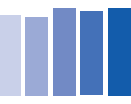
статистическите източници, като допълват информацията чрез съчетаване на данни от различни източници на данните; 3) осигуряване на напълно нови статистически числа, които могат да допълват и да се интегрират с наличната статистическа информация, което е значително по-добрият начин за тяхното съвместно използване. Първите два начина вероятно биха могли да доведат до намаляване на разходите и натоварването на респондентите, но това, от своя страна, ще доведе до нови задачи за адаптиране, съчетаване и хармонизиране на различни структури от данни към вече утвърдени и общоприети статистически концепции, дефиниции и класификации. Логично погледнато, големите данни не могат да заменят напълно или частично статистическите източници в краткосрочен план и това би било твърде скъпо по отношение на времето, финансовите и човешките ресурси. Наред с това на този етап от глобализирането на света между статистическите и големите данни се наблюдават моментни процеси на конвергенция, които са необходими за управлението на бизнеса. Фирмите от частния сектор, произвеждащи статистика на основата на големите данни, следват третия път и не се сблъскват с подобни проблеми.

Предизвикателства пред официалната статистика

Срещата на големите данни с официалната статистика води до много и различни предизвикателства.

Едно от най-големите предизвикателства пред статистиците при използването на големите данни засяга **методологията**. При провеждане на едно традиционно изследване статистиците дефинират генерална съвкупност, разработват дизайн на извадката, събират данните и т.н. Много източници на големи данни като например съобщенията в социалните мрежи нямат добре дефинирана генерална съвкупност, структура и качество. Това прави трудно прилагането на традиционните статистически методи, основани на теорията на извадковите изследвания. Основната особеност е, че при работа с големи данни първо идват събраните данни и след това статистиците трябва да приложат специфични методи за тяхната обработка и анализ (визуализационни методи, техники за извличане на знания от данни или други методи, които да „направят големите данни малки“). Наред с методологията важен елемент на официалната статистика е осигуряването на информация с високо **качество**. Дали сегашните принципи и стандарти за качество са пряко приложими за големите данни, или е необходимо да бъдат подходящо адаптирани?

Поверителността и правните въпроси са друго предизвикателство. Защитата от разкриване на идентичността на индивидите е задължителна, но това е трудно да се осигури, когато е свързано с големите данни. Проблемът с големите данни е, че често потребителите на услуги и устройства, генериращи данни, не са запознати, че правят това и/или за какво тези данни могат да бъдат използвани впоследствие. Друг правен въпрос е свързан с авторските права и собствеността на данните. Дори ако данните могат легално да бъдат използвани, това не предполага че е разумно или подходящо да се направи с оглед нарушаване на тяхната конфиденциалност. Например в някои случаи може да бъде полезно да се приложи подходът на информираното съгласие. Някои договори за абонамент на мобилна



услуга включват клауза за използване на данни от договора за други цели освен за предоставяне на самата услуга.

Обработката, съхранението и трансферът на големи масиви от данни създават предпоставки за възникване на чисто **технологично** предизвикателство. Технологичният напредък - нарастването на компютърната мощност, по-големи ИТ устройства за съхранение и високоскоростни канали за данни, може частично да реши тези въпроси. Събирането на данни в реално време отваря нови възможности за комбиниране на административни данни с високоскоростни големи данни, идващи от различни източници, като търговски данни (транзакции с кредитни карти, онлайн разплащания, продажби и др.), мобилни устройства и сензори (мобилни телефони, GPS, камери, метеорологични сензори, сензори за замърсяване на въздуха и др.), социални медии (Twitter, Facebook, Google) и други обществено достъпни данни.

Друго предизвикателство е **възможната променливост и съпоставимост** на източниците на големи данни, имайки предвид факта, че официалната статистика поддържа анализ на динамичните редове. За много потребители продължаването на тези динамични редове е от особена важност и това не може да бъде пренебрегнато.

Не на последно място е размерът на **финансовите ресурси**, които статистическите институти трябва да заделят, за да придобият права на собственост върху големи данни, чиито собственици са компании от частния сектор. Още повече, че съгласно настоящото законодателство статистическите институти придобиват данни от държавни институции, вкл. от административни регистри и респонденти, безвъзмездно.

За да използва ефективно големите данни, официалната статистика се нуждае от експерти с **различно мислене и нови умения**, които да могат да извличат ценно „познание“ от данните. Такива специалисти са т.нар. изследователи на данни („data scientists“).

Големите данни за официалната статистика означават и по-голям обем информация, която е предмет на дефиниране на **нови политики и директиви** за управление и защита на тази информация.

Всички изброени предизвикателства пораждат основателния въпрос: не е ли дошло време за големите данни и тяхното място в официалната статистика да се мисли неконвенционално и в нова перспектива? Очевидно е, че към момента мисленето изостава по отношение на информационните технологии и големите данни.

Методите на статистиката и големите данни

Възниква фундаменталният въпрос: могат ли големите данни да се обработват и анализират с методите на статистиката? По-конкретно, това означава изследователите да могат да използват корелационния, дисперсионния и хи квадрат анализ, когато проверяват своите научни тези. В допълнение това означава още, че те могат да се възползват от целия инструментариум за проверка на хипотези. Очевидно е, че към традиционния инструментариум от теорията на статистиката ще бъдат включени и други подходи за изследване и анализ. Това обстоятелство на практика откроява две насоки за развитие на научноизследователската работа:

- Използване на добре известните статистически методи за анализ. Това създава условия за съчетаване на данните от официално провежданите статистически изследвания с потоците от големи данни.

- Въвеждане на други подходи, начини и методи за самостоятелно или съчетано използване на големите данни.

Възниква също въпросът: какво печели официалната статистика от използването на големите данни за аналитични цели? Очакванията в тази посока са най-вече в две направления:

- Намаляване на натоварването на респондентите. Това е особено важно, тъй като в последните години статистическите изследвания се развиват особено екстензивно. Потребностите от информация нарастват лавинообразно. Нарастват също и изискванията по отношение на качеството, бързината и детайлизацията на информационния продукт. Все повече внимание се обръща на информационната осигуреност на малки териториални пространства. Стремехът да се дефинират най-добрите и оптимални управленски решения налага разработването на нови показатели и подходи за съчетаване на данните от текущата статистика и големите данни.

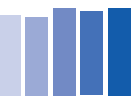
- Намаляване на цената на информационния продукт. Известно е, че статистическите изследвания имат своя цена, която се повишава в зависимост от поставените условия за представителност, точност и достоверност на статистическите оценки. Колкото по-високи са изискванията в това отношение, толкова по-големи ще бъдат разходите за направените изследвания. Стойностите могат да бъдат екстремални. Например ако се постави изискването точността на оценките от едно извадково статистическо изследване да нарасне два пъти, то обемът на извадката трябва да нарасне четири пъти. Това означава, че разходите за изследването (отпечатване на документи, заплати, транспортни разходи, консумативи, обучение на по-голям брой анкетьори, евентуално заплащане на респондентите и т.н.) също нарастват четири пъти при равни други условия.

Най-големите опасности от технологична гледна точка, които следва да се имат предвид, когато се съчетават данните на официалната статистика и големите данни, са:

- Достоверност на големите данни. Разгледани в този аспект, големите данни изключително много зависят от източника на тяхното генериране. Използването на индикатори, гарантиращи достоверността на източниците, е от първостепенно значение за качеството на големите данни.

- Представителност на големите данни. От съществено значение за представителността на големите данни е тяхното селектиране, избор и систематизация. Това означава, че субективният подход за тяхното генериране трябва да бъде елиминиран. Принципите и условията на рандомизация следва да бъдат водещи при определяне на данните, които се използват в съчетание със статистическите данни, отговарящи на същите условия.

- Точност на големите данни. Обемът на големите данни трябва да бъде достатъчен, за да се редуцира стохастичната грешка. Това означава, че големите данни имат висока аналитична значимост. Тази предпоставка предопределя качеството на изводите и посланията към потребителите.



На практика това са фундаментални понятия от теорията на извадковите статистически изследвания, които следва да бъдат ключът към успешно, целево и аналитично съчетаване на данните от официалната статистика и големите данни.

Как големите данни могат да бъдат използвани в статистическия бизнес процес?

Използването на източниците на големи данни в реалното производство на статистика е все още предмет на предварителни проучвания. Някои европейски страни вече са стартирали задълбочени проучвания в тази насока, такъв пример е Нидерландия по отношение на **статистиката от социалните медии**. Социалните медии са потенциален източник на големи данни. В тях хората доброволно споделят информация, дискутират интересуващи ги теми и общуват със семействата и приятелите си. Около един милион съобщения от социалните медии се генерират ежедневно в Нидерландия и са достъпни за всеки, който използва интернет. Нидерландският статистически институт изучава тези съобщения от две гледни точки: като съдържание и като мнение. Изучаването на съдържанието на съобщенията напр. в Twitter показва, че 50% от тях са с безсмислено съдържание и не носят никаква информация. Останалите са свързани предимно с дейностите през свободното време (10%), работата (7%), телевизията и радиото (5%) и политиката (3%). Изразяването на субективно мнение в социалните мрежи разкрива интересни потенциални възможности на тези източници за статистически цели. Мненията са тясно корелирани с потребителското доверие и в частност с мненията по отношение на икономическата ситуация. Оказва се, че тази корелационна връзка е стабилна на месечна и седмична база. Ежедневните данни не са толкова надеждни, тъй като показват висок дял на променливо потребителско поведение. Всичко това доказва, че е възможно производството на седмични и/или месечни статистически индикатори за потребителското доверие за първия работен ден след наблюдаваната седмица, като се отчита възможността за постигане на бързи резултати.

Успоредно с националната практика Евростат също стартира някои инициативи по отношение на големите данни. В годишните работни програми на Евростат са дефинирани някои предварителни проучвания и изучаването на възможности за използване на източниците на големи данни в следните статистически области:

Статистика на цените - използване и анализ на цените, събирани от интернет. Това е 24-месечен проект (стартирал през януари 2013 г.) за разработване на специфичен софтуер, подпомагаш специалистите при автоматизирано събиране на цените от интернет за изчисляване на индекс на потребителските цени. След приключване на проекта софтуерът ще бъде тестван методологично и технологично в пет европейски страни. Софтуерът, който ще бъде разработен в рамките на проекта, ще бъде предоставен и на други статистически организации за работа под лиценз (EURL).

Статистика на туризма - предварително проучване за използване на мобилно позиционирани данни за статистиката на туризма. През януари 2013 г. стартира 15-месечен проект, който изследва резултатите от използването на мобилно позиционирани данни за статистиката на туризма (и свързаните с него области) и оценява предимствата и недостатъците. Въпросите, които са предмет на проекта, са достъпът (и непрекъснатостта на достъпа), доверието (на

производителите и потребителите на статистика), разходите, понятията (транслиране на съществуващите концепции за статистиката на туризма към новите източници на данни) и други методологични въпроси (например представителност и обем на извадката и др.). Възможността за обхващане на големи файлове от данни, генерирани от мобилните оператори, се счита за предизвикателството, което трябва да се преодолее в бъдеще като следствие от успешно завършилия проект.

Използване на ИКТ - предварително проучване за използването на трафика на интернет потоците за събиране на статистика за информационното общество. Чрез този проект Евростат цели да изучи пилотно и да оцени приложимостта на потребителски ориентирани и уебориентирани измерителни подходи от техническа, методологическа, разходна, правна и социално-политическа перспектива. В рамките на проекта са планирани три отчета, които са ценни за практиката на националните статистически институти: 1) Как да се разработи официална процедура за събиране на статистика за използването на ИКТ от източници на големи данни; 2) Методологично и технологично ръководство за внедряване; 3) Тестване на концепцията „обединени отворени данни“⁴, отнасяща се за споделен набор от големи данни от частния сектор, които ще бъдат отворени за използване от статистическите институти.

През периода 2015 - 2020 г. се очаква развитието на множество проекти с цел проучване и използване на големите данни като достоверен и надежден източник за производство на официална статистика.

Бъдещи стъпки - сътрудничество: с кого и как?

Исправяйки се пред тези предизвикателства, официалната статистика ясно осъзнава необходимостта да не работи в изолация, а в сътрудничество с всички останали заинтересовани страни. Целта на това сътрудничество се състои обикновено в обмяна на знания, опит и добри практики.

Един възможен и необходим партньор са **потенциалните доставчици на големи данни**: ако те не гарантират достъпа до техните данни заради правилата за поверителност, историята ще приключи, преди да е започнала. Но тъй като източниците на големи данни не са пригодени за статистическо използване, такова сътрудничество е от съществено значение, за да се постигне добро познание за произхода на тези източници. Освен това за статистическото производство ще бъде по-ефективно да обработва данните на мястото на тяхното събиране и съхранение, т.е. данните няма да бъдат осигурени безплатно. От друга страна, статистиците могат да извършват анализи на големите данни, което ще помогне на техните собственици да ги разбират и тълкуват по-добре. По този начин взаимоотношенията с доставчиците на данни биха могли да преминат в истинско и ползотворно сътрудничество от взаимен интерес. В допълнение, официалната статистика може да играе и специфична роля на доверителна трета страна. На пазара конкурентите ще бъдат недоверчиви да споделят чувствителни данни един с друг. Но те могат да споделят тези данни със статистическите институти, събиращи статистическа информация, която е от полза за всички. Водещото правило в тази ситуация е създаване на статистическа култура и статистическо мислене. Работата в тази посока ще доведе до покриване на белите полета между експертите.

⁴ „Federated open data”.

Сътрудничеството между официалната статистика и **академичната общност** може да се разшири по отношение на големите данни - например при решаването на методологични проблеми, разработване на технически решения и обучение на бъдещи изследователи на данни (data scientists). В дългосрочен план новите знания и умения за големите данни ще изискват адаптиране на университетските учебни програми и планове за обучение по нови дисциплини - например за изследователи на данни, или надграждане на дисциплините в специалност „Статистика” (някои европейски университети вече предлагат подобни курсове на своите студенти). В краткосрочен и средносрочен план е необходимо статистическите институти да предложат на своя експертен състав специализирани обучения по темата за големите данни и по този начин да изградят вътрешен аналитичен капацитет. Международното сътрудничество в тази насока ще бъде много полезно за общността на официалната статистика. Такова сътрудничество също може да бъде подкрепено от разработването и финансирането на различни проекти, които подпомагат научноизследователската и иновативната дейност. Натрупването на знания е ключов момент за постигане на кохерентност между потоците от информация.

Има много **търговски партньори от частния сектор**, с които статистическите институти биха могли да си сътрудничат. Социалните мрежи Google и Facebook са два ясни примера, за които големите данни формират сърцевината на техния бизнес модел. Тяхното познание и данните, до които имат достъп, могат да бъдат и съотносими за официалната статистика. ИТ компаниите също притежават съответстващи знания за обработката, съхранението, сигурността и облачните технологии на големите данни.

Евростат и Европейската комисия са предприели различни инициативи в рамките на ЕСС, имащи за цел да изследват пълния потенциал на големите данни за европейската икономика, общество и обществени услуги.

Други международни организации (ОИСР, Статистическата комисия на ООН, Световната банка, UNECE) също са въввлечени в развитието на темата за големите данни. Показателен пример за това е проектът на стратегическата международна група по отношение на ролята на големите данни в модернизиранието на статистическия производствен процес. Целта на проекта е да направи възможно статистическите организации да действат единно чрез стандартизиран подход по методологически въпроси, свързани с големите данни и производството на статистика, да демонстрират приложимост на използването на източниците на големи данни за официална статистика и да подпомагат обмяната на знания, експертиза и средства между статистическите организации.

Някои **национални статистически офиси** вече осъществяват проекти и са запознати добре с големите данни. Техният опит от успешното използване на големите данни може да бъде изучаван и споделян с други страни с цел извличане на ценни познания и прилагане на добри практики по отношение на източниците на големи данни. Освен това националните статистически организации са окуражавани да включат официално въпросите за големите данни в техните годишни програми и стратегически документи чрез осъществяване на изследователски и пилотни проекти в избрани области и чрез разпределяне на подходящите ресурси за тези цели.

Интегриране на големите данни в ЕСС

За да посрещне предизвикателствата и да създаде подходящи условия за сътрудничество между частните единици и държавните статистически институти за достъп до нови източници на данни, Евростат е разработил пътна карта и план за действие. Всички заинтересовани страни от политиката, науката, статистиката и бизнеса, както и обществото като цяло, ще бъдат привлечени.

Целта на пътната карта е да направи възможно постепенното интегриране на източниците на големи данни в производството на европейска и национална статистика в дългосрочен, средносрочен и краткосрочен аспект, като по този начин ще допринесе значително за постигане на целите на „Визия 2020“ на ЕСС.

В дългосрочната визия (след 2020 г.) е залегнала идеята, че източниците на големи данни трябва да бъдат окончателно интегрирани в производствения бизнес процес на официалната статистика в ЕСС. Очаква се след 2020 г. източниците на големи данни да бъдат налични за широк кръг от статистически продукти и съответното законодателство да претърпи съществени промени, така че да позволи използване на източници на големи данни за официална статистика и да се работи в условия, когато обществото ще „признае“ използването на тези източници за статистически цели.

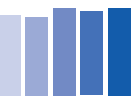
Средносрочните цели (до 2020 г.) включват пълна интеграция на стратегията за големите данни в националните стратегии, интегриране на подходящи ИТ инфраструктури, методологии, избрани приложения в статистическото производство, публично-частно партньорство с доставчиците на данни за целите на официалната статистика и увеличаване на броя на статистиците, занимаващи се с големи данни в ЕСС.

Краткосрочните перспективи (до края на 2016 г.) са насочени към дейности, свързани предимно с анализиране и подготовка на условия и инфраструктура за използването на големите данни в официалната статистика и стартиране на конкретни пилотни проекти за придобиване на знания и опит. Тези проекти включват дейности по анализ на източниците на големи данни, изследване на потенциала за партньорство с доставчиците на данни, развитие на обучителни програми, идентифициране на изследователските нужди, интегриране на стратегията за големите данни в цялостната стратегия на европейско ниво, доразвиване и внедряване на комуникационна стратегия.

Планът за действие е свързан с хоризонталните дейности, както и с провеждането на пилотни проекти. Целта на пилотните проекти е да се добият специфични умения при използване на големите данни в контекста на официалната статистика. Това би могло да включва идентифициране, анализиране и решаване на проблеми от хоризонтално естество или теми, както и обследване и развитие на бъдещи бизнес модели за производството на статистически данни, свързани със специфични източници на данни или статистически продукти. Предложеният подход изисква координация между пилотните проекти, свързани с източниците на данни и дейностите по теми. Евростат идентифицира следните групи по теми:

- **Политика**

Стратегията за големите данни в официалната статистика трябва да бъде вградена в управленските стратегии на национално и европейско ниво. Големите



данни трябва да бъдат част от осигуряването на информация за социално-икономическите програми за управление.

- **Комуникация**

Някои от източниците на големи данни съдържат чувствителна информация. Използването на тези източници за целите на официалната статистика може да предизвика негативни възприятия в обществото и заинтересованите лица. Ето защо е необходимо да се дефинират целите, процедурите и резултатите от използването на големи данни съгласно Кодекса на европейската статистическа практика, като фокусът е върху етичните принципи - например поверителността.

Наред с това в процеса на комуникация се верифицират данните от различните източници на информация, включително и информацията на официалната статистика.

Трябва да се има предвид, че медиите са ключов посредник между статистическите офиси и широката общественост и включването им в цялостната комуникационна стратегия е от съществено значение за успеха на плана за действие.

- **Източници на големите данни**

Броят на източниците на големи данни нараства лавинообразно. Разнообразието и размерът им определят и разширяването на потенциала на големите данни за производството на официална статистика. За да се вземат информирани решения за действие и за пилотните проекти, е важно да се работи за създаването и разработването на наръчник и за класифициране на източниците на големите данни. По този начин ще се осъществява мониторинг по отношение на тяхната автентичност, качество и достоверност.

- **Приложения/пилотни проекти**

На европейско и международно ниво вече са предприети действия по отношение на големите данни - например създаването на Task Force група в Евростат и High Level група на ниво Статистическа комисия на ООН. С тяхната дейност стартират процеси, които ще се разклоняват във времето и пространството.

- **Методи**

Използването на източници на големи данни изисква приложение на нови методи при анализа и обработката на данни. В същото време методите са зависими от източниците на данни, например дали тези източници съдържат структурирани данни, или просто текстова информация. Очевидна е необходимостта от ревизия на досега използваните методи за обработка и анализ на информацията, както и от нови подходи в тази насока. Това е продължителен процес с основна цел да се постигне ефективност при разкриването на закономерности в информационните потоци, формиращи големите данни, което е от първостепенно значение за доброто управление и натрупването на знания.

- **Качество**

Осигуряването на информация с високо качество е един от най-важните елементи на официалната статистика. Статистическата информация, която ще се произвежда от източниците на големи данни, е необходимо да е пригодна и да отговаря на стандартите за качество на ЕСС, за да бъде използвана от потребителите. Това е гаранция за доброто интегриране на данните от различни източници.

• ИТ инфраструктура

Характеристиките на големите данни, включително техният обем, скорост и разнообразие, оказват влияние върху ИТ системите и инфраструктурата. Бъдещите ИТ инфраструктури ще бъдат определени предимно от новите бизнес модели, които ще се внедрят за производство на статистика от големите данни. Това е процес, който е характерен за глобалната икономика.

• Умения

Достъпът, управлението, обработката и анализът на големите данни изисква специфични нови умения или комбинация от умения, които към настоящия момент не са част от официалната статистика. Тези нови умения се идентифицират с понятието „изследовател на данни” („data scientist”) и стремежът е да има все повече такива специалисти в официалните статистически организации. Наред с това успехът на аналитичното мислене повече от всякога ще зависи от умението да се екстрахират най-важните данни от много източници за кратко време, като се има предвид динамиката на изменение в информационните потоци. Уникалността на експертите в това отношение ще зависи от тяхната интуиция и сензитивност към информационните единици.

• Обмяна на опит

Важен елемент е споделянето на опит за проекти, приложения, пилотни проекти и източници на големи данни в ЕСС. Обмяната на добри практики е важен момент за ускоряване на процесите по опознаване и рационално използване на големите данни.

• Законодателство

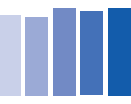
Законодателството е от основно значение при определянето на законова рамка за достъп, обработка и разпространение на статистически продукти от източници на големи данни. Това означава, че то трябва да бъде насочено към осигуряване на среда, където доминират точно установен ред, регламенти и информационна защита. Сложността на законодателството при тези обстоятелства зависи и от определянето на правилата, по които функционират информационните системи, контролирани и използвани от различни органи на управление.

• Управление

Планът за действие и пътната карта са интегрална и неделима част от „Визия 2020” на ЕСС. Дейностите, свързани с големите данни, са ключова област, дефинирана в Плана за внедряване на „Визия 2020”. Предизвикателствата при използването на големите данни на практика са неделима част от „Визия 2020”. Традиционните методи за управление, независимо дали вършат добра работа, се нуждаят от обновление. Успешното управление се осъществява единствено и само въз основа на знание. Въпросът е как това знание да се добие от информационните потоци, циркулиращи в общественото пространство.

Заклучение

Бъдещето на официалната статистика в ерата на големите данни все още е въпрос на обсъждане и експериментиране. Необходимо е цялата международна и европейска статистическа общност да се адаптира към новата реалност и да отговори адекватно на възможностите и предизвикателствата, които тази реалност предлага. За да се случи това, е нужно широко и ползотворно сътрудничество с различни „играчи” във и извън статистическата общност чрез формиране на



разнообразни мрежи, които могат да изградят нови пътища за генериране на статистически данни. Това е гаранция за успешно обществено развитие в условията на глобалната икономика.

За всички, които се занимават със статистика, ерата на големите данни е много вълнуващо време. Време за натрупване на нови знания с поглед към бъдещето. В обозримия хоризонт информационните потоци ще играят все по-осезаема роля в управлението на световните пазари. Технологията за използване на големите данни е в началния етап на своето развитие. Все още се дефинират рамките, мащабът и обхватът на дейностите. Това, което със сигурност се знае, е, че процесите са обективни и необратими. Контролът и управлението са невъзможни без качествено използване на потоците информация, които се генерират лавинообразно.

ЦИТИРАНА ЛИТЕРАТУРА:

Big data - an opportunity or a threat to official statistics? Paper prepared by Eurostat for the April 2014 plenary session of the Conference of European Statisticians as part of the seminar entitled „What is the value of official statistics and how do we communicate that value?“

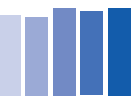
(<http://www1.unece.org/stat/platform/pages/viewpage.action?pageId=100307471>).

Big Data (and official statistics). Piet Daas and Mark van der Loo, Statistics Netherlands.

ESS Big Data Action Plan and Roadmap 1.0 Work Programme. Eurostat paper - room document.

Official statistics and Big Data, Peter Struijs, Barteld Braaksma and Piet JH Daas Big Data & Society 2014 1: DOI: 10.1177/2053951714538417.

What does „Big Data“ mean for official statistics? Michael Glasson (Australia), Julie Trepanier (Canada), Vincenzo Patrino (Italy), Piet Daas (Netherlands), Michail Skaliotis (Eurostat) and Anjum Khan (UNECE)
(<http://www1.unece.org/stat/platform/display/hlgbas>).



ГОЛЕМИТЕ ДАННИ - ВЪЗМОЖНОСТ, ПРЕДИЗВИКАТЕЛСТВО ИЛИ ЗАПЛАХА ПРЕД ОФИЦИАЛНАТА СТАТИСТИКА

*Галя Статева**

РЕЗЮМЕ В статията се разглежда ролята на големите данни в официалната статистика. Акцентът е поставен върху въпроса какво се случва, когато официалната статистика срещне „големите данни“. Дефинирани са основните възможности и предизвикателства пред официалната статистика в ерата „големи данни“: „нова“ методология, високо качество, поверителност и правни въпроси, ИТ технологии, финансови ресурси, необходимост от различно мислене и умения, нови политики и директиви.

Дискутира се как големите данни могат да бъдат използвани за производство на официална статистика: като заменят изцяло статистическите източници; като допълват информацията чрез съчетаване на данни от различни източници или като осигуряват напълно нови статистически числа, които могат да допълват наличната статистическа информация.

Описани са методите на статистиката и тяхното приложение по отношение на големите данни, като са посочени някои примери от европейската практика за използването на големите данни в статистическия производствен процес.

Изправяйки се пред възможности и предизвикателства, официалната статистика ясно осъзнава необходимостта да не работи в изолация, а в тясно сътрудничество с академичната общност, търговски партньори от частния сектор, Евростат и Европейската комисия, други международни организации и национални статистически офиси. В статията подробно са разяснени механизмът и инструментите за интеграция в рамките на Европейската статистическа система в краткосрочен, средносрочен и дългосрочен хоризонт.

* Държавен експерт в отдел „Обща методология и анализ на статистическите изследвания“, НСИ;
e-mail: gstateva@nsi.bg.

„БОЛЬШИЕ ДАННЫЕ” - ВОЗМОЖНОСТЬ, ВЫЗОВ ИЛИ УГРОЗА ДЛЯ ОФИЦИАЛЬНОЙ СТАТИСТИКИ

*Галя Статева**

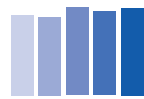
РЕЗЮМЕ В статье в основном рассматривается роль „больших данных” в официальной статистике. Акцент ставится на вопрос, что происходит когда официальная статистика встречается с так называемыми „большими данными”. Определены основные возможности и вызовы перед официальной статистикой в эру „больших данных”, таких как: „новая” методология, высокое качество, конфиденциальность и правовые вопросы, ИТ-технологии, финансовые ресурсы, потребность в различном мышлении и навыках, новые политики и директивы.

Обсуждается как „большие данные” могут быть использованы для производства официальной статистики различными способами: путем полной замены статистических источников; дополнением информации объединением данных из различных источников или полным обеспечением новых статистических чисел, будущих в состоянии дополнять имеющуюся статистическую информацию.

Описываются также методы статистики и их применение по отношению к „большим данным”, при этом упоминая некоторые примеры из европейской практики в использовании „больших данных” в статистическом производственном процессе.

Столкнувшись лицом с возможностями и проблемами, официальная статистика четко осознает необходимость в том, чтобы не работала в изоляции, а в близком сотрудничестве с академической общностью, коммерческими партнерами из частного сектора, Евростатом и Европейской Комиссией, другими международными организациями и национальными статистическими ведомствами. В статье детально изложены механизм и инструменты для интеграции в рамках Европейской статистической системы в краткосрочной, среднесрочной и долгосрочной перспективе.

* Държавен експерт в отдел „Обща методология и анализ на статистическите изследвания”, НСИ;
e-mail: gstateva@nsi.bg.



BIG DATA - OPPORTUNITIES, CHALLENGES OR THREATS BEFORE THE OFFICIAL STATISTICS

*Galia Stateva**

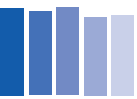
SUMMARY The article mainly discusses the role of big data in official statistics. Emphasis is placed on the issue of what happens when official statistics meets the ‘big data’. Defined are the main opportunities and challenges before the official statistics in the ‘big data’ era, ‘new’ methodology, high quality, confidentiality and legal issues, IT technologies, financial resources, the necessity for different thinking and skills, new policies and directives.

Discussed is how big data can be used for production of official statistics in various ways: by totally replacing statistical sources by complementing information by combining data from different sources or provision of completely new statistical figures that can complement the available statistical information.

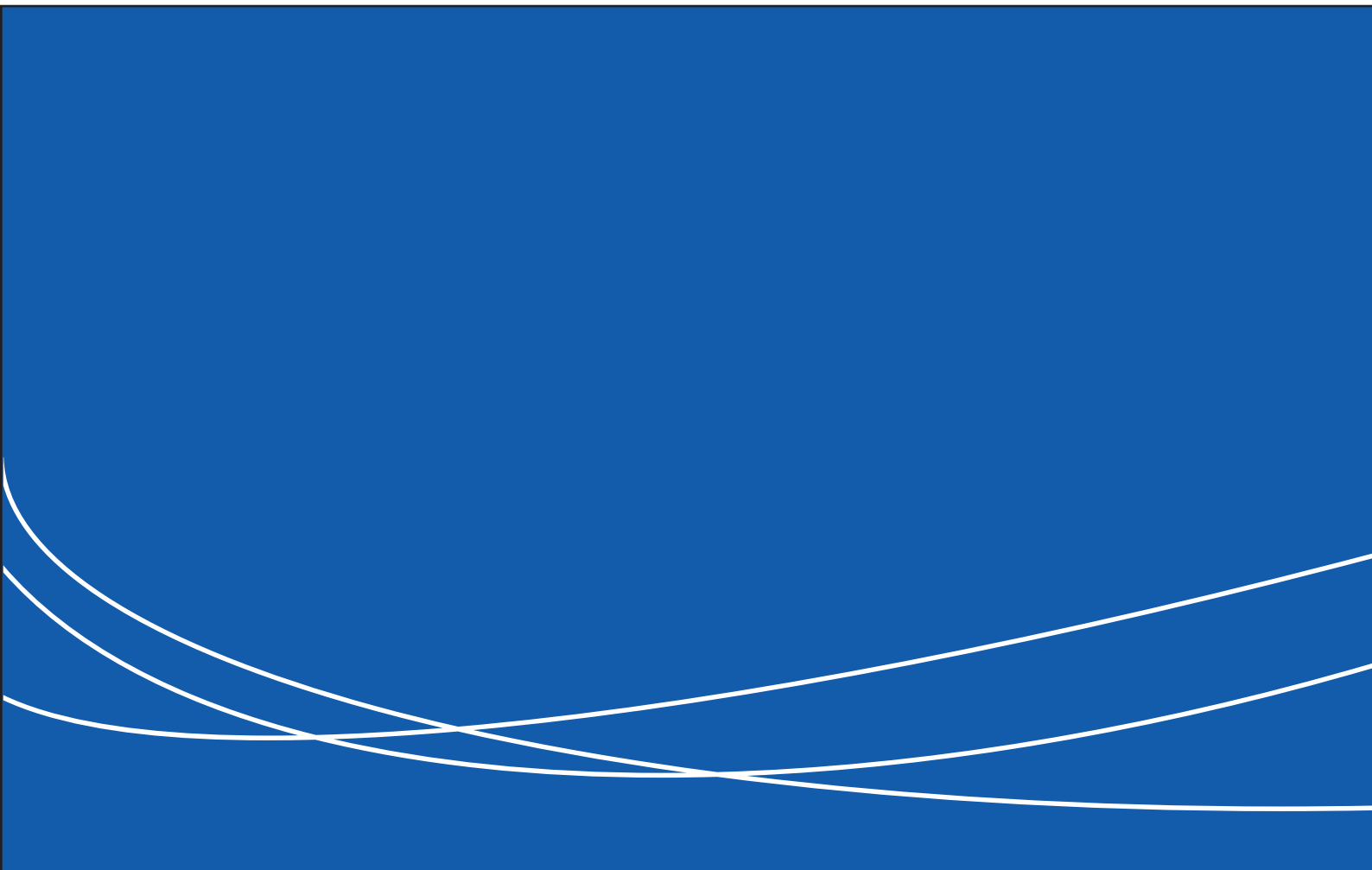
Described are methods of statistics and their application regarding big data, referring to some examples of European practice for the use of big data in the statistical production process.

The official statistics clearly notes the need not work in isolation but in close cooperation with academia, commercial partners from the private sector, Eurostat and the European Commission, other international organizations and national statistical offices. The article explains in detail the mechanism and tools for integration within the European Statistical System in short, medium and long term.

* State Expert ‘General Methodology and Analysis of Statistical Surveys’ Department, NSI;
e-mail: gstateva@nsi.bg.



**ИСТОРИЧЕСКО РАЗВИТИЕ НА
СТАТИСТИЧЕСКАТА ТЕОРИЯ И ПРАКТИКА**



ИЗМЕНЕНИЯ В БРОЯ НА НАСЕЛЕНИЕТО ПО БЪЛГАРСКИТЕ ЗЕМИ В СЪСТАВА НА ОСМАНСКАТА ИМПЕРИЯ

*Димитър Аркадиев**



В края на XIV век българският народ губи своята държавност почти за пет столетия. Българските земи са завладени от непознат дотогава завоевател - държавата на османците. Това се отразява твърде неблагоприятно върху растежа на населението на завладените територии или поне на по-голямата част от тях. В следващите векове се променят и някои негови структури и преди всичко тази по религиозна принадлежност. Българският народ се откъсва от семейството на повечето европейски народи, които през това време постигат забележителни успехи в своето развитие. Най-важното е, че в българските земи се налага източен (неевропейски) цивилизационен модел, отличаващ се с голяма консервативност и обективно забавящ общественото развитие.

Българските земи се намирали на кръстопътя при завоевателните походи на османците. Обширни гъсто заселени преди области като Тракийската низина, Дунавската равнина и др. запустели.

За това положение съществуват изобилни свидетелства на съвременници от тази епоха. Ето две от тях. Византийският хронист Дука (началото на XV век) съобщава, че „турците от Мала Азия вървят пеша до Дунава, за да поробват християни. Защото, въпреки че десетки хиляди от тях нахлуват в някоя провинция, все пак те навлизат като разбойници и след като я ограбят, избягват. Тези нашествия са превърнали в пустиня цяла Тракия чак до Далмация”¹. В писмото на Бартоломео де Яно до ръководителя на йоанитския орден в Ерусалим от 1443 г. се посочва: „От шест години насам турците са отвлекли от християнска земя повече от 400 000 християни, заробили всички, погубили и унищожили множество от царствата и областите на Сърбия, Словения, Далмация, Трансилвания, България, Босна и особено кралствата Унгария и Саксония, като не смятаме старите и болните, които убивали и обезглавявали, понеже не можели да ги

* Д-р, професор в Стопанския факултет на Великотърновския университет „Св.св. Кирил и Методий”; e-mail: arkadiyev@abv.bg.

¹ Петров, П., В. Гюзелев, (1978) Христоматия по история на България. Том 2. Същинско средновековие - краят на XII-XIV век. С., Наука и изкуство, с.377.

откарат, като не вземем под внимание градове, крепости и селища с цялата им околност, които разориха с огън и меч. Ние го видяхме дори със собствените си очи и го наблюдавахме и сега”².

Това показва, че в епохата на първоначалното създаване на Османската империя условията за растеж на населението в завладените земи са били меко казано неблагоприятни. Някои от факторите, действащи по посока на намаляване на неговия брой, са отвлечане на част от населението (предимно от младите възрасти) като военна плячка, пряко физическо унищожаване, оттегляне на население в по-спокойни места или изселване в по-далечни територии, възможна поява на епидемии (чумни и др.) и т.н. В нотариалните книги на венецианския нотариус Антонио Брешиано са открити записи за продажбата на огромно количество българи роби на пазарите на остров Крит³. Така че народните предания за „трите синджира роби”, не са само фолклор.

Тук е мястото да се изкажат някои съображения за османското завоюване на българските земи и на целия Балкански полуостров, за което са създадени и се разпространяват в публичното пространство много митове и легенди. То не продължава месец-два, дори няколко години, а цяло столетие.

Какво представлява Османската държава в средата на XIV век? Отговорът на този въпрос обяснява много от действията на османците, а донякъде и поведението на балканските държави.

В средата на XIV век османската държава е заключена между Мраморно, Черно и Егейско море, а на изток има сухоземна граница. Общата ѝ територия е с дължина около 400 км, широка е около 150 км, а площта ѝ е около 50 хил. км². Тя обхваща приблизително съвременните вилаети Коджаели, Сакария, Бурса, Балъкесир, Чанаккале и Беледжик. Понастоящем повечето от тях са доста гъсто населени, но въпреки това в тях живее по-малко от 10% от съвременното население на Турция (и на азиатската ѝ част). Вероятно броят на населението на османската държава по това време не е превишавал 250 - 300 хил. души. Броят на възрастното мъжко население не е достигал повече от 60 - 75 хил. души. Само част от тях са могли да участват във военни действия.

Тези съображения обясняват защо османците не си позволяват мащабни военни действия, а завоюват парче по парче територии от балканските държави в продължение на десетилетия, а също така използват и начини за тяхното завладяване, различни от военните. Това се отнася и за задържането на вече завоюваните територии. Османците не са имали достатъчен човешки ресурс да ги населят в задоволителна степен, за да поддържат властта си.

Това обяснява донякъде и поведението на балканските държави по същото време. Османската държава не се е различавала от множеството по-големи или по-малки независими владения на феодали, отделили се от по-големите държави майки като Византия, България, Сърбия и т.н. Непрекъснатото съперничество и противоречия между по-големите държави и противоборствата вътре в тях не позволяват на техните владетели да видят опасността, която се задава от сравнително малката държава на османците.

Във всеки случай това все още не е могъщата империя от следващите векове с площ колкото половин Европа (над 5 млн. км²).

Казаното се потвърждава от някои разработки по темата.

За османските методи (или поне за част от тях) за завоюване и задържане на завладените земи съобщават и някои съвременни турски автори. Един от тях е Халил

² Георгиева, Цв., Д. Цанев (1986). Документи за историята на българския народ XV - XIX век. С., Народна просвета, с. 7.

³ Петров, П., В. Гюзелев (1978). Цит. съч., с. 196 - 198.



Иналджък⁴. Според него първият е методът на постепенното завладяване. Той се провежда в два етапа: 1) установяване на сюзеренитет над съседните държави; 2) установяване на пряк контрол върху съседните държави чрез премахване на местните династии.

Прекият контрол от страна на османците означава основно прилагане на системата за описи, които се основават на методично записване на населението и ресурсите на страните в дефтери (официални регистри). Създаването на опис-система не означава непременно революционна промяна в първоначалния социален и икономически ред. Тя била в действителност консервативно помирение на местните условия с османските институции и имала за цел постепенната асимилация. Използването на тези два етапа при османските завоевания може да се открие още в началото на османската история.

По-нататък е търсен повод васалите да направят някаква грешка, за да бъдат отстранени от власт, и страната им да бъде окончателно завладяна.

Един от военните методи е след завладяването на някакви крепости повечето от тях да бъдат разрушени по специална заповед на султана. Оставят се няколко неразрушени на стратегически места с малки гарнизони. Тази мярка, която е често прилагана от османците, е взета, от една страна, за да се избегне необходимостта от поддържане на военни сили в тях, и от друга страна, за да се предотврати възникването на центрове на съпротива при местните господари. След това като правило на конниците спахии, от които се състои основната сила на османската войска, са дадени тимари в селата на цялата новозавладяна страна. Другият начин е привличането на местните християнски владетели и подчиненото им население в османската войска при завоевателните походи. Тези владетели приемат исляма лично или в следващите поколения.

Необходимо е да се даде отговор на един важен въпрос. Ако броят на населението в съвременните граници на България преди османското нашествие е 2 - 2.5 млн. души, възможно ли е за около един век той да намалее с повече от 1 млн. души?

Това число обикновено се оспорва. Но незнаяно защо не се взема предвид, че то не е получено в резултат на еднократен акт, а е последица от действия и събития, продължили 100 - 150, а може би и повече години. Ако разделим посоченото число (1 млн.) на 100, ще получим средногодишно 10 хил., а на 150 - около 7 хиляди. Относително това прави средногодишно около 5% или дори по-малко. В условия, когато обикновено смъртността е достигала 35 - 40%, посоченото число е пренебрежимо малко. Това означава, че целият естествен прираст и малко повече е „изяждан“ от описаните събития.

Неслучайно Хр. Матанов отбелязва: „Впрочем известно е, че периодът от втората половина на XIV и целия XV век се характеризира с демографска стагнация“⁵.

Незнаяно защо това число се свързва с избиване на населението. Всъщност основната причина за намаляването на неговия брой е външната миграция - или самостоятелна към по-спокойни места, или насилствена към райони в Мала Азия, опустошени в резултат на няколко века военни действия, грабежи и разбойничества.

По този повод византийският император Михаил II Палеолог, участвайки в поход в източните части на Мала Азия през 1391 г. като васал на султан Баязид, пише в свое писмо: „Без съмнение римляните са имали название за малката долина, където се намираме, когато са живеели и управлявали тук... Тук има много градове, но не им

⁴ Inalcik, Halil (1954). Ottoman Methods of Conquest. Studia Islamica, No. 2, pp. 103 - 129.

⁵ Матанов, Хр. (2000). Възникване и облик на Кюстендилския санджак през XV - XVI в. София, ИФ-94, с. 113 - 114.

достига това, от което се състои истинското великолепие на града..., а именно, хора. Мнозинството лежи сега в руини... не са запазили дори техните названия...”⁶.

Но и в двата случая миграцията е предизвикана от завоевателите.

Направленията на миграциите в ранната стратегия на Османската империя се обсъжда в изследването на Пол Хупър „Принудителни прехвърляния на населението в ранната османска имперска стратегия”⁷. Основното направление на преселване от източните предели на Мала Азия е на номадски групи като юруци и други. Тези заселници се настаняват на свободни земи на Балканите и не търсят някакво по-особено устройване. Обратният поток е към пределите на Мала Азия на земеделско население с опит в това занятие от Балканите. С това се решават най-малко две задачи: по-гъсто заселване и обработване на тамошните земи за осигуряване на добиви от селскостопански култури и по-лесно приемане от населението на ислямската религия.

Преселваните са лица предимно в млада възраст, вкл. и омъжени млади жени с техните деца. С други думи, това са тези контингенти на населението, от които зависи разширеното възпроизводство на неговия брой. Колкото по-малко са те в оставащото на място население, толкова по-малък е неговият естествен прираст. В много случаи той може да бъде и отрицателен.

Роля играе и „кръвният данък” (девширме), чрез който са били отнемани деца - момчета от немюсюлманското население. Те са приемали исляма и след обучение са попълвали предимно войските на империята, тъй като първоначалното мюсюлманско население не е могло да компенсира изцяло загубите при почти непрекъснатите военни действия.

Влияние оказват и разпространените по това време чумни епидемии. Те започват от средата на XIV век - времето, когато започва и настъплението на османците на Балканския полуостров. В Западна Европа те отнемат 30 - 40% и повече от броя на населението. Макар че отсъстват сигурни сведения, възможно е рядката населеност на североизточните български земи, на райони в Тракия и Дунавската равнина да е предизвикана от чумата. Това без съмнение улеснява завоевателните походи на османците. Кратко обобщение на тази причина прави Ивелин Иванов⁸.

Друга причина е настъпването на продължителен период на промени в климата, довеждащи до засушавания и оттам - до по-ниски добиви от селскостопанските култури, недостатъчна паша за селскостопанските животни и в крайна сметка - трудно изхранване на по-многобройно население.

Наред с други причини посоченото съчетаване на неблагоприятни обстоятелства води до намаляване на броя на населението.

Трябва да отбележим, че в настоящето само за 25 години (1989 - 2014 г.) България загуби около 1.8 млн. души население, или над 70 хил. души средногодишно (8 - 10%), без да е имало войни, а само външна миграция и естествени възпроизводствени процеси (по-висока смъртност от промяната на възрастовия състав и по-ниска раждаемост).

В други периоди - например Балканските войни (1912 - 1913), Първата световна война (1914 - 1918), а и Руско-турската освободителна война (1877 - 1878) и други

⁶ Финкелъ, К. (2010). История Османской империи. Видение Османа (перев. с англ.). Москва, АСТ, с. 51 - 52.

⁷ Hopper, P. L. (2003). Forced Population Transfers in Early Ottoman Imperial Strategy: a Comparative Approach. A senior thesis submit to the Department of Near Eastern Studies of Princeton University in partial fulfillment of the requirements for the degree of Bachelor of Arts. Princeton. http://www.unm.edu/~phooper/thesis_condensed.pdf.

⁸ Иванов, Ив. (2003). Чумата в Европа и българските земи в края на средновековието. Електронно списание Liternet, бр. 6 (43), юни/2003; http://liternet.bg/publish8/ivelin_ivanov/chumata.htm.

драматични събития, загубите са много големи предвид краткия период, през който се случват.

Отговорът на този въпрос трябва да се търси и в сведенията, които са оставени от самите османци чрез описите на домакинствата по тогавашното административно деление. Известно е, че отделните територии не са придобити само чрез военни действия, а и по други начини, един от които е чрез доброволно присъединяване под различни форми. По този начин най-общо могат да се сравняват територии, където не е имало военни действия (или поне не често), и такива, където е имало. Разликите в броя и най-вече в гъстотата на населението би могла да покаже наличието на загубите му в резултат на военните действия и последиците, предизвикани от тях.

Един бегъл поглед върху събитията показва, че през целия XIV и XV век османците предприемат почти непрекъснато походи към съседните държави⁹.

В повечето случаи османците не са създавали нови административни единици, а са наследявали териториалното деление до епохата на завладяването или доброволното присъединяване. По такъв начин Видинският санджак отговаря (възможно приблизително) на територията на Видинското царство към края на XIV век, Никополският санджак - на Търновското царство, Кюстендилският - на владенията на Константин Драгаш, и т.н.

Системата за създаване на описи на завладените земи, за която споменава Х. Иналджък, е покривала всички завзети територии. Тя се налага във времето след превземането на Константинопол (1453 г.), когато османската държава се заема по-сериозно с устройството на завладяната територия. През периода 1543 - 1546 г. се установява първото османско „кануннаме“ (основни правила) за положението и задълженията на поданиците на империята¹⁰. Доколкото е известно, първите тимарски регистри, открити в архивите, са направени около средата на XV век¹¹.

Открити са много документи, съдържащи описи на домакинствата по това време по повод на събирането на различни данъци, раздаването на тимари и т.н.

Един такъв опис е обнародван и обсъждан от Н. Тодоров¹². Той съдържа данни за броя на джезие-ханета (християнски домакинства, плащащи данък „джезие“) по санджаци за 1490 - 1491 година.

Какво прави впечатление при преглеждането на тези данни? На първо място, най-голяма гъстота на домакинствата има в тези територии, които доброволно са се присъединили към османската държава и са й останали верни (включително са поддържали военната ѝ експанзия). Повечето от тях се намират в планински и полупланински места. На второ място, най-малка гъстота има в териториите, които са били васални и не са останали верни. Преобладаващата част от площта им се намира в равнинни места. Подобно е положението и в територии, където са се водили военни действия или са били на кръстопътя на завоевателните походи. На трето място, средна гъстота се наблюдава в територии, които са оказали съпротива, но се намират в труднодостъпни места.

Основният извод, който се налага, е следният: Първата категория територии след османското нашествие са запазили, общо взето, първоначалния брой на населението си

⁹ Хитцелъ, Фр. (2006). Османская империя. Перев. с франц. Москва, Вече, с. 29 - 45.

¹⁰ Хитцелъ, Фр. (2006). Цит. съч., с. 32.

¹¹ Ковачев, Р. (2005). Още за етническите граници на българската народност според османските регистри от XV и XVI в. - Известия на катедра „Българска история и археология“ и катедра „Обща история“ - ЮЗУ Неофит Рилски“- Благоевград, т. 5.

¹² Тодоров, Н. (1959/60). За демографското състояние на Балканския полуостров през XV - XVI в. - Годишник на Софийския университет, Философско-исторически факултет, т. I - III, кн. 2, с. 191 - 232.

или за загубили съвсем малка част от него. Втората категория територии са загубили по-голямата част от населението си. Третата категория територии са загубили значителна част от населението си.

До сходни изводи стига и Хр. Матанов¹³. Той сравнява населението на Кюстендилския санджак, който спада по нашата класификация към първата категория, и други санджаци като Видинския и Никополския, които спадат към втората категория. Установява се голяма разлика при гъстотата на домакинствата (съответно населението) между тях. Авторът смята, че броят на населението на Кюстендилския санджак към средата на XIV век е същият както към края на XV век.

По данните на Н. Тодоров средната гъстота на домакинствата на квадратен километър за Кюстендилския санджак е 2.16 бр., за Видинския - 0.94 бр., за Никополския - 0.48 бр., и за Силистренския - 0.32 броя. За по-голяма яснота при приемане на 5 лица средно на домакинство приблизително гъстотата на населението е: Кюстендилски санджак - 11 бр./км², Видински - 5 бр./км², Никополски - 2.5 бр./км², и Силистренски - около 2 бр./км².

От тази гъстота следва, че тя е най-голяма в преобладаващо планинския и полупланинския терен на Кюстендилския санджак, а по-малка - в преобладаващо равнинния терен на Видинския и Никополския санджак. Обикновено при нормални условия положението е обратното. Не е възможно гъстотата на населението в плодородните равнинни райони да е по-малка отколкото в планинските (поне за европейските условия).

Нека да сравним какъв ще бъде броят на населението, ако се приеме средната гъстота на Кюстендилския санджак - 11 бр./км² (табл. 1).

1. Условен и действителен брой на населението в някои санджаци на османската държава към края на XV век

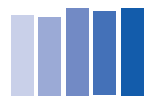
Санджаци	Население - хил. бр.			Отношение - условно/действително население
	условно при 11 бр./км ²	действително	абсолютна разлика	
Кюстендилски	220	220	0	1.0
Видински	120	50	70	2.4
Никополски	400	90	310	4.4
Силистренски	420	60	360	7.0
Софийски	130	95	35	1.4
Общо	1290	515	775	2.5

Дори от тези груби оценки и добавянето на Софийския санджак се изяснява, че населението в сравнение с това към средата на XIV век е намаляло повече от два пъти. Сравненията тук са правени при гъстота 11 бр./км², но в равнинните райони тя е била по-голяма. Така че тази хипотеза е приемлива.

Обобщено мнение за последиците от османското нашествие споделят много автори.

В. Гюзелев посочва, че за намаляването на българското население през XII - XIV век има няколко причини. Една от тях е османското нашествие на Балканския полуостров (1352 - 1453 г.). Състои се в избиване на населението, отвлечане в робство,

¹³ Матанов, Хр. (2000). Цит. съч., с. 110 - 135.



емиграция в Сърбия, Влахия, Молдова, Русия, Унгария, Полша, Италия и други. В резултат населението е намаляло почти наполовина¹⁴.

Цветана Георгиева въпреки по-мекото изразяване също стига до подобни изводи. Според нея продължителността на османското завладяване е една от основните причини за демографските загуби¹⁵.

До подобни изводи достига и Павел Мурджев¹⁶.

Някои изследвания потвърждават голямата емиграция на българско население към съседни земи. Смята се, че в Румъния за целия период на османското владичество са се изселили 0.8 - 1 млн. души¹⁷. Необходимо е да се подчертае, че това са съседните земи на едни от най-рядко населените санджаци през XV век - Видински, Никополски и Силистренски.

Трябва да се отбележи, че има и противоположни мнения, появили се у нас след началото на 90-те години на XX век. Но затова е необходимо да се отговори на въпроса на Цицерон: „Кому е изгодно?“, за да се разбере причината за тях. Все пак е желателно да се вярва повече на свидетелите и свидетелствата от онази епоха, отколкото да се правят изводи и прилагат виждания и теории от съвременността.

След успокояване на обстановката на Балканския полуостров през следващите векове българският народ не само успява да преодолее демографската катастрофа, но и да възвърне положението си на най-многоброен етнос на полуострова.

По отношение на източниците на данни за населението обсъжданият период коренно се различава от предишните.

На първо място, съществуват много документи и продължават да се откриват и разработват нови, които характеризират пряко или косвено броя и състава на населението.

На второ място, с определяне на този брой и състав са се занимавали много автори - както съвременници на съответната епоха, така и от по-ново време.

Затова тук ще бъдат набелязани само някои основни моменти с оглед на целта и задачите на настоящото изследване.

По отношение на документите за броя на населението и на неговия състав изследваният период може да се раздели на два подпериода - първи, до началото на XIX век, и втори - от 30-те години на XIX век до 1878 година (Освобождението на България).

През първия подпериод основен източник на данни, с помощта на които може да се характеризира броят на населението по българските земи (а и в останалата част на Балканския полуостров), са различни описи на домакинствата. Те са съставяни за фискални цели. Част от тях са напълно или частично запазени и пригодени за използване, а друга част все още се намира в архивите и постепенно се обработва, за да стане достояние на съвременните изследователи на тази епоха.

Тези документи се отличават с някои особености.

На първо място, като основна облагаема единица е възприета „хане“ (буквално къща), която включва едно, а понякога и повече домакинства. На второ място, в първичните описи са отбелязвани само мъжете над определена възраст (15 години).

¹⁴ Гюзелев, В. (1981б). Средновековна България (история, държава, църква, народност и култура). В: История, изкуство и култура на средновековна България. С., Народна просвета, с. 9 - 29.

¹⁵ Георгиева, Цв., Генчев, Н. (1999). История на България XV - XIX век. История на България в три тома, том II. София, Аноубис, с. 57 - 63.

¹⁶ Murdzhev, P. (2008). The Medieval Town in Bulgaria, Thirteenth to Fourteenth Century. A PhD Dissertation, University of Florida, pp. 54 - 70, http://etd.fcla.edu/UF/UFE0024053/murdzhev_p.pdf.

¹⁷ Трайков, В., Н. Жечев (1986). Българската емиграция в Румъния XIV век - 1878 година и участието ѝ в стопанския, обществено-политическия и културния живот на румънския народ. С., Наука и изкуство, с. 151 - 155.

Женското население е пропускано с изключение на вдовиците - глави на домакинства. Отбелязвани са отделно и неженените мъже.

На трето място, забелязва се известна манипулация в данните. Вероятно това е с цел облагодетелстване на служебно ангажираните със събирането на данъци лица.

През втория подпериод основен източник на данни са започналите в периода на реформите (танзимат) преброявания на населението в Османската империя. Началото им е през 30-те години на XIX век.

Резултатите за някои от тях с различна степен на подробност се поместват в започналите да излизат от 1847 г. в империята годишници (Салнаме), а от 1865 г. - и за отделните вилаети със статистически, географски и други данни и сведения¹⁸.

Втори източник са наблюденията на място, извършвани от различни чуждестранни учени и пътешественици, дипломати, стопански военни дейци, журналисти и други. В техните съчинения, доклади, дописки и т.н. са отразени и данни за населението и неговия състав.

Трети източник са данните, събирани от различни организации на българското население - екархийски списъци и други.

Преброяванията на населението в Османската империя са извършвани през различни години, като са обхващали изцяло всички или само отделни области на империята. Първото успешно всеобщо преброяване е през 1831 г., а следващото - през 1844 година. Някои не завършват (1854) или не се провеждат (1870). Други обхващат части на империята (1856, 1866, 1874). Преброяванията продължават до окончателното ѝ разпадане след Първата световна война.

Какво е характерно за посочените преброявания като основен източник?

Първо, данните или не са публикувани (в съответното време), или това е направено частично.

Второ, публикуваните данни не включват обобщаване на всички събирани при регистрацията сведения - отразява се само облагаемото мъжко население, или в най-добрия случай цялото мъжко население.

Трето, методологията на събиране на данните, категориите население, които са включени, или никога не се изясняват, или това се прави частично.

Четвърто, регистрацията на първичните сведения не се прави към определен критичен момент (дата), а дори понякога продължава по няколко години.

Посочените съображения налагат много внимателно да се използват споменатите данни.

Тук авторите, разработващи темата, могат да бъдат разделени условно на няколко групи.

В първата група са авторите, които правят оценки за броя и състава на населението в съвременните граници на България. Към тази група принадлежат Ан. Тотев, П. Стоев, П. Петров, Д. Аркадиев, К. МакЕведи и Р. Джоунс, А. Мадисън, М. Паларе и други.

Във втората група са тези автори, които се изказват за населението на Османската империя, нейни части, други държави на Балканския полуостров до 1878 г., а и след това. Такива са Ж. А. А. Юбичини, Ами Буе, В. Теплов и други.

Към третата група спадат автори, които са развили отделни теми, свързани с демографията на българските земи в тази епоха. Тази група е най-многобройна. Към нея могат да се причислят М. Дринов, Хр. Гандев, Н. Тодоров, М. Тодорова, Хр. Матанов, Антон Минков и много други.

¹⁸ Duman, H. (2000). A Bibliography and Union Catalogue of Ottoman Yerrbooks. - Fondation for Information and Documentation Servisies, Vol. I, pp. 19 - 38, Ankara.

Към четвъртата група спадат автори, които са събрали статистически данни за броя и състава на населението въз основа на официални източници, архивни документи и други. В тази група са Й. Л. Баркан, К. Карпат, П. Буте, Сл. Драганова, Н. Тодоров, М. Тодорова, Р. Мор и други.

С оценки на броя и състава на населението по българските земи по това време са се занимавали извънредно много чуждестранни автори. По-голямата част от тях са съвременници на съответната епоха. Интересът им е предизвикан преди всичко от събитията около Априлското въстание (1876), последвалата го Руско-турска война (1877 - 1878) и Освобождението на България. Повечето от направените оценки на броя на населението не са съпоставими по териториален обхват и затова няма да бъдат предмет на обсъждане освен в специални случаи.

Извадки от техните съчинения, отнасящи се до обсъжданата тема, са събрани в капиталния труд на Никола Михов (1877 - 1962) „Населението на Турция и България през XVIII и XIX век“¹⁹. Също така част от тях са коментирани и обнародвани и от Кемал Карпат^{20, 21}.

Някои от най-известните автори са Ж. А. Абдолоним Юбичини (Jean Henri Abdolonyme Ubicini) (1818 - 1884), Гийом Лежан (Guillaume Lejean) (1828 - 1871), Ами Буе (Ami Boué) (1794 - 1881), Феликс Каниц (Felix Kanitz) (1829 - 1904), Пол Буте (Paul Boutet), Юрий Ив. Венелин (1802 - 1839), Владимир Ал. Теплов и много други.

Редица български автори обсъждат изменението на населението като цяло за българската народност на Балканския полуостров в днешните териториални граници на българската държава, в отделни райони и части на страната.

Един от първите български автори, който прави опит да определи броя на населението върху част от българските земи след османското нашествие, е Марин Дринов (1838 - 1906). Поводът за това са резултатите от първото преброяване на населението в Княжество България към 1.01.1881 г. и разпределението му по матерен език. Тези резултати показват значителна концентрация на турско (и мюсюлманско население) в източната част на Княжеството - 482 349 турци срещу 477 132 българи. В два броя (VII и VIII) на „Периодическо списание на Българското Книжовно Дружество“ през 1884 г. той публикува студията си „Историческо осветление върху статистиката на народностите в източната част на Българското княжество“²². Поставя си за цел да проследи исторически и да обясни защо се е получило това положение. Започва своите изследвания по отделни векове (XV, XVI и т.н.) за районите с най-голяма концентрация на турско население - Добруджа, Делиорман (Лудогорие) и Герлово. Въз основа на различни исторически факти доказва, че през XV век турците там са били нищожна част от населението. В края на XVI век той стига до извода, че християнското (българското) население е 400 - 500 хил., турското - 150 хил. души. Съответно към началото на 1881 г. българското население в същите райони е 230 хил., а турското - 430 хиляди. Той дава и обяснение за полученото през вековете обратно съотношение между българското и турското население - от 2 - 3 към 1 през XVI век до 1 към 2 за 1881 година. За съжаление, изследванията му за следващите векове не са завършени.

¹⁹ Михов, Н. (1915 - 1967). Населението на Турция и България пред XVIII и XIX в. Библиографски издирвания със статистични и етнографски данни. В 5 тома (т. 1 - 1915, т. 2 - 1924, т. 3 - 1929, т. 4 - 1935, т. 5 - 1967). С., БАН.

²⁰ Karpat, K. (1985). Ottoman Population 1830 - 1914. Demographic and Social Characteristics. Madison, University of Wisconsin Press, pp. 18 - 86.

²¹ Karpat, K. (2003). Osmanlı nüfusu (1830 - 1914). Demografik ve sosyal özellikleri. İstanbul, Tarih Vakfı Yurt Yayınları, s. 40 - 144.

²² Дринов, М. (1909). Историческо осветление върху статистиката на народностите в източната част на Българското княжество. В: Съчинения, т. I. Трудове по Българска и Славянска история. С., Държавна печатница, с. 521 - 548.

Единственият български автор, който дава цялостна оценка за общия брой на населението в съвременни граници, а също и обобщение на направеното преди него, е Анастас Тотев (1906 - 2000) в редица негови статии и други произведения^{23, 24}. Има отделна оценка за 1875 година. Тя е направена въз основа на материалите за населението по повод Цариградската конференция (1876 г.) за решаване на българския въпрос след Априлското въстание чрез създаване на две автономни области²⁵. За тази оценка вече беше споменато. Населението на България според нея към 1875 г. е 4 380 хил. души. Според мен при тази оценка броят на населението е преувеличен.

Посочените оценки са поместени в табл. 2. Те са използвани за същия период от време от Димитър Аркадиев²⁶.

Оценки за населението в тази епоха прави и Павел Стоев²⁷, които се отнасят за етническите граници, защото през този период не съществува българска държава. Населението е изчислено общо и само българи. Оценка са поместени в табл. 2. Те са приведени от мен към сегашната територия на страната съобразно средната гъстота на населението.

Оценките, които прави Петър Петров²⁸, също са поместени в табл. 2.

От чуждестранните автори най-важни са оценките на Колин МакЕведи и Ричард Джоунс²⁹.

Оценки помества и Ангъс Мадисън (1926 - 2010)³⁰. За 1500, 1600, 1700, 1850 и за 1900 г. те са същите както на К. МакЕведи и Р. Джоунс. Населението за 1820 г. е 2 187 хил., за 1870 г. - 2 586 хил., за 1890 г. - 3 445 хил. души.

2. Оценка за броя на населението в днешния териториален обхват на България от различни автори за периода 1500 - 1900 година

Година	(Хиляди)			
	Тотев (1968, 1973)	Стоев (1969)	Петров (1980)	McEvedy; Jhones (1978)
1500	1179	1066	980	800
1550	1386
1600	1631	1436	1673	1250
1650	1990
1700	2257	1936	2362	1250
1750	2463
1800	2745	2610	3053	2000
1850	2984	2500
1875	3402	3000
1900	4315	3526	3744	4000

Източник: 1. Тотев, 1968, 1973; 2. Стоев, 1969; 3. Петров, 1980; 4. McEvedy; Jhones, 1978, p. 109.

²³ Тотев, Ан. (1968). Населението на България. - Годишник на Софийския университет, Юридически факултет, т. LIX, кн. 2, с. 3 - 45.

²⁴ Тотев, Ан. (1973). Демографски сведения за българите преди 1880 г. - Статистика, кн. 4, с. 25 - 47.

²⁵ Тотев, Ан. (1982). Ценен документ за историческата етническа демография на Балканския полуостров. - Исторически преглед, кн. 5, с. 105 - 113.

²⁶ Аркадиев, Д. (1988). Изменения в броя на населението и на домакинствата (семействата) по българските земи. - Население, кн. 4, с. 41 - 57.

²⁷ Стоев, П. (1969). Електронно пътешествие из нашата история. Някои данни за историческото минало на България, получени чрез икономически модел на развитие. - Младеч, бр. 1, с. 12 - 17.

²⁸ Петров, П. (1980). Основни периоди и закономерности в демографските промени на населението в НРБългария. - В кн.: Етнография на България, том I. С., БАН, с. 252 - 258.

²⁹ McEvedy, C., R. Jones (1978). Atlas of the World Population History. New York, Penguin Books, p. 109.

³⁰ Maddison, An. (2003). The World Economy Historical Statistics. Paris, OECD, pp. 95 - 96.

Обобщение и коментари върху написаното от български автори по темата до края на 80-те години на XX век прави Пантелей Зарев (1911 - 1997) в том от поредицата доклади на БАН със заглавие „България през социализма“³¹.

Повечето от изследванията на български историци от по-ново време докъм края на 80-те години на XX век са обобщени от Светлана Иванова³². След това и досега има изключително много разработки, които основно или по определен повод засягат броя и състава на населението за обсъжданата епоха.

Опит за определяне на броя на населението и неговите загуби в резултат на османското нашествие през XV век прави Христо Гандев (1907 - 1987) в книгата си „Българската народност през 15-и век“^{33, 34}. За целта той използва основно данните, публикувани от Н. Тодоров, за броя на домакинствата (хане) по българските земи към 1490 - 1491 година. Освен това се позовава на много други материали и публикации. Основната му идея се опира на виждането му, че има много запустели места (мезри), които според него са бивши населени места, разрушени по време на османското нашествие. Въз основа на много свидетелства стига до извода, че техният брой е 652 за около една четвърт от българските земи, или общо 2 608 селища. Установява, че за цялата наблюдавана територия средният брой домакинства в едно селище е 43. За всичките 2 608 селища това прави общо 112 144 домакинства. Приемайки, че домакинството има средно по петима членове, изчислява, че общият брой загубено българско население от тях е 560 720 души³⁵. Освен това той смята, че част от градското население също е изгубено. Затова го прибавя към горното число и получава общо загуби от 60-те години на XIV век до 80-те години на XV век от 680 720 души, или 136 144 домакинства³⁶. Общият брой на населението (немюсюлманско) по българските земи Гандев изчислява на 1 127 244 души³⁷. От това число той изважда 80 хил. небългарско население и получава 1 047 244 души³⁸. След редица други съображения той приспада още 157 086 лица и в крайна сметка получава общия брой на българската народност към 1490 г. - 890 155 души в 178 031 домакинства³⁹. Основният му извод е, че това съотношение - 890 000 към 680 000, или 100:75, не може да служи за изчисляване на броя на населението през XIV век⁴⁰.

Във второто издание на книгата си (1989 г.), отчитайки критиката по отношение на неговите изчисления, той намалява броя на населените места с изгубено население от 2 608 на 1 304 (получено от 652 умножено по 2). Тогава домакинствата в тези селища стават 55 072, а броят на изгубеното население - 280 360 души. Добавяйки загубите на градското население, се получават общите загуби - 363 360 души в 68 072 домакинства. За общия брой на населението и другите числа, свързани с него, не се прави промяна. Има добавка за оценка на броя на помюсюлманчените и техните потомци. Той смята, че към края на столетието те са общо към 50 000 домакинства, или 250 000 души, а за четири столетия - минимум 1 300 000 души.

³¹ Зарев, П. (1989). България през социализма. Доклади, т. 9. С., БАН, с. 45 - 65.

³² Иванова, Св. (1987) Етнодемографски изследвания за периода XV - XVIII в. в съвременната българска историография. В: България през XV - XVIII в. Историографски изследвания, т. 1. С., БАН, с. 155 - 170.

³³ Гандев, Хр. (1972). Българската народност през 15-и век. Демографско и етнографско изследване. С., Наука и изкуство.

³⁴ Гандев, Хр. (1989). Цит. съч., второ издание.

³⁵ Гандев, Хр. (1972). Цит. съч., с. 55 - 56.

³⁶ Гандев, Хр. (1972). Цит. съч., с. 100.

³⁷ Гандев, Хр. (1972). Цит. съч., с. 118.

³⁸ Гандев, Хр. (1972). Цит. съч., с. 121.

³⁹ Гандев, Хр. (1972). Цит. съч., с. 129.

⁴⁰ Гандев, Хр. (1972). Цит. съч., с. 131.

Освен някои от изчисленията спор предизвиква и броят на лицата в домакинството. Според Гандев средният брой на членовете в едно домакинство е не повече от пет души, даже с половин единица по-малко. Числото 5 той приема за „желязна норма”, която не може да се превишава⁴¹.

Срещу подобна категорична постановка биха могли да се направят някои принципици възражения. От една страна, по първичните описи са установени само родствените връзки баща - син или синове, но не са взети предвид зетьовете, а също съвместното живеене на двама или повече братя. От друга страна, пресилено е да се приеме, че домакинствата на вдовиците се състоят само от двама души, тъй като овдовяването обикновено е в по-късна възраст, когато вече се е реализирала възможността за раждане на деца. Така че би следвало тази категория да има средно по 4 лица. На трето място, неясно е положението с т.нар. неженени - ако те живеят самостоятелно, явно са глави на домакинства, в които и двамата родители са починали; ако е обратното, те са част от домакинствата на бащите си и по такъв начин се включват и увеличават броя на членовете в домакинството.

Същият автор привежда данни (като относителни дялове) за броя на мъжете в домакинството без децата, старците и недъгавите, които не са били обложени с данък⁴².

Въз основа на посочените данни е построена табл. 3.

3. Домакинства по брой на мъжете над 15 години по българските земи, 1490 - 1491 година

Мъже - брой	Относителен дял - %
1	47.5
2 - 3	52.0
4 - 5	0.5
Общо	100.0

Източник: Гандев, 1972, с. 181.

Според тези данни броят на мъжете над 15 години е достигал най-малко 1.5 и най-много 2.1 средно на домакинство, или средно 1.8. Според това средният брой на членовете в едно домакинство се е движел между 6 и 8. При отчитане на броя на вдовиците числото 5 може да се приеме като минимално, а не като максимално за броя на членовете в домакинството. Подобно становище се поддържа и от Е. Грозданова^{43, 44}. Затова в изданието от 1989 г. тази позиция на Гандев не е толкова категорична.

Изтъкнатите доводи навеждат на мисълта, че средният брой на членовете на домакинството може да бъде по-голям от сакралното число 5. Тогава оценките за броя на населението ще го променят към по-голям, отколкото се смяташе досега.

Гандев приема (както беше отбелязано), че общият брой на немюсюлманското население по българските земи е 1 127 244, или кръгло около 1 130 хил. души. При среден брой членове 5 това прави около 225 хил. домакинства. При по-голям среден брой членове на домакинство се получава следното: а) 6 бр. - 1 350 хил.; б) 7 бр. - 1 575 хил.; в) 8 бр. - 1 800 хиляди. Освен това трябва да се добавят и 10 - 20% или повече

⁴¹ Гандев, Хр. (1972). Цит. съч., с. 180 - 187.

⁴² Гандев, Хр. (1972). Цит. съч., с. 181.

⁴³ Грозданова, Е. (1972). За данъчната единица хане в демографските проучвания. - Исторически преглед, кн. 3, с. 81 - 91.

⁴⁴ Грозданова, Е. (1989). Българската народност през XVII век. Демографско изследване. С., Наука и изкуство, с. 60.

мюсюлманско население, част от което е придошло от Мала Азия, а друга - приела исляма и по този начин намалила местното немюсюлманско население.

При това положение разликата между броя на населението от средата на XIV и към края на XV век не изглежда толкова голяма.

Съществуват данни за общия брой на домакинствата, платци на данъка „джезие“, в балканските владения на Османската империя⁴⁵. Те са поместени в табл. 4.

4. Домакинства, платци на данък „джезие“, в балканските владения на Османската империя през XV век

Години	Абсолютен брой	Темп на растеж - %	
		верижна база	база 1488 г. = 100
1488	639119	100.0	100.0
1489	646550	101.2	101.2
1490	621508	96.1	97.2
1491	639387	102.9	100.0

Източник: Inalcik, Quataert, 1997, p. 28.

Данните се показателни - броят на домакинствата почти не се е изменял. Това означава, че има почти нулев или минимален растеж на немюсюлманското население, а може би и намаление. Възможните причини за това състояние са процесът на ислямизация, преместването на населението в други части на империята и т.н.

Освен това има известен брой домакинства, които плащат въпросния данък в намален размер. През 1488 г. те са 19 079, през 1489 г. - 34 902, през 1491 г. - 34 970. Тази категория домакинства са извън българските земи.

Както забелязва и самият Гандев, сравнението между ведомостите на описите за две съседни години (напр. 1490 и 1491 г.) дават нереални числа за смъртността, за новонамерените домакинства, за вдовиците и други⁴⁶.

Няма никаква гаранция, че регистраторът на домакинствата дава точни данни за техния брой. Не трябва да се забравя, че тези описи се правят за данъчно облагане и има лица в данъчната йерархия, които са заинтересовани от показване на по-малък брой от действителния. За да изглежда всичко наред, понякога този брой се коригира и това довежда до забележими отклонения. Регистраторите са далеч от разбиранията за раждаемост, смъртност и естествен прираст и други демографски категории. Това води до систематично подценяване на действителния брой на домакинствата и оттам - на населението.

Изводът, който се налага, е, че посочените данни от данъчното облагане не трябва да се абсолютизират при изчисляването на броя на населението и характеризирането на демографските процеси.

В тази светлина посочените загуби на население от османското нашествие е възможно да не са толкова големи (във всеки случай са по-малки).

Въпреки всичко трябва да се отдаде заслуженото на изследването на Гандев като първи опит за определяне на загубите на българския народ от тази епоха въз основа на изключително много документи, обработката на които е колосален труд.

⁴⁵ Inalcik, H., D. Quataert (1997). An Economic and Social History of the Ottoman Empire, Vol. I: 1300 - 1600. Cambridge, Cambridge University Press, p. 28.

⁴⁶ Гандев, Хр. (1972). Цит. съч., с. 119 - 129.

Положението през XVII век описва Елена Грозданова в книгата си „Българската народност през XVII век“⁴⁷. Изходният материал са сведения за 4 444 селища от българската етническа територия, от които 3 537 са идентифицирани и съществуват понастоящем. За съжаление, те не са равномерно представени по отделни години. Въз основа на обработката на огромен фактически материал тя стига до извода, че числеността на българската народност през XVII век достига около 2 млн. души, от които около 1.5 млн. християни и 0.5 млн. мюсюлмани. Също така прави много други комбинации от различни признаци на наличните данни за XVII век. Въз основа на поместения богат фактически материал е възможно да се направят допълнителни разработки за броя и състава на населението в днешния териториален обхват. Един от най-важните процеси, които анализира Е. Грозданова, е процесът на ислямизация на местното население през XVII век, започнал още през първите години на завоюването на българските земи и описан и от Гандев в посоченото произведение, като за 100 - 150 години този процес доста е напреднал. Въз основа на многобройния материал за населените места тя установява, че спрямо началото на века християнското население е намаляло до 30-те - 40-те години общо с 12.5%, между тях и 50-те - 90-те години на века - с 30.7%, а между тях и началото на века - общо с 39.4%⁴⁸.

Процесът на ислямизация на немюсюлманското население се потвърждава и от данните на О. Л. Баркан за началото на XVII век (1520 - 1535 г.) за домакинствата, обложени с данъка „авариз“ (извънреден данък)⁴⁹. Той се плаща от всички домакинства. На Балканския полуостров относителният дял на немюсюлманските домакинства е общо 77.8%, а на мюсюлманските - 22.2%. Делът на мюсюлманските домакинства не може да бъде обяснен само с броя на мигрантите от азиатската част на империята.

Николай Годоров е авторът, който отделя много място в своите произведения на демографското развитие на българските земи през тази епоха. Той проследява също и процеса на ислямизация. Посочва, че броят на преселниците от азиатската част на империята не е достатъчно голям, за да обясни големия относителен дял на мюсюлманското население на Балканите към първата четвърт на XVII век. Основната част от тях са юруци - около 40 хил., спахии, държавни чиновници и други категории - 40 - 50 хиляди⁵⁰. Това означава, че общо посочените категории не са превишавали 100 хил. души. Ако се приеме, че мюсюлманското население на Балканите към 1520 - 1535 г. наброява общо 1 млн. души, то тези преселници са 10% от неговия брой и само 2% от броя на цялото население (5 млн. души). С това той отговаря и на тезите на съвременни турски автори от 40-те и 50-те години на XX век, че османците са победили, защото в борбата на две цивилизации по-мощната се е наложила. Това е съчетавано с добре обмислена миграция на големи маси мюсюлманско население на Балканите. Посочените виждания, макар и в по-мек вариант, продължават да се поддържат и досега въпреки доказаната им несъстоятелност. Те се опровергават и от много други факти.

Славка Драганова събира и обнародва различни запазени в архивите описи за населението и неговия състав по различни признаци в Дунавския вилает^{51,52}. Направени

⁴⁷ Грозданова, Е. (1989). Цит. съч.

⁴⁸ Грозданова, Е. (1989). Цит. съч., с. 89 - 526.

⁴⁹ Barkan, O. L. (1957). Essai sur les données statistiques des registres de recensement dans L'Empire Ottoman aux XVe et XVIe siècles. - Journal of the Economic and Social History of the Orient, Vol. 1, No. 1 (Aug., 1957), pp. 9 - 36.

⁵⁰ Годоров, Н. (1972). Балканският град - XV - XIX век. С., Наука и изкуство, с. 44 - 49.

⁵¹ Драганова, Сл. (1980). Материали за Дунавския вилает. Русенска, Силистренска, Шуменска и Тутраканска каза, през 50-те - 70-те години на XIX в. С., БАН.

⁵² Драганова, Сл. (2005). Селското население на Дунавския вилает. С., Авангард-Прима.

са групировки на населението по възраст (обособени три възрастови групи), етнически състав и други признаци в редица населени места.

Мария Тодорова също прави въз основа на редица материали обобщения за семейството и домакинството и някои демографски процеси по българските земи от тази епоха⁵³. Освен на собствени проучвания се опира и на данни, публикувани от Сл. Драганова.

За съжаление, все пак в някои от тези разработки се забелязва известна едностранчивост. Една от тях е, че се обобщава българската етническа територия, а не се търси съвременната. Крайно време е да разберем, че ние разполагаме със 111 хил. км² и трябва да направим всичко възможно те да бъдат запазени. Важно е също така и че на тази територия съществуват и други етноси. Затова трябва да се проследява и тяхното развитие.

Един от авторите, който разработва част от настоящата тема, е Антон Минков (от български произход, но работещ в Канада) в книгата си „Приемане на исляма на Балканите“⁵⁴. Авторът обобщава всички данни, публикувани от други преди него, включително някои от цитираните в настоящия текст. Като цяло се разглежда процесът на ислямизация на части от населението на Балканите след завладяването им от османците и в следващите векове. В тази светлина процесът се илюстрира с наличните числови данни.

Има и много чужди автори, които разработват темата. По-специално, трябва да бъдат отделени тези, които предоставят различни данни, извлечени от първични документи, пазени в архивите, или обобщени от официални източници от епохата (напр. годишниците Салнаме). Такива са О. Л. Баркан, К. Карпат и други.

Баркан е първият автор, който публикува данни, извлечени от османските архиви, отразяващи данъчно обложените лица, вкл. по религиозен признак, в рамките на Османската империя в началото на XVI век. В тях са включени и области, влизащи в обхвата на днешната българска държава⁵⁵. Тези данни могат да служат като основа за оценка на общия брой на населението и състава му по вероизповедание. Те са коментирани и от Дж. Ръсел⁵⁶.

Много важно място трябва да се отдели на Кемал Карпат. Капиталният му труд „Населението на Османската империя (1830 - 1914)“ е издаден през 1985 г. в САЩ. През 2003 г. има ново издание в Турция^{57,58}, в което са включени някои допълнителни материали. В това произведение К. Карпат събира всички налични данни за преброяванията на населението на Османската империя през разглеждания период, които са били обнародвани в статистическите годишници на империята и на отделните вилаети (Салнаме). Освен това са отразени данни и оценки от това време и по-късно, публикувани в различни книги, статии, документи и др. от учени, пътешественици, дипломати, журналисти и т.н. За някои от тях вече беше споменато. В книгата се правят анализи и коментари на методологията за събиране на данните, на различни теми

⁵³ Todorova, M. (2006). *Balkan Family Structure and the European Pattern: Demographic Developments in Ottoman Bulgaria*. Budapest-New York, Central European University Press.

⁵⁴ Minkov, A. (2004). *Conversion to Islam in the Balkans: Kisve Bahas, Petitions and Ottoman Social Life, 1670 - 1730*. Leiden-Boston, BRILL, pp. 28 - 63.

⁵⁵ Barkan, O. L. (1957). *Op.cit.*

⁵⁶ Russell, J. C. (1960). *Late Medieval Balkan and Asia Minor Population*. - *Journal of the Economic and Social History of the Orient*, Vol. 3, No. 3 (Oct. 1960), pp. 265 - 274.

⁵⁷ Karpat, K. (1985). *Ottoman Population 1830 - 1914. Demographic and Social Characteristics*. Madison, University of Wisconsin Press.

⁵⁸ Karpat, K. (2003). *Osmanlı nüfusu (1830 - 1914). Demografik ve sosyal özellikleri*. İstanbul, Tarih Vakfı Yurt Yayınları.

относно броя и състава на населението по различни признаци и други. Безспорно най-ценната част са таблиците с данните. Въз основа на тях могат да се направят оценки за броя на населението в съвременните граници на България (а също и в етническите), за териториалното му разпределение, разпределение по етнорелигиозни категории и някои други. Етнорелигиозните категории започват да се поместват с данните от 1881 - 1882 г. нататък. Но все пак трябва да се подхожда по-внимателно при боравенето с всички данни, тъй като са възможни грешки, вкл. и печатни, в първоизточниците, а и съзнателни манипулации и нагласявания.

Ценни в това отношение са и други публикации. Една от тях е „Етнография на виалетите Одрин, Манастир и Солун“ от 1878 г., която е извадка от обнародвани материали във френския вестник „Le Courier d'Orient“, излизал в столицата на Османската империя през втората половина на XIX век⁵⁹. В нея се дават данни по отделни административни единици от различен ред и по населени места. В тези структури са поместени данни за броя на къщите, броя на населението по етно-религиозни категории - мюсюлмани, българи, гърци, цигани и др. (възможно е да е даден само броят на мъжете). За да се използва тази информация по населени места, те трябва да бъдат идентифицирани, тъй като наименованията са от това време. Дадена е и обобщена информация за всеки вилает поотделно по етнорелигиозни показатели. Съставителите на публикацията посочват, че тя е в отговор на гръцките претенции към цяла Македония. Дори един бегъл поглед върху названията на населените места (предимно села) в разглежданите вилаети (Солун и Манастир) недвусмислено показва, че те са български.

Подобна е публикацията на Пол Буте „Османската империя. Статистически документи“ от 1877 г.⁶⁰, но тя е с по-малко подробности⁶⁰.

Оценки като цяло за Османската империя и в някои случаи и за българите и българските земи прави и Енгин Акарлъ в произведението си „Османско население в Европа през 19-и век: неговият териториален, расов и религиозен състав“⁶¹.

Авторът се опира на многобройни извори, изразяващи се предимно в публикации в годишниците Салнаме, разработки на учени, дипломати, журналисти, посланици, политици и др., съвременници на епохата и други от по-ново време.

За оценките в първите три десетилетия на XIX век, когато отсъстват официални данни за броя на населението, авторът прилага статистическия подход. В случая той се изразява в определяне на основните средни величини - средна аритметична, медиана и мода от оценките на населението, публикувани от различни автори (нещо като експертни оценки). След това той взема средния аритметичен брой, обикновено изразен в хиляди. Нагажда го чрез закръгляване на последната от цифрите на хилядите на 0 или 5.

Този подход се прилага и по-късно, когато има налице официални данни - изцяло или частично запазени. Това крие опасност от преувеличаване или намаляване на действителния брой на населението и изопачаване на структурата му по някои признаци (етнически и религиозен състав). Няма пълна гаранция, че авторите са напълно коректни при своите оценки, както и за сигурността на техните източници на данни и т.н. Това се отнася най-много за оценките на етническия и религиозния състав, тъй като националистически настроени автори обикновено преувеличават

⁵⁹ Ethnographie des vilayets d'Andrinople, de Monastir et de Salonique (1878). Extrait du Courier d'Orient. Constantinople.

⁶⁰ Boutet, P. (1877). La population de la Turquie. L'Exploration, Tome 1, p.40 - 46, Paris.

⁶¹ Akarli, E. D. (1972). Ottoman population in Europe in the 19th century: its territorial, racial, and religious composition. Madison, University of Wisconsin.

действителния брой на даден етнос или религиозна група. В други случаи е възможно той да бъде подценен поради недостиг на източници.

Най-ясно такова несъответствие се очертава в обобщаващите таблици за етническия и религиозния състав на населението в европейската част на Османската империя. Оценкаите са отнесени към 1820, 1840, 1870 и 1890 година.

Вероятно оценкаите за броя на населението върху съвременната територия на България, направени от А. Мадисън, се опират на тези данни.

В таблица XXXIII са оценкаите за етническия състав⁶². От нея следва, че броят на българите нараства от 615 хил. към 1820 г. на 3 000 хил. към 1840 г. до 3 451 хил. към 1870 година. Съответните относителни дялове от общото население са 6.0, 19.2 и 34.0%.

По този повод Х. Иналчък отбелязва: „От края на 18-и век светът вече не знаеше за нация, наречена българи ... някои от тях се смятат за турцизирани, а други за гърцизирани... но през втората половина на (19) век целият свят разпознава българите като една от най-многобройните нации на Балканите”⁶³.

Същевременно броят на турците е: 1820 г. - 2 755 хил., 1840 г. - 2 116 хил., 1870 г. - 1 827 хиляди. Съответните относителни дялове са: 1820 г. - 27.0%, 1840 г. - 13.7%, 1870 г. - 18.0%.

Посоченият брой на българите към 1820 г. (615 хил. души) трябва да се възприема като израз на недостатъчното познаване на действителното положение. Такъв бурен растеж на българското население (почти 5 пъти) между 1820 и 1840 г. (само за 20 години) е просто невъзможен. Същото се отнася и за „намалването” на абсолютния брой и относителния дял на турското население в европейските владения на империята.

За съжаление, Е. Акарлъ не прави оценка за съвременната територия на България. Посочени са само данни за Княжество България и Източна Румелия към 1877 година. За Княжеството броят на населението е 1 874 638, а за автономната провинция - 613 822⁶⁴.

Развитието на населението по това време в балканските страни проследява и Майкъл Паларе в книгата си „Балканските икономики 1800 - 1914. Еволюция без развитие”⁶⁵. Първата глава е посветена специално на населението. В нея се прави критичен анализ на източниците на данни за броя на населението в Османската империя и в частност на владенията ѝ на Балканите. Отбелязва се, че има сериозни пропуски в знанията ни относно демографското състояние на отделните провинции.

Единствените оценки на населението Майкъл Паларе дава за територията на Княжество България и Източна Румелия. Те се отнасят за 1870 година. За Княжеството броят на населението е 2 022 хил., а за автономната провинция - 747 хил. души⁶⁶. Данните за следващите години са от официални преброявания и регистрации.

Въз основа на приведените от първоизточниците данни от някои автори е възможно да бъде направен опит за определяне на броя на населението на българските земи в сегашния им териториален обхват.

⁶² Akarli, E. D. (1972). *Op.cit.*, p. 114.

⁶³ Inalcık, H. (1992). *Tanzimat ve Bulgar Meselesi. Le Tanzimat et la Question Bulgare (Thèse de doctorat)*. Ankara, Türk Tarih Kurumu Bastmevi, 1943. In: Halil Inalcık. *Tanzimat ve Bulgar Meselesi. (Doktora Tezi'nin 50. Yılı, s. 17), 1942 - 1992. Istanbul, EREN.*

⁶⁴ Akarli, E. D. (1972). *Op.cit.*, p. 98.

⁶⁵ Palairat, M. (2002). *The Balkan Economies, c. 1800 - 1914. Evolution without development*. Cambridge, Cambridge University Press. Има и българско издание - Паларе, Майкъл (2005). *Балканските икономики 1800 - 1914: еволюция без развитие*. Превод от англ. С., Апострофи.

⁶⁶ Palairat, M. (2002). *Op.cit.*, p. 20.

Данните от първото преброяване на населението в Османската империя се отнасят за 1831 година. Те се поместват от К. Карпат, Е. Акарлъ и др., за които беше споменато.

За съжаление, те са дадени по тогавашното административно деление, което към втората половина на XIX век е променено. Привеждането им към по-късното административно устройство е направено от Н. Тодоров⁶⁷. Вилаетите, които покриват съвременната българска територия, имат следното мъжко население: Дунавски - 477 862; Одрински - 421 721; Солунски - 240 411, или общо за трите - 1 139 994.

⁶⁷ Тодоров, Н. (1972). Цит. съч., с. 298.

Следвайки публикацията на К. Карпат^{68,69}, определяме следния брой на мъжкото население за българските земи в тогавашния момент, покривани от еялетите Румелия и Силистра. За каазите в Румелия се получават 362 469, а за тези в Силистра - 134 275, или общо 496 744 души. Етнорелигиозният състав може да се проследи в табл. 5 и е графично изобразен на фиг. 1.

**5. Етнорелигиозен състав на населението върху
съвременната територия на България според
преброяването в Османската империя към 1831 година**

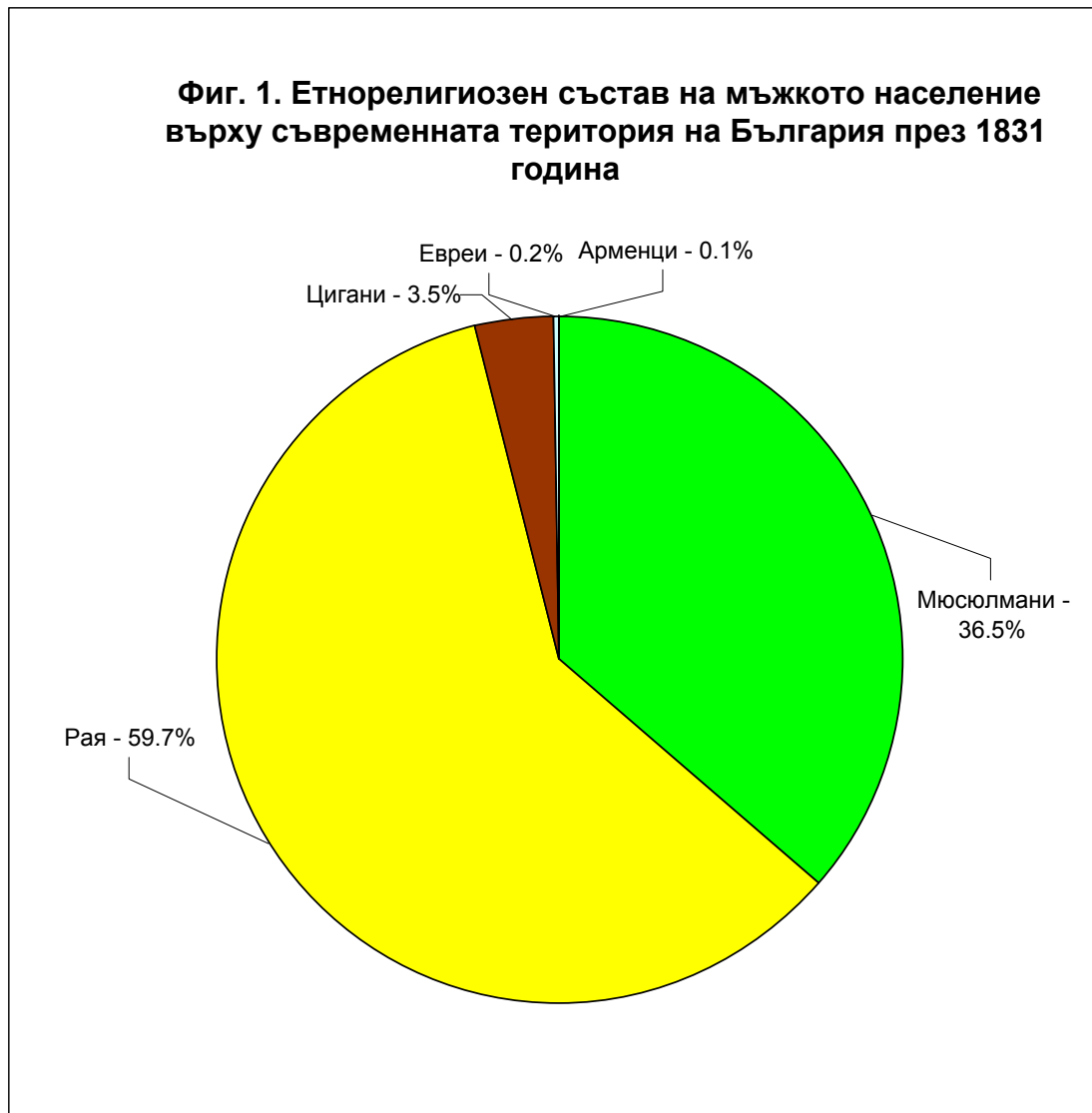
Етнорелигиозни категории	Абсолютен брой	Относителен дял - %
Мюсюлмани	181455	36.5
Рая	296769	59.7
Цигани	17474	3.5
Евреи	702	0.2
Арменци	344	0.1
Общо	496744	100.0

Източник: Карпат, 1984, pp. 109 - 110; 2003, s.149 - 150.

⁶⁸ Карпат, К. (1985). Op.cit., pp. 109 - 110.

⁶⁹ Карпат, К. (2003). Bkz. s. 149 - 150.

Фиг. 1. Етнорелигиозен състав на мъжкото население върху съвременната територия на България през 1831 година



Веднага прави впечатление твърде малкият брой на населението. Това е знак, че този брой подценява действителния. Той отразява само част от мъжкото население. Първо, това са лица на възраст над 15 години. Второ, от тях трябва да се приспаднат недъгави, болни и други подобни категории, освободени от данъци, а при мюсюлманите това са военноподължените. Категориите са добре описани от К. Карпат⁷⁰. Той не посочва какъв е техният дял. Според Н. Тодоров тези категории са между 20 и 40%⁷¹.

Дори и тази поправка не променя много резултата. Данни за възрастовия състав на населението не са публикувани, но такива са извлечени от първични записи за села в части от Дунавския вилает от Сл. Драганова за 1872 година⁷².

Обобщени са резултати за 16 107 лица в 69 населени места в три възрастови групи мъжко население. Общо се установява, че относителният дял на лицата от 0 до 14 години е 43.0%, от 15 до 60 години - 49.5%, и над 60 години - 7.5%. Мъжете са разпределени и по народност - българи, турци, цигани, други (молдовани, татари, черкези). Най-младо се оказва българското население.

⁷⁰ Карпат, К. (1985). *Op.cit.*, pp. 20 - 21.

⁷¹ Тодоров, Н. (1972). *Цит. съч.*, с. 296 - 297.

⁷² Драганова, Сл. (1980). *Цит. съч.*, с. 328 - 438, табл. 327.

Ако приемем, че това са лицата между 15 и 60 години, то мъжкото население ще бъде: $(496\,744 : 49.5) \times 100 = 1\,003\,523$, или коефициентът за умножение трябва да бъде $(100 : 49.5) = 2.02$. Ако се добави и броят на жените, тогава общият брой на населението е около 2 млн. души. Но и това число предизвиква съмнение за подценяване на действителния брой.

Вероятно обнародваните резултати за преброяването през 1831 г. се отнасят само за броя на облагаемото мъжко население, чийто относителен дял е по-малък от използвания.

В потвърждение на казаното можем да цитираме текст от вестник „Дунав“, бр. 914 от 18 октомври 1874 г., издаван в гр. Русе - център на Дунавския вилает, по повод започналото преброяване на населението през 1866 година. Въпросният текст е (по съвременната литературна норма): „Ако ограничим върху нашия вилает думите, които изказахме тука общо, ще кажем, че преди тридесет години до деня на въвеждането на тази нова система за преброяване, понеже не беше се направило едно точно преброяване на жителите, тия се брояха предположително. Следователно не можеше да се повярва, че истинското число на жителите в Дунавската област е до тази степен, която се доказва в сегашното преброяване, а особено някои от разните народи се представяха с преувеличени цифри, а някои се представяха по-малко, отколкото са, щото не можеше да се узнае истинското им число“⁷³.

Затова други автори използват по-голям коефициент - 5.905. Един от тях е Брус МакГоуан в книгата си „Икономическият живот в Османска Европа“, където той подробно коментира част от засегнатите въпроси и стига до подобни изводи⁷⁴. Към края на XVII век (1691 г.) Османската империя променя данъчната си система. Преминва се от данъчно облагане на домакинствата към индивидуално облагане на мъжете от възрастното население (15 и повече години).

Обобщена таблица за броя на индивидуалните платци на данъка „джезие“ помества М. Палайрет⁷⁵. Тя има следния вид (табл. 6).

На табл. 4 и 6 се вижда, че всъщност в края на XVII век броят на индивидуалните данъкоплатци почти съвпада с броя на домакинствата за същия данък от края на XV век. Това означава, че до началото на XVIII век процесът на ислямизация на немюсюлманското население (наред с други процеси) забавя неговия растеж. Едва през XVIII и началото на XIX век броят му започва осезаемо да нараства.

6. Данъкоплатци на джезие и оценка на населението в Османска Европа по години

Години	Данъкоплатци на джезие - бр.	Оценка на населението - хил. бр.
1700	635835	3755
1718	683316	4035
1740	756949	4470
1788	868648	5130
1815	932322	5506

Източник: Palairret, 2002, p. 6, tab. 1.1.

⁷³ В. „Дунав“ (1874). Бр. 914.

⁷⁴ McGowan, Br. (2010). *Economic Life in Ottoman Europe. Taxation, Trade and the Struggle for Land, 1600 - 1800*. New York, Cambridge University Press, pp. 80 - 114.

⁷⁵ Palairret, M. (2002). *Op.cit.*, p. 6, tab. 1.1.

За оценка на броя на населението се прилага коефициент (5.905), с който се умножава броят на данъкоплатците.

Ако този коефициент се приложи към данните за мъжкото население през 1831 г., за съвременната територия на България ще се получи следното: $(496\,744 \times 5.905) = 2\,933\,273$. Население от близо 3 млн. души е доста по-приемливо число за тогавашните условия.

До 1878 г. най-често се дават данни за населението на Дунавския вилает, но с различна степен на подробност.

За целта са взети данните, обнародвани от К. Карпат^{76,77}. Данните се разпростират върху общата територия на вилаета и територията, върху която е създадено Княжество България (без санджаците Тулча и Ниш). Резултатите могат да се проследят в табл. 7.

7. Дунавски вилает и територия на Княжество България по брой на мъжкото население, 1868 - 1877/1878 година

Години	Дунавски вилает - общо	Княжество България	
		абсолютен брой	относителен дял - %
1868	1023309	811312	79.3
1869	1033740*	810624*	78.4
1874	996039**	887994	89.2
1877 - 1878	1247954	947329	75.9

Забележка. * - Броят на населението на Варненския санджак вероятно е сгрешен. В сравнение с 1868 г. само той намалява.

** - Не включва Нишкия санджак.

Източник: Karpat, 1984, pp. 116 - 118; 2003, s. 156 - 158.

За съвременната територия на България най-голям интерес представляват данните, обнародвани във връзка с Цариградската конференция през 1876 г. за решаване на българския въпрос. Те са публикувани изцяло и коментирани от Ан. Тотев⁷⁸. Съдържат данни за броя на мъжкото население общо и по религиозен признак, по вилаети, санджаци и каази към 1875 година. Териториалният обхват съвпада до голяма степен с този за българската етническа територия. Те също имат някои забележими недостатъци - не са налице данни по каази за Нишкия, Сереския и Скопския санджак.

Разпределихме териториалните единици приблизително върху съвременната територия на България. За каазите от Серския санджак (Мелник, Неврокоп, Петрич), попадащи в този обхват, бяха използвани данни от публикацията на вестник „Куриер де Ориент“⁷⁹. Резултатите са поместени в табл. 8.

⁷⁶ Karpat, K. (1985). Op.cit., pp. 116 - 118.

⁷⁷ Karpat, K. (2003). Op.cit., pp. 156 - 158.

⁷⁸ Тотев, Ан. (1982). Цит. съч.

⁷⁹ Ethnographie... (1878), pp. 29 - 34, 37 - 39, 41 - 42, 65.

8. Мъжко население на съвременната територия на България през 1875 г. по вилаети на Османската империя

Вилаети	Немюсюлмани	Мюсюлмани	(Брой)
			Общо
Дунавски	712842	451680	1164522
Одрински	293546	194263	487809
Солунски	46999	42055	89054
Общо	1053387	687998	1741385

Източник: Тотев, 1982; Ethnographie...(1878), pp. 29 - 34, 37 - 39, 41 - 42, 65.

Относителният дял на немюсюлманското мъжко население е: Дунавски вилает - 61.2%, Одрински - 60.2%, Солунски - 52.8%, и общо - 60.5%. Почти цялото немюсюлманско население се състои от лица, принадлежащи към българския етнос. При мюсюлманското население значителна част също принадлежи към българския етнос.

За територията на бъдещото Княжество България (частта от Дунавския вилает) съответните данни са следните: немюсюлмани - 712 842, или 61.2%; мюсюлмани - 451 680, или 38.8%; общо - 1 164 522.

Тук е мястото да се определи и общият брой на населението. Повечето автори умножават броя на мъжете по 2 и го получават. Но действителното съотношение между мъже и жени не е 50:50%. По данни на Н. Тодоров за 16 града на Дунавския вилает през 1866 г. на 1 000 мъже се падат по 956 жени⁸⁰. С други думи, жените са 95.6% от броя на мъжете. При това съотношение общият брой на населението е 3 406 149, от които 1 741 385 мъже и 1 664 764 жени. За бъдещото Княжество България това прави 2 277 805 (мъже - 1 164 522, жени - 1 113 283). Общият брой почти съвпада с оценката на Ан. Тотев - 3 402 066⁸¹.

Числото 3.4 млн. приемаме за приблизителния общ брой на населението върху територията на България към 1875 година. Това число се съгласува с резултатите от първите преброявания на населението в България и събитията, повлияли върху него около Руско-турската война от 1877 - 1878 г. (миграции, жертви сред мирното население и др.).

Ако се приеме относителният дял на жените за горното изчисление, то за територията на бъдещото Княжество България от табл. 7 общият брой на населението ще бъде: 1868 г. - 1 586 926; 1869 г. - 1 585 580; 1874 г. - 1 736 916; 1877 - 1878 г. - 1 852 976.

Всичко това подсказва систематично подценяване на действителния брой на населението в Османската империя през XIX век. Има нагаждане, неспазване на единна методология и начин на събиране на данните и публикуване на резултатите, грешки и т.н. Империята изостава от останалите водещи европейски страни и САЩ и в това отношение. Първото преброяване от модерен тип е проведено в САЩ през 1790 г., във Франция и Великобритания - през 1801 г., и т.н. Така че изоставането тук е с 30 - 40 години (1831 г.), а по методология - с над 100 години.

Изглежда най-близо до действителността са оценките за броя на населението на Ан. Тотев, поместени в табл. 2. Графичното им изображение може да се види на фиг. 2.

⁸⁰ Тодоров, Н. (1972). Цит. съч., с. 343.

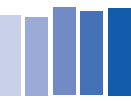
⁸¹ Тотев, Ан. (1973). Цит. съч.



За бъдещото Княжество България е по-трудно да се представи етническата принадлежност на населението по няколко причини, които ще бъдат изяснени по-нататък. През 1866 г. започва преброяване на населението в Дунавския вилает, което завършва през 1873 година. При него се наблюдава религиозната и етническата принадлежност на населението. Общо жителите на вилаета са разделени на мюсюлмани и християни. Всяка една от тези религиозни категории е разделена от своя страна на етнически. Категорията „мюсюлмани“ включва (например за Тулча): стари мюсюлмани, татари, ногайци, черкези, хаджеми (общо название на талиши и мазандеренци, живеещи предимно в Иран около Каспийско море), арнаути, бошнаци, кюрди, мюсюлмански и татарски цигани. Категорията „християни“ включва: гърци, българи, българи гагаузи, българи шопи, българи арнаути, молдовани, арменци, арменци католици, италианци, французи, англичани, унгарци, власи, сърби, черногорци, казаци, казаци зебури, немци, липовани, липовани безпович, нимулаци, провеслави, липовани съботници и същи липовани. Отделно са посочени евреите: американски евреи, турски евреи и цигански евреи⁸².

Данните от въпросното преброяване са представени в обобщаваща таблица, но етнорелигиозните категории в нея са уедрени. Те могат да се видят в табл. 9 в малко изменено представяне в сравнение с оригинала (тук са поправени някои грешки, допуснати в него).

⁸² В. „Дунав“ (1974). Цит. съч.



9. Мъжко население на Дунавския вилает по етнорелигиозни категории, 1866 - 1873 година

Етнорелигиозни категории	Абсолютен брой	Относителен дял - %
Мюсюлмани	481798	42.2
- Стари жители	392369	34.4
- Преселници	64398	5.6
- Мюсюлмански цигани	25031	2.2
Християни	646215	56.6
- Българи	592573	51.9
- Гърци	7655	0.7
- Арменци	2128	0.2
- Католици	3556	0.3
- Други	40303	3.5
Евреи	5375	0.5
Немюсюлмански цигани	7663	0.7
Общо	1141051	100.0

Източник: В. „Дунав“, 1874.

Ако прибавим и броя на жените, общият брой на населението е 2 231 895. Но този брой включва Тулчанския санджак (а може би и Нишкия).

Според А. Юбичини населението на цялото Княжество преди войната според Салнаме 1294 година (1877 - 1878) е 1 914 638⁸³. За отделните санджаци то е представено в табл. 10.

10. Население на територията на бъдещото Княжество България през 1877 - 1878 г. по санджаци

Санджаци	Абсолютен брой	Относителен дял - %
Русчук	649594	33.9
Варна	171610	9.0
Видин	227646	11.9
Търново	491788	25.7
София	374000	19.5
Общо	1914638	100.0

Източник: Ubicini, A, 1879, pp. 81 - 100.

Станислас Г. Б. Сен Клер (1835 - 1887) в брошурата си „Български ужаси и източната политика на мистър Гладстоун“ обосновава част от тезите си с данни за етно-религиозния състав на населението във вилаета Русчук (Дунавския вилает)⁸⁴.

⁸³ Ubicini, A. (1879). La principauté de Bulgarie. Revue de Géographie, tome 5, Paris, pp. 81 - 100.

⁸⁴ Clair, S. G. B. St. (1876). Bulgarian Horrors! and Mr. Gladstone's Eastern Policy. London, Blanchard & Sons, p. 19.

Авторът е изключително антибългарски настроен и затова тези данни са поместени без промени⁸⁵. Добавени са само относителните дялове на отделните категории и са представени в табл. 11.

11. „Това, което той нарича България, е вилаетът Русчук, в който ние намираме (1875 - 1876) мъжко население от:”

	Абсолютен брой	Относителен дял - %
Истински турци	457018	35.9
Други мюсюлмани	104639	8.2
Мюсюлмани (мъже)	561647	44.1
Арменци	2128	0.2
Власи и гърци	56647	4.4
Цигани	8220	0.6
Евреи	5847	0.5
Небългари	634469	49.8
Срещу т.нар. българско население	639813	50.2
Общо	1274282	100.0

Източник: Clair, S. G. В. St., 1876, p. 19.

Между двете групи данни има съвпадения и различия. Съвпада напълно броят на арменците и донякъде този на евреите и немюсюлманските цигани. Преувеличава се относителният дял на мюсюлманите (вкл. на „истинските турци”), а се подценява този на българите и други.

За бъдещата Източна Румелия етническият състав също може да бъде определен приблизително. Голяма помощ в това отношение представлява книгата на Робърт Дж. Мор (Robert Jasper More) (1836 - 1903) „По Балканите. Бележки за посещение в окръг Филиппополис през 1876 г.”⁸⁶. Там се дават данни за мъжкото население на Пловдивския санджак по каази и етнорелигиозни категории. Също така, но не толкова подробно, това се прави и за Сливенския санджак. От двата санджака по-късно се създава Източна Румелия. Годината, за която се отнасят данните, не е спомената, но броят на населението за Пловдивския санджак съвпада с данните на Ан. Тотев за 1875 година. Затова приемаме, че източникът им е един и същ. Данните за Сливенския санджак се различават с 11 хил. ($116\,579 - 105\,601 = 10\,978$), но повече подробности не са известни. Вероятно тук е включена някаква друга териториална единица. Но в случая подобна неточност е неизбежна.

От данните за Пловдивския санджак е необходимо да се приспадат тези за каазите Ахъчелеби и Султанери. В крайна сметка се получава поместеното в табл. 12.

Общият брой на населението на бъдещата Източна Румелия към 1875 г. ще бъде 833 152.

⁸⁵ Това е британски агент, известен у нас под името Сенклер. Той е организатор на метежа на мюсюлманското население в Родопите след края на Руско-турската война 1877 - 1878 година. Заедно с Чарлс Брофи издава книгата „Пребиваване в България. Бележки по ресурсите и администрацията на Турция“, Лондон, 1869 година. В нея двамата автори пишат отвратителни неща за българския народ и неговото битие.

⁸⁶ More, R. J. (1877). Under the Balkans. Notes of a visit to the district of Philippopolis in 1876. London, Henry S. King & Co., pp. 16 - 18.

И в двата случая - бъдещото Княжество България и бъдещата автономна област Източна Румелия, се установява, че българското християнско население е повече от половината от общото население (52 - 60%). Дори по явно манипулираните данни на С. Сен Клер то все пак е над 50%.

12. Мъжко население на бъдещата Източна Румелия по санджаци и етнорелигиозни категории през 1875 година

Етнорелигиозни категории	Абсолютен брой	Относителен дял - %
Пловдивски санджак		
Мюсюлмани	117662	36.7
-- Турци	105727	33.0
-- Цигани	11935	3.7
Немюсюлмани	202684	63.3
-Християни	200625	62.6
-- Българи	194727	60.8
-- Цигани	1421	0.4
-- Гърци	4000	1.2
-- Арменци	477	0.2
-Евреи	2059	0.7
Общо	320346	x
Сливенски санджак		
Мюсюлмани	44747	42.4
Немюсюлмани	60854	57.6
от тях българи	60000	56.8
Общо	105601	100.0
Бъдеща Източна Румелия		
Мюсюлмани	162409	38.1
Немюсюлмани	263538	61.9
от тях българи	254727	59.8
Всичко	425947	100.0

Източник: Тотев, 1982; More, 1877, pp. 16 - 18.

Ако се приеме, че населението на условната територия, върху която се разполага съвременна България, към 1500 г. е 1.2 млн., а към 1875 г. - 3.4 млн., то това означава, че за този период то се е увеличило 2.8 - 3 пъти.

По същото време населението на Англия (част от Великобритания) се оценява съответно на 2.8 млн. и на 24.5 млн. души. Тук увеличението е над 8 пъти.

Въз основа на изложеното дотук могат да се направят следните обобщения:

1. В епохата на османското завоюване на българските земи броят на населението на условната територия, върху която днес е разположена българската държава, е намалял с повече от 1 млн. души, или повече от два пъти.

2. Причините за намаляването на броя на населението са комплексни:
а) непосредствено физическо унищожение в резултат на военните действия;
б) отвеждане на част от населението на завладените територии като военна плячка;
в) миграции на местното население - към незавладени от османците територии (например към днешната територия на Румъния и други страни) и преднамерено

извършвани от османската държава към Мала Азия; г) наличие на чумни епидемии през тази епоха; д) загуби на местно население от участие във военни действия на страната на османците; е) дългосрочни промени в природните условия на някои райони и други.

3. Причина за относителното намаляване на броя на българското християнско население е процесът на ислямизация, извършван с различни методи доброволно или насилствено. Този процес протича с различна интензивност през целия период на османското владичество. Особено силно е изразен през XVII век.

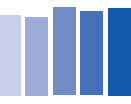
4. Като цяло населението на съвременната територия на България между 1500 и 1875 г. се е увеличило около три пъти - от 1.2 млн. до 3.4 млн. души.

5. Не е възможно да се направят много точни оценки за броя на населението и неговите структури по различни признаци през изучавания период от време.

6. Основните причини за недостатъчно точните оценки е наличната изворова база. До 30-те години на XIX век тези оценки са на базата на описи, използвани за данъчно облагане, с всичките им недостатъци. От 30-те години на XIX век се основават на започналите преброявания на населението, които в методологично и организационно отношение изостават от състоянието в напредналите страни.

7. Анализът на данните от преброяванията показва, че има систематично недооценяване на броя на населението и неговите структури по различни признаци.

8. Въпреки всички недостатъци този период е несравнимо по-благоприятен за оценки за броя на населението и неговите структури в сравнение със Средновековието и Античността.



ИЗМЕНЕНИЯ В БРОЯ НА НАСЕЛЕНИЕТО ПО БЪЛГАРСКИТЕ ЗЕМИ В СЪСТАВА НА ОСМАНСКАТА ИМПЕРИЯ

*Димитър Аркадиев**

РЕЗЮМЕ В статията се обсъжда растежът на населението и изменението на някои негови структури приблизително в съвременните граници на България през периода 1500 - 1878 година. По това време тези територии са в рамките на Османската империя. В сравнение със средата на XIV век населението намалява с повече от 1 млн. души, или повече от два пъти. Основните причини за промените в броя на населението са: пряко физическо унищожение в резултат на османското нашествие, отвлечане като военна плячка, миграции, чумни епидемии и други. Промени настъпват и в религиозната и етническата структура на населението. Главната причина за това е процесът на ислямизация - доброволна и насилствена.

Направените оценки са на базата на данни от данъчните регистри на домакинствата и преброяванията на населението, започнали от началото на 30-те години на XIX век. Всеки от тези източници има определени недостатъци. Констатира се систематично подценяване на действителния брой на населението и изопачаване на някои негови структури. В методологично и организационно отношение преброяванията в Османската империя изостават с десетилетия от напредналите европейски страни и САЩ. Установява се, че населението нараства от 1.2 млн. през 1500 г. до 3.4 млн. души през 1875 г., или около три пъти.

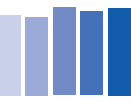
* Д-р, професор в Стопанския факултет на Великотърновския университет „Св.св. Кирил и Методий“;
e-mail: arkadiev@abv.bg.

ИЗМЕНЕНИЯ В ЧИСЛЕННОСТИ НАСЕЛЕНИЯ В БОЛГАРСКИХ ЗЕМЛЯХ В СОСТАВЕ ОСМАНСКОЙ ИМПЕРИИ

*Димитр Аркадиев**

РЕЗЮМЕ Обсуждается рост населения и изменение некоторых его структур приблизительно в современных границах Болгарии в период с 1500 по 1878 год. Тогда эти территории находились в рамках Османской империи. По сравнению со серединой XIV-ого столетия население уменьшилось больше чем на миллион, т.е. больше чем в два раза. Приводятся причины изменений в численности населения. Основными среди них являются: прямое физическое уничтожение в результате османского нашествия, похищение в качестве военной добычи, миграции, эпидемии чумы и другие. Основные изменения происходят в религиозной и этнической структуре населения. Главной причиной этого является процесс исламизации - добровольной и по принуждению. Оценки основываются главным образом на данных с налоговых регистров домашних хозяйств и переписей населения, начавшихся с начала тридцатых годов XIX-ого столетия. Каждый из упомянутых источников имеет определенные недостатки. Замечается систематическое недооценивание действительной численности населения и искажение некоторых его структур. В методологическом и организационном отношении переписи в Османской империи отстают на десятилетия от передовых стран Европы и США. Установлено, что население увеличивается с 1.2 миллиона в 1500 году на 3.4 миллиона в 1875 году, т.е. приблизительно в три раза.

* Д-р, профессор в ВТУ „Св. св. Кирилл и Мефодий” - г. Велико Тырново; электронная почта: arkadiev@abv.bg.



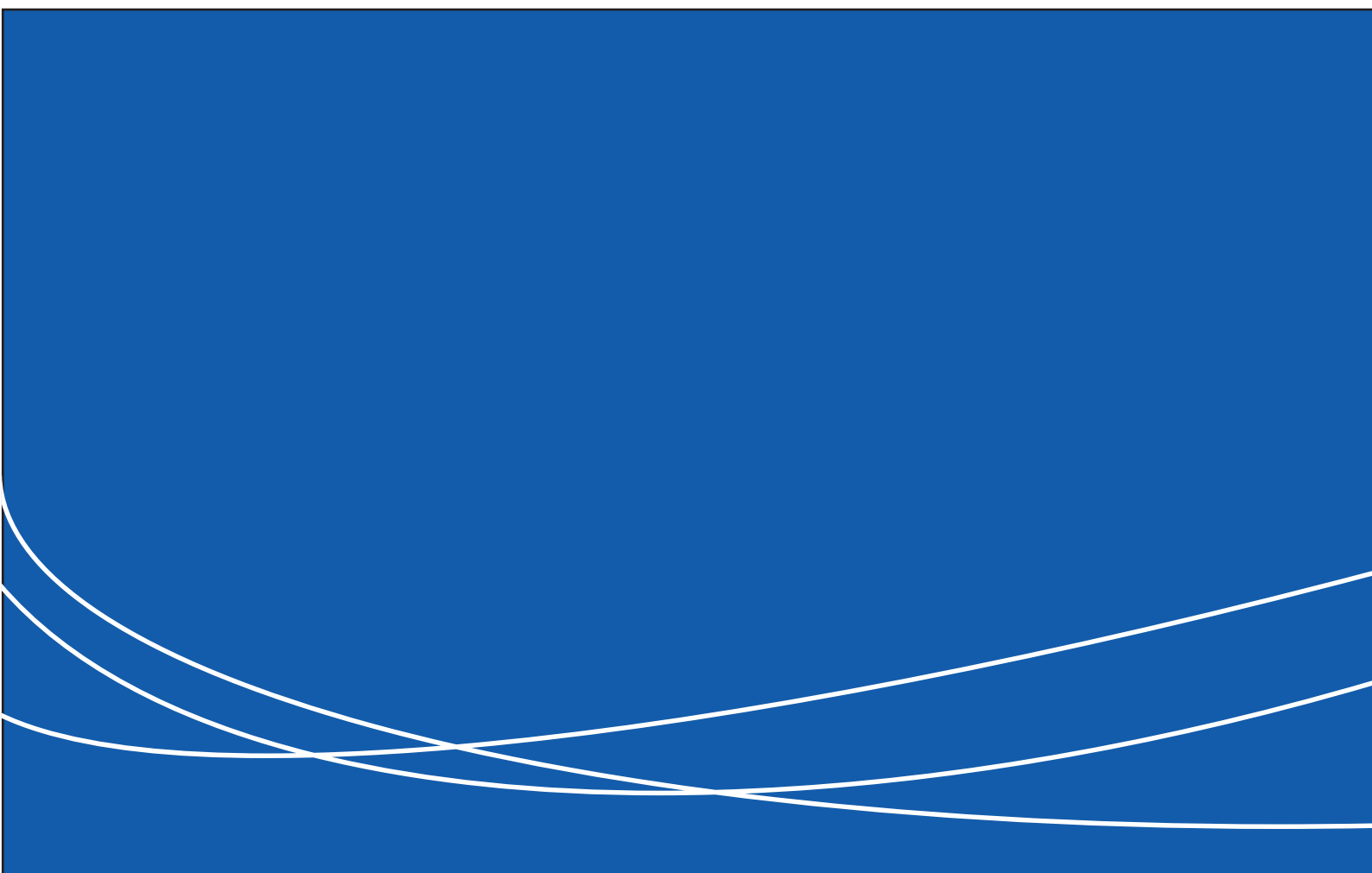
CHANGES IN THE NUMBER OF POPULATION IN BULGARIAN LANDS OF THE OTTOMAN EMPIRE

*Dimitar Arkadiev**

SUMMARY The population growth and the changes in some of its structures approximately in modern borders of Bulgaria during the period 1500-1878 are discussed. At that time these territories are part of the Ottoman Empire. Compared to the mid-14th century the population decreased by more than 1 million, or more than twice. The reasons for changes in population numbers are reflected. The main ones are: direct physical extermination as a result of the Ottoman invasion, kidnapping as a military prey, migrations, plague epidemics and others. Major changes occur in the religious and ethnic structure of the population. The main reason for that is the process of Islamization - voluntary and forceful. The assessments are made mainly of data from tax records of households and population censuses since the beginning of the 30s of the 19th century. Each of these sources has certain disadvantages. Systematical underestimation the actual number of population and distortion of some of its structures was found out. In methodological and organizational terms the censuses in the Ottoman Empire fall decades behind from the advanced European countries and the USA. It was found that the population grew from 1.2 million in 1500 to 3.4 million in 1875, or about 3 times.

Dr., Professor in the 'St. St. Cyril and Methodius' University - Veliko Tarnovo; e-mail: arkadiev@abv.bg.

**ИНФОРМАЦИИ, РЕЦЕНЗИИ,
КОНСУЛТАЦИИ**



ИНФОРМАЦИЯ ЗА ПРОЕКТ „ПРИЛАГАНЕ НА ТЕХНИЧЕСКИ И СТАТИСТИЧЕСКИ SDMX¹ СТАНДАРТИ В НАЦИОНАЛНАТА СТАТИСТИЧЕСКА СИСТЕМА НА БЪЛГАРИЯ”

Галя Статева*, Костадин Георгиев**



През декември 2012 г. Националният статистически институт стартира работата по проект „Хоризонтална и вертикална интеграция: Прилагане на технически и статистически стандарти в Европейската статистическа система” съгласно Споразумение за субсидия на Европейската комисия № 60502.2012.003-2012.098. Проектът беше с продължителност 24 месеца (декември 2012 - ноември 2014 г.) и на обща стойност 113 000 евро. Неговото успешно изпълнение се осъществи съвместно от статистически и ИТ експерти в НСИ.

Основните цели на проекта са прилагане и популяризиране на техническия и статистически SDMX стандарт за обмен на статистически данни и метаданни в националната статистическа практика и постигане на вертикална интеграция между Националната статистическа система и Европейската статистическа система.

За изпълнение на целите проектът е конструиран в няколко **компонента**:

- 1. Прилагане на ЕСС стандарти за метаданни и разработване на национална референтна база данни, интегрираща структурите ESMS² и ESQRS³**
- 2. Разработване на база данни за разпространение в съответствие с SDMX технически стандарт 2.1**
- 3. Прилагане на SDMX-RI⁴ и свързани ИТ софтуерни средства.**

* Държавен експерт в отдел „Обща методология и анализ на статистическите изследвания”, НСИ; e-mail: gstateva@nsi.bg.

** Главен експерт в отдел „Връзки с потребителите и е-услуги”, НСИ; e-mail: kgeorgiev@nsi.bg.

¹ Statistical Data and Metadata eXchange.

² Euro SDMX Metadata Structure.

³ ESS Standard for Quality Reports Structure.

⁴ SDMX – Reference Infrastructure.

В рамките на всеки подкомпонент е извършен набор от високоспециализирани дейности, които спомагат за постигане на крайните резултати от проекта.

Необходимостта от този проект се обуславя от обективни обстоятелства. В НСИ нямаше изградена национална референтна база данни. Метаданни информацията и докладите за качество се съхраняваха в Word или Excel файлове, разпространявайки се на сайта на НСИ в PDF и HTML формат. Структурата за метаданни, която се използваше в предходните години не е стандартна и не включва всичките 21 понятия на ESMS, като понятията за качеството изобщо липсват. Докладите за качеството не са ориентирани към потребителите и не се публикуват в структурата ESQRS.

Статистическата информация, предназначена за разпространение, се съхранява в Excel или HTML файлове. Тези файлове нямат стандартна структура, което ги прави трудни за конвертиране в различни формати за разпространение. Изграждането на релационна база данни по проекта подобрява разпространението и съхранението на статистическа информация и ще улесни разпространението на информация в различни формати като HTML, SDMX, SVG, KML, DSPL, PX-AXIS, CSV, Excel, PDF и други.

В НСИ е инсталирана SDMX-RI 2.0 като част от проект „Census HUB“. За да се използва своевременно, SDMX-RI трябва да бъде актуализирана до версия 2.1, когато Евростат може да я предостави.

Стандартът SDMX е модел за описание на статистически данни и метаданни чрез технология за поддържане на стандартизирани ИТ средства. Състои се от информационен модел и статистически стандарти. SDMX информационният модел е изграден от технически стандарт 2.0/2.1, XML формат за обмен на SDMX структурирани данни и метаданни, дефиниции за структурата на данни/метаданни (DSD/MSD) за различни статистически области и SDMX регистри (напр. Европейски SDMX регистър). Статистическите стандарти съдържат следните компоненти: Ръководство, ориентирано към съдържанието, статистически концепции и кодови списъци (напр. честота, статус, формат за време, измерителна единица), списък на тематичните статистически области и общ речник за метаданни (MCV).

В рамките на Компонент 1 чрез специално създадени за целта работни групи бяха постигнати следните резултати:

- 380 файла с национални справочни метаданни на български и английски език в Европейската SDMX структура за метаданни (ESMS), разпространени на сайта на НСИ.
- Разработено DSD⁵ за краткосрочната бизнес статистика (КБС).

⁵ Data Structure Definition.

- Разработени DSD за структурната бизнес статистика за нефинансовите предприятия и статистиката на туризма чрез използване на ИТ средството Data Structure Wizard и прилагане на адаптирани концепции и стандартни кодови списъци за нефинансовите предприятия и статистиката на туризма на български и английски език.
- ИТ средство за създаване и съхранение на национални справочни метаданни, в което е заредена наличната информация за метаданни в структурите ESMS и ESQRS (версии на английски и български език).

ИТ средството за създаване и съхранение на национални справочни метаданни е разработено на базата на свободни софтуерни решения с отворен код. Използваният уебсервър е Apache, софтуерът за приложения е PHP, базата данни е MySQL и съдържанието се управлява с модулната система Drupal.

Националните справочни метаданни се описват в 116 понятия от Европейската интегрирана структура за метаданни (EIMS), включваща понятията за справочни метаданни на ESMS и ESQRS. За редактиране на описанията на метаданните в рамките на проекта беше създаден модул „Статистически данни и метаданни”, интегриращ EIMS за системата Drupal. Освен съхранение и редактиране на метаданни модулът предлага и дефиниране, и извличане на съобщения за метаданни в SDMX формат за произволно подмножество от понятия на EIMS. Подмножествата от понятия за метаданни на SDMX и ESQRS са предварително дефинирани в модула.

ИТ средството е оформено в инсталационен пакет, включващ Drupal (основни модули и допълнителни модули за разширяване на функционалност) и модул „Статистически данни и метаданни”. В инсталационния пакет са включени наличните 26 езикови пакета за превод на интерфейса на Drupal на използваните езици в Европейския съюз.

В рамките на Компонент 2 беше разработена национална база данни за разпространение в съответствие с SDMX технически стандарт. С цел бързодействие и яснота полетата от таблиците в базата данни съответстват на понятията за статистически данни в съответните DSD, разработени по Компонент 1. Базата данни за разпространение всъщност включва две бази данни - за конфиденциални данни и за неконфиденциални данни. Достъпът до базата с неконфиденциални данни е свободен, а достъпът до базата с конфиденциални данни е оторизиран и по защитен канал.

Достъпът до данните в базата данни за разпространение беше тестван за бързодействие. Използваният софтуер за провеждането на тестове беше Apache JMeter. Отчитайки данните за разглеждане на статистическа информация, на сайта на НСИ бяха дефинирани и изпълнени различни тестови сценарии. Тестовите симулираха директен SQL достъп до данните, достъп през уебслужба със SOAP, достъп и

визуализация с HTML и достъп и визуализация с външна библиотека за графики Google Charts.

В рамките на Компонент 3 беше изградена референтна инфраструктура SDMX-RI, което прави възможен обмен на информация за три статистически области, предварително дефинирани по проекта: краткосрочна бизнес статистика, статистика на туризма и статистика на нефинансовите предприятия (данни за БНБ). По този начин работещата референтна инфраструктура SDMX-RI улеснява обмена на данни между НСИ, Евростат и БНБ. SDMX-RI е универсална рамка за модерно предоставяне на данни, базирана на използването на SDMX стандарта. Съгласно визията на Евростат това е софтуер, който осигурява хоризонтална и вертикална интеграция и е изграден чрез модулен подход (от отделни стандартни софтуерни компоненти). SDMX-RI е разработена на Java и .NET и се използва за ефикасен обмен и споделяне на статистически данни и метаданни.

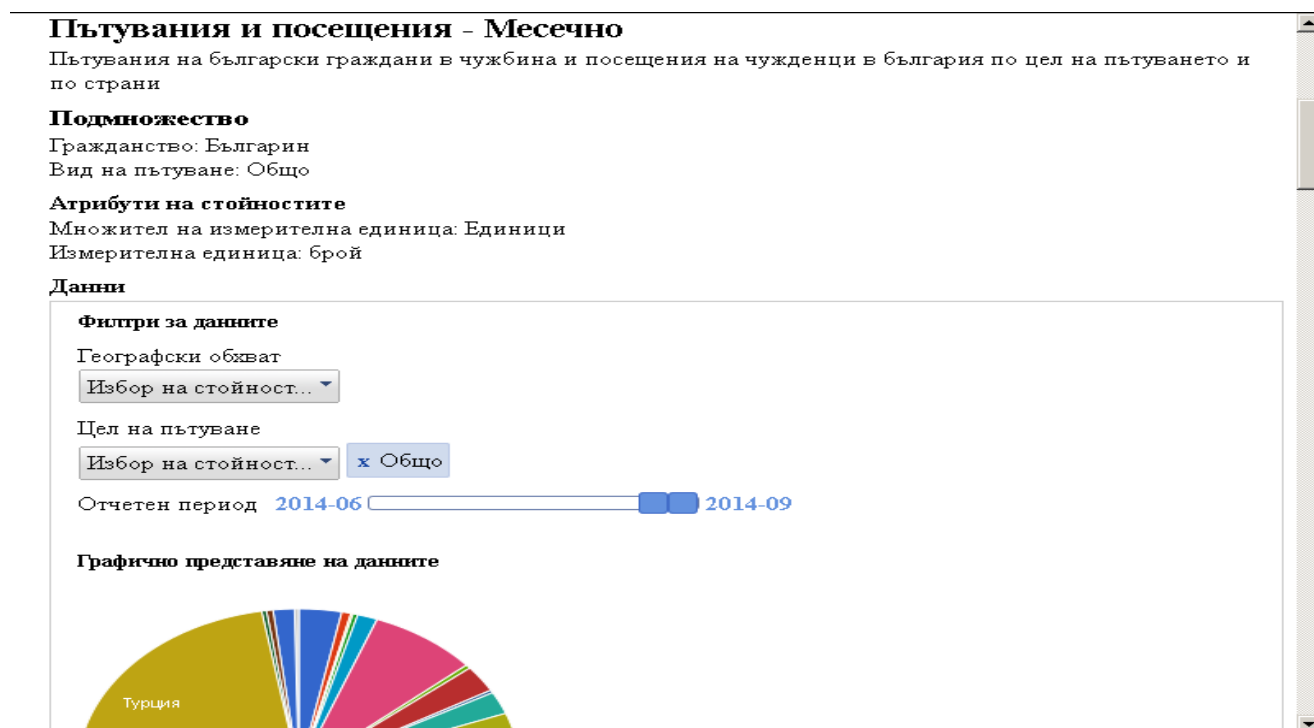
На основата на вече изградената база данни за разпространение от Компонент 2 и специално създадено приложение за манипулация на данните в базата бяха заредени статистически данни за краткосрочна статистика (за STS_INDICATOR = PROD – месечни данни за периода октомври 2012 - юни 2014 г.), туризъм (месечни данни за периода октомври 2012 - септември 2014 г.), данни за нефинансовите предприятия за обмен с БНБ (годишни данни за 2010, 2011, 2012 и 2013). Описанието на данните за разпространение, тяхното съдържание и местоположение се записват в база Mapping Store посредством използването на десктоп приложението Mapping Assistant. Процесът включва дефиниране на конекции към двете SQL бази данни - за разпространение и Mapping Store, зареждане на дефинираните DSD за трите статистически области и дефиниране на съответствието между полетата, описани в DSD, и полетата на таблиците от базата данни за разпространение (mapping).

Извличането на данни от базата данни за разпространение се осъществява чрез уебслужба, която използва информацията, съхранена в базата Mapping Store. Услугата получава SDMX заявка за данни, трансформира SDMX заявката в SQL заявка посредством информацията в Mapping Store, получава данни от базата за разпространение след изпълнението на SQL заявката, трансформира данните в SDMX съобщение с данни, използвайки информацията в Mapping Store, и връща отговор в SDMX формат.

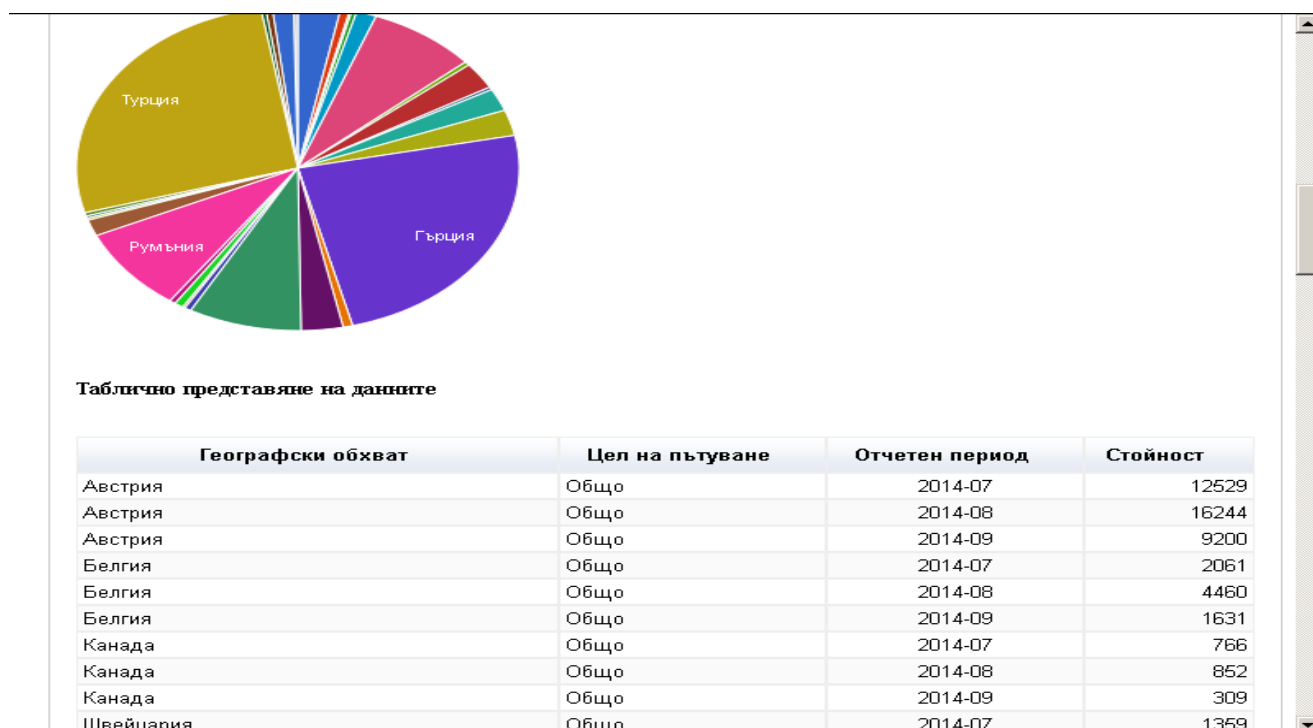
За презентиране на неконфиденциалните данни за туризма от базата данни за разпространение на сайта на НСИ беше разработено приложение, използващо външна библиотека за графики - Google Charts. Приложението показва данните за „Пътуване на български граждани в чужбина и посещения на чужденци в България” от статистическа област „Туризъм” в таблица, графика тип „торта” и карта на държавите. Потребителите имат възможност да прилагат различни филтри върху данните - географски обхват, цел

на пътуване и отчетен период и по този начин да изберат само информацията, която ги интересува. Филтрираната информация може да бъде извлечена и съхранена в CSV формат. Потребителите могат да извлекат и сахранят и цялата информация в SDMX компактно съобщение с данни, генерирано въз основа на SDMX заявката от уебслужбата и предоставено за презентиране от приложението (фиг. 1, 2, 3, 4 и 5).

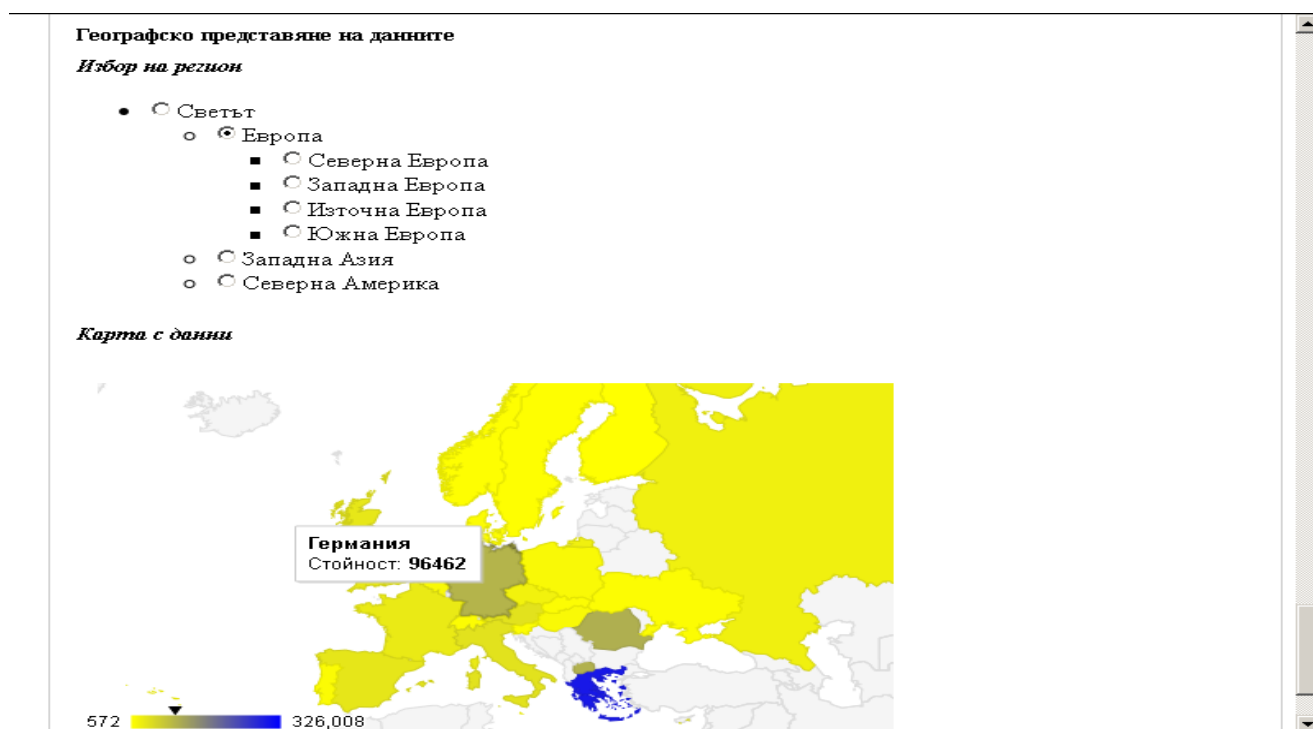
Фиг. 1. Презентиране на данните



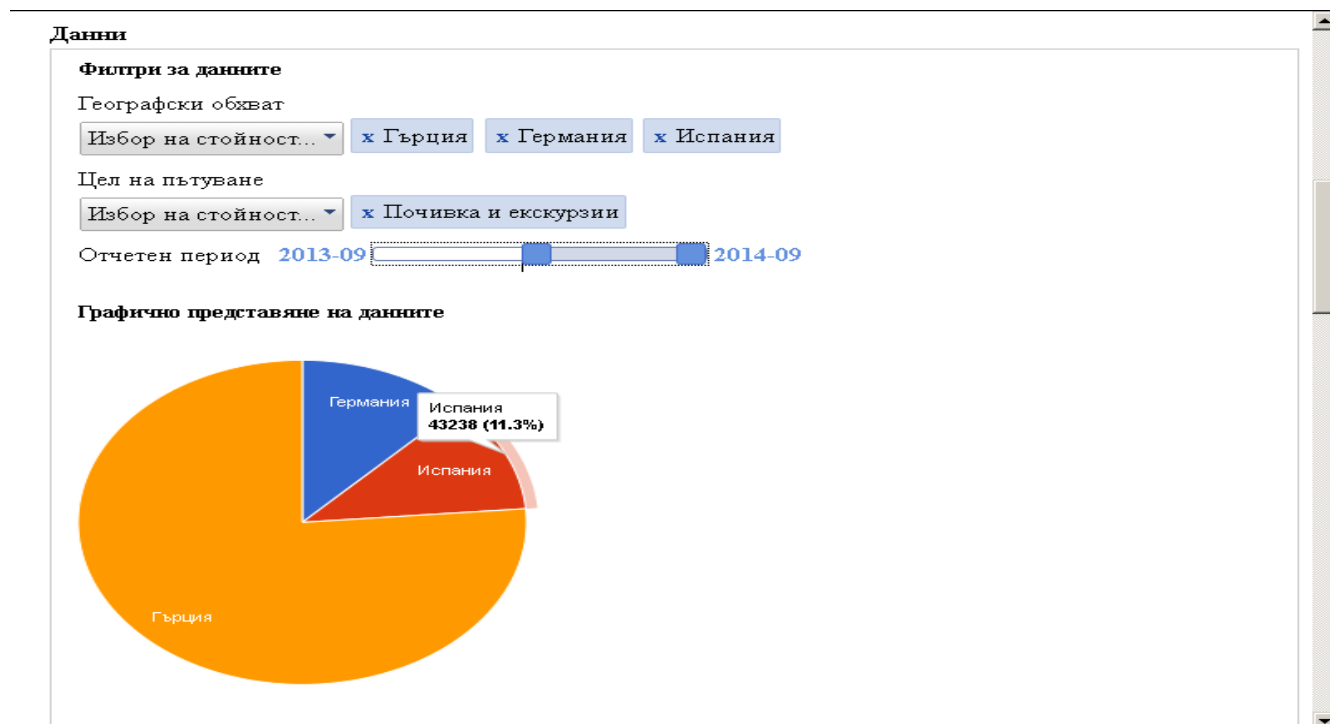
Фиг. 2. Графика тип „торта” и таблица „Географски обхват”



Фиг. 3. Карта „Географско представяне на данните”



Фиг. 4. Филтриране на данните



Фиг. 5. SDMX компактно съобщение с данни, върнато от уебслужбата

```

- <CompactData xsi:schemaLocation="http://www.SDMX.org/resources/SDMXML/schemas/v2_0/message
SDMXMessage.xsd urn:estat:sdmx:infomodel:keyfamily.KeyFamily=BNSI:BNSI_TOURISM_TABLES:1.0:compact
BNSI_BNSI_TOURISM_TABLES_Compact.xsd">
  - <Header>
    <ID>BNSI_TOURISM_TABLES</ID>
    <Test>>false</Test>
    <Name xml:lang="en">Trips and Visits</Name>
    <Prepared>2014-11-25T17:38:40Z</Prepared>
  - <Sender id="BNSI">
    <Name xml:lang="en">Bulgarian National Statistical Institute</Name>
    - <Contact>
      <Email>webmaster@nsi.bg</Email>
    </Contact>
  </Sender>
  <KeyFamilyRef>BNSI_TOURISM_TABLES</KeyFamilyRef>
  <KeyFamilyAgency>BNSI</KeyFamilyAgency>
  <DataSetAgency>BNSI</DataSetAgency>
  <DataSetID>BNSI_TOURISM_TABLES_BNSI_2014-11-25T17:38:40</DataSetID>
  <DataSetAction>Information</DataSetAction>
  <Extracted>2014-11-25T17:38:40</Extracted>
</Header>
  - <bns:DataSet UNIT_MULT="0" UNIT="NBR">
    - <bns:Series FREQ="M" REF_AREA="AT" PURPOSE="PERS_HOL" CITIZENSHIP="NAT"
      TRIP_TYPE="TOTAL" TIME_FORMAT="P1M">
      <bns:Obs TIME_PERIOD="2012-10" OBS_VALUE="1386" OBS_STATUS="A"/>
      <bns:Obs TIME_PERIOD="2012-11" OBS_VALUE="1141" OBS_STATUS="A"/>
      <bns:Obs TIME_PERIOD="2012-12" OBS_VALUE="960" OBS_STATUS="A"/>
      <bns:Obs TIME_PERIOD="2013-01" OBS_VALUE="621" OBS_STATUS="A"/>
      <bns:Obs TIME_PERIOD="2013-02" OBS_VALUE="830" OBS_STATUS="A"/>
    
```

Резултатите от проекта ще бъдат полезни не само за НСИ, а и за Националната статистическа система като цяло. Изградената SDMX-RI и наличието на уебслужби за базата данни за разпространение може да бъде допълвана с нови статистически области с цел обмен на данни и метаданни на национално и международно ниво. В процеса на изпълнение на проекта са приложени най-актуалните средства за обмен и разпространение на статистически данни и метаданни в Европейската статистическа система чрез прилагане на SDMX стандарта.

Авторите изказват благодарност към целия екип на проекта за техния принос и висок професионализъм при изпълнението на дейностите и постигането на поставените цели.

ИНФОРМАЦИЯ ЗА ПРОЕКТ „ПРИЛАГАНЕ НА ТЕХНИЧЕСКИ И СТАТИСТИЧЕСКИ SDMX СТАНДАРТИ В НАЦИОНАЛНАТА СТАТИСТИЧЕСКА СИСТЕМА НА БЪЛГАРИЯ”

Галя Статева, Костадин Георгиев***

РЕЗЮМЕ В информацията са описани дейностите и постигнатите резултати по изпълнението на проект „Прилагане на технически и статистически SDMX стандарти в Националната статистическа система на България”. Проектът беше с продължителност 24 месеца (декември 2012 - ноември 2014 г.) и на обща стойност 113 000 евро. Неговото успешно изпълнение се осъществи съвместно от статистически и ИТ експерти в НСИ.

Основните цели на проекта са прилагане и популяризиране на техническия и статистически SDMX стандарт за обмен на статистически данни и метаданни в националната статистическа практика и постигане на вертикална интеграция между Националната статистическа система и Европейската статистическа система.

Резултатите от проекта ще бъдат полезни не само за НСИ, а и за Националната статистическа система като цяло. Изградената SDMX-RI и наличието на уебслужби за базата данни за разпространение може да бъде допълвана с нови статистически области с цел обмен на данни и метаданни на национално и международно ниво. В процеса на изпълнение на проекта са приложени най-актуалните средства за обмен и разпространение на статистически данни и метаданни в Европейската статистическа система чрез прилагане на SDMX стандарта.

* Държавен експерт в отдел „Обща методология и анализ на статистическите изследвания”, НСИ; e-mail: gstateva@nsi.bg.

** Главен експерт в отдел „Връзки с потребителите и е-услуги”, НСИ; e-mail: kgeorgiev@nsi.bg.

ИНФОРМАЦИЯ О ПРОЕКТЕ „ПРИМЕНЕНИЕ ТЕХНИЧЕСКИХ И СТАТИСТИЧЕСКИХ SDMX СТАНДАРТОВ В НАЦИОНАЛЬНОЙ СТАТИСТИЧЕСКОЙ СИСТЕМЕ БОЛГАРИИ”

*Галя Статева *, Костадин Георгиев***

РЕЗЮМЕ В информации описываются работы и достигнутые результаты в реализации проекта „Применение технических и статистических стандартов SDMX в национальной статистической системе Болгарии”. Проект длился 24 месяца (Декабрь 2012 - Ноябрь 2014 года), на общую сумму € 113 000. Его успешная реализация осуществляется совместными усилиями статистических и ИТ-специалистов в НСИ.

Основные цели проекта - это применение и популяризация технического и статистического SDMX стандарта для обмена статистическими данными и метаданными в национальной статистической практике, а также и достижение вертикальной интеграции между национальной статистической системой и Европейской статистической системой.

Результаты проекта будут полезны не только для НСИ, но и для национальной статистической системы в целом. Созданные SDMX-RI и наличие веб-сервисов для базы данных по распространению, могут быть дополнены новыми областями статистики в целях обмена данными и метаданными на национальном и международном уровне. В процессе реализации проекта использованы самые современные средства для обмена и распространения статистических данных и метаданных в Европейской статистической системе путем применения стандарта SDMX.

* Государственный эксперт в отделе „Общая методология и анализ статистических исследований”, НСИ; электронная почта: gstateva@nsi.bg.

** Главный эксперт в отделе „Связи с потребителями и электронные услуги”, НСИ; электронная почта: kgeorgiev@nsi.bg.

INFORMATION ABOUT THE 'IMPLEMENTATION OF SDMX TECHNICAL AND STATISTICAL STANDARDS IN THE NATIONAL STATISTICAL SYSTEM OF BULGARIA' PROJECT

*Galya Stateva**, *Kostadin Georgiev***

SUMMARY The information describes the activities and results achieved in the implementation of the 'Implementation of SDMX technical and statistical standards in the National Statistical System of Bulgaria' project. The project had a duration of 24 months (December 2012 - November 2014), totalling € 113,000. Its successful implementation was carried out jointly by statistical and IT experts at the NSI.

The main objectives of the project are implementation and promotion of SDMX technical and statistical standard for exchange of statistical data and metadata in the national statistical practice and achievement of vertical integration between the National Statistical System and the European Statistical System.

The results of the project will be useful not only for the NSI but also for the National Statistical System as a whole. The established SDMX-RI and the availability of web services for the distribution database can be supplemented with new statistical areas in order to exchange data and metadata at national and international level. In the process of implementation of the project are embedded the latest tools for exchange and dissemination of statistical data and metadata in the European Statistical System by applying the SDMX standard.

* State Expert 'General Methodology and Analysis of Statistical Surveys' Department, NSI;
e-mail: gstateva@nsi.bg.

** Chief expert in 'Relations with Users and e-Services Department' Department, NSI;
e-mail: kgeorgiev@nsi.bg.

НА ВНИМАНИЕТО НА АВТОРИТЕ

Статиите, предложени на редакцията на сп. „Статистика”, трябва да бъдат оригинални, непубликувани в други издания.

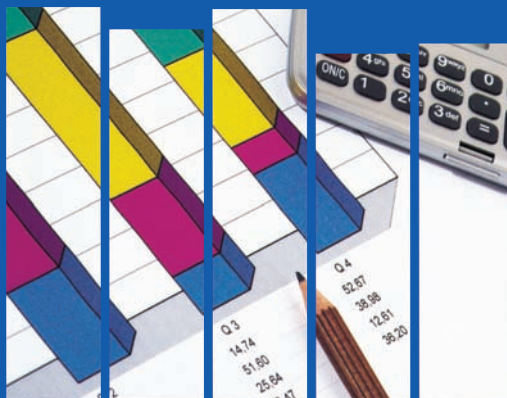
По решение на Редколегията на авторите се предоставя възможност статиите им да бъдат публикувани или на български, или на английски език. За целта е необходимо авторът, който желае статията му да бъде публикувана на английски език, да представи в редакцията ръкописи и съответни файлове и на български, и на английски език при спазване на посочените изисквания и за двата ръкописа. Авторите трябва да имат предвид, че английският текст ще бъде публикуван без редакционна намеса, което означава, че те носят пълна отговорност за коректността на превода.

Статиите, придружени от кратко резюме на български и на английски език и номерата на научната област от JEL класификацията, се представят в редакцията на сп. „Статистика”. Текстът трябва да бъде на Word, а графиките - Excel, със задължително приложени данни. Формулите в текста и тези на самостоятелен ред (без числата в тях) трябва да бъдат задължително курсивни (*Italic*). Ако графиките са дадени отделно, в текста трябва да се посочи точното им място. Авторите носят отговорност за коректността на данните и цитатите. Ръкописите трябва да са придружени от трите имена, научната степен и звание, точния адрес, телефона, длъжността, мястото на работа и електронния адрес (e-mail) на автора.

Поместването на статия на страниците на списанието не означава, че предложенията и евентуалните критични бележки, които авторът отправя, се споделят от Редакционната колегия или от ръководството на Националния статистически институт.

Редакцията на сп. „Статистика” би приветствала различни инициативи на читателската аудитория под формата на дискусии на страниците на списанието, научно-професионални „реплики” по повод на отпечатана статия, самостоятелни кореферати по повдигнатата тема - изобщо всяка инициативност по теми и проблеми на науката и практиката на статистическите изследвания от сферата на българската и европейската статистика.

Редакция на списание „Статистика”



СТАТИСТИКА STATISTICS

3/2014

www.nsi.bg