



ЮГОЗАПАДЕН УНИВЕРСИТЕТ
„НЕОФИТ РИЛСКИ” БЛАГОЕВГРАД
ФИЛОСОФСКИ ФАКУЛТЕТ
КАТЕДРА „СОЦИОЛОГИЯ”

ВЕРА СВЕТОСЛАВОВА ВЕЛЕВА

*КАЛИБРАЦИЯ НА ДАННИ ОТ
СОЦИОЛОГИЧЕСКИ ИЗСЛЕДВАНИЯ*

А В Т О Р Е Ф Е Р А Т

на дисертационен труд за присъждане на образователна
и научна степен „доктор”, специалност „Социология”,
професионално направление

3.1 Социология, антропология и науки за културата

Научен ръководител:
Доц. д-р Йордан Калчев

Благоевград, 2014

Дисертационният труд се състои от увод, три глави, заключение, списък на използваната и цитираната литература. Съдържа 192 страници.

Дисертационният труд на тема **„Калибрация на данни от социологически изследвания”** е предложен и обсъден за публична защита на заседание на катедра „Социология” към Философски факултет при Югозападен университет „Неофит Рилски” - Благоевград, проведено на 24.02.2014г.

Защитата на дисертационния труд ще се състои на 28.04.2014 година от 10.30 ч. в Югозападен университет „Неофит Рилски” - Благоевград, УК-1.

Материалите по защитата са на разположение на интересуващите се в кабинет 349, УК-1, на Югозападен университет „Неофит Рилски”, ул. „Иван Михайлов” 66, гр. Благоевград.

Съдържание дисертационния труд:

У В О Д.....	3
ГЛАВА ПЪРВА.....	10
СОЦИОЛОГИЧЕСКИТЕ ЗНАНИЯ В ИЗУЧАВАНЕ НА ОБЩЕСТВЕНИТЕ ПРОЦЕСИ	10
I. ЗА ЕМПИРИЧНОТО НАЧАЛО В РАЗВИТИЕТО НА СОЦИОЛОГИЯТА	10
II. РОЛЯ НА СОЦИОЛОГИЧЕСКИТЕ ИЗСЛЕДВАНИЯ В УПРАВЛЕНИЕТО НА СОЦИАЛНИТЕ ПРОЦЕСИ.....	17
1. Социалните процеси като обект на изследване и управление	17
2. Функции на емпиричното социологическо изследване	21
ГЛАВА ВТОРА.....	25
„ИЗСЛЕДОВАТЕЛСКИ ПРОЦЕС”. ЕТАПИЗАЦИЯ И ПРОЦЕДУРИ НА ИЗСЛЕДВАНЕТО.....	25
I. СЪЩНОСТ НА СОЦИОЛОГИЧЕСКОТО ИЗСЛЕДВАНЕ. ЕТАПИЗАЦИЯ И ПРОЦЕДУРИ НА ИЗСЛЕДВАНЕТО.....	25
1. Специфика на емпиричните социологически изследвания	25
2. Социологическото изследване като изследователска процедура (познавателни дейности)	32
3. Социологическото изследване като статистически производствен процес (технологични дейности).....	40
II. ИЗИСКВАНИЯ ЗА КАЧЕСТВОТО НА ИНФОРМАЦИЯТА ОТ СОЦИОЛОГИЧЕСКИТЕ ИЗСЛЕДВАНИЯ.....	60
1. Същност и основни функции на показателите за качеството на информацията	60
2. Някои основни характеристики за качеството на социологическата информация.....	62
ГЛАВА ТРЕТА.....	71
ПРИЛОЖЕНИЕ НА СТАТИСТИЧЕСКИТЕ МЕТОДИ В ЕМПИРИЧНИТЕ СОЦИОЛОГИЧЕСКИ ИЗСЛЕДВАНИЯ. КАЛИБРАЦИЯ НА ДАННИТЕ ОТ ИЗВАДКОВИ ИЗСЛЕДВАНИЯ	71
I. СТАТИСТИЧЕСКИЯТ ПОДХОД В СОЦИОЛОГИЧЕСКИТЕ ИЗСЛЕДВАНИЯ.....	71
1. Накратко за методологията на социологическите изследвания..	71
2. Проблемът за количествените и качествените методи в социологическите изследвания	74

3. Основни направления в приложението на статистическите методи в социологическите изследвания	84
4. Проблеми на измерванията в социологическите изследвания.....	87
II. ИЗВАДКОВИЯТ МЕТОД В СОЦИОЛОГИЧЕСКИТЕ ИЗСЛЕДВАНИЯ	92
1. Същност и приложение на извадковия метод в социологическите изследвания.....	92
2. Статистически извод. Свойства на получените статистически оценки от извадкови изследвания	99
3. Източници и видове грешки в извадкови изследвания	103
4. Влияние на грешките върху извадкови резултати.....	106
5. Техники за намаляване на ефектите от стохастичните грешки и подобряване на обхвата на единиците в извадката	112
III. КАЛИБРАЦИЯ НА ДАННИТЕ ОТ ИЗВАДКОВИ СОЦИОЛОГИЧЕСКИ ИЗСЛЕДВАНИЯ.....	120
1. Проблемът за липсващите данни за отделни признаци и методи за неговото решаване.	120
2. Проблемът за липсващи данни поради необхват на единиците в извадката.	125
3. Влияние на необхвата върху точността на оценките	130
4. Претегляне на данни от извадката. Необходимост от претегляне	135
5. Корекция на данните чрез методите на постстратификацията .	140
6. Калибрация на данните от извадкови изследвания. Същност. Нов системен подход за претегляне на извадкови данни.....	142
7. Аналитично представяне на концепцията за калибрацията на данните.....	147
8. Калибрация на данни от емпирични изследвания	151
ЗАКЛЮЧЕНИЕ.....	178
СПРАВКА ЗА ОСНОВНИТЕ ПРИНОСИ В ДИСЕРТАЦИОННИЯ ТРУД	184
СПИСЪК НА ИЗПОЛЗВАНАТА И ЦИТИРАНАТА ЛИТЕРАТУРА	186
ПУБЛИКАЦИИ ВЪВ ВРЪЗКА С ДИСЕРТАЦИОННИЯ ТРУД	192

Социологическата информация, получена в резултат на проведени емпирични изследвания е основен инструмент в управлението на всички страни от обществения живот. За да изпълни тези важни функции социологията, като теория и практика, следва на базата на теоретично и методично добре разработени проекти на социологическите изследвания да осигурява необходимата **качествена информация за развитие на научното знание и научно управление на обществото.**

В **Уводната** част са определени изследователските граници на дисертационния труд.

Обект и предмет на изследването

Обект на изследване в дисертационният труд е социологическата теория и практика при провеждането на емпирични изследвания чрез приложение на репрезентативни извадкови методи за осигуряване на социологическа информация и знания за обществените процеси.

Предмет на изследване в работата са възможностите за приложение на съвременни математически методи и софтуерни процедури за подобряване качеството на получаваните данни от извадкови (репрезентативни) социологически изследвания в аспект на повишаване на тяхната надеждност, валидност и достоверност при отразяване на изследваните социални процеси и ролята им в социалното управление.

Цел и задачи на изследването

Калибрацията на данни от емпиричните изследвания е нова област в математическата статистика и в разработвания специализиран софтуер. Неговата обосновка и вариантност изисква задълбочени познания в областта на математическото моделиране и програмирането. Затова целта и задачите се определят и свързват изцяло с необходимостта и потреблението на тези научни и технологични достижения.

Целта на дисертационния труд е въз основа на литературни източници и практическата дейност на различни

български и чуждестранни институции да се проучат функциите и възможностите за приложение на някои основни методи (процедури) за калибрацията на данните от представителни социологически извадкови проучвания и представянето им като научно обоснован подход за повишаване качеството и достоверността на информацията от емпиричните социологически изследвания.

В смисъла на така формулираната цел се поставят следните **изследователски задачи**:

1. Въз основа на същността, функциите и мястото на социологическите изследвания в управлението на обществото да се разгледат и представят някои критерии и изисквания за качеството на получаваната от тях информация.

2. Да се приложи нов подход в разглеждането на емпиричното социологическо изследване като изследователски и производствен процес и очертаване на дейностите при неговата реализация.

3. Приложение на репрезентативните извадкови методи при изследване на социалните процеси в съвременната действителност и посочване на някои важни проблеми, рискови фактори и грешки, водещи до влошаване качеството на социологическите данни от емпиричните изследвания.

4. Обосновка на необходимостта от калибриране на данните от социологическите изследвания в контекста на значението на социологическата информация при управлението на социалните процеси и научното познание.

5. Дискутиране на общите теоретични принципи и определяне на практически рамки за същността на калибрацията на данните от социологическите изследвания.

6. Изследване на опита на специализирани институции у нас в използването на процедурите за калибрацията на данни от извадкови социални изследвания и анализиране на резултатите от калибрацията.

7. Калибрацията на данни от конкретни социологически изследвания с цел анализиране на получените резултати и установяване на информационните ефекти върху надеждността и точността на данните от социологическото изследване. Формулиране на изводи и заключения относно приложението на процедурите за калибрацията на данните от емпиричните социологически изследвания.

Изследователска теза

Изследователската теза на дисертационния труд се обосновава от обективното ситуиране на мястото и значението на получаваната и анализирана информация от социологическите изследвания в управлението на социалните процеси на всички административни нива. От тук произтичат и се налагат все повече **високи изисквания при осигуряването на актуална, точна, надеждна и съдържателна социологическа информация.**

Осигуряването на тези характеристики на социологическата информация днес е възпрепятствано от различни рискови фактори. Най-общо те се изразяват в констатацията, че всеки (макар и на високо професионално ниво) разработен теоретичен модел и последващата го добра операционализация практически е трудно напълно да бъде реализиран на теренните условия.

Все по-трудно (по различни причини: всекидневна и териториална мобилност, нежелание за сътрудничество, право за защита на лични данни, трудна достъпност и т.н. на респондентите) се реализират проектираните модели на извадките, които определят представителността, точността и надеждността на получаваната социологическа информация, независимо от предприетите мерки и търсенето на различни дизайн-ефекти.

Тези обстоятелства налагат задължително, наред с редактирането на данните по същество, да се предприемат всички възможни процедури за постигане на необходимите качествени параметри на извадкови данни. Едно от възможните направ-

ления на работа в това отношение е калибрацията на данните от емпиричните социологически изследвания. Най-общо нейният смисъл за приложение се изразява в следното:

1. Принципно при всяко извадково изследване е налице определена степен на изместеност и неточност на оценките спрямо параметрите на генералната съвкупност;

2. Чрез приложението на алгоритмите за калибрацията се минимизират получаваните (пресметнатите) несходства (разлики) по предварително проектирания модел на изследването (извадката). Разликите се подлагат на калибриране по изходните параметри на изследваната съвкупност или чрез регресионни модели установяващи зависимости със стабилни признаци.

3. Подобряване на статистическите оценки от извадковото изследване и приближаването им към изходните данни от генералната съвкупност. В разработката тази процедура се разглежда в ограничени рамки.

Теоретико-методологични основи

Теоретико-методологичните основи на дисертационния труд са изградени върху анализи и обобщения от литературни източници на български и чуждестранни учени и изследователи в областта на теорията на социологията и практиката на социологическите проучвания, на изследователи, занимаващи се с приложението на математико-статистически методи.

Използвани са официално публикувани и разработвани документи на национални и международни организации, регламентиращи изискванията за параметрите на получаваните данни от провеждани емпирични извадкови изследвания. Представени са в кратък вид описания на разработени модели и готови програмни продукти и технологии за калибрацията на данните и резултати от опита на някои национални статистически офиси в прилагането на тези методи и технологии.

Логическата постановка на изложението започва с показването на полезността на социологическата информация и нейното изключително важно място в обогатяване на теорията

за обществото и неговото реално управление, след което е насочена към разглеждането на емпиричното социологическо изследване като научна процедура и осъществяването ѝ като статистически производствен процес, до извеждането на критериите и изискванията за качеството на получаваните данни.

Крайната цел е показването на възможностите на разработени методи на математическата статистика и информационни технологии за подобряване на комплекса от качествени характеристики на социологическата информация.

Ограничения в изследването

В широк и задълбочен план поставената тематика в дисертационния труд в теоретично отношение е изключително сложна и на високо специализирано научно ниво. Поради това в целта и задачите не се включват за обсъждане проблеми, отнасящи се до теоретичните основи на калибрацията на данните и тяхното моделиране, да се съпоставят и оценяват предимствата на един или друг метод. Обсъждането на тази проблематика в работата е много по-ограничено. Тук, изследователските задачи са насочени изцяло потребителски, от позицията на „неквалифицирания потребител” към най-общо представяне - показване на калибрацията като методология и инструмент за подобряване на качеството на емпиричните данни.

За постигане на дефинираните цел и задачи, работата е насочена към достигане на констатации относно значението на информацията от социологическите изследвания, обсъждането на нейните качествени параметри и разкриване на възможностите за тяхното подобряване. Въз основа на проучен, макар и все още малък, практически опит, се доказва необходимостта от проследяване и внедряване на съвременните методи и информационни технологии в практиката на социологическите изследвания, които са насочени към решаването на проблемите с липсващи данни (импутиране на данни) и подобряване на

статистическите оценки от използваните извадки чрез претеглянето, респ. калибрацията на данните.

В **Първа глава** на дисертационния труд се разглежда същността и ролята на социологическите изследвания в управлението на социалните процеси. Предимствата на социологическия подход при емпиричното изследване на социалните системи и техните елементи се проявяват в постигането на интегритет в събирането, обработката и анализирането на информацията за обществените процеси. Затова в тези изследвания социологическият подход се колаборира със статистическите способности и методи за изучаване на социалната действителност. Това поставя социалната статистика и емпиричната социология в позиция с особена важност в механизмите на социалното управление.

Направеният исторически преглед в развитието на основни социални теории и концепции за процесите в обществото показва накратко формирането на социологическата наука като система от знания за обществото като цяло и неговите съставни структури. През целия период на своето утвърждаване и развитие социологията е свързана с решаването както на теоретични, така и на практически задачи. Много мащабни социологическите изследвания са основа за проверка и разработване на специални социологически теории. В същото време тези теории помагат за тълкуването на събраните данни и опосредстват връзката между общите социологически позиции и конкретните емпирични изследвания на процесите в социалния живот.

В рамките на **Втора глава** от дисертационния труд изложението е насочено към изследване и анализ на спецификата на конкретните социологически изследвания, техните особености и функции при осигуряването на качествена информация за изследваните процеси.

Социологическото изследване се разглежда като комплексно изучаване на дефинирани социални проблеми чрез специално разработена методика и технология за събиране и обработка на първични сведения от единиците на изследвания обект.

Конкретните социологически изследвания се опират на специализирани социологически теории, които изкрystalизират в методики с формулирани теоретични и операционализирани понятия, хипотези, модели и т.н., а също и включването на методи за събиране и обработка на данните. Следователно социологическото изследване се обосновава и реализира чрез използването на статистически процедури и методи.

При разглеждане на същността и задачите, които си поставя всяко социологическо изследване се дефинират различен по брой и с различно съдържание, етапи, фази, технологични процедури и дейности, които включват различни елементи от процеса на изследването.

Разглеждането на емпиричното социологическо изследване като една изследователска процедура, включваща поредица от познавателни елементи насочва и към една друга страна на представяне, а именно организационно-процедурната. И докато съдържателната страна на емпиричното изследване се определя от постиженията в социологическите знания, то неговата теренна реализация се основава на научните принципи, разработени от други науки и преди всичко на статистиката, математическата статистика и теорията на вероятностите.

Социологическите изследвания основно се осъществяват на базата на извадковия метод. Тези изследвания „могат да се използват за получаването както на статистическа (статистически разпределения, статистически оценки), така и на социологическа информация.”¹ Или, с други думи, този вид

¹ Калчев, Й., (2005) Социологически траектории, УИ „Св. Кл. Охридски“, С., стр. 69

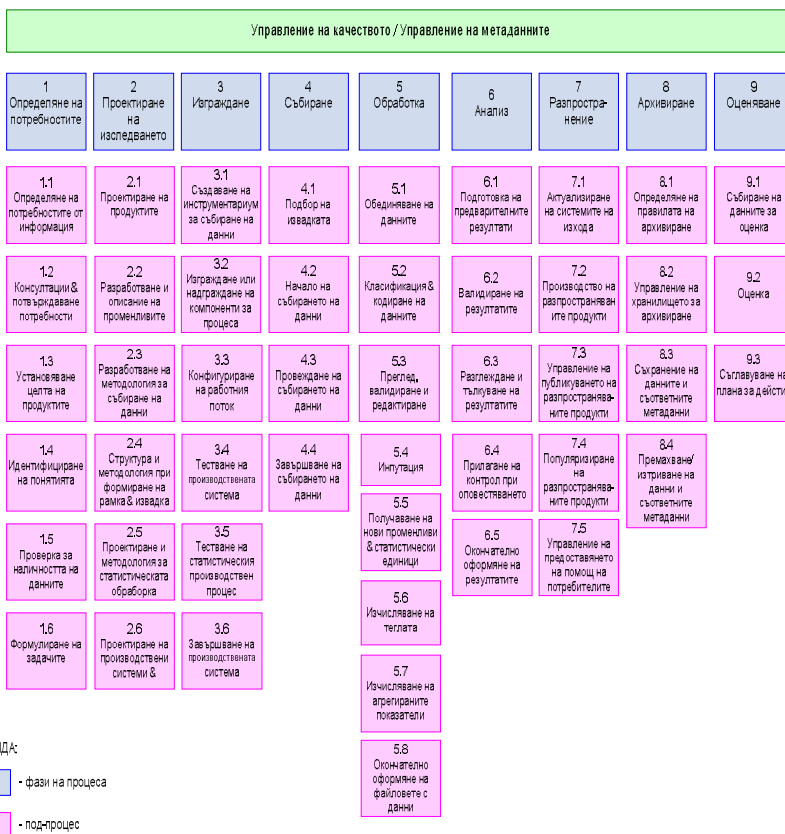
социологически изследвания погледнати като познавателна процедура са социологически, а като технология на осъществяване са статистически.

Основанията за такъв подход при разглеждането на изследванията в обкръжаващата ни социална действителност (в случая социологически и статистически) е, че винаги изследванията се извършват върху един обект и той се дефинира еднозначно. Независимо от предмета на изследването, използваните методи за събиране на информацията са възможно приложими при различни по теми изследвания, методите за обработка и анализ на получените данни не са свързани с тяхната тематична област и имат универсален характер.

Тъй като качеството на данните от социологическите изследвания в голяма степен зависят от технологиите за събиране, регистрация и обработка на данните в тази глава е представен един международно обсъждан модел на целия статистически производствен процес. Той не се лансира като твърда рамка, в която всички стъпки трябва да бъдат следвани в строг ред, а по-скоро като модел, който посочва стъпките в статистическия бизнес процес и взаимозависимостите между тях.

В съответствие с рамките на представения модел (GSBPM) са разгледани същностните страни и спецификата на емпирично социологическо изследване (ЕСИ) като последователност от изследователски (познавателни) процедури и като технологични процедури в статистическия производствен процес.

Схема 1: Общ модел на статистическия производствен процес (GSBPM)²



В заключителната част се обсъждат някои важни критерии и характеристики, които трябва да се прилагат при оценяването на качеството на информацията от социологическите изследвания. Интерпретирана е същността на изиск-

² Общият модел на статистическия производствен процес (GSBPM), подготвен от секретариата на UNECE, се намира на адрес: <http://www1.unecce.org/stat/platform/display/metis/The+Generic+Statistical+Business+Process+Model>

ванията за: приложимост, своевременност, сравнимост и съгласуваност, достъпност и яснота (вкл. конфиденциалност), интерпретивност (разбираемост). Изброените до тук характеристики имат предимно номинален смисъл. Съществуват обаче и други характеристики, които имат своя теоретичен статистически смисъл и могат да се оценяват - достоверност, представителност, надеждност, точност.

В **Третата глава** са разгледани различни концепции за самостоятелността в научното познание на количествените и качествените методи и се достига до заключението, че видът на метода (количествен; качествен) би следвало да не се определя от броя на наблюдаваните респонденти, а от наблюдаваните факти, които качествено се обобщават въз основа на тяхното количествено натрупване. Препоръчва се в тяхното приложение да се търсят общометодологическите основи на методите за изследване в тяхната неделимост, респ. съвместимост при използването им.

Приложение на извадковия подход в ЕСИ. Проблемът на необхванатите единици

Важна част от тази глава на дисертационния труд е приложението на извадковия метод в социологическите изследвания. Представени са нови съвременни практически процедури при проектирането на извадките с цел подобряване на дизайн-ефекта и преодоляване, още в тази фаза при проектирането на ЕСИ, на проблемите, свързани с неизпълнението на извадките.

Направена е кратка систематизация и описание на грешките, предизвикани от липсващи данни за отделни признаци и грешки, породени от необхват на единиците в извадката и тяхното влияние върху стохастичната точност на получаваната информация. Те са разгледани през призмата на свойствата на оценките, получени от представителни извадки - състоятелност, неизместеност, ефективност и достъпност.

Стохастичните грешки и нестохастичните грешки на извадката възникват поради естеството на извадковия метод, но и от неправилни процедури при проектирането и излъчването на извадката.

Общото отклонение (грешка) за даден параметър е сума от всички отклонения. Освен това тяхното проявление има различна посока (положителна и отрицателна), поради което може да се очаква, че част от тях могат взаимно да се компенсират. Независимо от това, естеството на стохастичните и нестохастичните грешки в даден случай може да доведе до намаляването на един вид грешки, но и до нарастването на друг вид грешки и оттам до нарастване на общото отклонение.

Изключително негативен проблем при съвременните изследвания **са грешките в извадкови оценки, породени от пропуските в обхвата на изследваните съвкупности.** Тежестта на проблема идва и от това, че непълнотата на обхвата влияе не само върху достоверността на получаваната извадкова информация, но и върху представителността на данните, а последното обезсмисля и прилагането на извадковия метод в емпиричните изследвания.

Преодоляването на този проблем е сериозна задача пред изследователя, тъй като отпадането на единиците от извадката е равносилно на изключването на част от генералната съвкупност в извадковата рамка. От тук и заключението, че извадката престава да бъде представителна за изследваната генерална съвкупност.

Теоретичните основи за проектиране на представителни извадки са добре известни, но в практиката на социалните изследвания винаги възникват проблеми по събирането на необходимите данни от представителната извадкова съвкупност. Така се оформя проблемът за липсващите (несъбрани) данни. В тази ситуация достигането до погрешни изводи е възможно, тъй като се нарушават принципите на вероятностните заключения и се получават изместени оценки.

Най-общо липсващите данни се установяват в две ситуации:

- когато липсват данни за значенията само на част от признаците за наблюдаваните единици (частична загуба на данни);
- когато неполучените данни изцяло са в резултат на необхват (отпаднали единици) от планираната извадка.

За решаването на тези проблеми е възможно да се приемат и приложат различни мерки и действия за намаляване на ефектите от неотговорилите и необхвата на единиците от извадката. Действията могат да бъдат превантивни от организационно-технически характер и методологически - използване на модели за последващо третиране на данните от реализираната вече извадка.

От тези с методологичен характер, най-често прилаганите са:

- присвояване (приписване) на значения (imputation) на липсващите данни;³
- традиционни методи за претегляне на данните от извадки;
- модели за калибрацията (Calibration) на данните от извадки.

Претегляне на данните от извадкови изследвания.

Постстратификация

Претеглянето на данните, получени от извадка е известна и често прилагана процедура за постигане на критериите за качество на информацията. Тя е особено необходима когато резултатите от проведеното емпирично изследване се разпростират (обобщават) за цялата генерална съвкупност. Претеглянето на данните се използва, за да се намали система-

³Присвояването на данни не е предмет на разглеждане в дисертационния труд. Трябва да се има предвид, че претеглянето на данните и преди всичко калибрацията на данните компенсира в много голяма степен липсата на данни за определена променлива.

тичното отклонение при оценките от извадковото изследване (нестохастичната грешка).

В статистиката се предлагат различни и известни решения за претегляне на данни от извадки, обединени под наименованието „методи за претегляне на извадкови данни“.

След събирането и редактирането на данните от извадковото изследване процедурата по претеглянето на данните включва изчисляване на тегла (W_i) обратно пропорционални на вероятността за включване на всяка единица в извадката (P_i) - (базови тегла): $W_i = \frac{1}{P_i}$

Тези тегла позволяват да се: а) отчетат параметрите на извадката и да се компенсират различията; б) да се направи корекция на неотговорилите в извадката (преразпределяне на теглата на неотговорилите сред респондентите); в) обвързване на оценките с известни параметри на генералната съвкупност⁴.

Известна стандартна процедура за претегляне на данните в социалните изследвания е изчисляване на постстратификационни тегла с отчитане на налични демографски и социални показатели.

Аналитично същността на постстратификацията може да се даде, чрез формулата:

$$Km = \frac{Nm}{\sum W(hi)_j \cdot \delta(j)}, \text{ където:}$$

Km - постстратификационен коефициент за m -та страта;

Nm - брой на единиците в m -та страта сред цялото население;

$W(hi)_j$ - тегло с поправка на неотговорилата единица j ;

$\delta(j)$ - множител със стойност 1 (едно), когато единицата принадлежи към дадената страта и 0 (нула), ако не принадлежи.

В изследователската практика за постстратификацията съществуват многовариантни предложения. Един от вариантите е наличието на контролна информация за изследваната съвкуп-

⁴ Виж. Kish, I., Survey sampling, N.Y. 1965

ност по два и повече признака. Когато са налични няколко контролни променливи могат да се използват и по-сложни методи като се изследват линейни зависимости, напр. регресионен анализ. В този смисъл постстратификацията може да се разглежда като един „несложен праволинеен метод за калибрацията на данни”⁵ от извадкови изследвания.

Калибрация на данни. Същност

През последното десетилетие се отделя значимо изследователско внимание на проблемите за непълните масиви от данни, получени от извадки. Разработени бяха и започнаха да се прилагат нови специфични методите на математическото моделиране, или т.нар. калибрация (Calibration) на данните в изходните извадкови масиви.

Идеята за калибрацията на данни от извадкови изследвания получава голям напредък след публикуването на поредица от статии⁶ от специалисти в областта на теорията и практиката на статистическите извадкови изучавания в края на миналия век.

В научната литература „калибрацията” се определя като метод за повторно претегляне и се използва, когато има достъп до няколко променливи, качествени и количествени, на които трябва да се извършва съвместно корекция. Друга дефиниция представя калибрирането като набор от тегла за единиците в извадката, които отговарят за калибриране на известни крайни обобщаващи за съвкупността величини, така, че оценителят

⁵ Руководство по подготовке статистических данных об использовании времени для оценки оплачиваемого и неоплачиваемого труда, Организация Объединенных Наций, Статистический отдел, Нью-Йорк, 2007, стр. 160

⁶ Deville, J.C. and Särndal, C.E. Calibration estimators in survey sampling. Journal of the American Statistical Association. 1992; Ekholm, A. and Laaksonen, S. Weighting via response modeling in the Finnish Household Budget Survey. Journal of Official Statistics. 1991; Lundstrom, S. and Särndal, C.E. Calibration as a standard method for treatment of nonresponse. Journal of Official Statistics. 1999; и много други.

(теглата) довежда до съответствие (проектиране в съответствие) със съвкупността. С други думи, целта е постигане на безпристрастно проектиране и изпълнение на извадката.

Крайната цел на калибрацията е намаляване на извадкови грешки и повишаване на точността и устойчивостта на извадкови оценки на статистическите показатели. „Калибрацията е процес на целенасочено изменение на параметрите на извадковия план, като извадкови тегла, ... за намаляване на извадкови грешки и повишаване на точността и устойчивостта на извадкови оценки на статистическите показатели”.⁷ Или, калибрирането е специфичен начин да се отчете налична за изследваната съвкупност помощна информация и чрез различни стандартни настройки тази информация да бъде включена в получаването на крайните оценки.

Аналитичен модел за калибрацията на данни

Трудно е да се направи универсално аналитично представяне на разработените модели за калибриране на данните, тъй като съществуват различни варианти на модели за калибрация. Един опростен вариант на модел за калибрацията на данни може да се придобие следния синтезиран вид:⁸

Нека приемем, че обект на изследване е налична генерална съвкупност U , която се състои от N елементи, като $U = \{1, \dots, k, \dots, N\}$. От проведено изследване е необходимо да се оцени значението на променливата J . Нека с $y_k = (y_{k1}, \dots, y_{kj})'$ означим вектор, значенията на променливата Jx_i за k -я елемент на съвкупността.

В изследването крайната цел е получаването на оценка на вектора от всички сумарни значения за променливата:

⁷ Степанов, С.В. Калибровка выборки, М. 2009, стр. 1

⁸ При представянето на модела са използвани следните публикации: Виж: Lundstrom, S. (Statistics Sweden), Särndal, C., (Statistics Canada), Calibration as a Method for Deriving Nonresponse Adjusted Weights Särndal, C. E.. The calibration approach in survey theory and practice, Statistics Statistique Canada, Vol.33, No2, 2007, p. 99-119

$T_y = \sum_{k \in U} y_k = Y'_{U'} 1_N$, където:
с $Y'_{U'}$ е означена матрица с $N \times I$ значения на y .

При формирането на извадката се допуска, че n единици от съвкупността U са включени в извадката s . За всяко извадково изследване крайната цел е да се получат оценки T_y . Ролята на „стандартна“ оценяваща формула на сумарните (общите - б.м) данни се изпълнява от оценъчната формула на Хорвиц - Томсън (H-T), която се определя като

$\check{T}_y = \sum_{k \in s} d_k y_k$, където:
 $d_k = 1/\delta_k$ - извадково тегло за единицата k ,
 δ_k е вероятност за включване на единицата k в извадката⁹.

Въпреки равните възможности за включване на всички единици от съвкупността U в извадката s , по различни причини липсват отговори (nonresponse).

Много често при проектирането на извадкови изследвания са налични променливи (x_k), които са обвързани с променливата y . Тази обвързаност и наличните данни за тези променливи се използват за подобряване на точността на оценките на изследвания параметър - T_y .

Цялата тази допълнителна информация може да се означаи и покаже като вектор X_k . За да продължи процедурата е необходимо да се състави т.нар. оценител¹⁰.

Същността на процеса на калибриране на данните накратко се изразява в следното. Приема се, че са известни сумарните (обобщаващите) оценки на съвкупността за променливите x и, че те трябва да се оценят по данни от извадката, като

⁹ Task Force and the Implemetation of NACE, Rev.2 - Neudlook on metoedological aspects elated to sampling designs and weeihts estimation, UNDE, Eurostat, 2006, Anex A, p.16

¹⁰ За по-подробно относно функциите и видовете оценители виж: Lundstrom, S. (Statistics Sweden), Särndal, C., (Statistics Canada), Calibration as a Method for Deriving Nonresponse Adjusted Weights

се използва оценяващата формула (Н-Т). Това предполага да се получат оценки T_x чрез:

$$\check{T}_x = \sum_{k \in S} d_k x_k.$$

Получените оценки \check{T}_x обаче, не винаги могат да осигурят точно съвпадение (точно оценяване) със съответните сумарни данни T_x на съвкупността, което се изразява в отклонението $\check{T}_x - T_x$.

За да се избегне тази разлика (грешка на калибрацията) оценяващата формула може да се промени като се заменят извадкови тегла d_k с нови тегла w_k , които се включват в калибровъчната формула:

$$\check{T}_{xC} = \sum_{k \in S} w_k x_k, \text{ където:}$$

(w_k, k_s) - са изследваните тегла без отклонения в калибровката и удовлетворяват равенството:

$$\check{T}_{xC} - T_x = \sum_{k \in S} w_k x_k - T_x = 0.$$

Смисълът на използването на този оценител е, ако използваните калибровъчни тегла (w_k) са в състояние да намалят, или ако е възможно да отстранят грешката при получаване на сумарните оценки на x , то те могат да се използват в оценъчната формула при калибрацията и да намалят грешката при получаването на сумарните оценки за y :

$$\check{T}_{xC} = \sum_{k \in S} w_k y_k.$$

Такова аналитично представяне на калибрацията на данни показва само общата идея на тази методология. Съществуват различни модели (оценяващи формули), които се прилагат в статистическата практика.

Най-често използваните калибрационни функции са:

- линейна функция: $f(x) = y = a + bx$
- експоненциална функция: $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$
- прекъснатата (ограничена, отрязана*) линейна

функция:

(*Функцията $f(x)$ е ограничена в интервала (a, b) , ако съществуват две числа A, B такива, че $A \leq f(x) \leq B$ за всяко $x \in (a, b)$).

- логит функция (ограничена) : $\text{logit}(p) = \log\left(\frac{p}{1-p}\right)$

Независимо от тяхното разнообразие и сложен математически апарат те са достъпни за по-широк кръг специализирани потребители, тъй като имат своите софтуерни аналози в готови програмни продукти.

Известни калибрационни програми са: CALMAR, GES, BASCULA, CLAN и g-CALIB¹¹. От посочените пет програми беше установено, че по-достъпна за приложение е програмата g-CALIB, която е базирана в софтуерния продукт – SPSS. Останалите основно са базирани в SAS софтуер, а той не е широко разпространен в нашата страна.

Калибрацията на данни от проведени емпирични извадкови изследвания

Ефектите от калибрацията на данни са показани чрез приложението на калибрационни процедури върху данни от извадкови изследвания. Представени са резултати от калибрацията на резултатите от национално представително изследване на НСИ - „Наблюдение на работната сила”, национално интернет изследване на МФВС и от „Епидемиологично изследване на насилие на деца в Балканските страни” (BECAN) - реализирано от екип в Югозападен университет „Неофит Рилски”.

Епидемиологичното изследване е проведено в три области: област Варна, област Велико Търново и област Благоевград.

Използваната извадка за определяне на единиците, обект на наблюдение, е излъчена пропорционално на броя на изследваната съвкупност във всяка област. За всяка област извадката е стратифицирана по местонамиране на училищата - (град, село) и по възраст на изследваните лица. Целта е

¹¹ Цялата математическа разработка и програмно осигуряване на тези подходи в сравнителен аспект могат да се видят в: Vanderhoeft, С., Generalised Calibration under SPSS, g-CALIB, StatisticsBelgium, 2002

избраният модел да осигури представителност и необходимата точност на резултатите общо, по отделните области и възрасти.

Обект на изследване са ученици, обхванати във формалното образование, на възраст 11, 13 и 16 години. Общият брой (генерална съвкупност) на учениците в трите области от посочените възрасти за изследвания период (2009-2010г.), по официални статистически данни е малко над 28 хиляди.

В представянето на този пример за калибрацията на данните се правят сравнения на резултатите получени пряко от изследването, от просто претегляне, с тегла от генералната съвкупност и резултатите след изпълнение на процедури за калибрацията на данните.

Изпълнението на процедурите включва:

1. Определяне на признаците, за които има данни за цялата генерална съвкупност и върху които могат да се приложат процедури по претегляне и калибрацията на резултатите. За претегляне на резултатите от изследването и изчисляване на калибрационните тегла са използвани статистически данни¹² за броя и разпределението на учениците по възраст в областите и по местонамиране на училището.

2. Изчисляване на относителните дялове по разновидностите на тези характеристики за общата съвкупност и за изваковата съвкупност.

3. Сравняване на тези характеристики и вземането на решение за претегляне и калибриране на събраните данни. Резултатите от направения анализ показваха, че е необходимо да бъдат предприети такива процедури с цел по-добро възпроизвеждане на основни структури на генералната съвкупност.

4. Избор на софтуер и модел за претегляне и калибриране на данните от изследването. С разполагаемата софтуерна среда у нас, единствено приложим е продукта g-CALIB, който е

¹² По данни на НСИ към м. октомври 2009 г.

базиран в SPSS. При калибрацията е използван модела на линейна-ограничена функция.

5. Въз основа на получените данни от извадковото изследване и наличните за генералната съвкупност беше подготвен информационен файл съобразно изискванията на избрания софтуер и беше включен в него.

6. Претегляне на данните от изследването и калибрацията на данните (изчисляване на калибрационни тегла).

Обикновеното претегляне на данните е извършено като е използвана наличната статистическа информация за разпределението на изследваната съвкупност по избраните признаци. Използвана е познатата процедура, която включва използването на тегла (W_i) обратно пропорционални на вероятността за включване на всяка единица в извадката - (P_i).

7. Претегляне на данните от извадката и генериране на кростабулации. След провеждане на процедурите по претеглянето и калибрирането на данните се произведени таблици, върху които е извършен анализ.

При сравняването на получените от извадката относителни дялове за възрастовите групи и местонамирането на училищата бяха установени различия. Например, делът на общо анкетираните ученици в градовете в отделните възрастови групи е:

<i>Възраст</i>	<i>Резултати от анкетата</i>	<i>Статистически данни</i>
11 години	72%	75%
13 години	69%	75%
16 години	73%	94%

Аналогични са резултатите и по отношение на учениците, които учат в селата.

Установените различия по тези стратифициращи признаци показват необходимостта от претегляне на данните (корекция на теглата).

Резултатите от изпълнените процедури по претеглянето на данните и калибрацията са представени в следващата таблица:

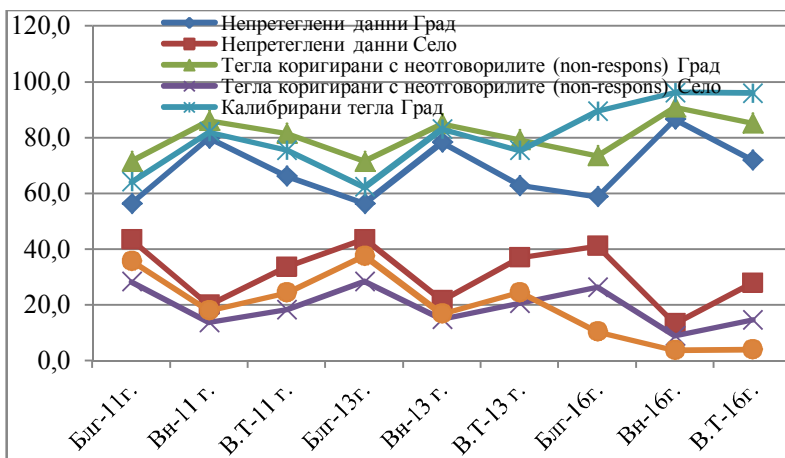
Кростабулация по области, местонамиране на училището и възраст

Възрастови групи			Непретеглени данни		Тегла коригирани с неотговорилите (non-response)		Калибриран и тегла		Общо
			Населено място		Населено място		Населено място		
			Град	Село	Град	Село	Град	Село	
а	б	в	1	2	3	4	5	6	7
11 год	Обл асти	Благоевград	56,5	43,5	71,6	28,4	64,2	35,8	100,0
		Варна	79,9	20,1	86,2	13,8	82,0	18,0	100,0
		В. Търново	66,2	33,8	81,6	18,4	75,5	24,5	100,0
	Общо		71,5	28,5	81,5	18,5	74,8	25,2	100,0
13 год	Обл асти	Благоевград	56,4	43,6	71,6	28,4	62,3	37,7	100,0
		Варна	78,3	21,7	85,1	14,9	83,1	16,9	100,0
		В. Търново	62,9	37,1	79,3	20,7	75,5	24,5	100,0
	Общо		68,8	31,2	79,4	20,6	74,6	25,4	100,0
16 год	Обл асти	Благоевград	58,8	41,2	73,5	26,5	89,5	10,5	100,0
		Варна	86,6	13,4	91,1	8,9	96,2	3,8	100,0
		В. Търново	72,0	28,0	85,3	14,7	95,9	4,1	100,0
	Общо		72,7	27,3	82,8	17,2	93,9	6,1	100,0
Об що	Обл асти	Благоевград	57,4	42,6	72,4	27,6	72,4	27,6	100,0
		Варна	81,2	18,8	87,2	12,8	87,2	12,8	100,0
		В. Търново	67,5	32,5	82,4	17,6	82,4	17,6	100,0
	Общо		71,0	29,0	81,3	18,7	81,3	18,7	100,0

В представената таблица в кол. 1 и 2 са посочени непрегеглените данни (относителни дялове - в %), получени от изследването за всяка възрастова група по отделно за всяка област, разпределени по местонамиране на училищата: град, село. В 3 и 4 кол., е извършено обикновено претегляне като са използвани данни на НСИ от статистиката на образованието. От същите официални статистически данни са изчислени калибрационни тегла, които са представени в колони 5 и 6.

При сравняването на получените относителни тегла (в %) се установява, че има различия в зависимост от начините на претегляне. Тук са представени общите изводи относно получените тегла и посоката на различията между тях.

Различията визуално са показани на следващата графика:



Съотношение на теглата на изследваните страти по възраст, области и тип населено място

От табличното и графично представяне на получените и изчислени тегла на отделните страти се установява, че и в трите области относителните тегла на непрегеглените данни за трите възрастовите групи 11, 13 и 16 години в градовете са система-

тично по-ниски от претеглените и калиброваните данни. Най-високи тегла са получени при обикновеното претегляне и в трите области за градовете за 11 и 13 годишните, в сравнение с непретеглените и калиброваните данни. За учениците на 16 години теглата от простото претегляне заемат междинни стойности между получените относителни дялове от анкетното изследване и калибрираните данни. За посочената, последна възрастова група от градовете, изчислените калибровъчни тегла имат най-високи стойности.

Разгледаните съотношения за стратите в селата са реципрочни. С най-големи относителни дялове за всички страти са тези получени от непретеглените данни. Средно положение по стойност заемат теглата получени в резултат на калибрацията за възрастовите страти 11 и 13 години и в трите области. За 16 годишните в селата калибровъчните тегла са с най-ниска стойност. За тази възрастова група теглата от простото претегляне имат междинни стойности спрямо теглата получени от анкетата и калибрацията, а тези от простото претегляне са най-ниски за възрастовите групи 11 и 13 години.

Направените сравнения показват, че при изпълнение на извадката в градовете не е постигнат проектирания обхват и респективно в селата се получава надценяване. Тези несъответствия, чрез корекция на теглата за всяка от 18-те страти, най-добре са решени чрез получените калибрационни тегла.

Предимството на калибрацията се заключава и в това, че присвоените индивидуални тегла на всяка единица по калибровъчните признаци едновременно дава възможност да се произведат аналитични таблици, без да е необходимо за всяка кростабулация да се извършва претегляне. Това е възможно, тъй като всяка от единиците в извадката е претегляна едновременно по всички използвани за калибрацията признаци.

В следващата таблица са представени данни за разпределението на изследваните лица по признака пол и възраст. За признака „пол” не е налична предварителна официална

статистическа информация и той не е използван при проектиране на извадката. След като са калибрирани данните по този признак, по-нататък произведените таблици ще включват и отразяват и половата структура на изследваната генерална съвкупност.

Кростабулация по пол и възраст

Непретеглени данни	Пол	Възрастови групи			Общо
		11 год.	13 год.	16 год.	
	Момичета	33,6	37,7	28,8	100,0
	Момчета	31,3	29,3	39,5	100,0
	Общо	32,5	33,6	34,0	100,0
Тегла коригирани с неотговорилите (non-response)	Пол	Възрастови групи			Общо
		11 год.	13 год.	16 год.	
	Момичета	33,4	37,3	29,3	100,0
	Момчета	30,5	28,5	41,0	100,0
	Общо	32,0	33,0	35,0	100,0
Калибрирани тегла	Пол	Възрастови групи			Общо
		11 год.	13 год.	16 год.	
	Момичета	36,5	33,7	29,9	100,0
	Момчета	34,2	26,8	38,9	100,0
	Общо	35,4	30,4	34,3	100,0

При сравняването на данните от таблицата може да се установи ефектът от допълнителните процедури по претеглянето на данните. В първата част на таблицата е представено полученото от анкетата разпределение на изследваните лица по тези два признака. Във втория box plot са претеглените резултати, като са коригирани теглата въз основа на информацията за възрастовите групи и са елиминирани неотговорилите. Фактически тук е отчетено разпределението на изследваните

единици в генералната съвкупност само по признака „възраст”. Получените тегла в резултат на калибрацията (в третия box plot) имат по-голямо различие спрямо тези от непретеглените и претеглените данни, но тук по-важното е, че едновременно са отчетени разпределенията на изследваните лица в трите области по пол и възрастови групи. Подобен анализ може да се направи и при получаване на разпределението на изследваните лица по възрастови групи и пол. По този начин ще се отрази половата структура във всяка възрастова група.

При прилагането на процедурите за калибрацията на данни от извадки е предвидено, ако получените резултати след калибрацията не удовлетворяват определени математически критерии, по преценка на изследователя е възможно процедурата да се приложи многократно, докато се избере подходящия модел, т.е. възможно е да се отстранят отдалечените стойности на g-теглата спрямо границите на определен интервал. Последният може да се определя, напр. като се използват медианата оценка и нейното кватилно отклонение, или други критерии. Следователно може да се каже, че калибрацията на данните от извадкови изследвания е един оптимизационен процес за разлика от останалите методи за просто претегляне на данни от извадки.

ИЗВОДИ И ЗАКЛЮЧЕНИЕ

Разглеждането на ЕСИ като научна процедура, чието изпълнение се представя като статистически производствен процес, дава възможност да се установят рисковите фактори и възникващите грешки през целия изследователски процес. Класификацията на възникващите грешки относно получаваната емпирична информация е начало за всеки изследователски екип в търсенето на различни, организационни, методологични и информационно-технологични средства за преодоляване и ограничаване на техните ефекти върху обективността на изследванията и качеството на получаваната информация.

Сред съвременните средства на математическото моделиране, пряко свързани с подобряването на качествените характеристики на данните от извадкови социологически изследвания, са моделите за калибриране на данните. Те са нов начин на мислене относно получаването и обработката на данните от извадкови изследвания и научно обоснован метод за неутрализиране влиянието на различни смущаващи фактори върху изследователския процес.

Трудните условия за набиране на социологическата информация и възможностите, които откриват съвременните комуникационни мрежи ще изискват и поставят калибрацията на данните от извадкови изследвания като противодействие на всички рискове и грешки, които възникват и се допускат в ЕСИ.

В нашата страна все още няма широко разпространен опит и специалисти в областта на приложението на методите за калибрацията и интерпретацията на резултатите от това приложение. Няма и задълбочена научно-практическа и изследователска дейност.

Надяваме се, че разгледаните проблеми в работата ще се възприемат като една малка, начална стъпка в разширяването на кръга за обсъждане на тази изключително важна процедура в приложението на извадкови социологически изследвания.

Въз основа на проведената изследователска работа могат да се направят следните обобщения:

1. Реализацията на социологическите изследвания в съвременни условия става все по-труден изследователски процес по отношение на осигуряването на пълнота на получаваната информация и планираните представителни извадки. Тези проблеми се превръщат в значими от позицията на научната обоснованост на резултатите и критериите на статистическата теория за качеството на оценките, получавани от извадкови изследвания.

2. В практиката на социологическите изследвания у нас не се отделя необходимото внимание и не се изследват пробле-

мите и ефектите, които се получават върху крайните резултати от ЕСИ поради непълноти в реализацията на планираните извадки. Най-често анализът се извършва само върху данни, получени от реализираната част от извадката, без да се държи сметка за настъпилите промени в нейните обем и структура.

3. Прилагането на организационно-технически мерки, осигуряващи провеждането на теренното изследване и на методични решения, като използването на напрегнати извадки, допринася за подобряване качеството на данните. Тяхното прилагане обаче, може да бъде ефективно само когато се осъществява в съответствие с концепциите за двойствената същност на социалното (социологическото) изследване, като научно-изследователска процедура и система от технологично свързани дейности (бизнес/производствен процес). Това означава, че е необходимо разработването и използването на адекватни мерки за всяка фаза и етап от изпълнението на емпиричното социологическо изследване.

4. Все по-голямо значение за развитието на ЕСИ трябва да придобият транслациите и използването на нови и утвърдени в работата на международни и национални институции подходи и изисквания по отношение на организацията на изследванията, критериите за качеството на получаваната информация и нейното разпространение. Направените, макар и ограничени проучвания, доказват необходимостта от проследяване и внедряване на съвременните методи и информационни технологии в практиката на социологическите изследвания. Разгледаните в дисертацията примери показват големите възможности, които разкриват методите за калибрацията на данните за преодоляване и минимизиране на генерираните грешки и неточности, свързани с неизпълнението на представителните извадки.

5. Нараства значението на проблема за необходимостта от експертен преглед на всяко ЕСИ по реда на процедурите по неговото изпълнение. Този подход ще даде основанията и

решенията за прилагането на различни технически и методологически процедури при реализацията на социологическите изследвания, включително и калибрацията на данните от тях. Така екипите ще се предпазят от ненужно (понякога) демонстриране на методологични и технологични процедури, които могат да доведат дори до изкривяването на данните.

6. Все още няма достатъчно натрупан методичен и изследователски опит за абсолютизиране на процедурите по претегляне на резултатите от извадките, а също и процедурите за калибрацията на данните. Безспорни са предимствата на тяхната строга научна обусловеност и ефектите, които те постигат за осигуряването на качествена информация от изследванията, но следва да се имат предвид и рисковете, които те крият при тяхното неправомерно прилагане. Изравняването на оценките от извадките (n) с параметрите на генералната съвкупност (N) не винаги води до добри резултати. Понякога съществува „скрит“, латентен ефект: напр. при някои процедури за претегляне може да повиши значението на систематичната грешка при слаб ефект върху стохастичната грешка.

7. Не е възможно еднозначно да се обоснове необходимостта и да се предложи един оптимизиращ подход за решаване на проблемите с липсващи данни и неизпълнението на планираните извадки чрез прилагането на процедури за претегляне и калибрацията на данните. За всеки конкретен случай това зависи от: 1) предмета, обекта и изследователските задачи на ЕСИ; 2) модела (дизайна) на извадката и регистрацията на данните; 3) наличната информация за параметрите на генералната съвкупност; 4) съществуващите връзки и зависимости между изследваните променливи; 5) целите на количествения анализ на данните от изследването; 6) наличния софтуер и експертен капацитет.

Изпълнението на поставените задачи по темата на дисертационния труд е опит за дефиниране и решаване на изключително важни проблеми относно провеждането и използването на информацията от емпиричните социологически изследвания. Прилагането на различните техники за претегляне, и преди всичко калибрацията на данните, в максимална степен редуцират проблемите в изследванията, предизвикани от непълен обхват на планираните извадки. В тази връзка не бива да се пренебрегва и прилагането на определени методични и организационно-технически мерки и дейности за преодоляване на пропуските в проектираните извадкови съвкупности.

Крайното обобщение е, че проектирането и изпълнението на ЕСИ следва да се разглежда като сложен изследователски и организационно-технологичен процес. За реализацията му е необходимо усвояване и натрупване на научни знания и опит, създаване на добра организация и управление на всички дейности в етапите на изследователския процес. Това означава дефиниране на проблема и разработване на проект за конкретно изследване, прилагане на технологии за регистрация и обработка на данните, получаване на крайни резултати, анализ и интерпретация до вид на качествен информационен продукт.

СПРАВКА ЗА ОСНОВНИТЕ ПРИНОСИ В ДИСЕРТАЦИОННИЯ ТРУД

1. Приложен е нов подход за разглеждане на същността и развитието на концепциите за емпиричните социални изследвания. Разкрива се двустранността на емпиричното социологическо изследване като научно-изследователска процедура и система от технологично свързани дейности (бизнес процес).

2. В методологичен план се разисква проблемът за силните и слаби страни на количествените и качествени методи (подходи) в социологическите изследвания като се стига до заключението, че видът на метода (количествен, качествен) би следвало да не се определя от броя на наблюдаваните респонденти, а от наблюдаваните факти, които качествено се обобщават въз основа на тяхното количествено натрупване. Препоръчва се в интерактивността на „качествено-количествения континуум“ да се търсят общометодологическите основи на методите за изследване в тяхната неделимост и съвместимост при използването им.

3. Адаптирано се предлага в практиката на ЕСИ да се използва моделът “Generic Statistical Business Process Model” (GSBPM) за класифициране на дейностите при осъществяването на социологическите изследвания като инструмент за планиране, проектиране и изпълнение на всяко мащабно социологическо изследване. Открояват се методологичните и технологичните дейности, които следва да се планират и осъществяват и отговорностите на изпълнителите за осигуряването на качествен информационен продукт от социологическите изследвания.

4. Извършена е обоснована систематизация на теоретичните критерии и практически изисквания за оценяване

на качеството на получаваната информацията от емпиричните социологически изследвания, провеждани на извадкова основа.

5. Изследвани и осветени са някои методични и практически проблеми в емпиричните социологически изследвания, свързани с липсата на данни в резултат на неизпълнение на проектираните извадки и нарушаване на тяхната представителност. Групирани и описани са видовете възможни грешки, които възникват при неизпълнение на извадките.

6. Доказана е необходимостта и ефективността от внедряване на съвременните методи и информационни технологии в практиката на социологическите изследвания, които са насочени към решаването на проблемите с липсващи данни и подобряване на статистическите оценки от използваните извадки чрез калибрацията на данните.

7. Представени са емпирични доказателства, на базата на реални изследвания, че калибрацията на данните от извадкови изследвания в максимална степен осигурява предствителността и постигането на статистическите критерии за качеството на статистическите оценки от извадкови изследвания, когато има необхват на извадкови единици.

8. Дисертационният труд дава своя скромнен принос в трансферирането на методология, подходи и практически приложения на калибрацията на данни от извадкови изследвания.

ПУБЛИКАЦИИ ВЪВ ВРЪЗКА С ДИСЕРТАЦИОННИЯ ТРУД

1. Велева, В., (2012). „Състоянието на населението в България: Динамика и тенденции”, сп. Социология и икономика, бр. 1, година II., Пловдив: Висше училище КИА
2. Велева, В., (2013). „Културният обмен като средство за общуване на народите”, сб. Културни паралели, Университетско издателство „Неофит Рилски”-Благоевград
3. Велева, В., (2013). „За съотношението на количествени и качествени методи в социалните изследвания”, УНСС
4. Велева, В., (в съавт.) (2013). „За необходимостта и калибрацията на данните в социалните изследвания”, сп. Статистика, бр. 3-4, (под печат)
5. Велева, В., (2014). „Оценка на резултати от интернет изследвания”, електронно списание „Проблеми на постмодерността”, (под печат)